

Fundamental limits on the rate of bacterial cell division

³ **Nathan M. Belliveau^{†, 1}, Griffin Chure^{†, 2}, Christina L. Hueschen³, Hernan G.**
⁴ **Garcia⁴, Jane Kondev⁵, Daniel S. Fisher⁶, Julie A. Theriot^{1, 7, *}, Rob Phillips^{8, 9, *}**

*For correspondence:

[†]These authors contributed equally to this work

⁵ ¹Department of Biology, University of Washington, Seattle, WA, USA; ²Department of
⁶ Applied Physics, California Institute of Technology, Pasadena, CA, USA; ³Department of
⁷ Chemical Engineering, Stanford University, Stanford, CA, USA; ⁴Department of
⁸ Molecular Cell Biology and Department of Physics, University of California Berkeley,
⁹ Berkeley, CA, USA; ⁵Department of Physics, Brandeis University, Waltham, MA, USA;
¹⁰ ⁶Department of Applied Physics, Stanford University, Stanford, CA, USA; ⁷Allen Institute
¹¹ for Cell Science, Seattle, WA, USA; ⁸Division of Biology and Biological Engineering,
¹² California Institute of Technology, Pasadena, CA, USA; ⁹Department of Physics,
¹³ California Institute of Technology, Pasadena, CA, USA; *Co-corresponding authors.
¹⁴ Address correspondence to phillips@pboc.caltech.edu and jtheriot@uw.edu

¹⁵

¹⁶ **Abstract** Recent years have seen a deluge of experiments dissecting the relationship between
¹⁷ bacterial growth rate, cell size, and protein content, quantifying the abundance of proteins across
¹⁸ growth conditions with unprecedented resolution. However, we still lack a rigorous
¹⁹ understanding of what sets the scale of these quantities and when protein abundances should
²⁰ (or should not) depend on growth rate. Here, we seek to quantitatively understand this
²¹ relationship across a collection of *Escherichia coli* proteomic data sets covering ≈ 4000 proteins
²² and 31 growth conditions. We estimate the basic requirements for steady-state growth by
²³ considering key processes in nutrient transport, energy generation, cell envelope biogenesis, and
²⁴ the central dogma, from which ribosome biogenesis emerges as a primary determinant of
²⁵ growth rate. We conclude by exploring a model of ribosomal regulation as a function of the
²⁶ nutrient supply, revealing a mechanism that ties cell size and growth rate to ribosomal content.

²⁷

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 *E. 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,

42 the study of bacterial growth has undergone a renaissance. Many of the key questions addressed
 43 by the pioneering efforts in the middle of the last century can be revisited by examining them
 44 through the lens of the increasingly refined molecular census that is available for bacteria such
 45 as the microbial workhorse *Escherichia coli*. In this work, we explore an amalgamation of recent
 46 proteomic data sets to explore fundamental limits of bacterial growth.

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

66 Throughout our estimates, we consider a modest growth rate of $\approx 0.5 \text{ hr}^{-1}$ corresponding to
 67 a doubling time of ≈ 5000 seconds, as the the data sets heavily sample this regime. While we for-
 68 mulate point estimates for the complex abundances at this division time, we consider how these
 69 values will vary at other growth rates due to changes in cell size, surface area, and chromosome
 70 copy number (*Taheri-Araghi et al., 2015*). Broadly, we find that for the majority of these estimates,
 71 the protein copy numbers appear well-tuned for the task of cell doubling at a given growth rate. It
 72 emerges that translation, particularly of ribosomal proteins, is the most plausible candidate for a
 73 molecular bottleneck. We reach this conclusion by considering that translation is 1) a rate limiting
 74 step for the *fastest* bacterial division, and 2) a major determinant of bacterial growth across the
 75 nutrient conditions we have considered under steady state, exponential growth. This enables us
 76 to suggest that the long-observed correlation between growth rate and cell size (*Schaechter et al.,*
77 1958; Si et al., 2017) can be simply attributed to the increased absolute number of ribosomes per
 78 cell under conditions supporting extremely rapid growth, a hypothesis which we formally mathe-
 79 matize and explore.

124 Uptake of Nutrients

125 We begin our series of estimates by considering the critical transport processes diagrammed in
 126 **Figure 1(A)**. In order to build new cellular mass, the molecular and elemental building blocks must
 127 be scavenged from the environment in different forms. Carbon, for example, is acquired via the
 128 transport of carbohydrates and sugar alcohols with some carbon sources receiving preferential
 129 treatment in their consumption (*Monod, 1947*). Phosphorus, sulfur, and nitrogen, on the other
 130 hand, are harvested primarily in the forms of inorganic salts, namely phosphate, sulfate, and am-
 131 monia (*Jun et al., 2018; Assentoft et al., 2016; Stasi et al., 2019; Antonenko et al., 1997; Rosenberg*
132 et al., 1977; Willsky et al., 1973). All of these compounds have different permeabilities across the
 133 cell membrane *Phillips (2018)* and most require some energetic investment either via ATP hydrol-
 134 ysis or through the proton electrochemical gradient to bring the material across the hydrophobic
 135 cell membrane. Given the diversity of biological transport mechanisms and the vast number of

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

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.

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. 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.

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.

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.

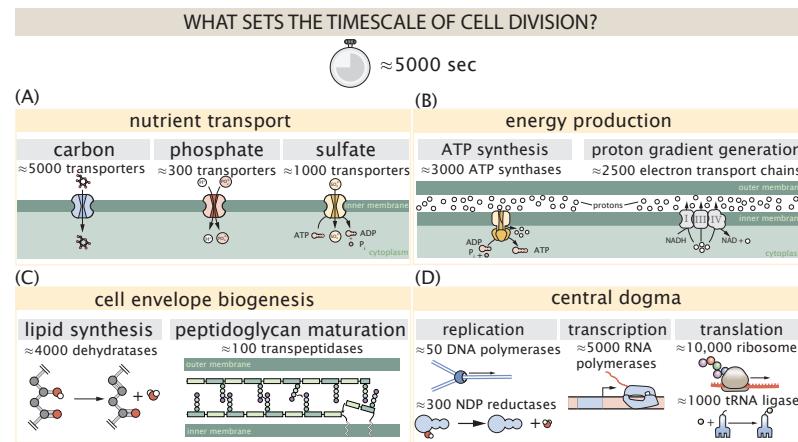


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP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

136 inputs needed to build a cell, we begin by considering transport of some of the most important
 137 cellular ingredients: carbon, nitrogen, oxygen, hydrogen, phosphorus, and sulfur.

138 The elemental composition of *E. coli* has received much quantitative attention over the past
 139 half century (Neidhardt et al., 1991; Taymaz-Nikerel et al., 2010; Heldal et al., 1985; Bauer and Ziv,
 140 1976), providing us with a starting point for estimating the copy numbers of various transporters.
 141 While there is some variability in the exact elemental percentages (with different uncertainties),
 142 we can estimate that the dry mass of a typical *E. coli* cell is $\approx 45\%$ carbon (BioNumber ID: 100649,
 143 see Box 1), $\approx 15\%$ nitrogen (BNID: 106666), $\approx 3\%$ phosphorus (BNID: 100653), and 1% sulfur (BNID:
 144 100655).

145 **Nitrogen Transport**

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

160 **Carbon Transport**

161 We begin with the most abundant element in *E. coli* by mass, carbon. Using $\approx 0.3 \text{ pg}$ as the typical
 162 *E. coli* dry mass (BNID: 103904), we estimate that $\sim 10^{10}$ carbon atoms must be brought into the
 163 cell in order to double all of the carbon-containing molecules (Figure 2(A, top)). Typical laboratory
 164 growth conditions provide carbon as a single class of sugar such as glucose, galactose, or xylose

165 to name a few. *E. coli* has evolved myriad mechanisms by which these sugars can be transported
 166 across the cell membrane. One such mechanism of transport is via the PTS system which is a
 167 highly modular system capable of transporting a diverse range of sugars (*Escalante et al., 2012*).
 168 The glucose-specific component of this system transports ≈ 200 glucose molecules per second
 169 per transporter (BNID: 114686). Making the assumption that this is a typical sugar transport rate,
 170 coupled with the need to transport $\sim 10^{10}$ carbon atoms, we arrive at the conclusion that on the
 171 order of 1000 transporters must be expressed in order to bring in enough carbon atoms to divide in
 172 5000 s, diagrammed in the top panel of *Figure 2(A)*. This estimate, along with the observed average
 173 number of the PTS system carbohydrate transporters present in the proteomic data, is shown in
 174 *Figure 2(A)*. While we estimate 1500 transporters are needed with a 5000 s division time, we can
 175 abstract this calculation to consider any particular growth rate given knowledge of the cell density
 176 and volume as a function of growth rate and direct the reader to the Appendix Extending Estimates
 177 to a Continuum of Growth Rates for more information. As revealed in *Figure 2(A)*, experimental
 178 measurements exceed the estimate by several fold, suggesting that transport of carbon into the
 179 cell is not rate limiting for cell division. Abstracting this point estimate at 5000 s to a continuum of
 180 growth rates (grey line in *Figure 2(A)*) reveals an excess of transporters even at faster growth rates.

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

192 Many metabolic operons are regulated with dual-input logic gates that are only expressed when
 193 glucose concentrations are low (mediated by cyclic-AMP receptor protein CRP) and the concentra-
 194 tion of other carbon sources are elevated (*Gama-Castro et al., 2016; Zhang et al., 2014b*). A famed
 195 example of such dual-input regulatory logic is in the regulation of the *lac* operon which is only
 196 activated in the absence of glucose and the presence of allolactose, an intermediate in lactose
 197 metabolism (*Jacob and Monod, 1961*), though we now know of many other such examples (*Ire-
 198 land et al., 2020; Gama-Castro et al., 2016; Belliveau et al., 2018*). Several examples are shown
 199 in *Figure 2(B)*. Points colored in red (labeled by red text-boxes) correspond to growth conditions
 200 in which the specific carbon source (glycerol, xylose, or fructose) is present. The grey lines in *Fig-
 201 ure 2(B)* show the estimated number of transporters needed at each growth rate to satisfy the
 202 cellular carbon requirement, adjusted for the specific carbon source. These plots show that, in
 203 the absence of the particular carbon source, expression of the transporters is maintained on the
 204 order of $\sim 10^2$ per cell. However, when the transport substrate is present, expression is induced
 205 and the transporters become highly-expressed. The low but non-zero abundances for many of
 206 these alternative across growth conditions may reflect the specific regulatory logic, requiring the
 207 cell to transport some minimal amount of an alternative carbon source in order to induce expres-
 208 sion of these alternative carbon-source systems. Together, this generic estimation and the specific
 209 examples of induced expression suggest that transport of carbon across the cell membrane, while
 210 critical for growth, is not the rate-limiting step of cell division.

211 **Phosphorus and Sulfur Transport**

212 We now turn our attention towards other essential elements, namely phosphorus and sulfur. Phos-
 213 phorus is critical to the cellular energy economy in the form of high-energy phosphodiester bonds
 214 making up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-translational

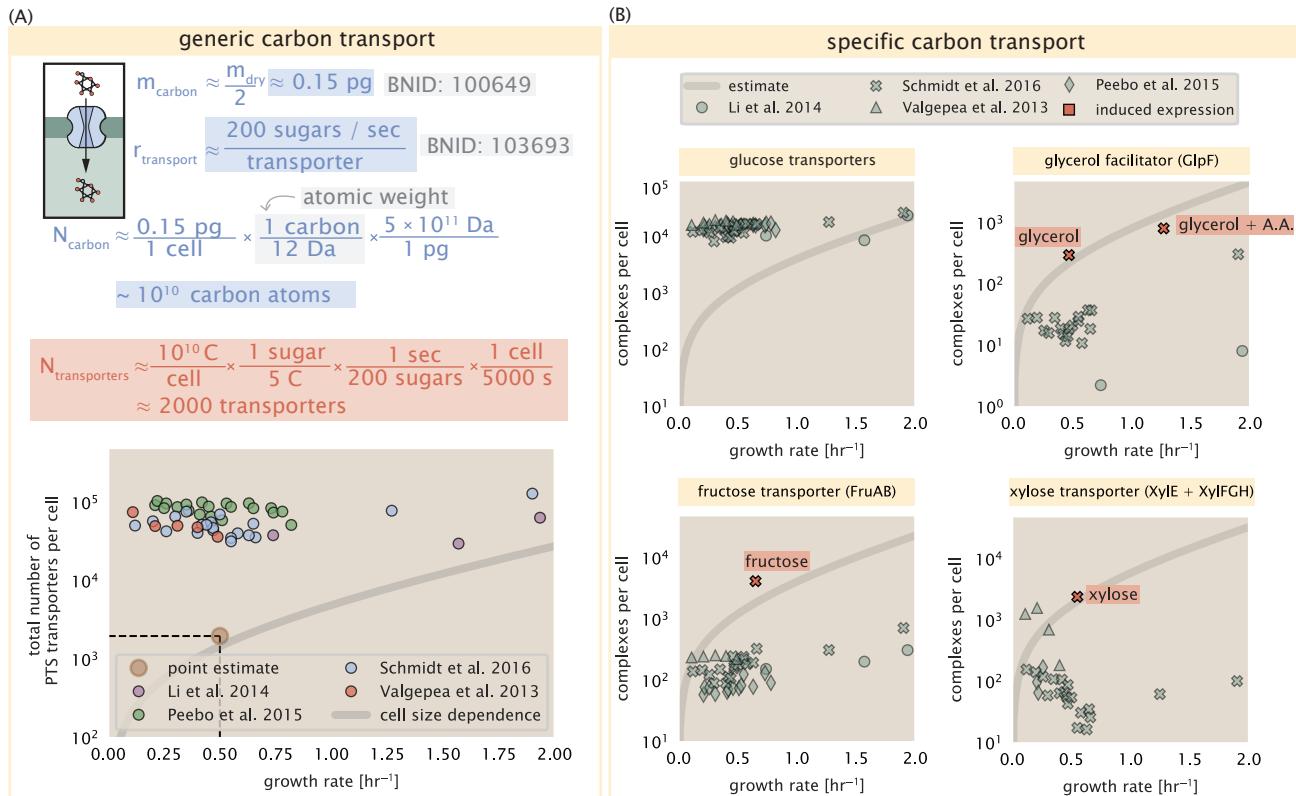


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules 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 PtsL, 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.

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

Satisfied that there are a sufficient number of phosphate transporters present in the cell, we now turn to sulfur transport as another potentially rate limiting process. Similar to phosphate, sulfate is highly-charged and not particularly membrane permeable, requiring active transport. While there exists a H⁺/sulfate symporter in *E. coli*, it is in relatively low abundance and is not well characterized (Zhang *et al.*, 2014a). Sulfate is predominantly acquired via the ATP-dependent ABC transporter CysUWA system which also plays an important role in selenium transport (Sekowska *et al.*, 2000; Sirko *et al.*, 1995). While specific kinetic details of this transport system are not readily available, generic ATP transport systems in prokaryotes transport on the order of 1 to 10 molecules per second (BNID: 109035). Combining this generic transport rate, measurement of sulfur comprising 1% of dry mass, and a 5000 second division time yields an estimate of ≈ 1000 CysUWA complexes per cell (**Figure 3(B)**). Once again, this estimate is in notable agreement with proteomic data sets, suggesting that there are sufficient transporters present to acquire the necessary sulfur. In a similar spirit of our estimate of phosphorus transport, we emphasize that this is likely an overestimate of the number of necessary transporters as we have neglected other sulfur scavenging systems that are in lower abundance.

250 **Limits on Transporter Expression**

251 So which, if any, of these processes may be rate limiting for growth? As suggested by **Figure 2(B)**,
 252 induced expression can lead to an order-of-magnitude (or more) increase in the amount of trans-
 253 porters needed to facilitate transport. Thus, if acquisition of nutrients was the limiting state in
 254 cell division, could expression simply be increased to accommodate faster growth? A way to ap-
 255 proach this question is to compute the amount of space in the bacterial membrane that could be
 256 occupied by nutrient transporters. Considering a rule-of-thumb for the surface area of *E. coli* of
 257 about 5 μm^2 (BNID: 101792), we expect an areal density for 1000 transporters to be approximately
 258 200 transporters/ μm^2 . For a typical transporter occupying about 50 nm^2/dimer , this amounts to
 259 about only 1 percent of the total inner membrane (Szenk *et al.*, 2017). In addition, bacterial cell
 260 membranes typically have densities of 10^5 proteins/ μm^2 (Phillips, 2018), implying that the cell could
 261 accommodate more transporters of a variety of species if it were rate limiting. As we will see in the
 262 next section, however, occupancy of the membrane can impose other limits on the rate of energy
 263 production.

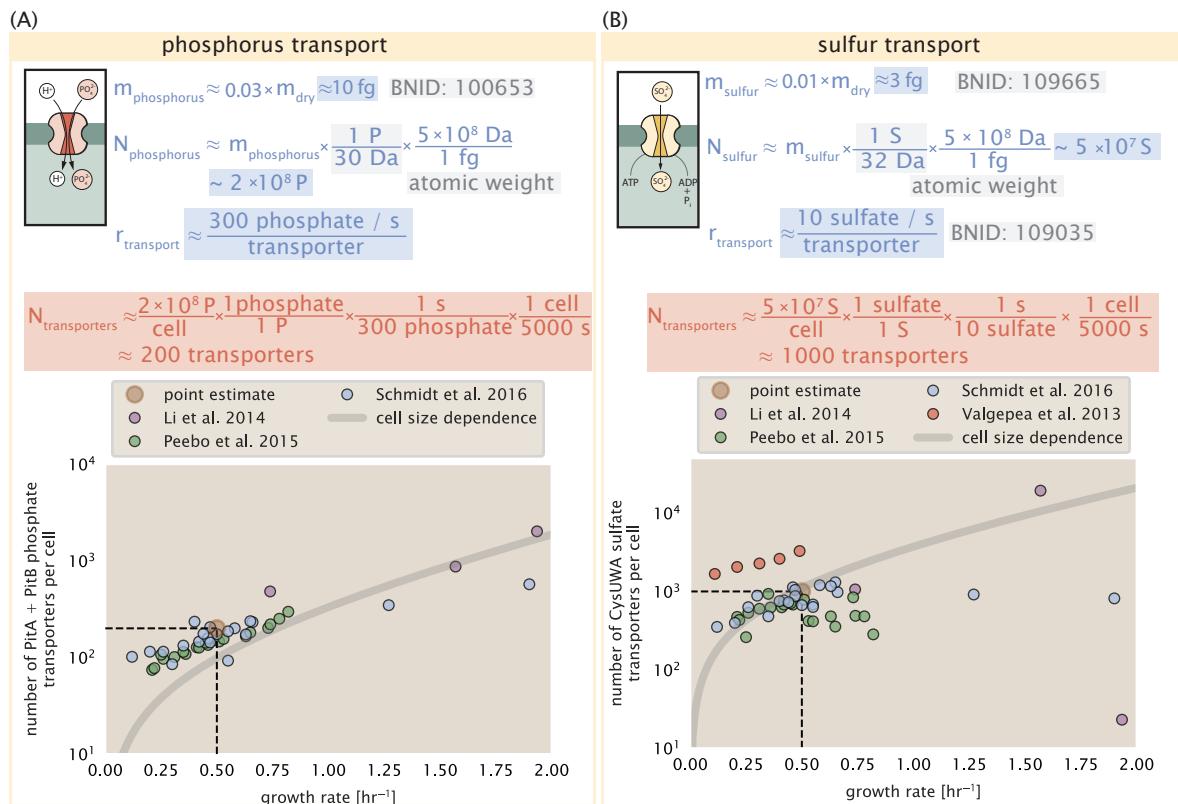


Figure 3. Estimates and measurements of phosphate and sulfate transport systems as a function of growth rate. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (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 [CysA]₂[CysU][CysW][Sbp/CysP]. Grey line in (A) and (B) represents the estimated number of transporters per cell at a continuum of growth rates.

264 Energy Production

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

271 ATP Synthesis

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

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

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

307 Generating the Proton Electrochemical Gradient

308 In order to produce ATP, the F₁-F₀ ATP synthase itself must consume energy. Rather than burn-
 309 ing through its own product (and violating thermodynamics), this intricate macromolecular ma-
 310 chine has evolved to exploit the electrochemical potential established across the inner membrane
 311 through cellular respiration. This electrochemical gradient is manifest by the pumping of protons
 312 into the intermembrane space via the electron transport chains as they reduce NADH. In *E. coli*,

313 this potential difference is ≈ -200 mV (BNID: 102120). A simple estimate of the inner membrane
 314 as a capacitor with a working voltage of -200 mV reveals that $\approx 2 \times 10^4$ protons must be present in
 315 the intermembrane space.

316 However, the constant rotation of the ATP synthases would rapidly abolish this potential difference
 317 if it were not being actively maintained. To undergo a complete rotation (and produce a single
 318 ATP), the F₁-F₀ ATP synthase must shuttle ≈ 4 protons across the membrane into the cytosol (BNID:
 319 103390). With ≈ 3000 ATP synthases each generating 300 ATP per second, the 2×10^4 protons estab-
 320 lishing the 200 mV potential would be consumed in only a few milliseconds. This brings us to our
 321 next estimate: how many electron transport complexes are needed to support the consumption
 322 rate of the ATP synthases?

323 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of
 324 intense biochemical and biophysical study over the past half century (*Ingledew and Poole, 1984;*
 325 *Khademian and Imlay, 2017; Cox et al., 1970; Henkel et al., 2014*). A recent work (*Szenk et al.,*
 326 *2017*) examined the respiratory capacity of the *E. coli* electron transport complexes using structural
 327 and biochemical data, revealing that each electron transport chain rapidly pumps protons into
 328 the intermembrane space at a rate of ≈ 1500 protons per second (BNID: 114704; 114687). Using
 329 our estimate of the number of ATP synthases required per cell [*Figure 4(A)*], coupled with these
 330 recent measurements, we estimate that ≈ 1000 electron transport complexes would be necessary
 331 to facilitate the $\sim 5 \times 10^6$ protons per second diet of the cellular ATP synthases. This estimate (along
 332 with a generalization to the entire range of observed growth rates) is in agreement with the number
 333 of complexes identified in the proteomic datasets (plot in *Figure 4(B)*). This suggests that every ATP
 334 synthase must be accompanied by ≈ 1 functional electron transport chain. Again, if this were a rate
 335 limiting process for bacterial growth, one must conclude that it is not possible for the cell to simply
 336 increase the production of both the number of electron transport chain complexes as well as ATP
 337 synthases. As both of these components only function bound to the inner membrane, we now
 338 turn our attention towards the available space in the membrane as well as surface-area-to-volume
 339 constraints.

340 Energy Production in a Crowded Membrane.

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

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

354 In our estimate of ATP production above we found that a cell demands about 5×10^9 ATP per cell
 355 cycle or 10^6 ATP/s. With a cell volume of roughly 1 fL, this corresponds to about 2×10^{10} ATP per fL
 356 of cell volume, in line with previous estimates (*Stouthamer and Bettenhausen, 1977; Szenk et al.,*
 357 *2017*). In *Figure 5* (A) we plot this ATP demand as a function of the S/V ratio in green, where we have
 358 considered a range of cell shapes from spherical to rod-shaped with an aspect ratio (length/width)
 359 equal to 4 (See appendix for calculations of cell volume and surface area). In order to consider the
 360 maximum power that could be produced, we consider the amount of ATP that can generated by a
 361 membrane filled with ATP synthase and electron transport complexes, which provides a maximal
 362 production of about 3 ATP / (nm²·s) (*Szenk et al., 2017*). This is shown in blue in *Figure 5*(A), which

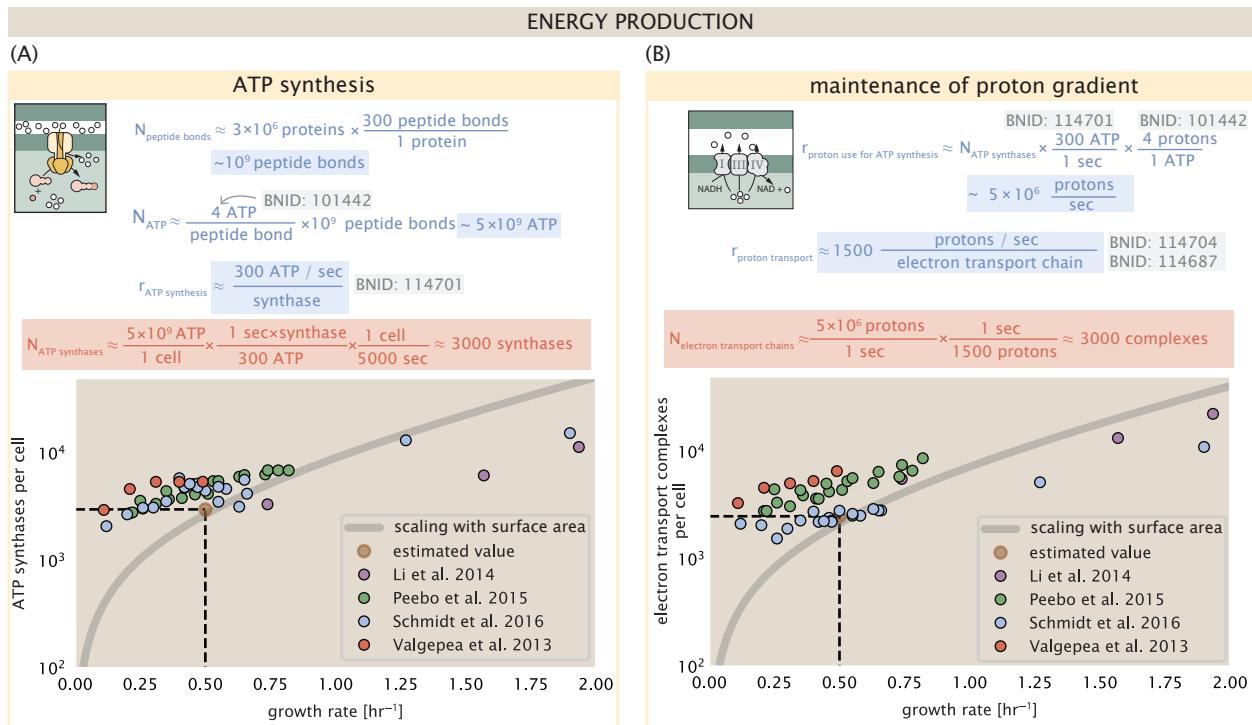


Figure 4. The abundance of F₁-F₀ ATP synthases and electron transport chain complexes as a function of growth rate. (A) Estimate of the number of F₁-F₀ ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes. Points in plot correspond to the mean number of complete F₁-F₀ ATP synthase complexes that can be formed given proteomic measurements and the subunit stoichiometry [AtpE]₁₀[AtpF]₂[AtpB][AtpC][AtpH][AtpA]₃[AtpG][AtpD]₃. (B) Estimate of the number of electron transport chain complexes needed to maintain a membrane potential of -200 mV given estimate of number of F₁-F₀ ATP synthases from (A). Points in plot correspond to the average number of complexes identified as being involved in aerobic respiration by the Gene Ontology identifier GO:0019646 that could be formed given proteomic observations. These complexes include cytochromes *bd1* ([CydA][CydB][CydX][CydH]), *bdII* ([AppC][AppB]), *bo3*,([CyoD][CyoA][CyoB][CyoC]) and NADH:quinone 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.

363 shows that at least for the growth rates observed, the energy demand is roughly an order of mag-
 364 nitude less. Interestingly, *Szenk et al.* (2017) also found that ATP production by respiration is less
 365 efficient than by fermentation per membrane area occupied due to the additional proteins of the
 366 electron transport chain. This suggests that, even under anaerobic growth, there will be sufficient
 367 membrane space for ATP production in general.

368 While the analysis in *Figure 5(A)* serves to highlight the diminishing capacity to provide re-
 369 sources to grow if the cell increases in size (and its S/V decreases), maximum energy production
 370 represents a somewhat unachievable limit since the inner membrane must also include other pro-
 371 teins including those required for lipid and membrane synthesis. We used the proteomic data to
 372 look at the distribution of proteins on the inner membrane, relying on the Gene Ontology (GO)
 373 annotations (*Ashburner et al.*, 2000; *The Gene Ontology Consortium*, 2018) to identify all proteins
 374 embedded or peripheral to the inner membrane (GO term: 0005886). Those associated but not
 375 membrane-bound include proteins like MreB and FtsZ, that traverse the inner membrane by tread-
 376 milling and must nonetheless be considered as a vital component occupying space on the mem-
 377 brane. In *Figure 5 (B)*, we find that the total protein mass per μm^2 is surprisingly constant across
 378 growth rates. Interestingly, when we consider the distribution of proteins grouped by their Clusters
 379 of Orthologous Groups (COG) (*Tatusov et al.*, 2000), the relative abundance for those in metabolism
 380 (including ATP synthesis via respiration) is also relatively constant across growth rates, suggesting
 381 that many other membrane associated proteins also increase in similar proportions to proteins
 382 devoted to energy production *Figure 5 (C)*.

383 Synthesis of the Cell Envelope

384 The subjects of our estimates thus far have been localized to the periphery of the cell, embedded
 385 within the hydrophobic lipid bilayer of the inner membrane. As outlined in *Figure 5*, cells could in
 386 principle increase the expression of the membrane-bound ATP synthases and electron transport
 387 chains to support a larger energy budget across a wide range of cell volumes and membrane sur-
 388 face areas. This ability, however, is contingent on the ability of the cell to expand the surface area
 389 of the cell by synthesizing new lipids and peptidoglycan for the cell wall. In this next class of esti-
 390 mates, we will turn our focus to these processes and consider the copy numbers of the relevant
 391 enzymes.

392 Lipid Synthesis

393 The cell envelopes of gram negative bacteria (such as *E. coli*) are composed of inner and outer
 394 phospholipid bilayer membranes separated by a $\approx 10 \text{ nm}$ periplasmic space (BNID: 100016, *Milo*
 395 *et al.* (2010)). As mentioned in our discussion of the surface area to volume constraints on energy
 396 production, *E. coli* is a rod-shaped bacterium with a 4:1 length-to-width aspect ratio. At modest
 397 growth rates, such as our stopwatch of 5000 s, the total cell surface area is $\approx 5 \mu\text{m}^2$ (BNID: 101792,
 398 *Milo et al.* (2010)). As there are two membranes, each of which composed of two lipid leaflets, the
 399 total membrane area is $\approx 20 \mu\text{m}^2$, a remarkable value compared to the $\approx 2 \mu\text{m}$ length of the cell.

400 While this represents the total area of the membrane, this does not mean that it is composed
 401 entirely of lipid molecules. Rather, the dense packing of the membrane with proteins means that
 402 only $\approx 40\%$ of the membrane area is occupied by lipids (BNID: 100078). Using a rule-of-thumb of
 403 0.5 nm^2 as the surface area of the typical lipid (BNID: 106993), we arrive at an estimate of $\sim 2 \times 10^7$
 404 lipids per cell, an estimate in close agreement with experimental measurements (BNID: 100071,
 405 102996).

406 The membranes of *E. coli* are composed of a variety of different lipids, each of which are unique
 407 in their structures and biosynthetic pathways (*Sohlenkamp and Geiger*, 2016). With such diversity
 408 in biosynthesis, it becomes difficult to identify which step(s) may be the rate-limiting, an objective
 409 further complicated by the sparsity of *in vivo* kinetic data. Recently, a combination of stochastic
 410 kinetic modeling (*Ruppe and Fox*, 2018) and *in vitro* kinetic measurements (*Ranganathan et al.*,
 411 2012; *Yu et al.*, 2011) have revealed remarkably slow steps in the fatty acid synthesis pathways

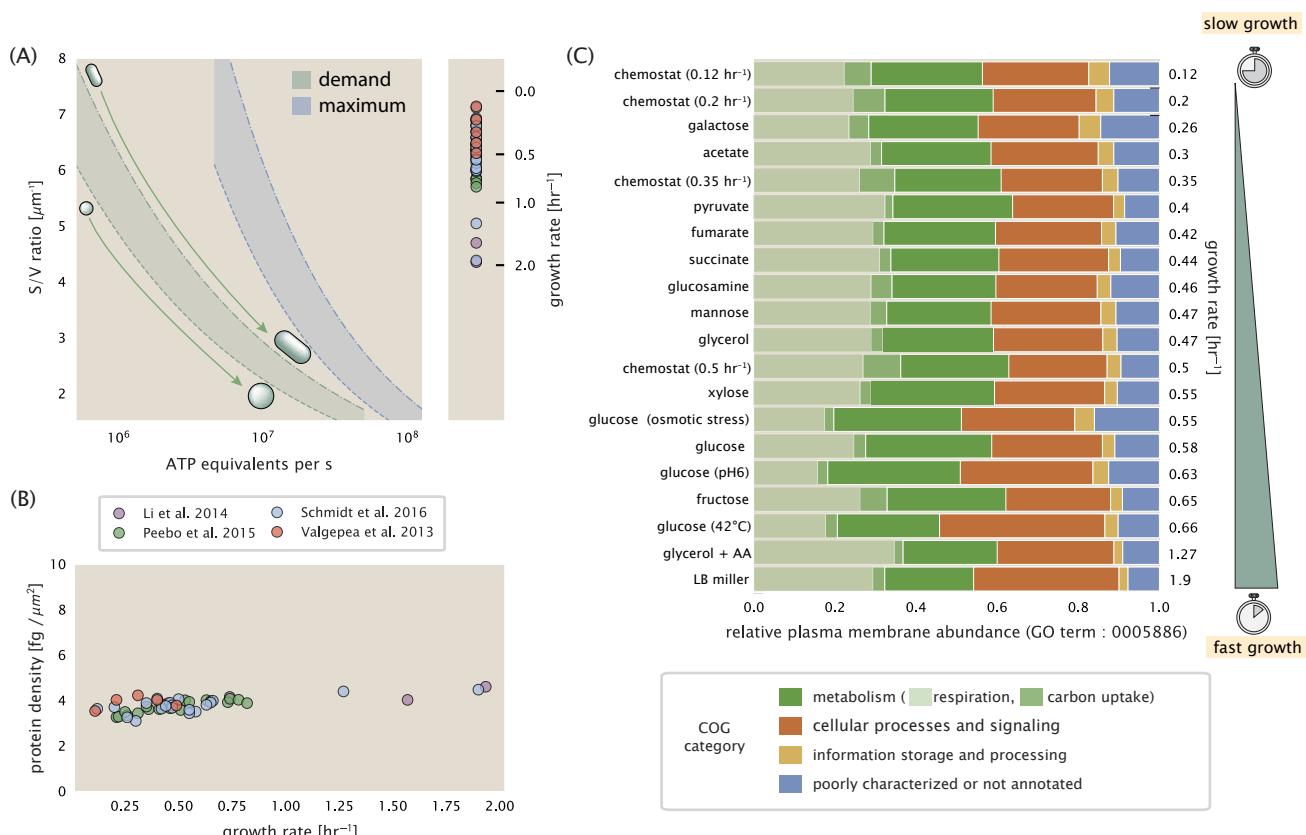


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A)

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 bo_3 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)).

412 which may serve as the rate limiting reactions. One such step is the removal of hydroxyl groups
 413 from the fatty-acid chain by ACP dehydratase that leads to the formation of carbon-carbon double
 414 bonds. This reaction, catalyzed by proteins FabZ and FabA in *E. coli* (*Yu et al., 2011*), have been
 415 estimated to have kinetic turnover rates of ≈ 1 dehydration per second per enzyme (*Ruppe and*
 416 *Fox, 2018*). Combined with this rate, our previous estimates for the number of lipids to be formed,
 417 and a 5000 second division yields an estimate that the cell requires ≈ 4000 ACP dehydratases. This
 418 is in reasonable agreement with the experimentally observed copy numbers of FabZ and FabA
 419 (*Figure 6(A)*). Furthermore, we can extend this estimate to account for the change in membrane
 420 surface area as a function of the growth rate (grey line in *Figure 6(A)*), which captures the observed
 421 growth rate dependent expression of these two enzymes.

422 Despite the slow catalytic rate of FabZ and FabA, we argue that the generation of fatty acids
 423 is not a bottleneck in cell division and is not the key process responsible for setting the bacterial
 424 growth rate. Experimental evidence has shown that the rate of fatty-acid synthesis can be dras-
 425 tically increased *in vitro* by increasing the concentration of FabZ *Yu et al. (2011)*. Stochastic simu-
 426 lations of the complete fatty acid synthesis pathway of *E. coli* further supports this experimental
 427 observation *Ruppe and Fox (2018)*. Thus, if this step was the determining factor in cell division,
 428 increasing growth rate could be as simple as increasing the number of ACP dehydratases per cell.
 429 With a proteome size of $\approx 3 \times 10^6$ proteins, a hypothetical increase in expression from 4000 to 40,000
 430 ACP dehydratases would result in a $\approx 1\%$ increase in the size of the proteome. As many other pro-
 431 teins are in much larger abundance than 4000 per cell (as we will see in the coming sections), it is
 432 unlikely that expression of ACP dehydratases couldn't be increased to facilitate faster growth.

433 Peptidoglycan Synthesis

434 While variation in cell size can vary substantially across growth conditions, bacterial cells demon-
 435 strate exquisite control over their cell shape. This is primarily due to the cell wall, a stiff meshwork
 436 of polymerized disaccharides interspersed with short peptide crosslinks termed the peptidogly-
 437 can. The cell wall is also a vital structural component that counteracts turgor pressure. In *E. coli*, this
 438 enormous peptidoglycan molecule is a few nanometers thick and resides within the periplasmic
 439 space between the inner and outer membrane. The formation of the peptidoglycan is an intri-
 440 cate process, involving the bacterial actin homolog MreB (*Shi et al., 2018*) along with a variety of
 441 membrane-bound and periplasmic enzymes (*Morgenstein et al., 2015*). The coordinated action
 442 of these components result in a highly-robust polymerized meshwork that maintains cell shape
 443 even in the face of large-scale perturbations and can restore rod-shaped morphology even after
 444 digestion of the peptidoglycan (*Harris and Theriot, 2018; Shi et al., 2018*).

445 In glucose-supported steady-state growth, the peptidoglycan alone comprises $\approx 3\%$ of the cellular
 446 dry mass (BNID: 101936), making it the most massive molecule in *E. coli*. The polymerized unit
 447 of the peptidoglycan is a N-acetylglucosamine and N-acetylmuramic acid disaccharide, of which
 448 the former is functionalized with a short pentapeptide. With a mass of ≈ 1000 Da, this unit, which
 449 we refer to as a murein monomer, is polymerized to form long strands in the periplasm which are
 450 then attached to each other via their peptide linkers. Using the aforementioned measurement that
 451 $\approx 3\%$ of the dry mass is peptidoglycan, it can be estimated that the peptidoglycan is composed of
 452 $\sim 5 \times 10^6$ murein monomers.

453 During growth, peptidoglycan is constantly being broken down to allow insertion of new murein
 454 monomers and cellular expansion. In order to maintain structural integrity these monomers must
 455 be crosslinked into the expanding cell wall, potentially limiting how quickly new material can be
 456 added and we consider this process as a possible rate-limiting step. In principle, each one of these
 457 murein monomers can be crosslinked to another glycan strand via the pentapeptide. In some
 458 species, such as in gram-positive bacterium *Staphylococcus aureus*, the extent of crosslinking can
 459 be large with $> 90\%$ of pentapeptides forming a connection between glycan strands. In *E. coli*, how-
 460 ever, a much smaller proportion ($\approx 20\%$) of the peptides are crosslinked, resulting in a weaker and
 461 more porous cell wall *Vollmer et al. (2008); Rogers et al. (1980)*. The formation of these crosslinks

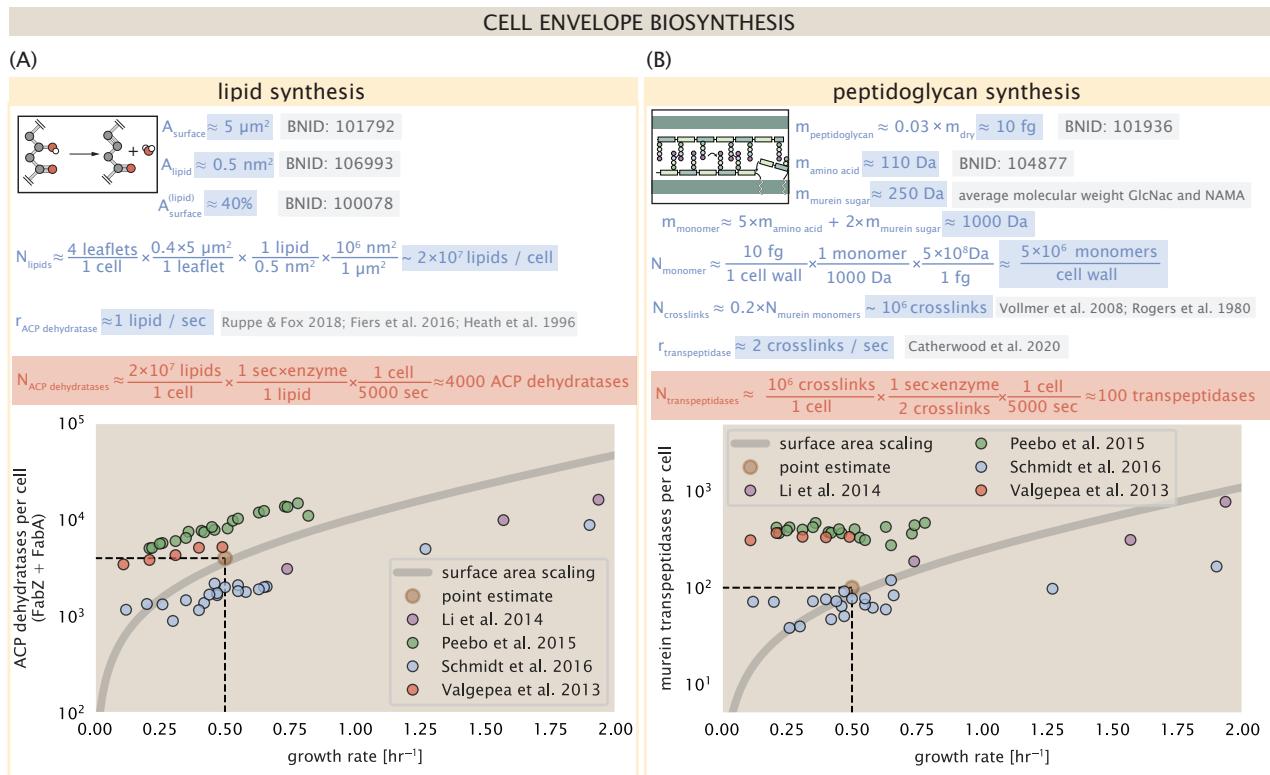


Figure 6. Estimation of the key components involved in cell envelope biosynthesis. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred 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 $[FabZ]_2$ and $[FabA]_2$. (B) An estimate for the number of peptidoglycan transpeptidases needed to complete maturation of the peptidoglycan. The mass of the murein monomer was estimated by approximating each amino acid in the pentapeptide chain as having a mass of 110 Da and each sugar in the disaccharide having a mass of ≈ 250 Da. The *in vivo* rate of transpeptidation in *E. coli* was taken from recent analysis by *Catherwood et al. (2020)*. The bottom panel shows experimental measurements of the transpeptidase complexes in *E. coli* following the stoichiometries $[MrcA]_2$, $[MrcB]_2$, $[MrdA]_1$, and $[MrdB]_1$. 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.

462 primarily occur during the polymerization of the murein monomers and is facilitated by a fam-
 463 ily of enzymes called transpeptidases. In *E. coli*, there are four primary transpeptidases that are
 464 involved in lateral and longitudinal extension of the peptidoglycan. These transpeptidases have
 465 only recently been quantitatively characterized *in vivo* via liquid chromatography mass spectrom-
 466 etry (*Catherwood et al., 2020*), which revealed a kinetic turnover rate of ≈ 2 crosslinking reactions
 467 formed per second per enzyme.

468 Pulling these measurements together permits us to make an estimate that on the order of ≈ 100
 469 transpeptidases are needed for complete maturation of the peptidoglycan, given a division time
 470 of ≈ 5000 seconds, a value that is closely aligned with the experimental observations (*Figure 6(B)*).
 471 Expanding this estimate to account for the changing volume of the peptidoglycan as a function of
 472 growth rates (grey line in *Figure 6(B)*) also qualitatively captures the observed dependence in the
 473 data, though systematic disagreements between the different data sets makes the comparison
 474 more difficult.

475 Much as in the case of fatty acid synthesis, we find it unlikely that the formation of peptidogly-
 476 can is a process which defines the rate of bacterial cell division. The estimate we have presented
 477 considered only the transpeptidase enzymes that are involved lateral and longitudinal elongation
 478 of the peptidoglycan (proteins MrdA, MrdB, MrcA, and MrcB). This neglects the presence of other
 479 transpeptidases that are present in the periplasm and also involved in remodeling and maturation
 480 of the peptidoglycan. It is therefore possible that if this was setting the speed limit for cell divi-
 481 sion, the simple expression of more transpeptidases may be sufficient to maintain the structural
 482 integrity of the cell wall.

483 Function of the Central Dogma

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

489 DNA

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

497 dNTP synthesis

498 We begin our exploration DNA replication by examining the production of the deoxyribonucleotide
 499 triphosphates (dNTPs). The four major dNTPs (dATP, dTTP, dCTP, and dGTP) are synthesized *de*
 500 *novo* in separate pathways, requiring different building blocks. However, a critical step present in
 501 all dNTP synthesis pathways is the conversion from ribonucleotide to deoxyribonucleotide via the
 502 removal of the 3' hydroxyl group of the ribose ring (*Rudd et al., 2016*). This reaction is mediated
 503 by a class of enzymes termed ribonucleotide reductases, of which *E. coli* possesses two aerobically
 504 active complexes (termed I and II) and a single anaerobically active enzyme. Due to their peculiar
 505 formation of a radical intermediate, these enzymes have received much biochemical, kinetic, and
 506 structural characterization. One such work (*Ge et al., 2003*) performed a detailed *in vitro* measure-
 507 ment of the steady-state kinetic rates of these complexes, revealing a turnover rate of ≈ 10 dNTP
 508 per second.

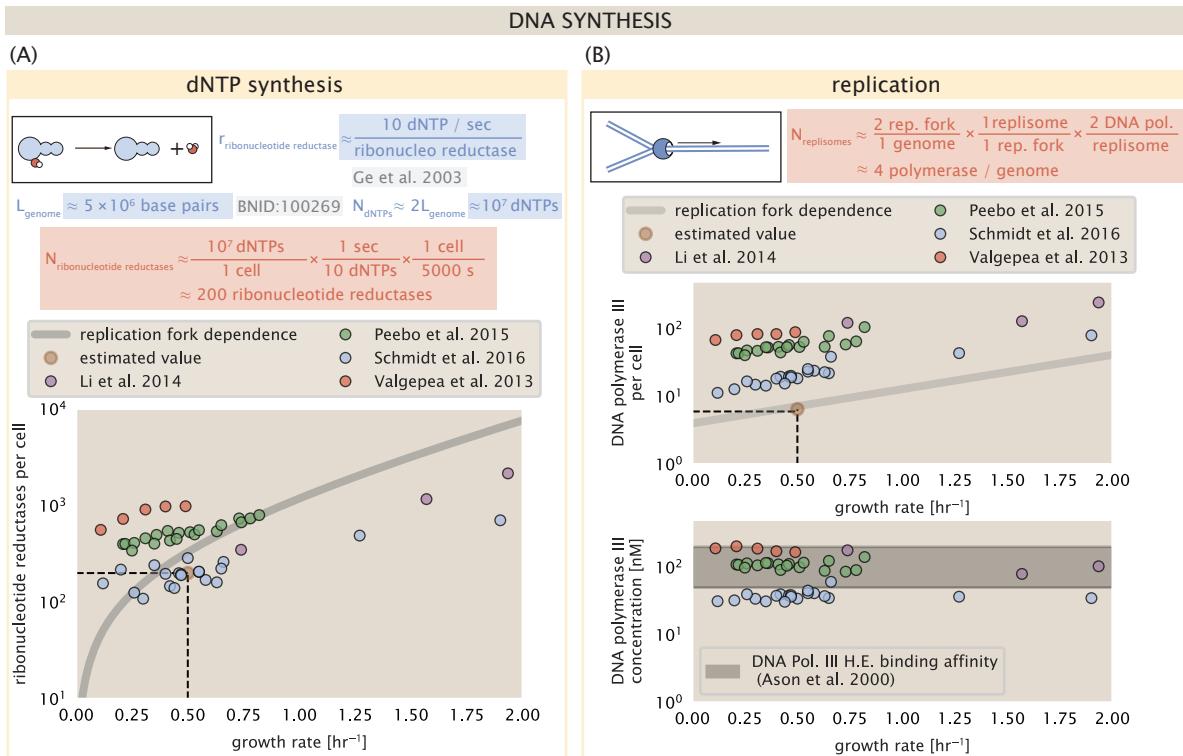


Figure 7. Complex abundance estimates for dNTP synthesis and DNA replication. (A) Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ($[NrdA]_2[NrdB]_2$) and ribonucleotide reductase II ($[NrdE]_2[NrdF]_2$) 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 ($[DnaE]_3[DnaQ]_3[Hole]_3[DnaX]_5[HolB]_2[HolA]_2[DnaN]_4[HolC]_4[Hold]_4$) 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.

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

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

524 DNA Replication

We now turn our focus to the integration of these dNTP building blocks into the replicated chromosome strand via the DNA polymerase. Replication is initiated at a single region of the chromosome

527 termed the *oriC* locus at which a pair of DNA polymerases bind and begin their high-fidelity repli-
 528 cation of the genome in opposite directions. Assuming equivalence between the two replication
 529 forks, this means that the two DNA polymerase complexes (termed replisomes) meet at the mid-
 530 way point of the circular chromosome termed the *ter* locus. The kinetics of the five types of DNA
 531 polymerases (I – V) have been intensely studied, revealing that DNA polymerase III performs the
 532 high fidelity processive replication of the genome with the other "accessory" polymerases playing
 533 auxiliary roles (*Fijalkowska et al., 2012*). *In vitro* measurements have shown that DNA Polymerase
 534 III copies DNA at a rate of \approx 600 nucleotides per second (BNID: 104120). Therefore, to replicate
 535 a single chromosome, two replisomes (containing two DNA polymerase III each) moving at their
 536 maximal rate would copy the entire genome in \approx 4000 s. Thus, with a division time of 5000 s (our
 537 "typical" growth rate for the purposes of this work), there is sufficient time for a pair of replisomes
 538 complexes to replicate the entire genome. However, this estimate implies that 4000 s would be
 539 the upper-limit time scale for bacterial division which is at odds with the familiar 20 minute (1200
 540 s) doubling time of *E. coli* in rich medium.

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

561 RNA Synthesis

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

571 rRNA

572 We begin with an estimation of the number of RNA polymerases needed to synthesize the rRNA
 573 that serve as catalytic and structural elements of the ribosome. Each ribosome contains three rRNA
 574 molecules of lengths 120, 1542, and 2904 nucleotides (BNID: 108093), meaning each ribosome
 575 contains \approx 4500 nucleotides. As the *E. coli* RNA polymerase transcribes DNA to RNA at a rate of \approx

576 40 nucleotides per second (BNID: 101904), it takes a single RNA polymerase \approx 100 s to synthesize
 577 the RNA needed to form a single functional ribosome. Therefore, in a 5000 s division time, a single
 578 RNA polymerase transcribing rRNA at a time would result in only \approx 50 functional ribosomal rRNA
 579 units – far below the observed number of \approx 10^4 ribosomes per cell.

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

589 *E. coli*, like many other bacterium, have evolved a clever mechanism to surpass this kinetic
 590 limit for the rate of rRNA production. Rather than having only one copy of each rRNA gene, *E. coli*
 591 has seven copies of the operon (BIND: 100352) four of which are localized directly adjacent to the
 592 origin of replication (*Birnbaum and Kaplan, 1971*). As fast growth also implies an increased gene
 593 dosage due to parallelized chromosomal replication, the total number of rRNA genes can be on
 594 the order of \approx 10 – 70 copies at moderate to fast growth rates (*Stevenson and Schmidt, 2004*).
 595 Given a 5000 second division time, we can make the lower-bound estimate that the typical cell will
 596 have \approx 7 copies of the rRNA operon. Synthesizing one functional rRNA unit per second per rRNA
 597 operon, a total of 5×10^4 rRNA units can be synthesized, comfortably above the observed number
 598 of ribosomes per cell.

599 How many RNA polymerases are then needed to constantly transcribe 7 copies of the rRNA
 600 genes? We approach this estimate by considering the maximum number of RNA polymerases
 601 tiled along the rRNA genes with a loading rate of 1 per second and a transcription rate of 40 nu-
 602 cleotides per second. Considering that a RNA polymerase has a physical footprint of approximately
 603 40 nucleotides (BNID: 107873), we can expect \approx 1 RNA polymerase per 80 nucleotides. With a total
 604 length of \approx 4500 nucleotides per operon and 7 operons per cell, the maximum number of RNA
 605 polymerases that can be transcribing rRNA at any given time is \approx 500. As we will see in the coming
 606 sections, the synthesis of rRNA is the dominant requirement of the RNA polymerase pool.

607 mRNA

608 To form a functional protein, all protein coding genes must first be transcribed from DNA to form an
 609 mRNA molecule. While each protein requires an mRNA blueprint, many copies of the protein can
 610 be synthesized from a single mRNA. Factors such as strength of the ribosomal binding site, mRNA
 611 stability, and rare codon usage frequency dictate the number of proteins that can be made from a
 612 single mRNA, with yields ranging from 10^1 to 10^4 (BNID: 104186; 100196; 106254). Computing the
 613 geometric mean of this range yields \approx 1000 proteins synthesized per mRNA, a value that agrees
 614 with experimental measurements of the number of proteins per cell ($\approx 3 \times 10^6$, BNID: 100088) and
 615 total number of mRNA per cell ($\approx 3 \times 10^3$, BNID: 100064).

616 This estimation captures the *steady-state* mRNA copy number, meaning that at any given time,
 617 there will exist approximately 3000 unique mRNA molecules. To determine the *total* number of
 618 mRNA that need to be synthesized over the cell's lifetime, we must consider degradation of the
 619 mRNA. In most bacteria, mRNAs are rather unstable with life times on the order of several minutes
 620 (BNID: 104324; 106253; 111927; 111998). For convenience, we assume that the typical mRNA in
 621 our cell of interest has a typical lifetime of \approx 300 seconds. Using this value, we can determine
 622 the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell.
 623 While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional
 624 dynamics, we state here that the total mRNA production rate must be on the order of \approx 15 mRNA
 625 per second. In *E. coli*, the average protein is \approx 300 amino acids in length (BNID: 108986), meaning

that the corresponding mRNA is \approx 900 nucleotides which we will further approximate as \approx 1000 nucleotides to account for the non-protein coding regions on the 5' and 3' ends. This means that the cell must have enough RNA polymerase molecules about to sustain a transcription rate of $\approx 1.5 \times 10^4$ nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA at a clip of 40 nucleotides per second, we arrive at a comfortable estimate of \approx 250 RNA polymerase complexes needed to satisfy the mRNA demands of the cell. It is worth noting that this number is approximately half of that required to synthesize enough rRNA, as we saw in the previous section. We find this to be a striking result as these 250 RNA polymerase molecules are responsible for the transcription of the \approx 4000 protein coding genes that are not ribosome associated.

635 tRNA

636 The final class of RNA molecules worthy of quantitative consideration are the tRNAs that are used
 637 during translation to map codon sequence on mRNA to specific amino acids. Unlike mRNA or
 638 rRNA, each individual tRNA is remarkably short, ranging from 70 to 95 nucleotides each (BNID:
 639 109645; 102340). What they lack in length, they make up for in abundance, with reported values
 640 ranging from $\approx 5 \times 10^4$ (BNID: 105280) to $\approx 5 \times 10^5$ (BNID: 108611). To test tRNA synthesis as a possible
 641 growth-rate limiting stage, we will err towards a higher abundance of $\approx 5 \times 10^5$ per cell. Combining
 642 the abundance and tRNA length measurements, we make the estimate that $\approx 5 \times 10^7$ nucleotides
 643 are sequestered in tRNA per cell. Unlike mRNA, tRNA is remarkably stable with typical lifetimes
 644 *in vivo* on the order of \approx 48 hours (*Abelson et al., 1974; Svenningsen et al., 2017*) – well beyond
 645 the timescale of division. Once again using our rule-of-thumb for the rate of transcription to be 40
 646 nucleotides per second and assuming a division time of \approx 5000 seconds, we arrive at an estimate
 647 of \approx 200 RNA polymerases to synthesize enough tRNA. This requirement pales in comparison to
 648 the number of polymerases needed to generate the rRNA and mRNA pools and can be neglected
 649 as a significant transcriptional burden.

650 RNA Polymerase and σ -factor Abundance

651 These estimates, summarized in *Figure 8* (A), reveal that synthesis of rRNA and mRNA are the
 652 dominant RNA species synthesized by RNA polymerase, suggesting the need for \approx 1000 RNA poly-
 653 merases per cell. As is revealed in *Figure 8* (B), this estimate is about an order of magnitude be-
 654 low the observed number of RNA polymerase complexes per cell (\approx 5000 - 7000). The difference
 655 between the estimated number of RNA polymerase needed for transcription and these ob-
 656 servations are consistent with recent literature revealing that \approx 80 % of RNA polymerases in *E.*
 657 *coli* are not transcriptionally active (*Patrick et al., 2015*). Our estimate ignores the possibility that
 658 some fraction is only nonspecifically bound to DNA, as well as the obstacles that RNA polymerase
 659 and DNA polymerase present for each other as they move along the DNA (*Finkelstein and Greene,*
 660 *2013*).

661 In addition, it is also vital to consider the role of σ -factors which help RNA polymerase identify
 662 and bind to transcriptional start sites (*Browning and Busby, 2016*). Here we consider σ^{70} (RpoD)
 663 which is the dominant "general-purpose" σ -factor in *E. coli*. While initially thought of as being solely
 664 involved in transcriptional initiation, the past two decades of single-molecule work has revealed
 665 a more multipurpose role for σ^{70} including facilitating transcriptional elongation (*Kapanidis et al.,*
 666 *2005; Goldman et al., 2015; Perdue and Roberts, 2011; Mooney and Landick, 2003; Mooney et al.,*
 667 *2005*). *Figure 8* (B) is suggestive of such a role as the number of σ^{70} proteins per cell is in close
 668 agreement with our estimate of the number of transcriptional complexes needed.

669 These estimates provide insight as to the observed magnitude of both RNA polymerase and the
 670 σ -70 factor. As we have done in the previous sections, and described in Appendix Extending Es-
 671 timates to a Continuum of Growth Rates, we can generalize these estimates across a wide range of
 672 growth rates (grey line in *Figure 8*(B)). While there remains some disagreement in the magnitude of
 673 the copy number, this estimate appears to very adequately describe the growth rate dependence
 674 of these complexes. Furthermore, these findings illustrate that transcription cannot be the rate

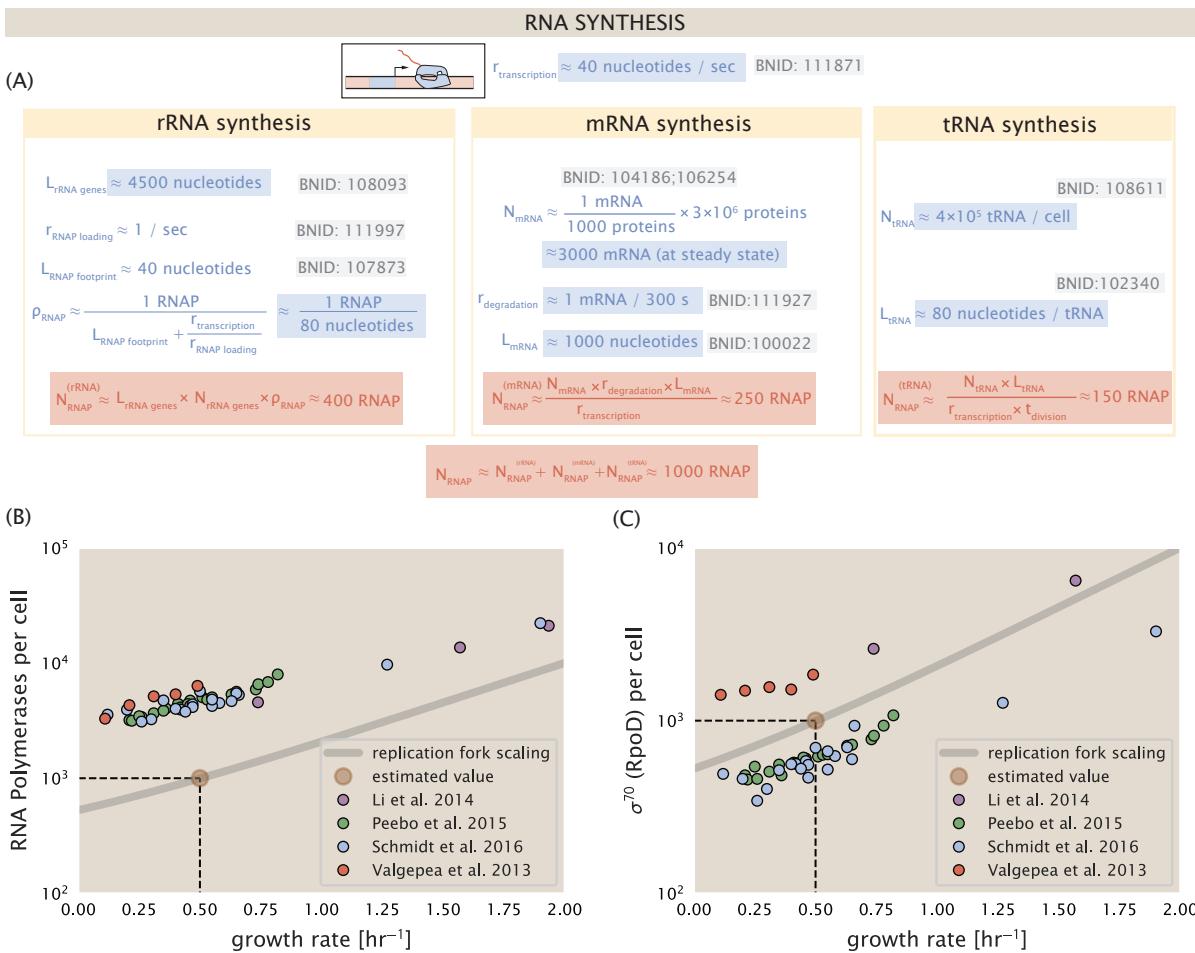


Figure 8. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. (B) The RNA polymerase core enzyme copy number as a function of 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} .

675 limiting step in bacterial division. **Figure 8** (A) reveals that the availability of RNA polymerase is not
 676 a limiting factor for cell division as the cell always has an apparent ~ 10 -fold excess than needed.
 677 Furthermore, if more transcriptional activity was needed to satisfy the cellular requirements, more
 678 σ^{70} -factors could be expressed to utilize a larger fraction of the RNA polymerase pool.

679 Translation and Ribosomal Synthesis

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

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

691 tRNA Synthetases

692 We begin by first estimating the number of tRNA synthetases in *E. coli* needed to convert free
 693 amino-acids to polypeptide chains. Again using an estimate of $\approx 3 \times 10^6$ proteins per cell at a 5000 s
 694 division time (BNID: 115702) and a typical protein length of ≈ 300 amino acids (BNID: 100017), we
 695 can estimate that a total of $\approx 10^9$ amino acids are stitched together by peptide bonds.

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

704 There are many processes which go into synthesizing a tRNA and ligating it with the appropriate
 705 amino acids. As we discussed previously, there appear to be more than enough RNA polymerases
 706 per cell to synthesize the needed pool of tRNAs. Without considering the many ways in which
 707 amino acids can be scavenged or synthesized *de novo*, we can explore ligation the as a potential
 708 rate limiting step. The enzymes which link the correct amino acid to the tRNA, known as tRNA
 709 synthetases or tRNA ligases, are incredible in their proofreading of substrates with the incorrect
 710 amino acid being ligated once out of every 10^4 to 10^5 events (BNID: 103469). This is due in part
 711 to the consumption of energy as well as a multi-step pathway to ligation. While the rate at which
 712 tRNA is ligated is highly dependent on the identity of the amino acid, it is reasonable to state that
 713 the typical tRNA synthetase has charging rate of ≈ 20 AA per tRNA synthetase per second (BNID:
 714 105279).

715 We can make an assumption that amino-acyl tRNAs are in steady-state where they are pro-
 716 duced at the same rate they are consumed, meaning that 2×10^5 tRNAs must be charged per second.
 717 Combining these estimates together, as shown schematically in **Figure 9(A)**, yields an estimate of
 718 $\sim 10^4$ tRNA synthetases per cell with a division time of 5000 s. This point estimate is in very close
 719 agreement with the observed number of synthetases (the sum of all 20 tRNA synthetases in *E. coli*).
 720 This estimation strategy seems to adequately describe the observed growth rate dependence of
 721 the tRNA synthetase copy number (shown as the grey line in **Figure 9(B)**), suggesting that the copy
 722 number scales with the cell volume.

723 In total, the estimated and observed $\sim 10^4$ tRNA synthetases occupy only a meager fraction of

724 the total cell proteome, around 0.5% by abundance. It is reasonable to assume that if tRNA charg-
 725 ing was a rate limiting process, cells would be able to increase their growth rate by devoting more
 726 cellular resources to making more tRNA synthases. As the synthesis of tRNAs and the correspond-
 727 ing charging can be highly parallelized, we can argue that tRNA charging is not a rate limiting step
 728 in cell division, at least for the growth conditions explored in this work.

729 Protein Synthesis

730 With the number of tRNA synthetases accounted for, we now consider the abundance of the pro-
 731 tein synthesis machines themselves, ribosomes. Ribosomes are enormous protein/rRNA com-
 732 plexes that facilitate the peptide bond formation between amino acids in the correct sequence
 733 as defined by the coding mRNA. Before we examine the synthesis of the ribosome proteins and
 734 the limits that may place on the observed bacterial growth rates, let's consider replication of the
 735 cellular proteome.

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

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

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

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

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

758 where λ is the cell growth rate in s^{-1} , r_t is the maximum elongation rate in $\text{AA}\cdot\text{s}^{-1}$, and R is the aver-
 759 age ribosome copy number per cell. The addition factor f_a refers to the fraction of actively translat-
 760 ing ribosomes, and allows us to account for the possibility of nonfunctional, immature ribosomes
 761 or active sequestration of ribosomes, mediated by the secondary-messenger molecule alarmones,
 762 guanosine pentaphosphate [(p)ppGpp] at slow growth (*Dennis et al., 2004; Dai et al., 2016*). Know-
 763 ing the number of peptide bonds formed per cell permits us to compute the translation-limited
 764 growth rate as

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

765 Alternatively, since N_{pep} is related to the total protein mass through the molecular weight of
 766 each protein, we can also consider the growth rate in terms of the fraction of the total proteome
 767 mass dedicated to ribosomal proteins. By making the approximation that an average amino acid
 768 has a molecular weight of 110 Da (BNID: 104877), the total protein mass m_{protein} is related to N_{AA}

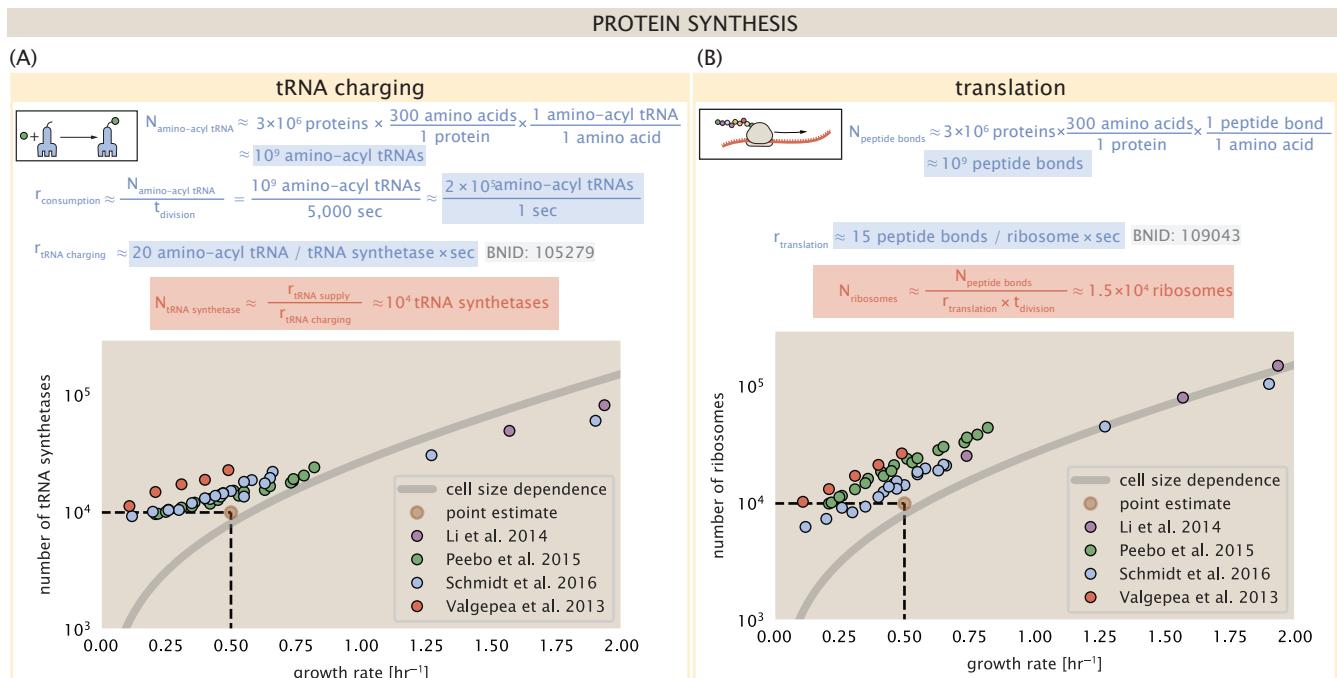


Figure 9. Estimation of the required tRNA synthetases and ribosomes. (A) Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [Glx], [IleS], [LeuS], [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 10^9 peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr^{-1} . Grey lines correspond to the estimated complex abundance calculated at different growth rates. See Appendix Extending Estimates to a Continuum of Growth Rates for a more detail description of this calculation.

769 by $(m_{\text{protein}}/110 \text{ Da}) \times N_A$, where N_A is Avogadro's number. Similarly, R is related to the ribosomal
 770 protein mass by $R \approx (m_R/800 \text{ Da}) \times N_A$, where 800 Da reflects the summed molecular weight of
 771 all ribosomal subunits. This allows us to approximate $R/N_{\text{pep}} \approx \Phi_R/L_R$, where Φ_R is the ribosomal
 772 mass fraction m_{protein}/m_R , and L_R the ratio of 800 kDa / 110 Da per amino acid or, alternatively, the
 773 total length in amino acids that make up a ribosome. The translation-limited growth rate can then
 774 be written in the form

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

775 This is plotted as a function of ribosomal fraction Φ_R in **Figure 10(B)**, where we take $L_R = 7459 \text{ AA}$,
 776 corresponding to the length in amino acids for all ribosomal subunits of the 50S and 30S complex
 777 (BNID: 101175), and $f_a = 1$.

778 The growth rate defined by **Equation 3** reflects mass-balance under steady-state growth and
 779 has long provided a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a
 780 function of growth rate (*Maaløe, 1979; Scott et al., 2010*). Here we see that there will be a maximum
 781 rate when $\Phi_R = 1$, only achieved if a cell contained only ribosomes. For an elongation rate of 17
 782 amino acids per second, this gives us $\lambda \approx 8 \text{ hr}^{-1}$ (**Figure 10(B)**, dashed line) and corresponds to
 783 the synthesis time of all ribosomal subunits, $L_R/r_t \approx 7 \text{ minutes}$ (*Dill et al., 2011*). Interestingly, this
 784 particular limit is independent of the absolute number of ribosomes and is simply given by the time
 785 to translate an entire ribosome. As shown in **Figure 10(A)**, we can reconcile this with the observation
 786 that in order to double the average number of ribosomes, each ribosome must produce a second
 787 ribosome and this process cannot be parallelized. Unless the elongation rate can increase, or cells
 788 trim their total ribosomal protein mass, this dependency limits both the maximum growth rate
 789 (when $\Phi_R = 1$), and the growth rate under more realistic values of Φ_R .

790 In recent work from *Dai et al. (2016)*, the authors made independent measurements of r_t , Φ_R
 791 (via RNA-to-protein ratios, and directly by mass spectrometry), and growth rate, enabling inference
 792 of the active fraction f_a across the entire range of growth rates considered here. In **Figure 10(C)**
 793 we use this measurement of f_a to estimate the active fraction of ribosomal protein across the
 794 proteomic data sets and number of other recent measurements. We see that cells are essentially
 795 skirting the limit in growth rate set by **Equation 3** as nutrient conditions vary.

796 Earlier, we considered rRNA synthesis (see Section **RNA Synthesis**), finding that, when the rRNA
 797 operons are maximally loaded with RNA polymerase, the cell can produce ≈ 1 functional rRNA
 798 unit per second per operon. While each *E. coli* genome has 7 copies of the rRNA operon (BNID:
 799 107866), parallelization of chromosomal replication can drastically change the effective number
 800 of rRNA operons. In the blue curve in **Figure 10(C)**, we assume that the number of rRNA operons
 801 increases in proportion to the number of origins of replication $\langle \# \text{ ori} \rangle$ and 1 functional rRNA unit
 802 per second per operon (solid blue line; with the calculation of $\langle \# \text{ ori} \rangle$ described in the next section).
 803 Although we expect this value to drastically overestimate rRNA abundance at slower growth rates
 804 ($\lambda < 0.5 \text{ hr}^{-1}$), it provides a useful reference alongside the proteomic measurements. For growth
 805 rates above about 1 hr^{-1} , we find that cells will need to transcribe rRNA near their maximal rate.
 806 As a counter example, if *E. coli* did not initiate multiple rounds of replication, they would be unable
 807 to make enough rRNA for the observed number of ribosomes (dashed blue curve in **Figure 10(C)**).
 808 The convergence between the maximum rRNA production and measured ribosome copy number
 809 suggests rRNA synthesis may begin to present a bottleneck in cell division at the fastest growth
 810 rates. Interestingly, while this strain of *E. coli* is rarely reported to grow faster than 2 hr^{-1} , other
 811 bacteria with more copies of rRNA genes have been found that surpass this growth rate (*Bremer*
 812 and *Dennis, 2008; Roller et al., 2016*).

813 Relationship Between Cell Size and Growth Rate

814 The relationship between cell size and growth rate has long been of interest in the study of bacterial
 815 physiology, particularly following the now six decade-old observation that cell volume appears to
 816 increase exponentially with growth rate; known as Schaechter's growth law (*Schaechter et al., 1958;*

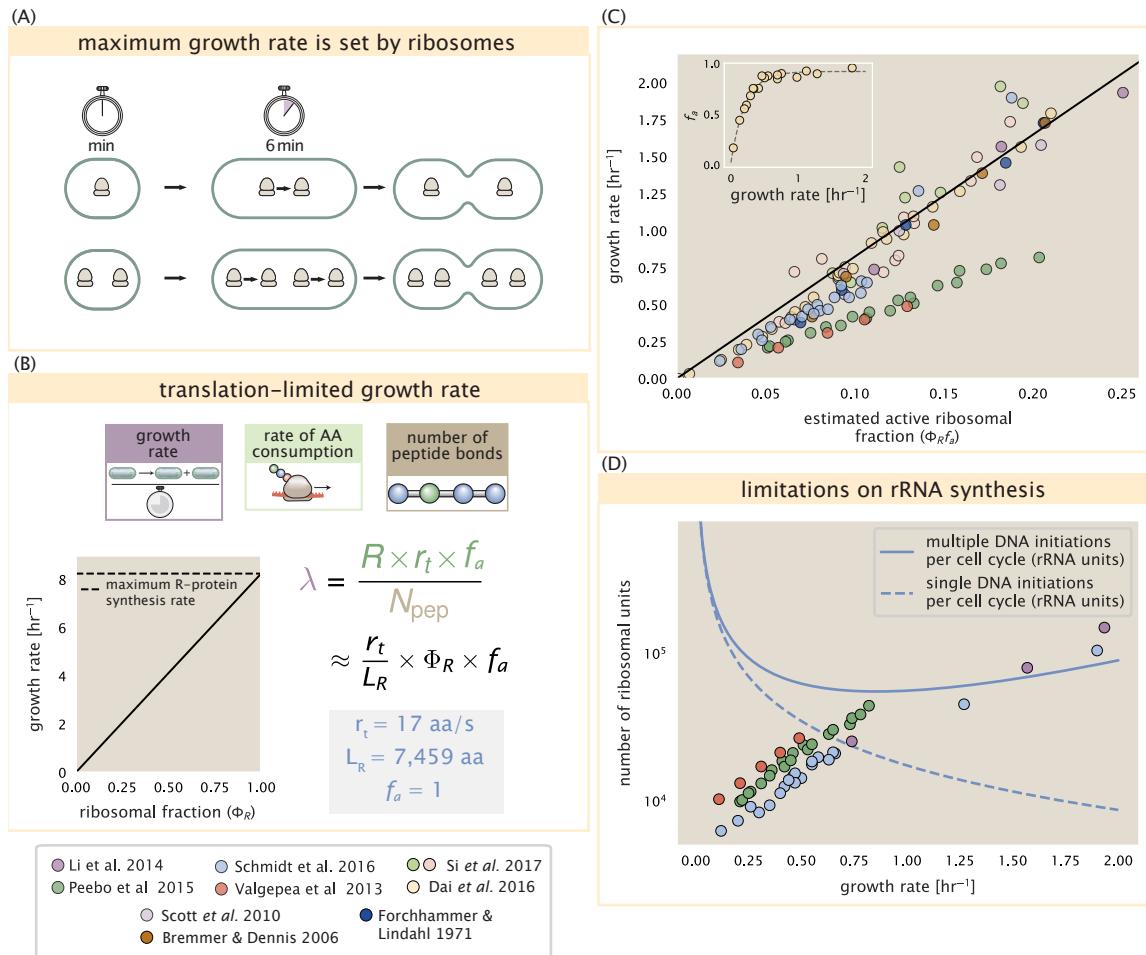


Figure 10. Translation-limited growth rate. (A) An inherent maximum growth rate is set by the time to synthesize all ribosomal subunits. This growth rate is given by r_t/L_R , where r_t is the elongation rate and L_R is the total number of amino acids that make up the entire ribosomal complex. This rate 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.

817 Taheri-Araghi et al., 2015. However, the mechanism that governs this relationship, and even the
818 question of whether the change in average cell size is truly exponential, has remained under debate
819 (**Harris and Theriot, 2018**). Given the importance of cell size in relating the total protein mass that
820 must be doubled (as well as in setting other parameters like the surface-area-to-volume ratio), we
821 examine the influence size may have in setting the scaling of protein abundance and growth rate
822 across the proteomic datasets.

823 As shown in **Figure 10(C)**, cells grow at a near-maximal rate dictated by their total ribosomal
824 mass fraction Φ_R , at least at moderate growth rates above 0.5 hr^{-1} (where f_a is close to 1). Here,
825 growth rate can be increased only by making more ribosomes in a way that increases Φ_R . As *E.*
826 *coli* grows faster, however, large swaths of the proteome also increase in absolute protein, and the
827 ability to add additional ribosomes is likely constrained by others factors such as crowding due to
828 their large size (**Delarue et al., 2018; Soler-Bistué et al., 2020**). It is now well-documented that *E.*
829 *coli* cells add a constant volume per origin of replication (termed a "unit cell" or "initiation mass"),
830 which is robust to a remarkable array of cellular perturbations (**Si et al., 2017**). To consider this
831 dependency in the context of the proteomic data, we used measurements from **Si et al. (2017)** for
832 wild-type *E. coli* cells grown in different nutrient conditions (**Figure 11(A)**) to estimate the average
833 number of origins per cell ($\langle \# \text{ ori} \rangle$) across the data.

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

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

836 where τ_{cyc} is the cell cycle time (referring to the time from replication initiation to cell division), and
837 τ is the cell doubling time. For ribosomal synthesis, we find an approximately linear correlation
838 between ribosome copy number and ($\langle \# \text{ ori} \rangle$) (**Figure 11(B)**). For a constant cell cycle time, observed
839 at growth rates above about 0.5 hr^{-1} (**Helmstetter and Cooper, 1968**), **Equation 4** states that ($\langle \# \text{ ori} \rangle$)
840 will need to increase exponentially with the growth rate in order to maintain steady-state growth.

841 Why does *E. coli* add a constant volume per ($\langle \# \text{ ori} \rangle$)? To consider how this trend pertains to
842 growth, we must consider how the proteome size and composition changes with respect to growth
843 rate. In **Figure 11(D)**, we analyze the position-dependent protein expression across the chromo-
844 some for each of the growth conditions from **Schmidt et al. (2016)**. Here, we have calculated a
845 running Gaussian average of protein copy number (20 kbp st. dev. averaging window) based on
846 each gene's transcriptional start site, which were then median-subtracted to account for the differ-
847 ences in total protein abundance with each growth condition. Importantly, we find that the major
848 deviations in protein copy number are largely restricted to regions of ribosomal protein genes, with
849 substantially higher deviations observed for cells with high ($\langle \# \text{ ori} \rangle$) (teal), as compared to those with
850 low ($\langle \# \text{ ori} \rangle$) (purple). This is particularly apparent for genes closer to the origin, where the major-
851 ity of ribosomal proteins are found. This suggests that in addition to the linear scaling between
852 protein abundance and ($\langle \# \text{ ori} \rangle$), the relative ribosomal abundance is tuned in proportion to ($\langle \# \text{ ori} \rangle$).
853 Given the increased rRNA gene dosage required at faster growth rates, additional rounds of DNA
854 replication have the effect of skewing DNA dosage in favor of additional ribosomal synthesis Since
855 growth rate depends specifically on the ribosomal fraction Φ_R , this result suggests that cells are
856 changing their size as a way to vary the absolute number of ribosomes per cell and tune Φ_R accord-
857 ing to better match available nutrient conditions.

858 Alarmone-Mediated Regulation Controls the Rate of Protein Synthesis

859 As we have seen, cell size, total proteomic content, and the number of ribosomes are all intercon-
860 nected and influence the achievable growth rate. The drastic change in these parameters across
861 different growth conditions also suggests that cells are tuning them to better match their biosyn-
862 thetic capacity to the specific environment. Take, as another illustration of this, the recent experi-
863 mental work by **Dai et al. (2016)**. In one set of experiments the authors considered growth in cells
864 whose primary glucose transport system was disrupted ($\Delta ptsG$). Unsurprisingly, the growth rate

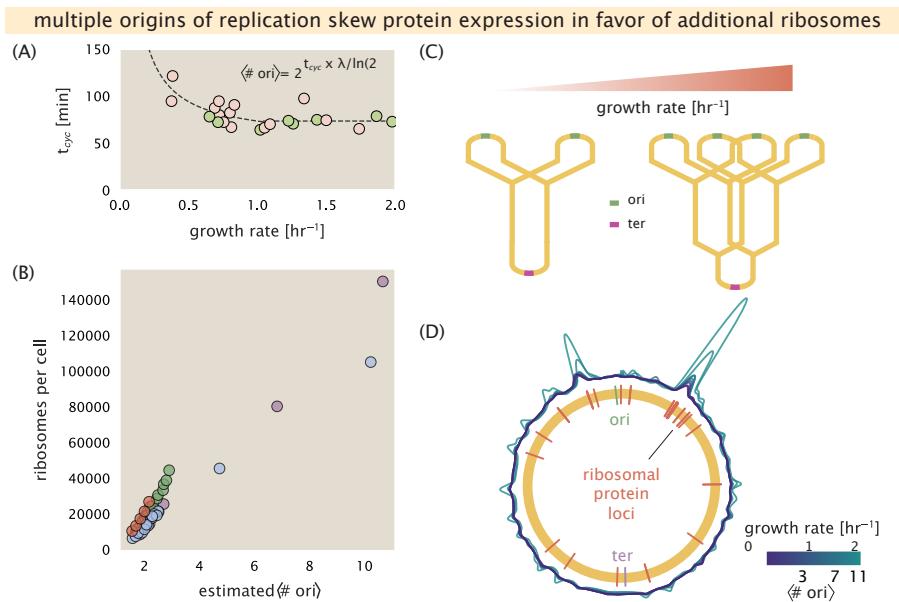


Figure 11. Cells increase absolute ribosome abundance with $\langle \# \text{ori} \rangle$. (A) Experimental data from Si *et al.* (2017). Dashed line shows fit to the data, which were used to estimate $\langle \# \text{ori} \rangle$. t_{cyc} was assumed to vary in proportion to τ for doubling times 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.

865 was reduced, and was measured at about two-fold slower than their wild-type line. This change,
 866 however, was not simply the result of now-limiting carbon uptake. Instead, cells accommodated this
 867 perturbation by also reducing their ribosomal mass fraction by a factor of two, which is still in line
 868 with **Equation 3** under translation-limited growth. In this final, we explore the interconnection be-
 869 tween cell size, ribosome content, and growth rate by formulating a minimal model of growth rate
 870 control. We use it to quantitatively show how tuning these parameters help cells maximize their
 871 growth rate.

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

879 Furthermore, (p)ppGpp can inhibit the initiation of DNA replication by mediating a change in
 880 transcriptional activity and the supercoiling state of the origin of replication (*Kraemer et al., 2019*).
 881 These observations all raise the possibility that it is through (p)ppGpp that cells mediate the growth-
 882 rate dependent changes in $\langle \# \text{ ori} \rangle$, cell size, and ribosomal abundance and activity (*Zhu and Dai,*
 883 *2019; Büke et al., 2020*). Indeed, recent work in a (p)ppGpp deficient strain of *E. coli* found that
 884 cells exhibited a high ratio of $\langle \# \text{ ori} \rangle$ to $\langle \# \text{ ter} \rangle$, and cell sizes that were more consistent with a fast
 885 growth state where (p)ppGpp levels are normally low (*Fernández-Coll et al., 2020*).

886 **Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid 887 Scarcity**

888 To better understand how cells maximize their growth rate across growth conditions, we consider
 889 a mode of regulation in which the rate of peptide elongation r_t depends only on the availability of
 890 amino acids (and, therefore, also amino-acyl tRNAs). It is through the elongation rate r_t that we
 891 assume cells adjust their ribosomal content (R, Φ_R) according to nutrient availability. As the rate of
 892 amino acid supply, denote by r_{AA} , decreases, the cell can tune the rate of amino acid consumption
 893 (mathematized as $r_t \times R \times f_a$) to remain in steady-state growth, shown schematically in **Figure 12(A)**.
 894 Under this model, other molecular players required for translation like elongation factors and GTP
 895 are considered in sufficient abundance, which appear to be valid assumptions given our analysis
 896 of the proteomic data and energy production thus far.

897 For simplicity, we consider all amino acids as a single species with an effective cellular con-
 898 centration $[AA]_{\text{eff}}$. The rate of elongation r_t will depend on how quickly the ribosomes can match
 899 codons with their correct amino-acyl tRNA, along with the subsequent steps of peptide bond for-
 900 mation and translocation. We therefore coarse-grain the steps of elongation to two time-scales,
 901 1) the time required to find and bind each correct amino-acyl tRNA, and 2) the remaining steps in
 902 peptide elongation that will not depend on the amino acid availability. The time to translate each
 903 codon is given by the inverse of the elongation rate r_t , which can be written as,

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

904 where we have assumed that the rate of binding by amino-acyl tRNA k_{on} is proportional to $[AA]_{\text{eff}}$
 905 by a constant α . The second term on the right-hand side reflects our assumption that other steps in
 906 peptide elongation are not rate-limiting, with a maximum elongation rate r_t^{\max} of about 17 amino
 907 acids per second (*Dai et al. (2016)*). This can be stated more succinctly in terms of an effective
 908 dissociation constant,

$$K_D = \frac{r_t^{\max}}{\alpha k_{on}}, \quad (6)$$

909 where the elongation rate r_t is now given by

$$r_t = \frac{r_t^{\max}}{1 + K_D/[AA]_{\text{eff}}}. \quad (7)$$

910 Under steady-state growth, the amino acid concentration is constant ($\frac{d[AA]_{\text{eff}}}{dt} = 0$), meaning that
 911 synthesis and consumption are matched. The effective amino acid concentration $[AA]_{\text{eff}}$ will relate
 912 to the rate of amino acid synthesis (or import, for rich media) and/or tRNA charging, as r_{AA} , and
 913 the rate of consumption, $r_t \times R \times f_a$ by,

$$\int_0^t \frac{d[AA]_{\text{eff}}}{dt} dt = \int_0^t ([r_{AA}] - [r_t \times R \times f_a]) dt, \quad (8)$$

914 where the time from 0 to t is an arbitrary length of time, and the square brackets indicate concen-
 915 trations per unit time. Integrating **Equation 8** yields.

$$[AA]_{\text{eff}} = t([r_{AA}] - [r_t \times R \times f_a]). \quad (9)$$

916 Alternatively, we can state this in terms of absolute ribosome copy number R by considering a
 917 unit volume V ,

$$[AA]_{\text{eff}} = \frac{t(r_{AA} - r_t \times R \times f_a)}{V}, \quad (10)$$

918 where r_{AA} is in units of AA per unit time and r_t is in units of AA per unit time per ribosome. With
 919 an expression for $[AA]_{\text{eff}}$ in hand, we can now solve **Equation 7** for r_t , which is a quadratic function
 920 with a physically-meaningful root of

$$r_t = \frac{t(r_{AA} + r_t^{\max} R f_a) + K_D V - \sqrt{(r_{AA} t + r_t^{\max} R f_a t + K_D V)^2 - 4(R f_a t)(r_t^{\max} r_{AA} t)}}{2 R f_a t}. \quad (11)$$

921 In **Figure 12(B)**, we illustrate how the elongation rate depends on the ribosomal copy number.
 922 Here, we have 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
 923 **Bennett et al. (2009)**), $f_a = 1$, and an arbitrarily chosen $r_{AA} = 5 \times 10^6 \text{ AA} \cdot \text{s}^{-1} \cdot \mu\text{m}^{-3}$. At low ribosome
 924 copy numbers, the observed elongation rate is dependent primarily on the ratio of $K_D/V r_{AA}$ [as
 925 $r_t^{\max} \times R \times f_a \ll r_{AA}$, point (1) in **Figure 12(B)**]. As the ribosome copy number is increased such
 926 that the amino acid supply rate and consumption rate are nearly equal [point (2) in **Figure 12(B)**],
 927 the observed elongation rate begins to decrease sharply. When the ribosome copy number is
 928 increased even further, consumption at the maximum elongation rate exceeds the supply rate,
 929 yielding a significantly reduced elongation rate [point (3) in **Figure 12B**]. While the elongation rate
 930 will always be dominated by the amino acid supply rate at sufficiently low ribosome copy numbers,
 931 the elongation rate at larger ribosome abundances can be increased by tuning f_a such that not all
 932 ribosomes are elongating, reducing the total consumption rate.

933 It is important to note that thus far, this model quantifies only the relationship between amino
 934 acid supply and consumption as a function of the ribosome copy number and states nothing about
 935 the cellular growth rate. With a sense of how elongation rate is tied to amino acid availability, we
 936 now turn to how this relates to the cellular growth rate.

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

939 To relate the elongation rate to growth rate, we constrain the set of parameters based on measured
 940 proteomic changes; namely, we will restrict the values of R , N_{pep} , and V to those associated with
 941 the amalgamated proteomic data. We will then consider how changes in the nutrient conditions,
 942 through the parameter r_{AA} , influence the maximum growth rate.

943 Earlier, we considered ribosome biosynthesis as the growth-rate determining cellular process
 944 in **Equation 2** by stating that the cellular growth rate λ was related to the ribosome abundance,
 945 elongation rate, active ribosome fraction, and the total number of peptide bonds to be formed,

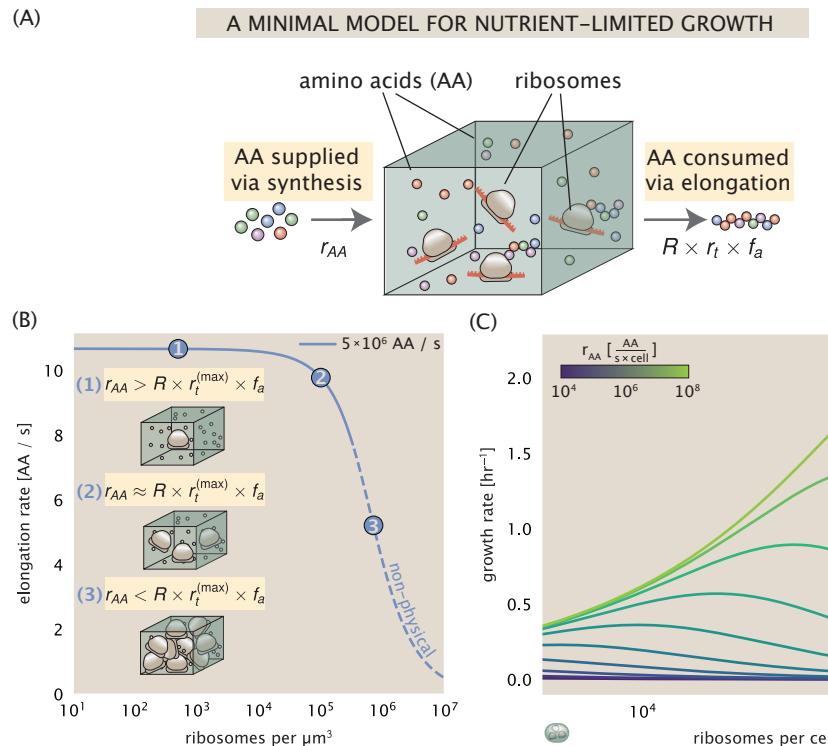


Figure 12. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times f_a$, 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 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.

946 N_{pep} . We return to this limit in light of our expression for a condition-dependent elongation rate
 947 r_t given by **Equation 11**. **Figure 12(C)** shows how the observed growth rate depends on the rate of
 948 amino acid supply r_{AA} as a function of the cellular ribosome copy number. A feature immediately
 949 apparent is the presence of a maximal growth rate whose dependence on R (and consequently,
 950 the cell volume) increases with increasing r_{AA} . Importantly, however, there is an optimum set of R ,
 951 N_{pep} , and V that are strictly dependent on the value of r_{AA} . Increasing the ribosomal concentration
 952 beyond the cell's metabolic capacity has the adverse consequence of depleting the supply of amino
 953 acids and a concomitant decrease in the elongation rate r_t , [**Figure 12(B)**].

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

962 Discussion

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

978 In general, the copy numbers of the complexes involved in these processes were reasonable
 979 agreement with our order-of-magnitude estimates. Since many of these estimates represent soft
 980 lower-bound quantities, this suggests that cells do not express proteins grossly in excess of what
 981 is needed for a particular growth rate. Several exceptions, however, also highlight the dichotomy
 982 between a proteome that appears to "optimize" expression according to growth rate and one that
 983 must be able to quickly adapt to environments of different nutritional quality. Take, for example,
 984 the expression of carbon transporters. Shown in **Figure 2(B)**, we find that cells always express
 985 a similar number of glucose transporters irrespective of growth condition. At the same time, it
 986 is interesting to note that many of the alternative carbon transporters are still expressed in low
 987 but non-zero numbers (≈ 10 -100 copies per cell) across growth conditions. This may relate to the
 988 regulatory configuration for many of these operons, which require the presence of a metabolite sig-
 989 nal in order for alternative carbon utilization operons to be induced (**Monod, 1949; Laxhuber et al.,**
 990 **2020**). Furthermore, upon induction, these transporters are expressed and present in abundances
 991 in close agreement with a simple estimate.

992 Of the processes illustrated in **Figure 1**, we arrive at a ribosome-centric view of cellular growth
 993 rate control. This is in some sense unsurprising given the long-held observation that *E. coli* and
 994 many other organisms vary their ribosomal abundance as a function of growth conditions and

996 growth rate *Scott et al. (2010); Metzl-Raz et al. (2017)*. However, through our dialogue with the pro-
 997 teomic data, two additional key points emerge. The first relates to our question of what process
 998 sets the absolute speed limit of bacterial growth. While a cell can parallelize many of its processes
 999 simply by increasing the abundance of specific proteins or firing multiple rounds of DNA replica-
 1000 tion, this is not so for synthesis of ribosomes (**Figure 10(A)**). The translation time for each ribosome
 1001 [≈ 6 min, *Dill et al. (2011)*] places an inherent limit on the growth rate that can only be surpassed if
 1002 the cell were to increase their polypeptide elongation rate, or if they could reduce the total protein
 1003 and rRNA mass of the ribosome. The second point relates to the long-observed correlations be-
 1004 tween growth rate and cell size (*Schaechter et al., 1958; Si et al., 2017*), and between growth rate
 1005 and ribosomal mass fraction. While both trends have sparked tremendous curiosity and driven
 1006 substantial amounts of research in their own regards, these relationships are themselves inter-
 1007 twined. In particular, it is the need for cells to increase their absolute number of ribosomes under
 1008 conditions of rapid growth that require cells to also grow in size. Further experiments are needed
 1009 to test the validity of this hypothesis. In particular, we believe that the change in growth rate in re-
 1010 sponse to translation-inhibitory drugs (such as chloramphenicol) could be quantitatively predicted,
 1011 given one had precision measurement of the relevant parameters, including the fraction of actively
 1012 translating ribosomes f_a and changes in the metabolic capacity of the cell (i.e. the parameter r_{AA}
 1013 in our minimal model) for a particular growth condition.

1014 While the generation of new ribosomes plays a dominant role in growth rate control, there exist
 1015 other physical limits to the function of cellular processes. One of the key motivations for consid-
 1016 ering energy production was the physical constraints on total volume and surface area as cells
 1017 vary their size (*Harris and Theriot, 2018; Ojkic et al., 2019*). While *E. coli* get larger as it expresses
 1018 more ribosomes, an additional constraint begins to arise in energy production due to a relative de-
 1019 crease in total surface area where ATP is predominantly produced (*Szenk et al., 2017*). Specifically,
 1020 the cell interior requires an amount of energy that scales cubically with cell size, but the available
 1021 surface area only grows quadratically (**Figure 5(A)**). While this threshold does not appear to be met
 1022 for *E. coli* cells growing at 2 hr^{-1} or less, it highlights an additional constraint on growth given the
 1023 apparent need to increase in cell size to grow faster. This limit is relevant even to eukaryotic or-
 1024 ganisms, whose mitochondria exhibit convoluted membrane structures that nevertheless remain
 1025 bacteria-sized organelles (*Guo et al., 2018*). In the context of bacteria growth and energy produc-
 1026 tion more generally, we have limited our analysis to the aerobic growth conditions associated with
 1027 the proteomic data and further consideration will be needed for anaerobic growth.

1028 This work is by no means meant to be an exhaustive dissection of bacterial growth rate control,
 1029 and there are many aspects of the bacterial proteome and growth that we neglected to consider.
 1030 For example, other recent work (*Liebermeister et al., 2014; Hui et al., 2015; Schmidt et al., 2016*)
 1031 has explored how the proteome is structured and how that structure depends on growth rate. In
 1032 the work of *Hui et al. (2015)*, the authors coarse-grained the proteome into six discrete categories
 1033 being related to either translation, catabolism, anabolism, and others related to signaling and core
 1034 metabolism. The relative mass fraction of the proteome occupied by each sector could be modu-
 1035 lated by external application of drugs or simply by changing the nutritional content of the medium.
 1036 While we have explored how the quantities of individual complexes are related to cell growth, we ac-
 1037 knowledge that higher-order interactions between groups of complexes or metabolic networks at
 1038 a systems-level may reveal additional insights into how these growth-rate dependences are mech-
 1039 anistically achieved. Furthermore, while we anticipate the conclusions summarized here are ap-
 1040 plicable to a wide collection of bacteria with similar lifestyles as *E. coli*, other bacteria and archaea
 1041 may have evolved other strategies that were not considered. Further experiments with the level of
 1042 rigor now possible in *E. coli* will need to be performed in a variety of microbial organisms to learn
 1043 more about how regulation of proteomic composition and growth rate control has evolved over
 1044 the past 3.5 billion years.

1045 Appendix for: Fundamental limits on 1046 the rate of bacterial cell division

1047 **Nathan M. Belliveau^{†, 1}, Griffin Chure^{†, 2}, Christina L. Hueschen³, Hernan G.**
1048 **Garcia⁴, Jane Kondev⁵, Daniel S. Fisher⁶, Julie A. Theriot^{1, 7, *}, Rob Phillips^{8, 9, *}**

*For correspondence:

[†]These authors contributed
equally to this work

1049 ¹Department of Biology, University of Washington, Seattle, WA, USA; ²Department of
1050 Applied Physics, California Institute of Technology, Pasadena, CA, USA; ³Department of
1051 Chemical Engineering, Stanford University, Stanford, CA, USA; ⁴Department of
1052 Molecular Cell Biology and Department of Physics, University of California Berkeley,
1053 Berkeley, CA, USA; ⁵Department of Physics, Brandeis University, Waltham, MA, USA;
1054 ⁶Department of Applied Physics, Stanford University, Stanford, CA, USA; ⁷Allen Institute
1055 for Cell Science, Seattle, WA, USA; ⁸Division of Biology and Biological Engineering,
1056 California Institute of Technology, Pasadena, CA, USA; ⁹Department of Physics,
1057 California Institute of Technology, Pasadena, CA, USA; *Co-corresponding authors.
1058 Address correspondence to phillips@pboc.caltech.edu and jtheriot@uw.edu

1059 Contents

1060 Introduction	1
1061 Uptake of Nutrients	2
1062 Nitrogen Transport	4
1063 Carbon Transport	4
1064 Phosphorus and Sulfur Transport	5
1065 Limits on Transporter Expression	7
1066 Energy Production	9
1067 ATP Synthesis	9
1068 Generating the Proton Electrochemical Gradient	9
1069 Energy Production in a Crowded Membrane.	10
1070 Synthesis of the Cell Envelope	12
1071 Lipid Synthesis	12
1072 Peptidoglycan Synthesis	14
1073 Function of the Central Dogma	16
1074 DNA	16
1075 dNTP synthesis	16
1076 DNA Replication	17
1077 RNA Synthesis	18
1078 rRNA	18
1079 mRNA	19
1080 tRNA	20
1081 RNA Polymerase and σ-factor Abundance	20
1082 Translation and Ribosomal Synthesis	22
1083 tRNA Synthetases	22
1084 Protein Synthesis	23
1085 Translation and Ribosomal Synthesis as a Growth-Rate Limiting Step	23
1086 Relationship Between Cell Size and Growth Rate	25
1087 Alarmone-Mediated Regulation Controls the Rate of Protein Synthesis	27
1088 Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid Scarcity	29
1089 Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and Metabolic Capacity.	30
1091 Discussion	32
1092 Experimental Details Behind Proteomic Data	37
1093 Fluorescence based measurements	37
1094 Ribosomal profiling measurements	37
1095 Mass spectrometry measurements	38
1096 Summary of Proteomic Data	38
1097 Estimation of Cell Size and Surface Area	39
1098 Estimation of Total Protein Content per Cell	40
1099 Estimating Volume and Number of Amino Acids from Ribosome Copy Number	41

1100	Additional Considerations of Schmidt <i>et al.</i> Data Set	41
1101	Effect of cell volume on reported absolute protein abundances	45
1102	Relaxing assumption of constant protein concentration across growth conditions	45
1103	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	46
1104	Calculation of Complex Abundance	46
1105	Extending Estimates to a Continuum of Growth Rates	47
1106	Estimation of the total cell mass	47
1107	Complex Abundance Scaling With Cell Volume	49
1108	A Relation for Complex Abundance Scaling With Surface Area	50
1109	Number of Lipids	50
1110	Number of Murein Monomers	50
1111	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	51
1112	Calculation of active ribosomal fraction.	51
1113	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	51

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).

1114 Experimental Details Behind Proteomic Data

1115 Here we provide a brief summary of the experiments behind each proteomic data set. The pur-
 1116 pose of this section is to identify how the authors arrived at absolute protein abundances. In
 1117 the following section (Section Summary of Proteomic Data) we will then provide a summary of
 1118 the final protein abundance measurements that were used throughout the main text. Table 1 pro-
 1119 vides an overview of the publications we considered. These are predominately mass spectrometry-
 1120 based, with the exception of the work from Li *et al.* (2014) which used ribosomal profiling, and the
 1121 fluorescence-based counting done in Taniguchi *et al.* (2010).

1122 Fluorescence based measurements

1123 In the work of Taniguchi *et al.* (2010), the authors used a chromosomal YFP fusion library where
 1124 individual strains have a specific gene tagged with a YFP-coding sequence. 1018 of their 1400
 1125 attempted strains were used in the work. A fluorescence microscope was used to collect cellular
 1126 YFP intensities across all these strains. Through automated image analysis, the authors normalized
 1127 intensity measurements by cell size to account for the change in size and expression variability
 1128 across the cell cycle. Following correction of YFP intensities for cellular autofluorescence, final
 1129 absolute protein levels were determined by a calibration curve with single-molecule fluorescence
 1130 intensities. This calibration experiment was performed separately using a purified YFP solution.

1131 Ribosomal profiling measurements

1132 The work of Li *et al.* (2014) takes a sequencing based approach to estimate protein abundance. Ri-
 1133 bosomal profiling, which refers to the deep sequencing of ribosome-protected mRNA fragments,
 1134 can provide a quantitative measurement of the protein synthesis rate. As long as the protein life-
 1135 time is long relative to the cell doubling time, it is possible to estimate absolute protein copy num-
 1136 bers. The absolute protein synthesis rate has units of proteins per generation, and for stable pro-
 1137 teins will also correspond to the protein copy number per cell.

1138 In the experiments, ribosome-protected mRNA is extracted from cell lysate and selected on
 1139 a denaturing polyacrylamide gel for deep sequencing (15–45 nt long fragments collected and se-
 1140 quenced by using an Illumina HiSeq 2000 in Li *et al.* (2014)). Counts of ribosome footprints from the
 1141 sequencing data were then corrected empirically for position-dependent biases in ribosomal den-
 1142 sity across each gene, as well as dependencies on specific sequences including the Shine-Dalgarno
 1143 sequence. These data-corrected ribosome densities represent relative protein synthesis rates. Ab-
 1144 solute protein synthesis rates are obtained by multiplying the relative rates by the total cellular

1145 protein per cell. The total protein per unit volume was determined with the Lowry method to
 1146 quantify total protein, calibrated against bovine serum albumin (BSA). By counting colony-forming
 1147 units following serial dilution of their cell cultures, they then calculated the total protein per cell.

1148 Mass spectrometry measurements

1149 Perhaps not surprisingly, the data is predominantly mass spectrometry based. This is largely due
 1150 to tremendous improvements in the sensitivity of mass spectrometers, as well as improvements in
 1151 sample preparation and data analysis pipelines. It is now a relatively routine task to extract protein
 1152 from a cell and quantify the majority of proteins present by shotgun proteomics. In general, this
 1153 involves lysing cells, enzymatically digesting the proteins into short peptide fragments, and then
 1154 introducing them into the mass spectrometer (e.g. with liquid chromatography and electrospray
 1155 ionization), which itself can have multiple rounds of detection and further fragmentation of the
 1156 peptides.

1157 Most quantitative experiments rely on labeling protein with stable isotopes, which allow multi-
 1158 ple samples to be measured together by the mass spectrometer. By measuring samples of known
 1159 total protein abundance simultaneously (i.e. one sample of interest, and one reference), it is pos-
 1160 sible to determine relative protein abundances. Absolute protein abundances can be estimated
 1161 following the same approach used above for ribosomal profiling, which is to multiply each relative
 1162 abundance measurement by the total cellular protein per cell. This is the approach taken by *Val-*
1163 gepea et al. (2013) and *Peebo et al. (2015)*, with relative protein abundances determined based on
 1164 the relative peptide intensities (label free quantification 'LFQ' intensities). For the data of *Valgepea*
1165 et al. (2013), total protein per cell was determined by measuring total protein by the Lowry method,
 1166 and counting colony-forming units following serial dilution. For the data from *Peebo et al. (2015)*,
 1167 the authors did not determine cell quantities and instead report the cellular protein abundances
 1168 in protein per unit volume by assuming a mass density of 1.1 g/ml, with a 30% dry mass fraction.

1169 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide
 1170 fragments of known abundance. These can serve as a direct way to calibrate mass spectrometry
 1171 signal intensities to absolute mass. This is the approach taken by *Schmidt et al. (2016)*. In addition
 1172 to a set of shotgun proteomic measurements to determine proteome-wide relative abundances,
 1173 the authors also performed absolute quantification of 41 proteins covering over four orders of
 1174 magnitude in cellular abundance. Here, a synthetic peptide was generated for each of the proteins,
 1175 doped into each protein sample, and used these to determine absolute protein abundances of the
 1176 41 proteins. These absolute measurements, determined for every growth condition, were then
 1177 used as a calibration curve to convert proteomic-wide relative abundances into absolute protein
 1178 abundance per cell. A more extensive discussion of the *Schmidt et al. (2016)* data set can be found
 1179 in Section Additional Considerations of Schmidt *et al.* Data Set.

1180 Summary of Proteomic Data

1181 In the work of the main text we only used the data from *Valgepea et al. (2013)*; *Li et al. (2014)*; *Peebo*
1182 et al. (2015); *Schmidt et al. (2016)*. As shown in *Figure 13(A)*, the reported total protein abundances
 1183 in the work of *Taniguchi et al. (2010)* and *Soufi et al. (2015)* differed quite substantially from the
 1184 other work. For the work of *Taniguchi et al. (2010)* this is in part due to a lower coverage in total
 1185 proteomic mass quantified, though we also noticed that most proteins appear undercounted when
 1186 compared to the other data.

1187 *Figure 13(B)* summarizes the total protein mass for each data point in our final compiled data
 1188 set. We note that protein abundances were all scaled so they followed a common growth rate-
 1189 dependent change in total protein mass. While our inclination initially was to leave reported copy
 1190 numbers untouched, a notable discrepancy in the scaling total protein per cell between *Schmidt*
 1191 *et al. (2016)* and the other data sets forced us to dig deeper into those measurements (compare
 1192 *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 13(A)*). The particular trend in *Schmidt et al.*
 1193 *(2016)* appears to be due to assumptions of cell size and we provide a more extensive discussion

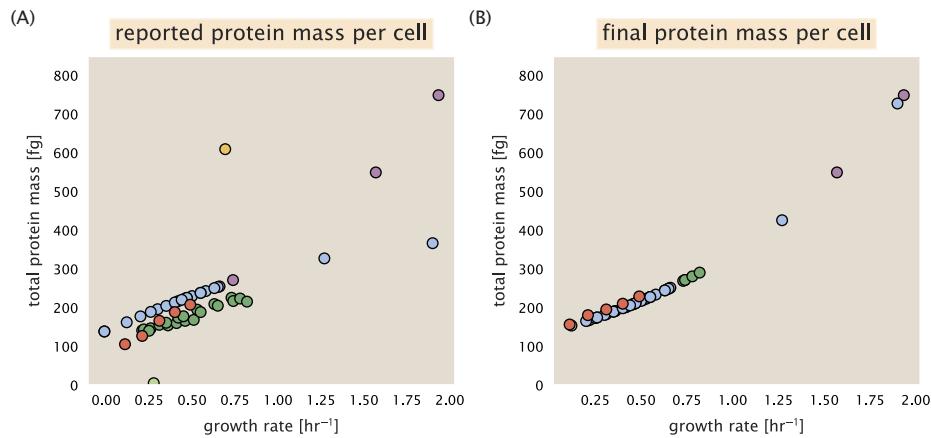


Figure 13. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as original reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al.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

and analysis of that data set in section Additional Considerations of Schmidt *et al.* Data Set. As a compromise, and in an effort to treat all data equally, we instead scaled all protein abundance values to a data-driven estimate of total protein per cell. Here we used cell size measurements from *Si et al. (2017, 2019)*, and an estimate of total protein content through expected dry mass. Total protein per cell was estimated using available data on total DNA, RNA, and protein from *Basan et al. (2015)*; *Dai et al. (2016)*, which account for the majority of dry mass in the cell. We consider these details in sections Estimation of Cell Size and Surface Area and Estimation of Total Protein Content per Cell that follows.

Lastly, in **Figure 14** we show the total proteomic coverage and overlap of proteins quantified across each data set. Here we have used an UpSet diagram (*Lex et al., 2014*) to compare the data. Overall, the overlap in quantified proteins is quite high, with 1157 proteins quantified across all data sets. The sequencing based approach of *Li et al. (2014)* has substantially higher coverage compared to the mass spectrometry data sets (3394 genes versus the 2041 genes quantified in the work of *Schmidt et al. (2016)*). However, in terms of total protein mass, the data from *Li et al. (2014)*; *Schmidt et al. (2016)*; *Peebo et al. (2015)* each quantify roughly equivalent total protein mass. An exception to this is in the data from *Valgepea et al. (2013)*, where we find that the total protein quantified in *Valgepea et al. (2013)* is 90-95 % of the total protein mass (when using the data from *Schmidt et al. (2016)* as a reference).

1212 Estimation of Cell Size and Surface Area

1213 Since most of the proteomic data sets lack cell size (i.e. volume) measurements, we chose instead
 1214 to use a common estimate of size for any analysis requiring cell size or surface area. Since each
 1215 of the data sets used either K-12 MG1655 or its derivative, BW25113 (from the lab of Barry L. Wan-
 1216 ner; the parent strain of the Keio collection (*Datsenko and Wanner, 2000*; *Baba et al., 2006*)), we
 1217 fit the MG1655 cell size data from the supplemental material of *Si et al. (2017, 2019)* using the
 1218 optimize.curve_fit function from the Scipy python package (*Virtanen et al., 2020*).

1219 The average size measurements from each of their experiments are shown in Figure **Figure 15**,
 1220 with cell length and width shown in (A) and (B), respectively. The length data was well described by

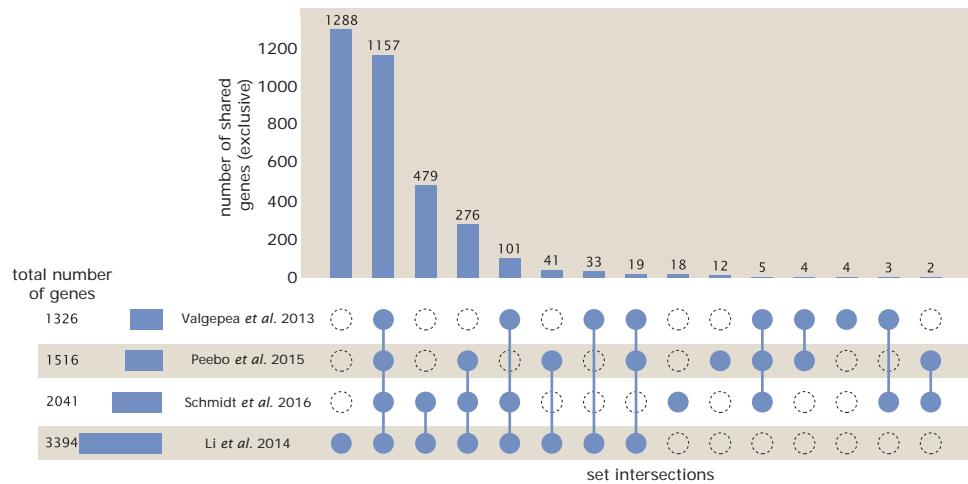


Figure 14. Comparison of proteomic coverage across different data sets. An UpSet diagram (Lex et al., 2014) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of Valgepea et al. (2013); Li et al. (2014); Schmidt et al. (2016); Peebo et al. (2015).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

the exponential function $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, while the width data was well described by $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. In order to estimate cell size we take the cell as a cylinders with two hemispherical ends (Si et al., 2017; Basan et al., 2015). Specifically, cell size is estimated from,

$$V = \pi \cdot r^2 \cdot (l - 2r/3), \quad (12)$$

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

$$S = \eta \cdot \pi \left(\frac{\eta \cdot \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V^{2/3}, \quad (13)$$

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

Estimation of Total Protein Content per Cell

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

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

Assuming that the protein, DNA, and RNA account for 90 % of the total dry mass, the protein mass can then determined by first subtracting the experimentally measured DNA mass, and then

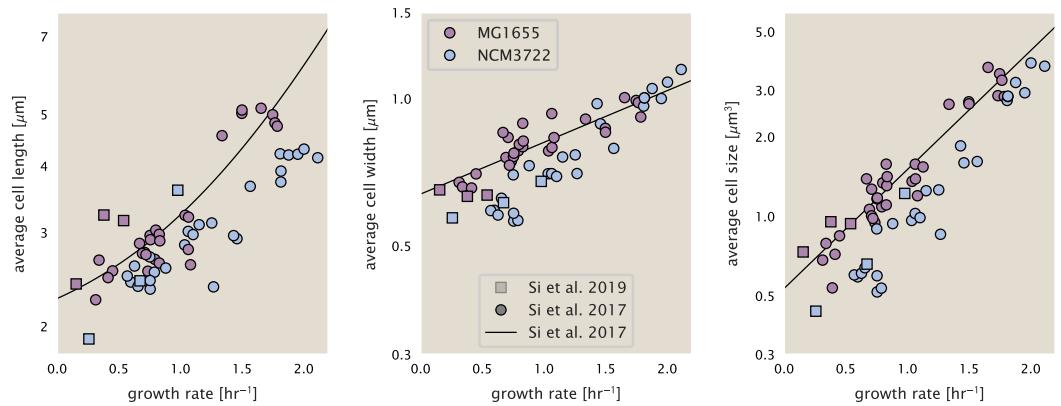


Figure 15. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 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: $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, width: $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. (C) Cell size, V , was calculated as cylinders with two hemispherical ends (Equation 12). The MG1655 strain data gave a best fit of $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$.

1246 using the experimental estimate of the RNA to protein ratio. The total protein per cell is will be
 1247 related to the summed RNA and protein mass by,

$$M_P = \frac{[M_P + M_{RNA}]}{1 + (RP_{ratio})}. \quad (14)$$

1248 (RP_{ratio}) refers to the RNA to protein ratio as measured by Dai et al.. In Figure **Figure 16(C)** we plot
 1249 the estimated cellular concentrations for protein, DNA, and RNA from these calculations, and in
 1250 Figure **Figure 16(D)** we plot their total expected mass per cell. This later quantity is the growth
 1251 rate-dependent total protein mass that was used to estimate total protein abundance across all
 1252 data sets (and summarized in **Figure 13(B)**).

1253 **Estimating Volume and Number of Amino Acids from Ribosome Copy Number**

1254 Towards the end of the main text, we examine a coarse-grained model of nutrient-limited growth. A
 1255 key point in our analysis was to consider how elongation rate r_i and growth rate λ vary with respect
 1256 to the experimentally observed changes in cell size, total number of peptide bonds per cell N_{pep} ,
 1257 and ribosomal content. In order to do maintain parameters in line with the experimental data, but
 1258 otherwise allow us to explore the model, we performed a phenomenological fit of N_{pep} and V as
 1259 a function of the measured ribosomal copy number R . As has been described in the preceding
 1260 sections of this supplement, we estimate cell volume for each growth condition using the size
 1261 measurements from **Si et al. (2017, 2019)**, and N_{pep} is approximated by taking the total protein
 1262 mass and dividing this number by the average mass of an amino acid, 110 Da (BNID: 104877).

1263 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression
 1264 of the log transform of these parameters as a function of the log transform of the ribosome copy
 1265 number. Using optimization by minimization, we estimated the best-fit values of the intercept and
 1266 slope for each regression. ?? shows the result of each regression as a dashed line, given in terms
 1267 of the linear-scaled data.

1268 **Additional Considerations of Schmidt et al. Data Set**

1269 While the data set from **Schmidt et al. (2016)** remains a heroic effort that our labs continue to
 1270 return to as a resource, there were steps taken in their calculation of protein copy number that
 1271 we felt needed further consideration. In particular, the authors made an assumption of constant

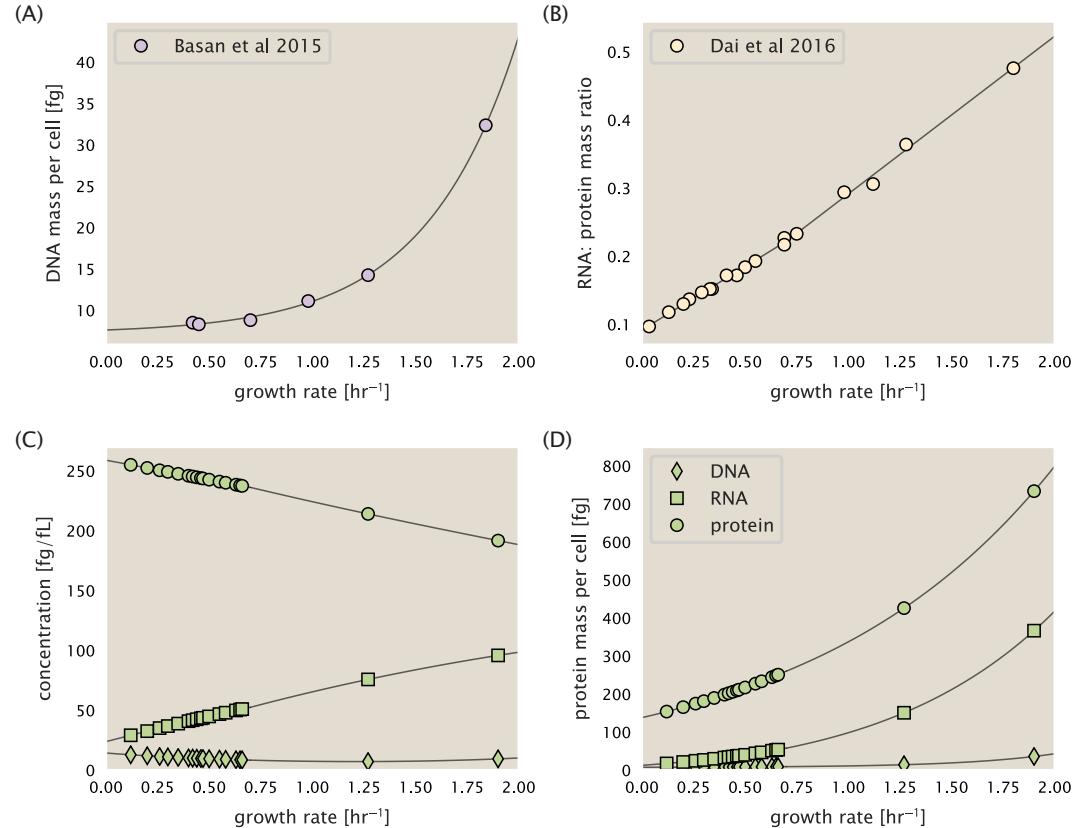


Figure 16. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23\lambda} + 7.2$ fg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

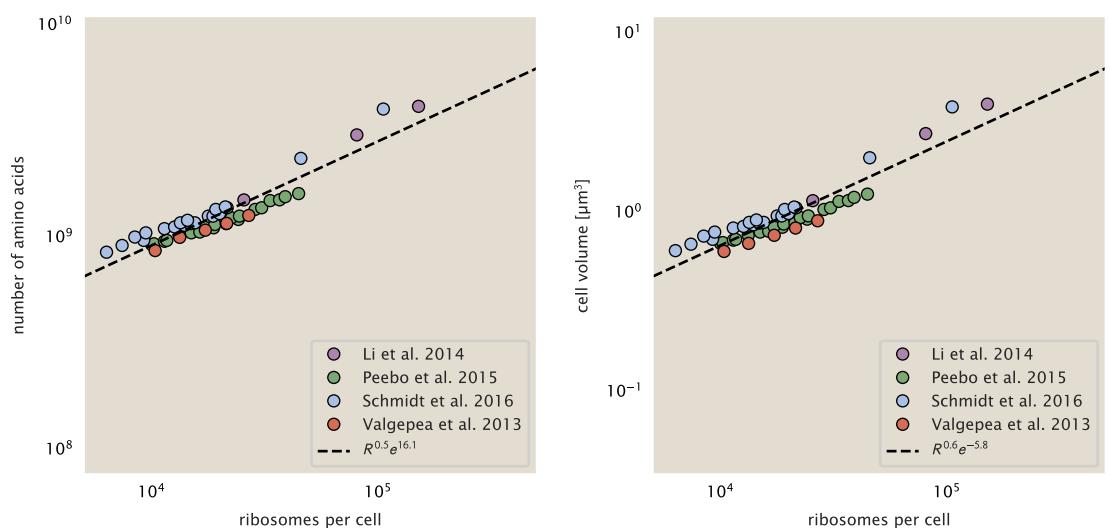


Figure 17. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows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

1272 cellular protein concentration across all growth conditions and used measurements of cell volume
1273 that appear inconsistent with an expected exponential scaling of cell size with growth rate that is
1274 well-documented in *E. coli* (**Schaechter et al. (1958); Taheri-Araghi et al. (2015); Si et al. (2017)**).

1275 We begin by looking at their cell volume measurements, which are shown in blue in Figure
1276 **Figure 18**. As a comparison, we also plot cell sizes reported in three other recent papers: measure-
1277 ments from Taheri-Araghi *et al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from
1278 Basan *et al.* come from the lab of Terence Hwa. Each set of measurements used microscopy and
1279 cell segmentation to determine the length and width, and then calculated cell size by treating the
1280 cell as a cylinder with two hemispherical ends, as we considered in the previous section. While there
1281 is notable discrepancy between the two research groups, which are both using strain NCM3722,
1282 Basan *et al.* found that this came specifically from uncertainty in determining the cell width. This
1283 is prone to inaccuracy given the small cell size and optical resolution limits (further described in
1284 their supplemental text). Perhaps the more concerning point is that while each of these alternative
1285 measurements show an exponential increase in cell size at faster growth rates, the measurements
1286 used by Schmidt *et al.* appear to plateau. This resulted in an analogous trend in their final re-
1287 ported total cellular protein per cell as shown in Figure **Figure 19** (purple data points), and is in
1288 disagreement with other measurements of total protein at these growth rates (**Basan et al. 2015**).

1289 Since it is not obvious how measurements of cell size influenced their reported protein abun-
1290 dances, in the following subsections we begin by considering this calculation. We then consider
1291 three different approaches to estimate the growth-rate dependent total protein mass to compare
1292 with those values reported from **Schmidt et al. (2016)**. The results of this are summarized in **Fig-**
1293 **ure 18(B)**, with the original values from both **Schmidt et al. (2016)** and **Li et al. (2014)** shown in
1294 **Figure 18(A)** for reference. For most growth conditions, we find that total protein per cell is still in
1295 reasonable agreement. However, for the fastest growth conditions, with glycerol + supplemented
1296 amino acids, and LB media, all estimates are substantially higher than those originally reported.
1297 This is the main reason why we chose to readjusted protein abundance as shown in **Figure 13(B)**
1298 (with the calculation described in section Estimation of Total Protein Content per Cell).

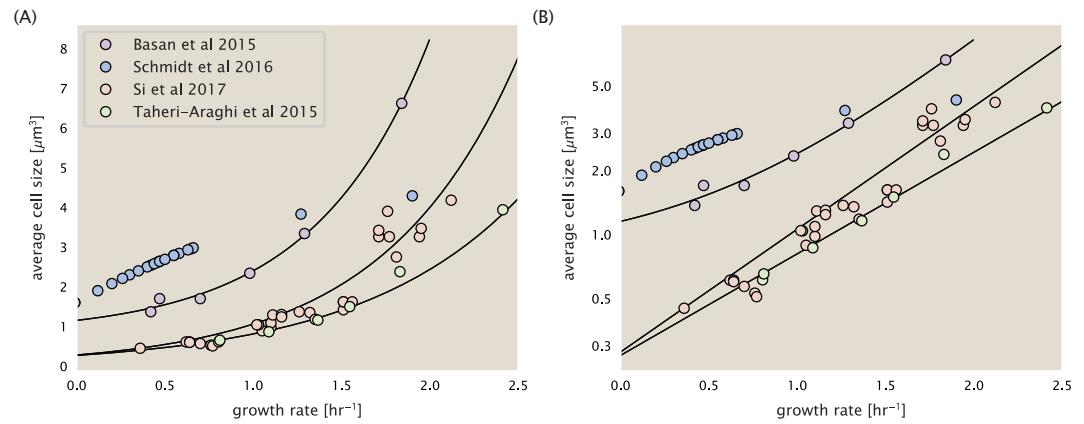


Figure 18. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are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

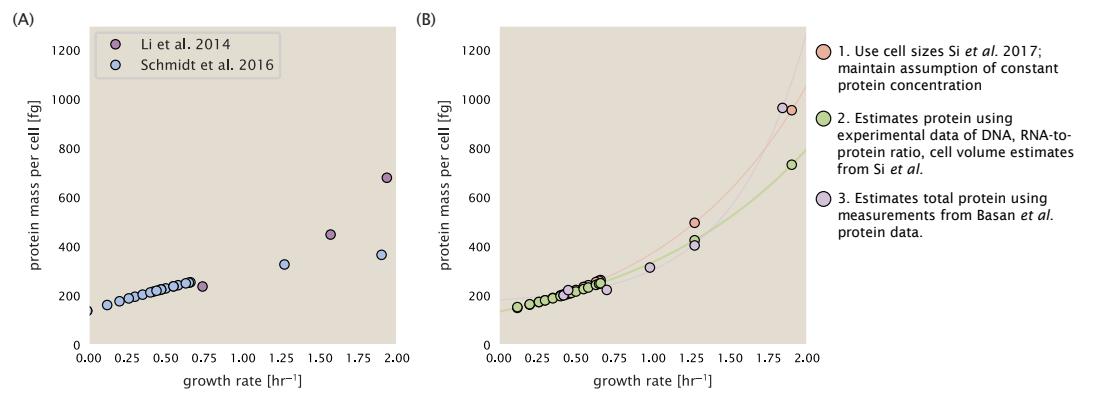


Figure 19. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt et al. (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. light red: Rescaling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

1299 Effect of cell volume on reported absolute protein abundances

1300 As noted in section Experimental Details Behind Proteomic Data, the authors calculated proteome-
1301 wide protein abundances by first determining absolute abundances of 41 pre-selected proteins,
1302 which relied on adding synthetic heavy reference peptides into their protein samples at known
1303 abundance. This absolute quantitation was performed in replicate for each growth condition. Sep-
1304 arately, the authors also performed a more conventional mass spectrometry measurement for
1305 samples from each growth condition, which attempted to maximize the number of quantified
1306 proteins but only provided relative abundances based on peptide intensities. Finally, using their
1307 41 proteins with absolute abundances already determined, they then created calibration curves
1308 with which to relate their relative intensity to absolute protein abundance for each growth con-
1309 dition. This allowed them to estimate absolute protein abundance for all proteins detected in their
1310 proteome-wide data set. Combined with their flow cytometry cell counts, they were then able to
1311 determine absolute abundance of each protein detected on a per cell basis.

1312 While this approach provided absolute abundances, another necessary step to arrive at total
1313 cellular protein was to account for any protein loss during their various protein extraction steps.
1314 Here the authors attempted to determine total protein separately using a BCA protein assay. In
1315 personal communications, it was noted that determining reasonable total protein abundances by
1316 BCA across their array of growth conditions was particularly troublesome. Instead, they noted
1317 confidence in their total protein measurements for cells grown in M9 minimal media + glucose
1318 and used this as a reference point with which to estimate the total protein for all other growth
1319 conditions.

1320 For cells grown in M9 minimal media + glucose an average total mass of $M_p = 240$ fg per cell was
1321 measured. Using their reported cell volume, reported as $V_{orig} = 2.84$ fl, a cellular protein concentra-
1322 tion of $[M_p]_{orig} = M_p/V_{orig} = 85$ fg/fl. Now, taking the assumption that cellular protein concentration
1323 is relatively independent of growth rate, they could then estimate the total protein mass for all
1324 other growth conditions from,

$$M_{P,i} = [M_p]_{orig} \cdot V_i \quad (15)$$

1325 where $M_{P,i}$ represents the total protein mass per cell and V_i is the cell volume for each growth
1326 condition i as measured in Volkmer and Heinemann, 2011. Here the thinking is that the values
1327 of $M_{P,i}$ reflects the total cellular protein for growth condition i , where any discrepancy from their
1328 absolute protein abundance is assumed to be due to protein loss during sample preparation. The
1329 protein abundances from their absolute abundance measurements noted above were therefore
1330 scaled to their estimates and are shown in Figure 19 (purple data points).

1331 If we instead consider the cell volumes predicted in the work of Si *et al.*, we again need to take
1332 growth in M9 minimal media + glucose as a reference with known total mass, but we can follow
1333 a similar approach to estimate total protein mass for all other growth conditions. Letting $V_{Si_glu} =$
1334 0.6 fl be the predicted cell volume, the cellular protein concentration becomes $[M_p]_{Si} = M_p/V_{Si_glu}$
1335 = 400 fg/fl. The new total protein mass per cell can then be calculated from,

$$M'_{P,i} = [M_p]_{Si} \cdot V_{Si,i} \quad (16)$$

1336 where $M'_{P,i}$ is the new protein mass prediction, and $V_{Si,i}$ refers to the new volume prediction for
1337 each condition i . These are shown as red data points in Figure 19(B).

1338 Relaxing assumption of constant protein concentration across growth conditions

1339 We next relax the assumption that cellular protein concentration is constant and instead, attempt
1340 to estimate it using experimental data. Here we use the estimation of total protein mass per cell
1341 detailed in section Estimation of Total Protein Content per Cell for all data points in the Schmidt
1342 *et al.* (2016) data set. The green data points in Figure 19(B) show this prediction, and this represents
1343 the approach used to estimate total protein per cell for all data sets.

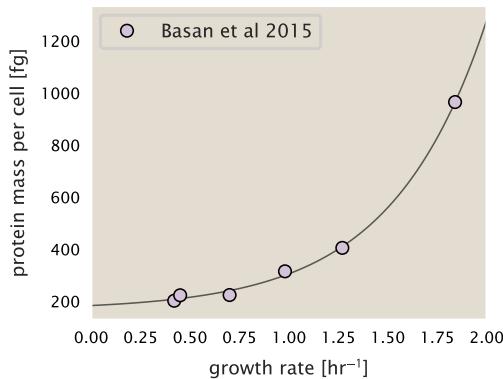


Figure 20. Total cellular protein reported in Basan *et al.* 2015. Measured protein mass as a function of growth rate as reproduced from Basan *et al.* 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given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

1344 Comparison with total protein measurements from Basan *et al.* 2015.

1345 One of the challenges in our estimates in the preceding sections is the need to estimate protein
1346 concentration and cell volumes. These are inherently difficult to accurately due to the small size
1347 of *E. coli*. Indeed, for all the additional measurements of cell volume included in Figure **Figure 18**, no
1348 measurements were performed for cells growing at rates below $0.5\ hr^{-1}$. It therefore remains to be
1349 determined whether our extrapolated cell volume estimates are appropriate, with the possibility
1350 that the logarithmic scaling of cell size might break down for slower growth.

1351 In our last approach we therefore attempt to estimate total protein using experimental data
1352 that required no estimates of concentration or cell volume. Specifically, in the work of Basan *et al*,
1353 the authors measured total protein per cell for a broad range of growth rates (reproduced in Figure
1354 **Figure 20**). These were determined by first measuring bulk protein from cell lysate, measured by
1355 the colorimetric Biuret method (*You et al. (2013)*), and then abundance per cell was calculated from
1356 cell counts from either plating cells or a Coulter counter. While it is unclear why Schmidt *et al.* was
1357 unable to take a similar approach, the results from Basan *et al* appear more consistent with our
1358 expectation that cell mass will increase exponentially with faster growth rates. In addition, although
1359 they do not consider growth rates below about $0.5\ hr^{-1}$, it is interesting to note that the protein
1360 mass per cell appears to plateau to a minimum value at slow growth. In contrast, our estimates
1361 using cell volume so far have predicted that total protein mass should continue to decrease slightly
1362 for slower growing cells. By fitting this data to an exponential function dependent on growth rate,
1363 we could then estimate the total protein per cell for each growth condition considered by *Schmidt*
1364 *et al. (2016)*. These are plotted as red data points in **Figure 19(B)**.

1365 Calculation of Complex Abundance

1366 All protein data quantified the abundance of individual proteins per cell. However, this work re-
1367 quires estimates on the abundance of individual protein *complexes*, rather than the copy number
1368 of individual proteins. In this section, we outline the approach we used to annotate proteins as
1369 being part of a macromolecular complex and how we computed their absolute abundances per
1370 cell.

1371 Protein complexes, and proteins individually, often have a variety of names, both longform and
1372 shorthand. As individual proteins can have a variety of different synonyms, we sought to ensure
1373 that each protein annotated in the data sets used the same synonym. To do use, we relied heavily
1374 on the EcoCyc Genomic Database (*Keseler et al., 2017*). Each protein in available data sets included
1375 an annotation of one of the gene name synonyms as well as an accession ID – either a UniProt

1376 or Blattner "b-number". We programmatically matched up individual accession IDs between the
 1377 proteins in different data sets. In cases where accession IDs matched but the gene names were
 1378 different, we manually verified that the gene product was the same between the datasets and
 1379 chose a single synonym. All code used in the data cleaning and unification procedures can be
 1380 found on the associated [GitHub repository] (DOI:XXX) associated with this paper as well as on the
 1381 associated [paper website](#).

1382 With each protein conforming to a single identification scheme, we then needed to identify the
 1383 molecular complexes each protein was a member of. Additionally, we needed to identify how many
 1384 copies of each protein were present in each complex (i.e. the subunit copy number) and compute
 1385 the estimated abundance complex that accounted for fluctuations in subunit stoichiometry. To
 1386 map proteins to complexes, we accessed the EcoCyc *E. coli* database *Keseler et al. (2017)* using
 1387 PathwayTools version 23.0 *Karp et al. (2019)*. With a license for PathWay Tools, we mapped each
 1388 unique protein to its annotated complexes via the BioCyc Python package. As we mapped each
 1389 protein with *all* of its complex annotations, there was redundancy in the dataset. For example,
 1390 ribosomal protein L20 (RplT) is annotated to be a component of the 50S ribosome (EcoCyc complex
 1391 CPLX-03962) as well as a component of the mature 70S ribosome (EcoCyc complex CPLX-03964).

1392 In addition to the annotated complex, we collected information on the stoichiometry of each
 1393 macromolecular complex. For a complex with N_{subunits} protein species, for each protein subunit i
 1394 we first calculate the number of complexes that *could* be formed given the measured protein copy
 1395 numbers per cell,

$$N_{\text{complex}}(\text{subunit } i) = \frac{P_{\text{subunit } i}^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (17)$$

1396 Here, $P_{\text{subunit } i}^{(\text{measured})}$ refers to the measured protein copy number of species i , and m refers to the num-
 1397 ber of monomers present for that protein in the complex. For example, the 70S mature ribosome
 1398 complex has 55 protein components, all of which are present in a single copy except L4 (RplL),
 1399 which is present in 4 copies ($m = 4$). For each ribosomal protein, we then calculate the maximum
 1400 number of complexes that could be formed using **Equation 17**. This example, along with example
 1401 from 5 other macromolecular complexes, can be seen in **Figure 21**.

1402 It is important to note that measurement noise, efficiency of protein extraction, and differences
 1403 in protein stability will mean that the precise value of each calculation will be different for each
 1404 component of a given complex. Thus, to report the total complex abundance, we use the arithmetic
 1405 mean of across all subunits in the complex,

$$\langle N_{\text{complex}} \rangle = \frac{1}{N_{\text{subunits}}} \sum_i^{N_{\text{subunits}}} \frac{P_i^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (18)$$

1406 in **Figure 21**, we show this mean value as a grey line for a variety of different complexes. Addi-
 1407 tionally, we have built an interactive figure accessible on the [paper website](#) where the validity of
 1408 this approach can be examined for any complex with more than two subunits (thus, excluding
 1409 monomers and dimers).

1410 Extending Estimates to a Continuum of Growth Rates

1411 In the main text, we considered a standard stopwatch of 5000 s to estimate the abundance of
 1412 the various protein complexes considered. In addition to point estimates, we also showed the
 1413 estimate as a function of growth rate as transparent grey curves. In this section, we elaborate
 1414 on this continuum estimate, giving examples of estimates that scale with either cell volume, cell
 1415 surface area, or number of origins of replication.

1416 Estimation of the total cell mass

1417 For many of the processes estimated in the main text we relied on a cellular dry mass of ≈ 300
 1418 fg from which we computed elemental and protein fractions using knowledge of fractional com-
 1419 position of the dry mass. At modest growth rates, such as the 5000 s doubling time used in the

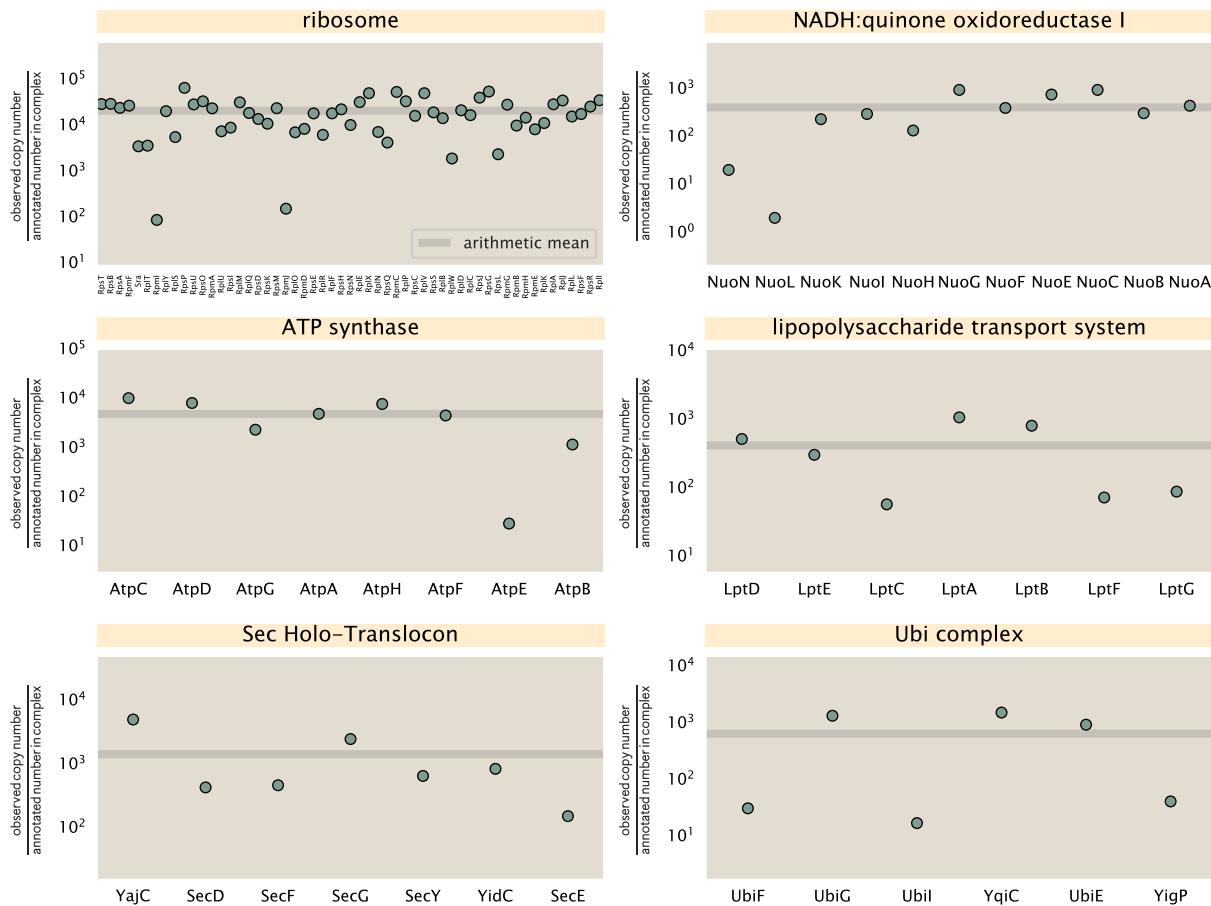


Figure 21. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds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

1420 main text, this is a reasonable number to use as the typical cell mass is ≈ 1 pg and *E. coli* cells can
 1421 approximated as 70% water by volume. However, as we have shown in the preceding sections, the
 1422 cell size is highly dependent on the growth rate. This means that a dry mass of 300 fg cannot be
 1423 used reliably across all growth rates.

1424 Rather, using the phenomenological description of cell volume scaling exponentially with growth
 1425 rate, and using a rule-of-thumb of a cell buoyant density of ≈ 1.1 pg / fL (BNID: 103875), we can
 1426 calculate the cell dry mass across a range of physiological growth rates as

$$m_{\text{cell}} \approx \rho V(\lambda) \approx pae^{\lambda * b} \quad (19)$$

1427 where a and b are constants with units of μm^3 and hr, respectively. The value of these constants can
 1428 be estimated from the careful volume measurements performed by *Siet al.* (2017?), as considered
 1429 in Appendix Estimation of Cell Size and Surface Area earlier.

1430 Complex Abundance Scaling With Cell Volume

1431 Several of the estimates performed in the main text are implicitly dependent on the cell volume.
 1432 This includes processes such as ATP utilization and, most prominently, the transport of nutrients,
 1433 whose demand will be proportional to the volume of the cell. Of the latter, we estimated the num-
 1434 ber of transporters that would be needed to shuttle enough carbon, phosphorus, and sulfur across
 1435 the membrane to build new cell mass. To do so, we used elemental composition measurements
 1436 combined with a 300 fg cell dry mass to make the point estimate. As we now have a means to esti-
 1437 mate the total cell mass as a function of volume, we can generalize these estimates across growth
 1438 rates.

1439 Rather than discussing the particular details of each transport system, we will derive this scaling
 1440 expression in very general terms. Consider that we wish to estimate the number of transporters
 1441 for some substance X , which has been measured to be made up some fraction of the dry mass,
 1442 θ_X . If we assume that, irrespective of growth rate, the cell dry mass is relatively constant (*Basan*
 1443 *et al.*, 2015) and $\approx 30\%$ of the total cell mass, we can state that the total mass of substance X as a
 1444 function of growth rate is

$$m_X \approx 0.3 \times \rho V(\lambda) \theta_X, \quad (20)$$

1445 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in **Equation 19**. To convert
 1446 this to the number of units N_X of substance X in the cell, we can use the formula weight w_X of a
 1447 single unit of X in conjunction with **Equation 20**,

$$N_X \approx \frac{m_X}{w_X}. \quad (21)$$

1448 To estimate the number of transporters needed, we make the approximation that loss of units
 1449 of X via diffusion through porins or due to the permeability of the membrane is negligible and that
 1450 a single transporter complex can transport substance X at a rate r_X . As this rate r_X is in units of
 1451 X per time per transporter, we must provide a time window over which the transport process can
 1452 occur. This is related to the cell doubling time τ , which can be calculated from the the growth rate
 1453 λ as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized transport scaling relation
 1454 of

$$N_{\text{transporters}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_X}{w_X r_X \tau}. \quad (22)$$

1455 This function is used to draw the continuum estimates for the number of transporters seen in
 1456 Figures 2 and 3 as transparent grey curves. Occasionally, this continuum scaling relationship will
 1457 not precisely agree with the point estimate outlined in the main text. This is due to the choice of \approx
 1458 300 fg total dry mass per cell for the point estimate, whereas we considered more precise values
 1459 of cell mass in the continuum estimate. We note, however, that both this scaling relation and the
 1460 point estimates are meant to describe the order-of-magnitude observed, and not the predict the
 1461 exact values of the abundances.

Equation 22 is a very general relation for processes where the cell volume is the "natural variable" of the problem. This means that, as the cell increases in volume, the requirements for substance X also scale with volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the dry mass attributable to the substance, and the formula mass of the substance is known, **Equation 22** can be used to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell, **Equation 22** can be slightly modified to the form

$$N_{\text{ATP synthases}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_{\text{protein}} N_{\text{ATP}}}{w_{AA} r_{\text{ATP}} \tau}, \quad (23)$$

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

A Relation for Complex Abundance Scaling With Surface Area

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

Number of Lipids

To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as 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 of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth rate. Using these approximations combined with **Equation 13**, and recognizing that each membrane is composed of two leaflets, we can compute the number of lipids as a function of growth rate as

$$N_{\text{lipids}}(\lambda) \approx \frac{4 \text{ leaflets} \times 0.4 \times \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{A_{\text{lipid}}} \quad (24)$$

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

Number of Murein Monomers

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

$$N_{\text{murein monomers}}(\lambda) \approx \frac{\rho_{\text{pg}} w \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{m_{\text{murein}}}, \quad (25)$$

where ρ_{pg} is the density of peptidoglycan.

1498 Complex Abundance Scaling With Number of Origins, and rRNA Synthesis

1499 While the majority of our estimates hinge on the total cell volume or surface area, processes related
1500 to the central dogma, namely DNA replication and synthesis of rRNA, depend on the number of
1501 chromosomes present in the cell. As discussed in the main text, the ability of *E. coli* to parallelize the
1502 replication of its chromosome by having multiple active origins of replication is critical to synthesize
1503 enough rRNA, especially at fast growth rates. Derived in *Si et al. (2017)* and reproduced in the main
1504 text and Appendix Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$ below, the average number of origins of
1505 replication at a given growth rate can be calculated as

$$\langle \#ori \rangle \approx 2^{t_{cyc} \lambda / \ln 2} \quad (26)$$

1506 where t_{cyc} is the total time of replication and division. We can make the approximation that $t_{cyc} \approx$
1507 70 min, which is the time it takes two replisomes to copy an entire chromosome.

1508 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of
1509 replication. Thus, at a given growth rate λ , the average dosage of rRNA operons per cell D_{rRNA} is

$$D_{rRNA}(\lambda) \approx N_{rRNA \text{ operons}} \times 2^{t_{cyc} \lambda / \ln 2}. \quad (27)$$

1510 This makes the approximation that *all* rRNA operons are localized around the origin. In reality,
1511 the operons are some distance away from the origin, making **Equation 27** an approximation
1512 (*Dennis et al., 2004*).

1513 In the main text, we stated that at a growth rate of 0.5 hr^{-1} , there is ≈ 1 chromosome per cell.
1514 While a fair approximation, **Equation 26** illustrates that is not precisely true, even at slow growth
1515 rates. In estimating the number of RNA polymerases as a function of growth rate, we consider that
1516 regardless of the number of rRNA operons, they are all sufficiently loaded with RNA polymerase
1517 such that each operon produces one rRNA per second. Thus, the total number of RNA polymerase
1518 as a function of the growth rate can be calculated as

$$N_{\text{RNA polymerase}}(\lambda) \approx L_{\text{operon}} D_{rRNA} \rho_{\text{RNA polymerase}}, \quad (28)$$

1519 where L_{operon} is the total length of an rRNA operon (≈ 4500 bp) and $\rho_{\text{RNA polymerase}}$ is packing density
1520 of RNA polymerase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

1521 Calculation of active ribosomal fraction.

1522 In the main text we used the active ribosomal fraction f_a that was reported in the work of *Dai*
1523 *et al. (2016)* to estimate the active ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We
1524 lacked any specific model to consider how f_a should vary with growth rate, and instead find that
1525 the data is well-approximated by fitting to an exponential curve ($f_a = -0.889 e^{4.6 \cdot \lambda} + 0.922$; dashed
1526 line in inset of **Figure 10(C)**). We use this function to estimate f_a for each of the data points shown
1527 in **Figure 10(C)**.

1528 Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.

1529 *E. coli* shows robust scaling of cell size with the average the number of origins $\langle \#ori \rangle$ per cell (*Si*
1530 *et al., 2017*). Since protein makes up a majority of the cell's dry mass, the change in cell size is also a
1531 reflection of the changes in proteomic composition and total abundance across growth conditions.
1532 Given the potential constraints on rRNA synthesis and changes in ribosomal copy number with
1533 $\langle \#ori \rangle$, it becomes important to also consider how protein copy numbers vary with the state of
1534 chromosomal replication. This is particularly true when trying to make sense of the changes in
1535 ribosomal fraction and growth-rate dependent changes in proteomic composition at a mechanistic
1536 level. As considered in the main text, it is becoming increasingly apparent that regulation through
1537 the secondary messengers (p)ppGpp may act to limit DNA replication and also reduce ribosomal
1538 activity in poorer nutrient conditions. In this context, both $\langle \#ori \rangle$, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio
1539 become important parameters to consider and keep tract of. An increase in $\langle \#ori \rangle / \langle \#ter \rangle$ ratio

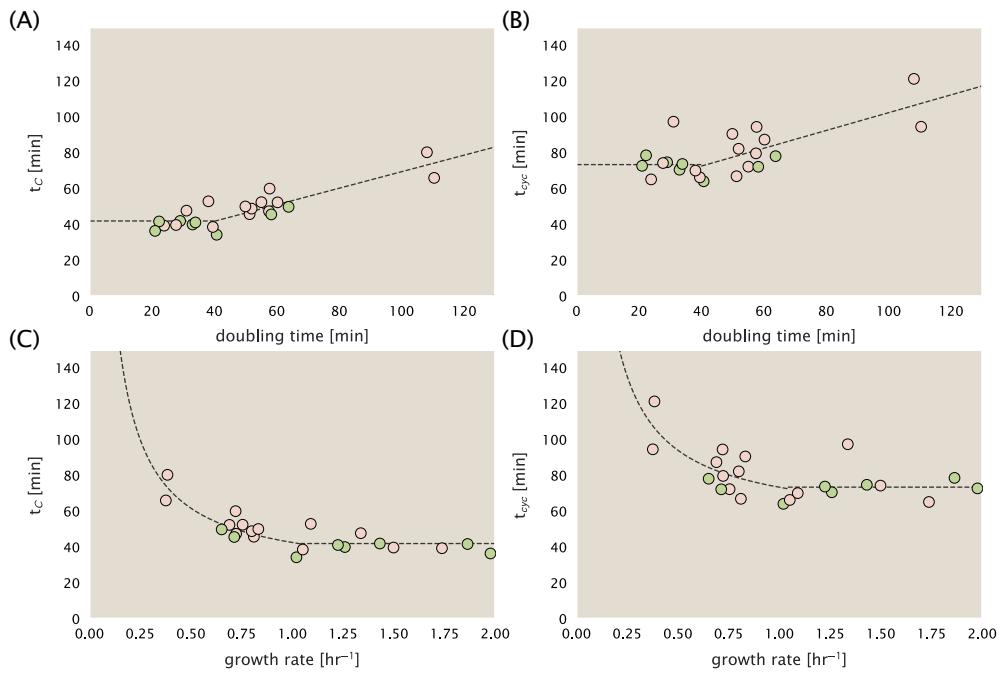


Figure 22. Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling times (rich media), t_C and t_{cyc} are assumed constant ($t_C = 42$ minutes, $t_{\text{cyc}} = 73$ 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, $t_{\text{cyc}} = 0.50 \tau + 52.7$ minutes). (C) and (D) plot the same data as in (A) and (B), but as a function of growth rate, given by $\lambda = \ln(2)/\tau$.

in particular, causes a relatively higher gene dosage in rRNA and r-protein genes due to skew in genes near the origin, where the majority of these are located

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

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 $t_{\text{cyc}} = t_C + t_D$ 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.

References

- 1563** Abelson, H., Johnson, L., Penman, S., and Green, H. (1974). Changes in RNA in relation to growth of the fibroblast:
1564 II. The lifetime of mRNA, rRNA, and tRNA in resting and growing cells. *Cell*, 1(4):161–165.
- 1565**
- 1566** Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose
1567 sugars in *Escherichia coli*. *BMC Systems Biology*, 8(1):133.
- 1568** Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes
1569 studied by ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1570** Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S.,
1571 Eppig, J. T., Harris, M. A., Hill, D. P., Isbel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E.,
1572 Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1573**
- 1574** Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B.
1575 (2000). A Model for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends:
1576 DNA Triggers A Change In Binding Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*,
1577 275(4):3006–3015.
- 1578** Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4
1579 as a NH₃ Channel. *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1580** Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L.,
1581 and Mori, H. (2006). Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio
1582 collection. *Molecular Systems Biology*, 2(1):2460.
- 1583** Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by
1584 increased protein synthesis. *Molecular Systems Biology*, 11(10):836.
- 1585** Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and
1586 Bioengineering*, 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/10.1002/bit.260180107>.
- 1587** Belliveau, N. M., Barnes, S. L., Ireland, W. T., Jones, D. L., Sweredoski, M. J., Moradian, A., Hess, S., Kinney, J. B.,
1588 and Phillips, R. (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional
1589 regulation in bacteria. *Proceedings of the National Academy of Sciences*, 115(21):E4796–E4805.
- 1590** Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009). Absolute
1591 metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical
1592 Biology*, 5(8):593–599.
- 1593** Birnbaum, L. S. and Kaplan, S. (1971). Localization of a Portion of the Ribosomal RNA Genes in *Escherichia coli*.
1594 *Proceedings of the National Academy of Sciences*, 68(5):925–929.
- 1595** Booth, I. R., Mitchell, W. J., and Hamilton, W. A. (1979). Quantitative analysis of proton-linked transport systems.
1596 The lactose permease of *Escherichia coli*. *Biochemical Journal*, 182(3):687–696.
- 1597** Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at
1598 Different Exponential Growth Rates. *EcoSal Plus*, 3(1).
- 1599** Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria.
1600 *Nature Reviews Microbiology*, 14(10):638–650.
- 1601** Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator.
1602 *bioRxiv*, 266:2020.06.16.154187.
- 1603** Catherwood, A. C., Lloyd, A. J., Tod, J. A., Chauhan, S., Slade, S. E., Walkowiak, G. P., Galley, N. F., Punekar, A. S.,
1604 Smart, K., Rea, D., Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stere-
1605 ochemical Control of Peptidoglycan Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of
1606 the American Chemical Society*, 142(11):5034–5048.
- 1607** Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in
1608 *Escherichia coli*. *Biochemical Journal*, 117(3):551–562.
- 1609** Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018).
1610 Slowdown of Translational Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*,
1611 9(1):281.

- 1612** Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P.,
1613 and Hwa, T. (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates
1614 during slow growth. *Nature Microbiology*, 2(2):16231.
- 1615** Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12
1616 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1617** Delarue, M., Brittingham, G. P., Pfeffer, S., Surovtsev, I. V., Pinglay, S., Kennedy, K. J., Schaffer, M., Gutierrez,
1618 J. I., Sang, D., Poterewicz, G., Chung, J. K., Plitzko, J. M., Groves, J. T., Jacobs-Wagner, C., Engel, B. D., and Holt,
1619 L. J. (2018). mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning
1620 Crowding. *Cell*, 174(2):338–349.e20.
- 1621** Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems
1622 Biology Approach. *Microbiology and Molecular Biology Reviews*, 68(4):639–668.
- 1623** Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National
1624 Academy of Sciences*, 108(44):17876–17882.
- 1625** Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli*
1626 phosphoenolpyruvate-carbohydrate phosphotransferase system: Peculiarities of regulation and impact on
1627 growth and product formation. *Applied Microbiology and Biotechnology*, 94(6):1483–1494.
- 1628** Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis,
1629 V., and Palsson, B. Ø. (2007). A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that
1630 accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology*, 3(1):121.
- 1631** Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M.,
1632 and Dunny, G. M. (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA
1633 Synthesis Independent of Growth Rates. *mBio*, 11(2):45.
- 1634** Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA
1635 polymerase affair. *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1636** Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–
1637 263.
- 1638** Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death.
1639 *Proceedings of the National Academy of Sciences*.
- 1640** Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S.,
1641 Alquicira-Hernández, K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-
1642 Lira, H., Bonavides-Martínez, C., Pérez-Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes,
1643 A., Hernández-Koutoucheva, A., Moral-Chávez, V. D., Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB
1644 version 9.0: High-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic
1645 Acids Research*, 44(D1):D133–D143.
- 1646** Ge, J., Yu, G., Ator, M. A., and Stubbe, J. (2003). Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class
1647 I Ribonucleotide Reductase. *Biochemistry*, 42(34):10071–10083.
- 1648** Godin, M., Delgado, F. F., Son, S., Grover, W. H., Bryan, A. K., Tzur, A., Jorgensen, P., Payer, K., Grossman, A. D.,
1649 Kirschner, M. W., and Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature
1650 Methods*, 7(5):387–390.
- 1651** Goldman, S. R., Nair, N. U., Wells, C. D., Nickels, B. E., and Hochschild, A. (2015). The primary σ factor in *Es-
1652 cherichia coli* can access the transcription elongation complex from solution *in vivo*. *eLife*, 4:e10514.
- 1653** 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.,
1654 Hu, J., Lippincott-Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal
1655 Interactions at Nanoscale Resolution on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- 1656** Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogen-
1657 esis. *Trends in microbiology*, 26(10):815–832.
- 1658** Harris, R. M., Webb, D. C., Howitt, S. M., and Cox, G. B. (2001). Characterization of PitA and PitB from *Escherichia
1659 coli*. *Journal of Bacteriology*, 183(17):5008–5014.

- 1660** Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into
1661 the role of (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- 1662** Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and
1663 elemental content of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1664** Helmstetter, C. E. and Cooper, S. (1968). DNA synthesis during the division cycle of rapidly growing Escherichia
1665 coli Br. *Journal of Molecular Biology*, 31(3):507–518.
- 1666** Henkel, S. G., Beek, A. T., Steinsiek, S., Stagge, S., Bettenbrock, K., de Mattos, M. J. T., Sauter, T., Sawodny, O.,
1667 and Ederer, M. (2014). Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying
1668 Oxygen Conditions. *PLoS ONE*, 9(9):e107640.
- 1669** Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015).
1670 Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular
1671 Systems Biology*, 11(2):e784–e784.
- 1672** Ingledew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*,
1673 48(3):222–271.
- 1674** Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B.,
1675 and Phillips, R. (2020). Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a
1676 time. *bioRxiv*.
- 1677** Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular
1678 Biology*, 3(3):318–356.
- 1679** Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*.
1680 *The EMBO journal*, 20(17):4952–4963.
- 1681** Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent
1682 progress, and the future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1683** Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonja, E., Mekler, V.,
1684 Ebright, R. H., and Weiss, S. (2005). Retention of Transcription Initiation Factor Σ70 in Transcription Elonga-
1685 tion: Single-Molecule Analysis. *Molecular Cell*, 20(3):347–356.
- 1686** Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M.,
1687 Midford, P. E., Ong, Q., Ong, W. K., Paley, S. M., and Subhraveti, P. (2019). The BioCyc collection of microbial
1688 genomes and metabolic pathways. *Briefings in Bioinformatics*, 20(4):1085–1093.
- 1689** Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., Assad-Garcia, N., Glass, J. I., and
1690 Covert, M. W. (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–
1691 401.
- 1692** Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-
1693 Castro, S., Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-
1694 Gil, M., Subhraveti, P., Velázquez-Ramírez, D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D.
1695 (2017). The EcoCyc database: reflecting new knowledge about *Escherichia coli*K-12. *Nucleic Acids Research*,
1696 45(D1):D543–D550.
- 1697** Khademi, S., O'Connell, J., Remis, J., Robles-Colmenares, Y., Miercke, L. J. W., and Stroud, R. M. (2004). Mechanism
1698 of Ammonia Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*, 305(5690):1587–1594.
- 1699** Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that
1700 enables the use of hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of
1701 Sciences*, 114(33):E6922–E6931.
- 1702** Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and
1703 proteome partition. *Current Opinion in Biotechnology*, 28:96–102.
- 1704** Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacte-
1705 ria. *Cell*, 139(7):1366–1375.
- 1706** Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical
1707 Review Letters*, 125(2):028103.

- 1708 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation
1709 in *E. coli* by Modulating Supercoiling of *oriC*. *mBio*, 10(4):822.
- 1710 Lascu, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and Biomembranes*, 32(3):237–246.
- 1712 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020).
1713 Theoretical investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1714 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting
1715 sets. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1983–1992.
- 1716 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates
1717 reveals principles underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1718 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein
1719 investment in cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1720 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs
1721 Reveal a Carbon Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–
1722 15927.
- 1723 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli*
1724 Glycerol Facilitator GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1725 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of
1726 Sciences*, 112(51):15690–15695.
- 1727 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On. Gene
1728 Expression*. Springer.
- 1729 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource
1730 allocation revealed by condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1731 Mikucki, J. A., Pearson, A., Johnston, D. T., Turchyn, A. V., Farquhar, J., Schrag, D. P., Anbar, A. D., Priscu, J. C., and
1732 Lee, P. A. (2009). A Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–
1733 400.
- 1734 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key num-
1735 bers in molecular and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1736 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and
1737 cellular differentiation. *Growth Symposium*, 9:223–289.
- 1738 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1739 Mooney, R. A., Darst, S. A., and Landick, R. (2005). Sigma and RNA Polymerase: An On-Again, Off-Again Rela-
1740 tionship? *Molecular Cell*, 20(3):335–345.
- 1741 Mooney, R. A. and Landick, R. (2003). Tethering Σ 70 to RNA polymerase reveals high *in vivo* activity of σ factors
1742 and Σ 70-dependent pausing at promoter-distal locations. *Genes & Development*, 17(22):2839–2851.
- 1743 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links
1744 MreB to cell wall synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National
1745 Academy of Sciences*, 112(40):12510–12515.
- 1746 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*,
1747 volume 1. Elsevier.
- 1748 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in
1749 rod-shaped bacteria. *eLife*, 8:642.
- 1750 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*.
1751 *Biochimie*, 119:80–91.
- 1752 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Es-
1753 cherichia coli* with increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.

- 1754 Perdue, S. A. and Roberts, J. W. (2011). σ^{70} -dependent Transcription Pausing in *Escherichia coli*. *Journal of Molecular Biology*, 412(5):782–792.
- 1756 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1758 Ramos, S. and Kaback, H. R. (1977). The relation between the electrochemical proton gradient and active transport in *Escherichia coli* membrane vesicles. *Biochemistry*, 16(5):854–859.
- 1760 Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., and Maranas, C. D. (2012). An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering*, 14(6):687–704.
- 1763 Rogers, H., Perkins, H., and Ward, J. (1980). *Microbial Cell Walls and Membranes*. Chapman and Hall, London.
- 1764 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature microbiology*, 1(11):1–7.
- 1766 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of Bacteriology*, 131(2):505–511.
- 1768 Rudd, S. G., Valerie, N. C. K., and Helleday, T. (2016). Pathways controlling dNTP pools to maintain genome stability. *DNA Repair*, 44:193–204.
- 1770 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–11734.
- 1772 Sánchez-Romero, M. A., Molina, F., and Jiménez-Sánchez, A. (2011). Organization of ribonucleoside diphosphate reductase during multifork chromosome replication in *Escherichia coli*. *Microbiology*, 157(8):2220–2225.
- 1774 Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- 1776 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann, M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 1779 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. *Science*, 330(6007):1099–1102.
- 1781 Sekowska, A., Kung, H.-F., and Danchin, A. (2000). Sulfur Metabolism in *Escherichia coli* and Related Bacteria: Facts and Fiction. *Journal of Molecular Microbiology and Biotechnology*, 2(2):34.
- 1783 Shi, H., Bratton, B. P., Gitai, Z., and Huang, K. C. (2018). How to Build a Bacterial Cell: MreB as the Foreman of *E. coli* Construction. *Cell*, 172(6):1294–1305.
- 1785 Si, F., Le Treut, G., Sauls, J. T., Vadia, S., Levin, P. A., and Jun, S. (2019). Mechanistic Origin of Cell-Size Control and Homeostasis in Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 1787 Si, F., Li, D., Cox, S. E., Sauls, J. T., Azizi, O., Sou, C., Schwartz, A. B., Erickstad, M. J., Jun, Y., Li, X., and Jun, S. (2017). Invariance of Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- 1789 Sirko, A., Zatyka, M., Sadowy, E., and Hulanicka, D. (1995). Sulfate and thiosulfate transport in *Escherichia coli* K-12: Evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins. *Journal of Bacteriology*, 177(14):4134–4136.
- 1792 Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews*, 40(1):133–159.
- 1794 Soler-Bistué, A., Aguilar-Pierlé, S., García-Garcérá, M., Val, M.-E., Sismeiro, O., Varet, H., Sieira, R., Krin, E., Skovgaard, O., Comerci, D. J., Eduardo P. C. Rocha, and Mazel, D. (2020). Macromolecular crowding links ribosomal protein gene dosage to growth rate in *Vibrio cholerae*. *BMC Biology*, 18(1):1–18.
- 1797 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1799 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*, 19.

- 1801** Stevenson, B. S. and Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in *Escherichia coli*. *Applied and Environmental Microbiology*, 70(11):6670–6677.
- 1803** Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van Leeuwenhoek*, 39(1):545–565.
- 1805** Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*. *Archives of Microbiology*, 113(3):185–189.
- 1807** Svennsgen, S. L., Kongstad, M., Stenum, T. S. n., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017). Transfer RNA is highly unstable during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Research*, 45(2):793–804.
- 1809** Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 1811** 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 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 1813** 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 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538.
- 1816** Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- 1818** Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*, 107(2):369–381. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1822** The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1):D330–D338.
- 1824** Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 1826** van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen Assimilation in *Escherichia coli*: Putting Molecular Data into a Systems Perspective. *Microbiology and Molecular Biology Reviews*, 77(4):628–695.
- 1828** Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., 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, E., Carey, C., Polat, i., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1834** Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.
- 1836** Vollmer, W., Blanot, D., and De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2):149–167.
- 1838** Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1FO-ATP synthase. *FEBS Letters*, 545(1):61–70.
- 1840** Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- 1843** You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C. W., Wang, Y.-P., Lenz, P., Yan, D., and Hwa, T. (2013). Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- 1845** Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 108(46):18643–18648.
- 1847** Zhang, L., Jiang, W., Nan, J., Almqvist, J., and Huang, Y. (2014a). The *Escherichia coli* CysZ is a pH dependent sulfate transporter that can be inhibited by sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(7):1809–1816.

- 1850** Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014b). Regulation of crp gene expression by the catabolite repres-
1851 sor/activator, cra, in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1852** Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource
1853 allocation in *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.