

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:

¹These authors contributed equally to this work

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

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

27

28 Introduction

29 The observed range of bacterial growth rates is enormously diverse. In natural environments, some ³⁰ microbial organisms might double only once per year (*Mikucki et al., 2009*) while in comfortable ³¹ laboratory conditions, growth can be rapid with several divisions per hour (*Schaechter et al., 1958*). ³² This six order-of-magnitude difference in time scales encompasses different microbial species and ³³ lifestyles, yet even for a single species such as *Escherichia coli*, the growth rate can be modulated ³⁴ over a similar scale by tuning the type and amount of nutrients in the growth medium. This ³⁵ remarkable flexibility in growth rate illustrates the intimate relationship between environmental ³⁶ conditions and the rates at which cells convert nutrients into new cellular material – a relationship ³⁷ that has remained a major topic of inquiry in bacterial physiology for over a century (*Jun et al., 2018*). ³⁸

39 As was noted by Jacques Monod, “the study of the growth of bacterial cultures does not constitute ⁴⁰ a specialized subject or branch of research, it is the basic method of Microbiology.” Those words ⁴¹ ring as true today as they did when they were written 70 years ago (*Monod, 1949*). Indeed, the

42 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 as
 45 the microbial workhorse *E. coli*. In this work, we explore an amalgamation of recent proteomic data
 46 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
 49 data. For example, what biological processes set the absolute speed limit for how fast bacterial
 50 cells can grow and reproduce? How do cells alter the absolute numbers and relative ratios of their
 51 molecular constituents as a function of changes in growth rate or nutrient availability? In this paper,
 52 we 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 a
 54 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
 56 over the past decade using either mass spectrometry (*Schmidt et al., 2016; Peebo et al., 2015;*
Valgepea et al., 2013) or ribosomal profiling (*Li et al., 2014*) across 31 unique growth conditions
 57 (see Appendix Experimental Details Behind Proteomic Data for further discussion of these data
 58 sets). Second, by compiling molecular turnover rate measurements for many of the fundamental
 59 processes associated with bacterial growth, we make quantitative estimates of key cellular pro-
 60 cesses (schematized in *Figure 1*) to determine whether the observed protein copy numbers under
 61 varying conditions appear to be in excess of what would be minimally required to support cell
 62 growth at the observed rates. The census, combined with these estimates, provide a window into
 63 the question of whether the rates of central processes such as energy generation or DNA synthesis
 64 are regulated systematically as a function of cell growth rate by altering protein copy number in
 65 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
 68 formulate 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.,*
1958; Si et al., 2017) can be simply attributed to the increased absolute number of ribosomes
 77 per cell under conditions supporting extremely rapid growth, a hypothesis which we formally
 78 mathematize and explore.

125 Uptake of Nutrients

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

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

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

87 **The rule of "one, few, and ten".** The philosophy behind order-of-magnitude estimates is
 88 to provide a estimate of the appropriate scale, not a prediction with infinite accuracy. As
 89 such, we define three different scales of precision in making estimates. The scale of "one" is
 90 reserved for values that range between 1 and 2. For example, If a particular process has been
 91 experimentally measured to transport 1.87 protons for a process to occur, we approximate
 92 this process to require 2 protons per event. The scale of "few" is reserved for values ranging
 93 between 3 and 5. For example, we will often use Avogadro's number to compute the number of
 94 molecules in a cell given a concentration and a volume. Rather than using Avogadro's number
 95 as 6.02214×10^{23} , we will approximate it as 5×10^{23} . Finally, the scale of "ten" is reserved for
 96 values which we know within an order of magnitude. If a particular protein complex is present
 97 at 883 copies per cell, we say that it is present in approximately 10^3 copies per cell. These
 98 different scales will be used to arrive at simple estimates that report the expected scale of the
 99 observed data. Therefore, the estimates presented here should not be viewed as hard-and-fast
 100 predictions of precise copy numbers, but as approximate lower (or upper) bounds for the
 101 number of complexes that may be needed to satisfy some cellular requirement.
 102 Furthermore, we use equality symbols (=) sparingly and frequently defer to approximation (\approx)
 103 or scaling (\sim) symbols when reporting an estimate. When \approx is used, we are implicitly stating
 104 that we are confident in this estimate within a factor of a few. When a scaling symbol \sim is used,
 105 we are stating that we are confident in our estimate to within an order of magnitude.

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

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

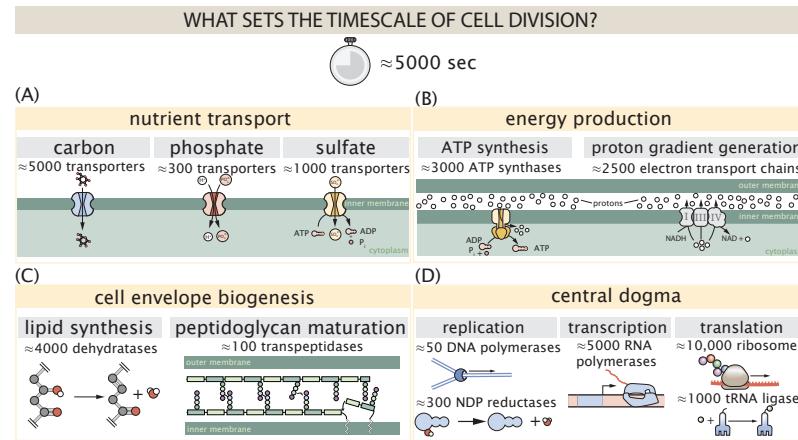


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP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

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

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

146 Nitrogen Transport

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

161 Carbon Transport

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

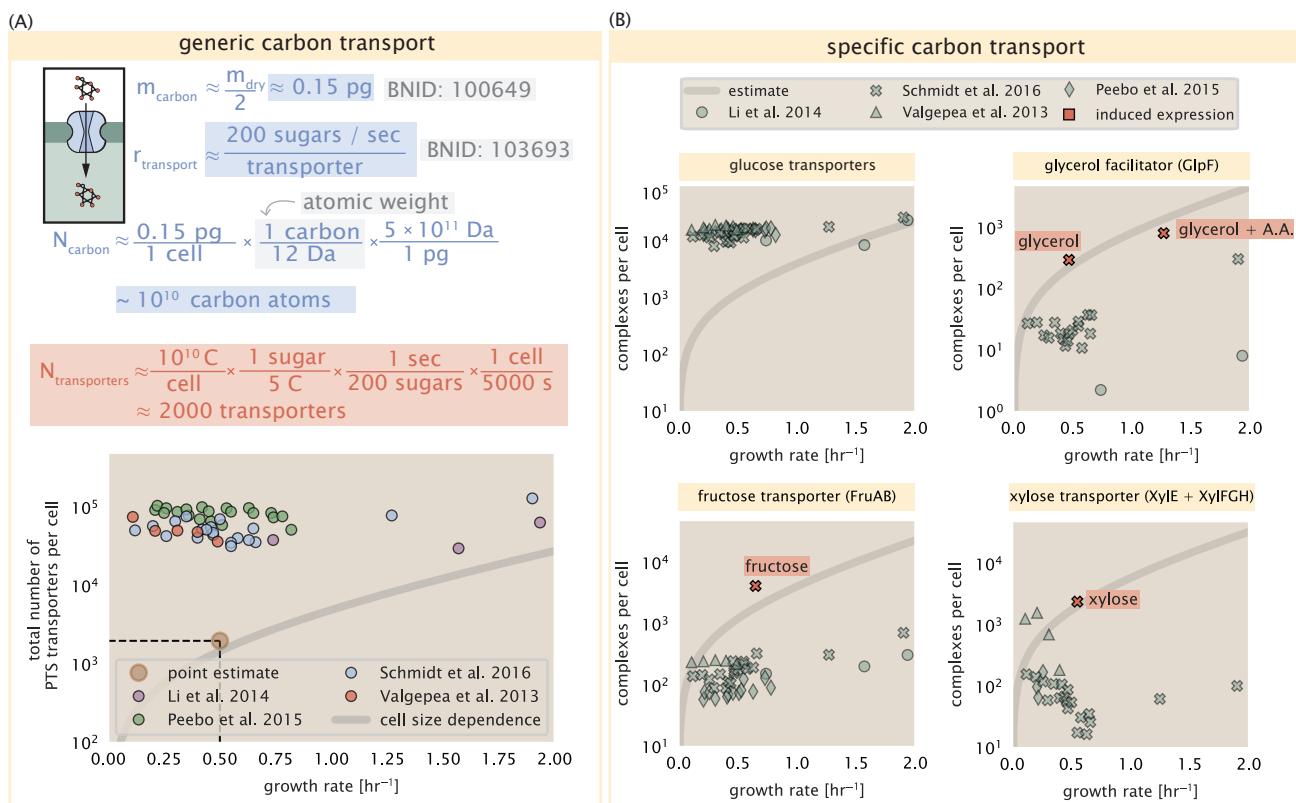


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules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

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

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

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

212 **Phosphorus and Sulfur Transport**

213 We now turn our attention towards other essential elements, namely phosphorus and sulfur.
 214 Phosphorus is critical to the cellular energy economy in the form of high-energy phosphodiester
 215 bonds making up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-
 216 translational modification of proteins and defining the polar-heads of lipids. In total, phosphorus
 217 makes up $\approx 3\%$ of the cellular dry mass which in typical experimental conditions is in the form of
 218 inorganic phosphate. The cell membrane has remarkably low permeability to this highly-charged
 219 and critical molecule, therefore requiring the expression of active transport systems. In *E. coli*, the
 220 proton electrochemical gradient across the inner membrane is leveraged to transport inorganic
 221 phosphate into the cell (**Rosenberg et al., 1977**). Proton-solute symporters are widespread in *E.*
 222 *coli* (**Ramos and Kaback, 1977; Booth et al., 1979**) and can have rapid transport rates of 50 to 100
 223 molecules per second for sugars and other solutes (BNID: 103159; 111777). As a more extreme
 224 example, the proton transporters in the F₁-F₀ ATP synthase, which use the proton electrochemical
 225 gradient for rotational motion, can shuttle protons across the membrane at a rate of ≈ 1000 per
 226 second (BNID: 104890; 103390). In *E. coli* the PitA phosphate transport system has been shown
 227 to be very tightly coupled with the proton electrochemical gradient with a 1:1 proton:phosphate
 228 stoichiometric ratio (**Harris et al., 2001; Feist et al., 2007**). Taking the geometric mean of the
 229 aforementioned estimates gives a plausible rate of phosphate transport on the order of 300
 230 per second. Illustrated in **Figure 3(A)**, we can estimate that ≈ 200 phosphate transporters are
 231 necessary to maintain an $\approx 3\%$ dry mass with a 5000 s division time. This estimate is consistent

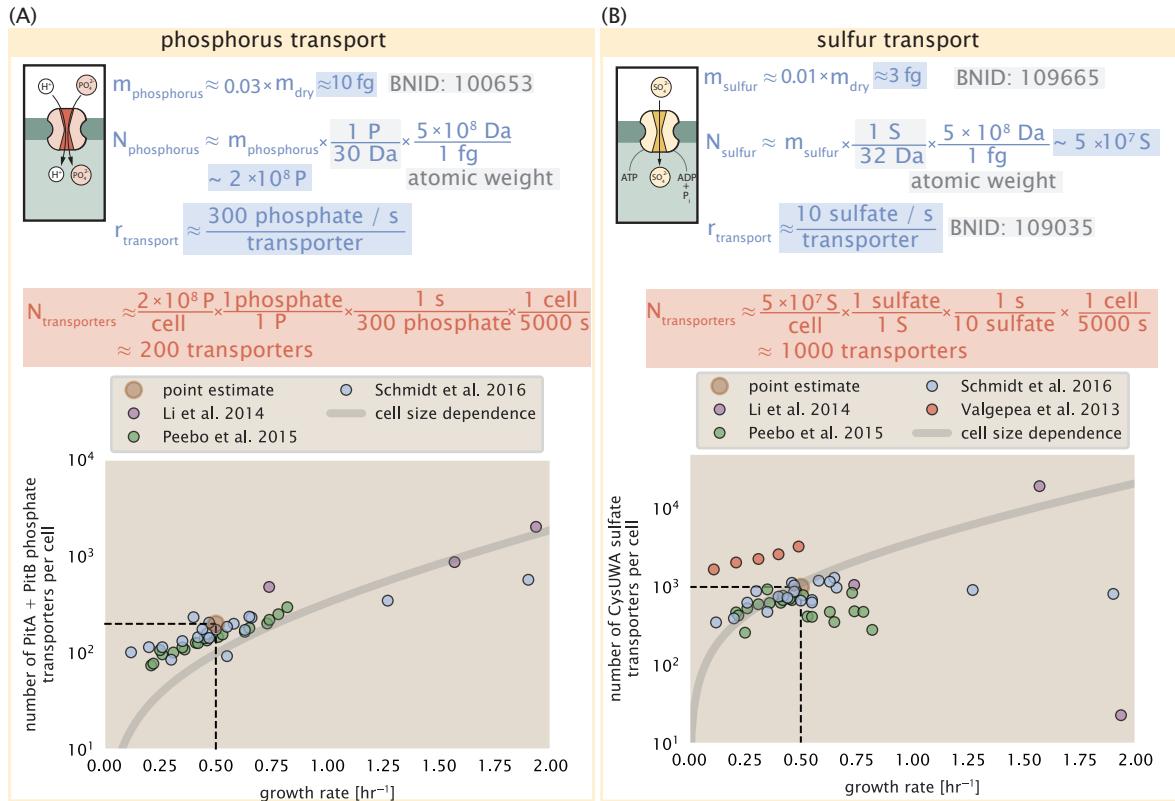


Figure 3. Estimates and measurements of phosphate and sulfate transport systems as a function of growth rate. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (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.

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.

252 **Limits on Transporter Expression**

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

266 **Energy Production**

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

273 **ATP Synthesis**

274 Hydrolysis of the terminal phosphodiester bond of ATP forming ADP (or alternatively GTP and GDP)
 275 and an inorganic phosphate is a kinetic driving force in a wide array of biochemical reactions. One
 276 such reaction is the formation of peptide bonds during translation which requires ≈ 2 ATPs for the
 277 charging of an amino acid to the tRNA and ≈ 2 GTP for the formation of the peptide bond between
 278 amino acids. Assuming the ATP costs associated with error correction and post-translational
 279 modifications of proteins are negligible, we can make the approximation that each peptide bond
 280 has a net cost of ≈ 4 ATP (BNID: 101442, *Milo et al. (2010)*). Formation of GTP from ATP is achieved
 281 via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy
 282 investment (*Lascu and Gonin, 2000*) and therefore consider all NTP requirements of the cell to
 283 be functionally equivalent to being exclusively ATP. In total, the energetic costs of peptide bond
 284 formation consume $\approx 80\%$ of the cells ATP budget (BNID: 107782; 106158; 101637; 111918, *Lynch*
 285 and *Marinov (2015)*; *Stouthamer (1973)*). The pool of ATP is produced by the F₁-F₀ ATP synthase – a
 286 membrane-bound rotary motor which under ideal conditions can yield ≈ 300 ATP per second (BNID:
 287 114701; *Weber and Senior (2003)*).

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

300 cytoplasm ($\rho \approx 1 \text{ pg} / \text{fL}$), and the empirical determination that approximately half of the dry mass is
 301 protein, we can compute the energy demand as a function of growth rte, indicated by the gray line
 302 in **Figure 4**.

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

309 **Generating the Proton Electrochemical Gradient**

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

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

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

342 **Energy Production in a Crowded Membrane.**

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

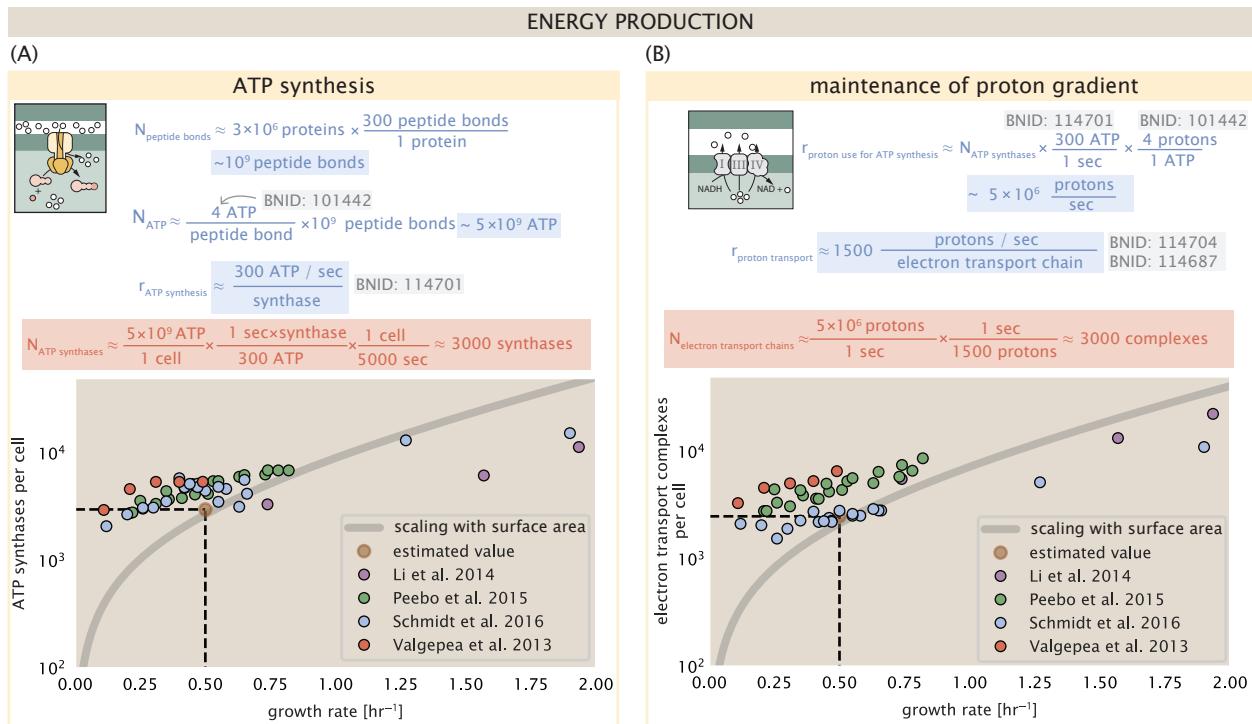


Figure 4. The abundance of F₁-F₀ ATP synthases and electron transport chain complexes as a function of growth rate. (A) Estimate of the number of F₁-F₀ ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes. Points in plot correspond to the mean number of 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 oxidoreductase 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.

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

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

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

385 Synthesis of the Cell Envelope

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

394 Lipid Synthesis

395 The cell envelopes of gram negative bacteria (such as *E. coli*) are composed of inner and outer
 396 phospholipid bilayer membranes separated by a ≈ 10 nm periplasmic space (BNID: 100016, *Milo*

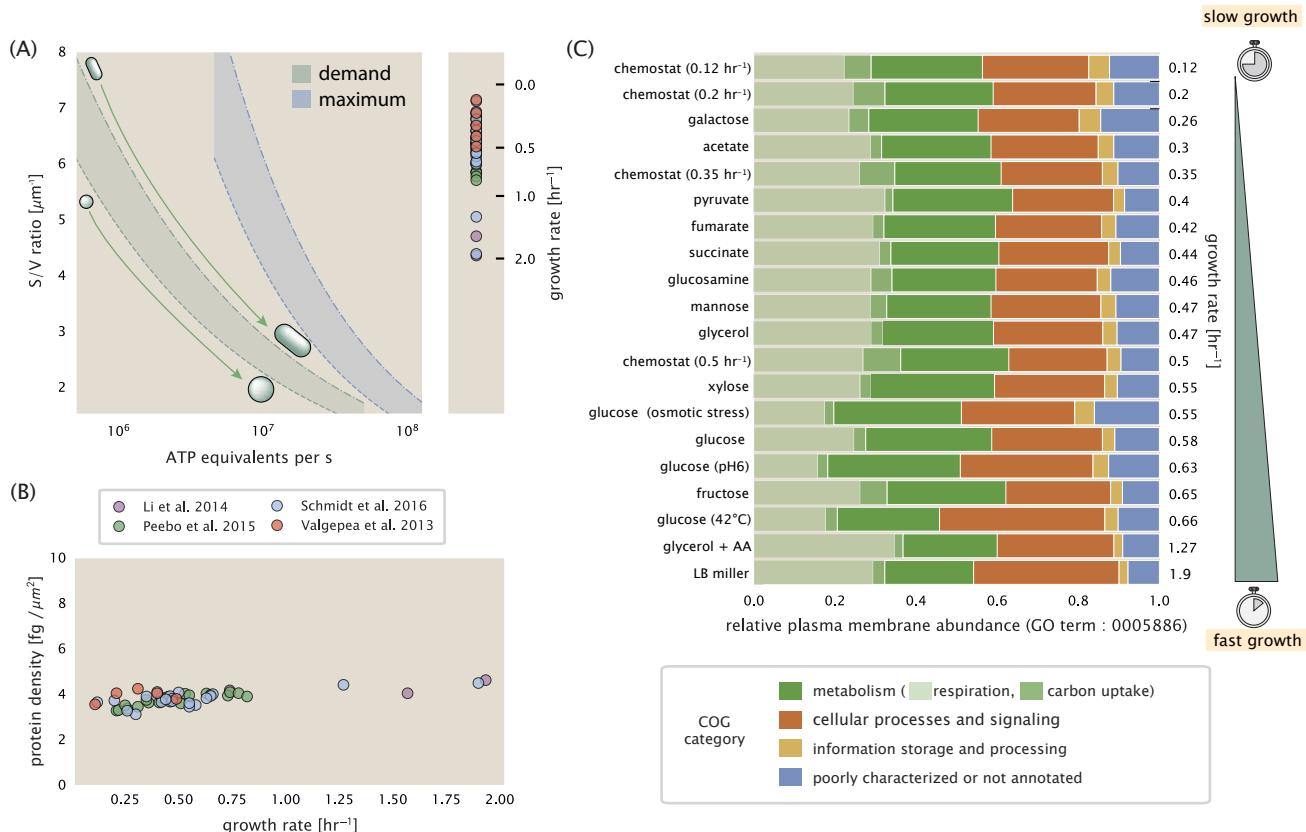


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (---) 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 b_0 , 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)).

397 **et al. (2010)**). As mentioned in our discussion of the surface area to volume constraints on energy
 398 production, *E. coli* is a rod-shaped bacterium with a 4:1 length-to-width aspect ratio. At modest
 399 growth rates, such as our stopwatch of 5000 s, the total cell surface area is $\approx 5 \mu\text{m}^2$ (BNID: 101792,
 400 **Milo et al. (2010)**). As there are two membranes, each of which composed of two lipid leaflets, the
 401 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.

402 While this represents the total area of the membrane, this does not mean that it is composed
 403 entirely of lipid molecules. Rather, the dense packing of the membrane with proteins means that
 404 only $\approx 40\%$ of the membrane area is occupied by lipids (BNID: 100078). Using a rule-of-thumb of
 405 0.5 nm^2 as the surface area of the typical lipid (BNID: 106993), we arrive at an estimate of $\sim 2 \times 10^7$
 406 lipids per cell, an estimate in close agreement with experimental measurements (BNID: 100071,
 407 102996).

408 The membranes of *E. coli* are composed of a variety of different lipids, each of which are unique
 409 in their structures and biosynthetic pathways (**Sohlenkamp and Geiger, 2016**). With such diversity
 410 in biosynthesis, it becomes difficult to identify which step(s) may be the rate-limiting, an objective
 411 further complicated by the sparsity of *in vivo* kinetic data. Recently, a combination of stochastic
 412 kinetic modeling (**Ruppe and Fox, 2018**) and *in vitro* kinetic measurements (**Ranganathan et al.,
 413 2012; Yu et al., 2011**) have revealed remarkably slow steps in the fatty acid synthesis pathways
 414 which may serve as the rate limiting reactions. One such step is the removal of hydroxyl groups
 415 from the fatty-acid chain by ACP dehydratase that leads to the formation of carbon-carbon double
 416 bonds. This reaction, catalyzed by proteins FabZ and FabA in *E. coli* (**Yu et al., 2011**), have been
 417 estimated to have kinetic turnover rates of ≈ 1 dehydration per second per enzyme (**Ruppe and
 418 Fox, 2018**). Combined with this rate, our previous estimates for the number of lipids to be formed,
 419 and a 5000 second division yields an estimate that the cell requires ≈ 4000 ACP dehydratases. This
 420 is in reasonable agreement with the experimentally observed copy numbers of FabZ and FabA
 421 (**Figure 6(A)**). Furthermore, we can extend this estimate to account for the change in membrane
 422 surface area as a function of the growth rate (grey line in **Figure 6(A)**), which captures the observed
 423 growth rate dependent expression of these two enzymes.

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

435 Peptidoglycan Synthesis

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

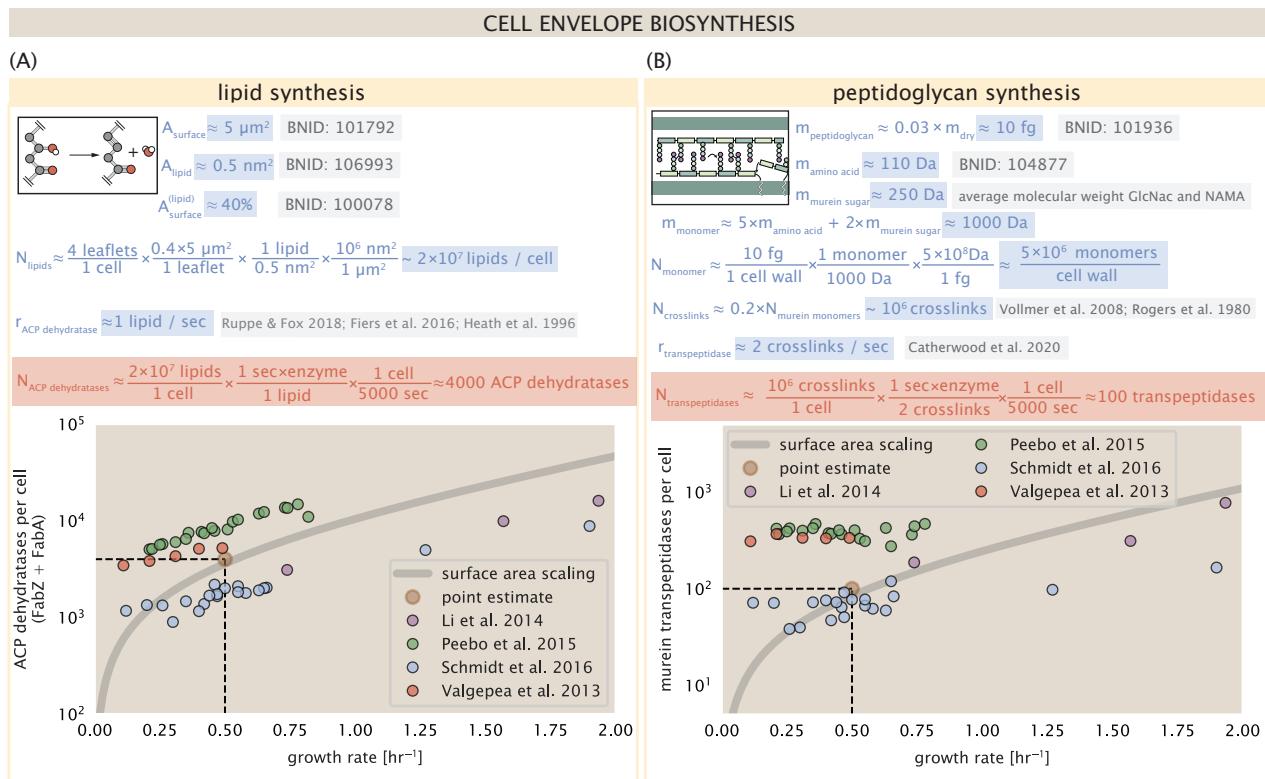


Figure 6. Estimation of the key components involved in cell envelope biosynthesis. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred from experimental measurements via a stochastic kinetic model described in [Ruppe and Fox \(2018\)](#). Bottom panel shows the experimentally observed complex copy numbers using the stoichiometries $[\text{FabA}]_2$ and $[\text{FabZ}]_2$. (B) An estimate for the number of peptidoglycan transpeptidases needed to complete maturation of the peptidoglycan. The mass of the murein monomer was estimated by approximating each amino acid in the pentapeptide chain as having a mass of 110 Da and each sugar in the disaccharide having a mass of ≈ 250 Da. The *in vivo* rate of transpeptidation in *E. coli* was taken from recent analysis by [Catherwood et al. \(2020\)](#). The bottom panel shows experimental measurements of the transpeptidase complexes in *E. coli* following the stoichiometries $[\text{MrcA}]_2$, $[\text{MrcB}]_2$, $[\text{MrdA}]_1$, and $[\text{MrdB}]_1$. Grey curves in each plot show the estimated number of complexes needed to satisfy the synthesis requirements scaled by the surface area as a function of growth rate. We direct the reader to the supplemental information for a more detailed discussion of this estimate.

447 In glucose-supported steady-state growth, the peptidoglycan alone comprises \approx 3% of the
 448 cellular dry mass (BNID: 101936), making it the most massive molecule in *E. coli*. The polymerized
 449 unit of the peptidoglycan is a N-acetylglucosamine and N-acetylmuramic acid disaccharide, of which
 450 the former is functionalized with a short pentapeptide. With a mass of \approx 1000 Da, this unit, which
 451 we refer to as a murein monomer, is polymerized to form long strands in the periplasm which are
 452 then attached to each other via their peptide linkers. Using the aforementioned measurement that
 453 \approx 3% of the dry mass is peptidoglycan, it can be estimated that the peptidoglycan is composed of \sim
 454 5×10^6 murein monomers.

455 During growth, peptidoglycan is constantly being broken down to allow insertion of new murein
 456 monomers and cellular expansion. In order to maintain structural integrity these monomers must
 457 be crosslinked into the expanding cell wall, potentially limiting how quickly new material can be
 458 added and we consider this process as a possible rate-limiting step. In principle, each one of
 459 these murein monomers can be crosslinked to another glycan strand via the pentapeptide. In
 460 some species, such as in gram-positive bacterium *Staphylococcus aureus*, the extent of crosslinking
 461 can be large with $>$ 90% of pentapeptides forming a connection between glycan strands. In *E.*
 462 *coli*, however, a much smaller proportion (\approx 20%) of the peptides are crosslinked, resulting in a
 463 weaker and more porous cell wall *Vollmer et al. (2008); Rogers et al. (1980)*. The formation of these
 464 crosslinks primarily occur during the polymerization of the murein monomers and is facilitated by a
 465 family of enzymes called transpeptidases. In *E. coli*, there are four primary transpeptidases that are
 466 involved in lateral and longitudinal extension of the peptidoglycan. These transpeptidases have
 467 only recently been quantitatively characterized *in vivo* via liquid chromatography mass spectrometry
 468 (*Catherwood et al., 2020*), which revealed a kinetic turnover rate of \approx 2 crosslinking reactions
 469 formed per second per enzyme.

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

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

485 Function of the Central Dogma

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

491 DNA

492 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal
 493 plasmids up to \sim 100 kbp in length. We will focus our quantitative thinking solely on the chromo-
 494 some of *E. coli* which harbors \approx 5000 genes and \approx 5×10^6 base pairs. To successfully divide and
 495 produce viable progeny, this chromosome must be faithfully replicated and segregated into each

496 nascent cell. We again rely on the near century of literature in molecular biology to provide some
 497 insight on the rates and mechanics of the replicative feat as well as the production of the required
 498 starting materials, dNTPs.

499 dNTP synthesis

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

510 Since this reaction is central to the synthesis of all dNTPs, it is reasonable to consider the
 511 abundance of these complexes is a measure of the total dNTP production in *E. coli*. Illustrated
 512 schematically in *Figure 7* (A), we consider the fact that to replicate the cell's genome, on the order of
 513 $\approx 10^7$ dNTPs must be synthesized. Assuming a production rate of 10 per second per ribonucleotide
 514 reductase complex and a cell division time of 5000 seconds, we arrive at an estimate of ≈ 200
 515 complexes needed per cell. As shown in the bottom panel of *Figure 7* (A), this estimate agrees
 516 with the experimental measurements of these complexes abundances within $\approx 1/2$ an order of
 517 magnitude. Extension of this estimate across a continuum of growth rate, including the fact that
 518 multiple chromosomes can be replicated at a given time, is shown as a grey transparent line in
 519 *Figure 7* (A). Similarly to our point estimate, this refinement agrees well with the data, accurately
 520 describing both the magnitude of the complex abundance and the dependence on growth rate.

521 Recent work has revealed that during replication, the ribonucleotide reductase complexes
 522 coalesce to form discrete foci colocalized with the DNA replisome complex (*Sánchez-Romero et al.,*
 523 *2011*). This is particularly pronounced in conditions where growth is slow, indicating that spatial
 524 organization and regulation of the activity of the complexes plays an important role.

525 DNA Replication

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

542 It is well known that *E. coli* can parallelize its DNA replication such that multiple chromosomes
 543 are being replicated at once, with as many as 10 - 12 replication forks at a given time (*Bremer*
 544 and *Dennis, 2008; Si et al., 2017*). Thus, even in rapidly growing cultures, we expect only a few

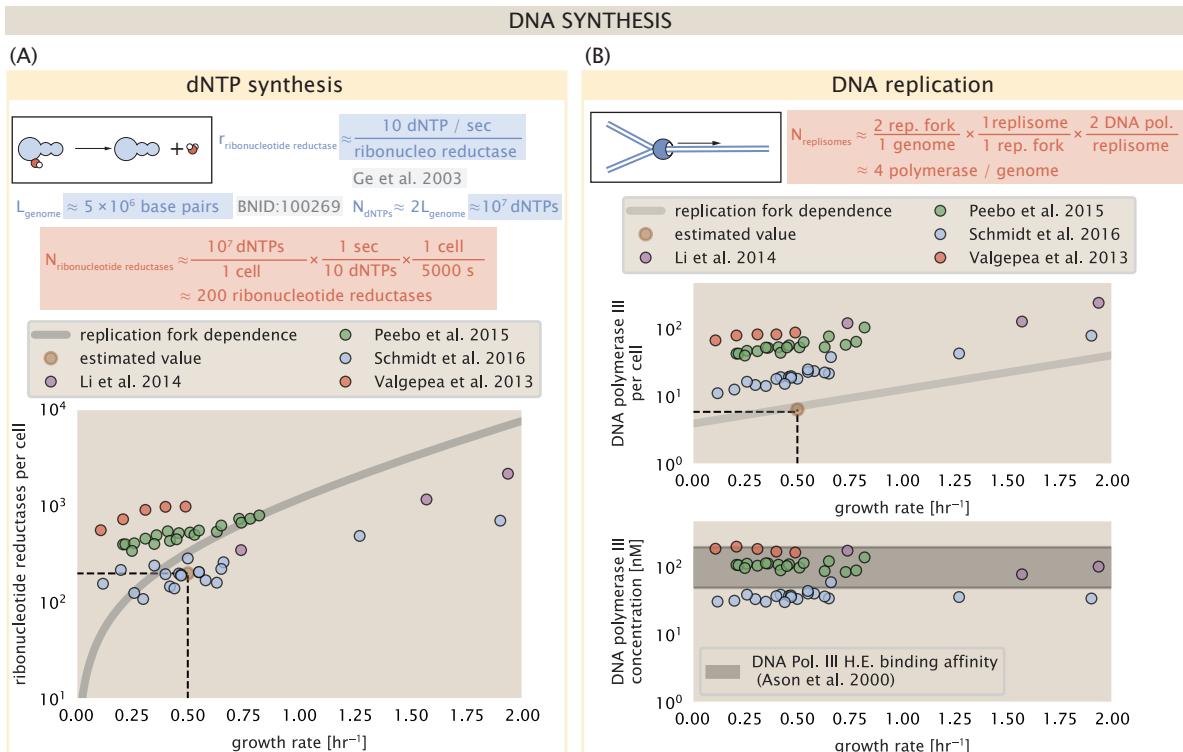


Figure 7. Complex abundance estimates for dNTP synthesis and DNA replication. (A) Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of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

545 polymerases (≈ 10) are needed to replicate the chromosome per cell doubling. However, as shown
 546 in **Figure 7(B)**, DNA polymerase III is nearly an order of magnitude more abundant. This
 547 discrepancy can be understood by considering its binding constant to DNA. DNA polymerase III
 548 is highly processive, facilitated by a strong affinity of the complex to the DNA. *In vitro* biochemical
 549 characterization has quantified the K_D of DNA polymerase III holoenzyme to single-stranded and
 550 double-stranded DNA to be 50 and 200 nM, respectively (Ason et al., 2000). The bottom plot in
 551 **Figure 7(B)** shows that the concentration of the DNA polymerase III across all data sets and growth
 552 conditions is within this range. Thus, while the copy number of the DNA polymerase III is in excess
 553 of the strict number required to replicate the genome, its copy number appears to vary such that its
 554 concentration is approximately equal to the dissociation constant to the DNA. While the processes
 555 regulating the initiation of DNA replication are complex and involve more than just the holoenzyme,
 556 these data indicate that the kinetics of replication rather than the explicit copy number of the DNA
 557 polymerase III holoenzyme is the more relevant feature of DNA replication to consider. In light
 558 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),
 560 the time to double their chromosome likely represents an upper limit to their growth rate.
 561

562 RNA Synthesis

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

572 rRNA

573 We begin with an estimation of the number of RNA polymerases needed to synthesize the rRNA
 574 that serve as catalytic and structural elements of the ribosome. Each ribosome contains three rRNA
 575 molecules of lengths 120, 1542, and 2904 nucleotides (BNID: 108093), meaning each ribosome
 576 contains ≈ 4500 nucleotides. As the *E. coli* RNA polymerase transcribes DNA to RNA at a rate of \approx
 577 40 nucleotides per second (BNID: 101904), it takes a single RNA polymerase ≈ 100 s to synthesize
 578 the RNA needed to form a single functional ribosome. Therefore, in a 5000 s division time, a single
 579 RNA polymerase transcribing rRNA at a time would result in only ≈ 50 functional ribosomal rRNA
 580 units – far below the observed number of $\approx 10^4$ ribosomes per cell.

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

590 *E. coli*, like many other bacterium, have evolved a clever mechanism to surpass this kinetic
 591 limit for the rate of rRNA production. Rather than having only one copy of each rRNA gene, *E.*
 592 *coli* has seven copies of the operon (BNID: 100352) four of which are localized directly adjacent
 593 to the origin of replication (Birnbaum and Kaplan, 1971). As fast growth also implies an increased

594 gene dosage due to parallelized chromosomal replication, the total number of rRNA genes can be
 595 on the order of $\approx 10 - 70$ copies at moderate to fast growth rates (**Stevenson and Schmidt, 2004**).
 596 Given a 5000 second division time, we can make the lower-bound estimate that the typical cell will
 597 have ≈ 7 copies of the rRNA operon. Synthesizing one functional rRNA unit per second per rRNA
 598 operon, a total of 5×10^4 rRNA units can be synthesized, comfortably above the observed number
 599 of ribosomes per cell.

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

608 mRNA

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

617 This estimation captures the *steady-state* mRNA copy number, meaning that at any given time,
 618 there will exist approximately 3000 unique mRNA molecules. To determine the *total* number of
 619 mRNA that need to be synthesized over the cell's lifetime, we must consider degradation of the
 620 mRNA. In most bacteria, mRNAs are rather unstable with life times on the order of several minutes
 621 (BNID: 104324; 106253; 111927; 111998). For convenience, we assume that the typical mRNA in
 622 our cell of interest has a typical lifetime of ≈ 300 seconds. Using this value, we can determine
 623 the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell.
 624 While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional
 625 dynamics, we state here that the total mRNA production rate must be on the order of ≈ 15 mRNA
 626 per second. In *E. coli*, the average protein is ≈ 300 amino acids in length (BNID: 108986), meaning
 627 that the corresponding mRNA is ≈ 900 nucleotides which we will further approximate as ≈ 1000
 628 nucleotides to account for the non-protein coding regions on the 5' and 3' ends. This means that
 629 the cell must have enough RNA polymerase molecules about to sustain a transcription rate of
 630 $\approx 1.5 \times 10^4$ nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA at a
 631 clip of 40 nucleotides per second, we arrive at a comfortable estimate of ≈ 250 RNA polymerase
 632 complexes needed to satisfy the mRNA demands of the cell. It is worth noting that this number is
 633 approximately half of that required to synthesize enough rRNA, as we saw in the previous section.
 634 We find this to be a striking result as these 250 RNA polymerase molecules are responsible for the
 635 transcription of the ≈ 4000 protein coding genes that are not ribosome associated.

636 tRNA

637 The final class of RNA molecules worthy of quantitative consideration are the tRNAs that are used
 638 during translation to map codon sequence on mRNA to specific amino acids. Unlike mRNA or
 639 rRNA, each individual tRNA is remarkably short, ranging from 70 to 95 nucleotides each (BNID:
 640 109645; 102340). What they lack in length, they make up for in abundance, with reported values
 641 ranging from $\approx 5 \times 10^4$ (BNID: 105280) to $\approx 5 \times 10^5$ (BNID: 108611). To test tRNA synthesis as a possible
 642 growth-rate limiting stage, we will err towards a higher abundance of $\approx 5 \times 10^5$ per cell. Combining

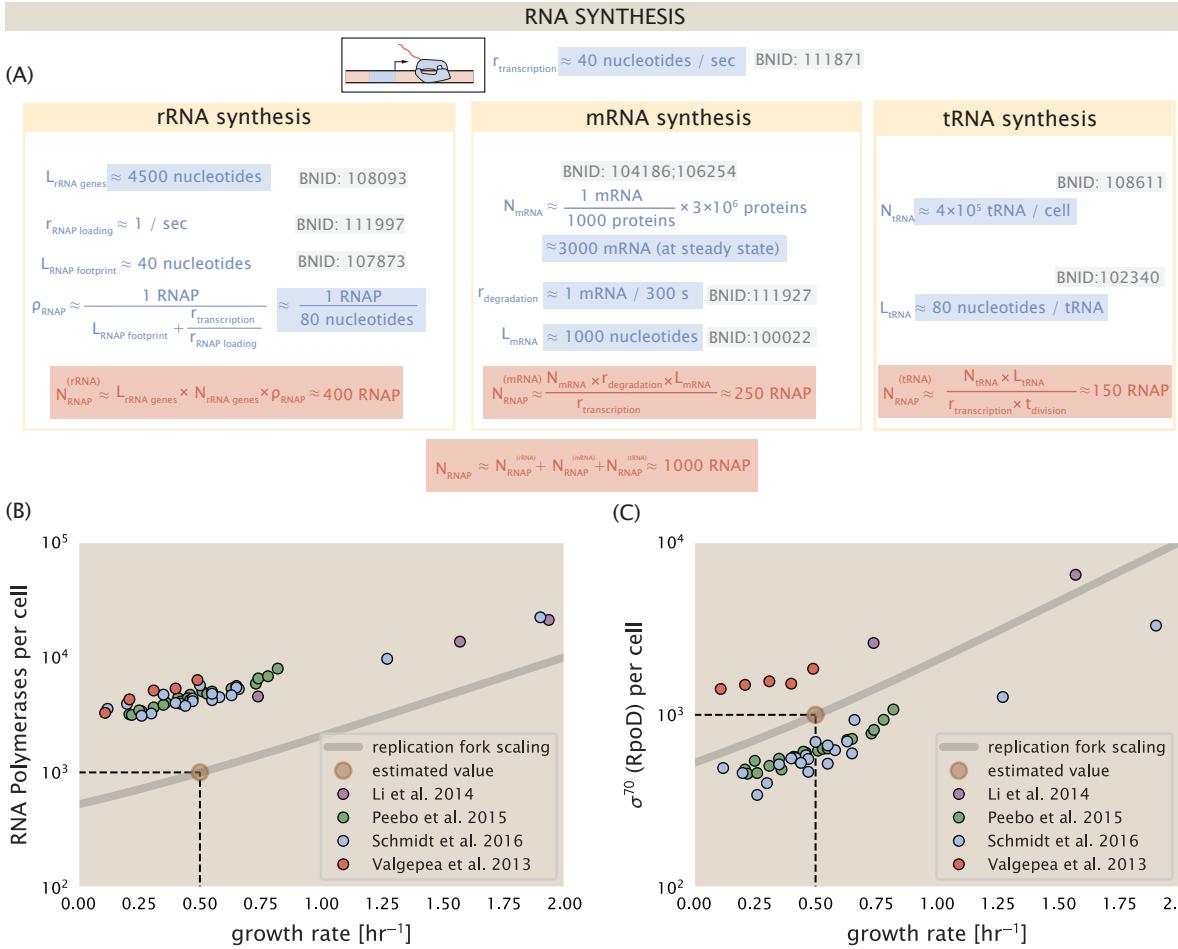


Figure 8. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. (B) The RNA polymerase core enzyme copy number as a function of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

the abundance and tRNA length measurements, we make the estimate that $\approx 5 \times 10^7$ nucleotides are sequestered in tRNA per cell. Unlike mRNA, tRNA is remarkably stable with typical lifetimes *in vivo* on the order of ≈ 48 hours (Abelson et al., 1974; Svenningsen et al., 2017) – well beyond the timescale of division. Once again using our rule-of-thumb for the rate of transcription to be 40 nucleotides per second and assuming a division time of ≈ 5000 seconds, we arrive at an estimate of ≈ 200 RNA polymerases to synthesize enough tRNA. This requirement pales in comparison to the number of polymerases needed to generate the rRNA and mRNA pools and can be neglected as a significant transcriptional burden.

651 RNA Polymerase and σ -factor Abundance

652 These estimates, summarized in **Figure 8 (A)**, reveal that synthesis of rRNA and mRNA are the dominant RNA species synthesized by RNA polymerase, suggesting the need for ≈ 1000 RNA polymerases per cell. As is revealed in **Figure 8 (B)**, this estimate is about an order of magnitude below the observed number of RNA polymerase complexes per cell ($\approx 5000 - 7000$). The difference between the estimated number of RNA polymerase needed for transcription and these observations are consistent with recent literature revealing that $\approx 80\%$ of RNA polymerases in *E. coli* are not transcriptionally active (Patrick et al., 2015). Our estimate ignores the possibility that some fraction

659 is only nonspecifically bound to DNA, as well as the obstacles that RNA polymerase and DNA
 660 polymerase present for each other as they move along the DNA (*Finkelstein and Greene, 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., 2005; Goldman et al., 2015; Perdue and Roberts, 2011; Mooney and Landick, 2003; Mooney et al., 2005*). **Figure 8** (B) is suggestive of such a role as the number of σ^{70} proteins per cell is in close
 666 agreement with our estimate of the number of transcriptional complexes needed.

667 These estimates provide insight as to the observed magnitude of both RNA polymerase and
 668 the σ -70 factor. As we have done in the previous sections, and described in Appendix Extending
 669 Estimates to a Continuum of Growth Rates, we can generalize these estimates across a wide range
 670 of growth rates (grey line in **Figure 8**(B)). While there remains some disagreement in the magnitude
 671 of the copy number, this estimate appears to very adequately describe the growth rate dependence
 672 of these complexes. Furthermore, these findings illustrate that transcription cannot be the rate
 673 limiting step in bacterial division. **Figure 8** (A) reveals that the availability of RNA polymerase is not
 674 a limiting factor for cell division as the cell always has an apparent \sim 10-fold excess than needed.
 675 Furthermore, if more transcriptional activity was needed to satisfy the cellular requirements, more
 676 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
 686 sections – 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
 699 of 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, we
 702 arrive at a requirement of $\approx 2 \times 10^5$ tRNA molecules to be consumed by the ribosome per second.

703 There are many processes which go into synthesizing a tRNA and ligating it with the appropriate
 704 amino acids. As we discussed previously, there appear to be more than enough RNA polymerases
 705 per cell to synthesize the needed pool of tRNAs. Without considering the many ways in which amino
 706 acids can be scavenged or synthesized *de novo*, we can explore ligation as a potential rate limiting
 707 step. The enzymes which link the correct amino acid to the tRNA, known as tRNA synthetases or

708 tRNA ligases, are incredible in their proofreading of substrates with the incorrect amino acid being
 709 ligated once out of every 10^4 to 10^5 events (BNID: 103469). This is due in part to the consumption
 710 of energy as well as a multi-step pathway to ligation. While the rate at which tRNA is ligated is
 711 highly dependent on the identity of the amino acid, it is reasonable to state that the typical tRNA
 712 synthetase has charging rate of ≈ 20 AA per tRNA synthetase per second (BNID: 105279).

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

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

727 Protein Synthesis

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

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

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

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

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

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

754 where λ is the cell growth rate in s^{-1} , r_t is the maximum elongation rate in $\text{AA}\cdot s^{-1}$, and R is the

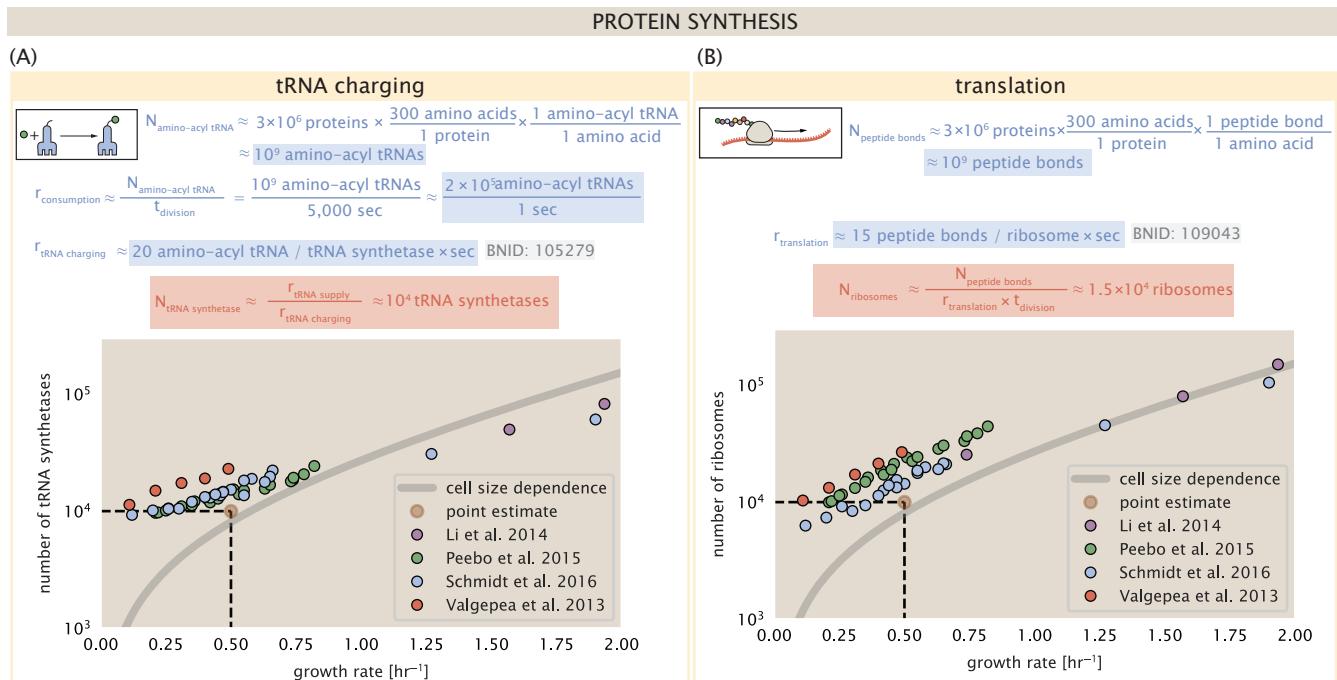


Figure 9. Estimation of the required tRNA synthetases and ribosomes. (A) Estimate for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [GltX], [IleS], [LeuS], [ValS], [AlaS]₂, [AsnS]₂, [AspS]₂, [TyrS]₂, [TrpS]₂, [ThrS]₂, [SerS]₂, [ProS]₂, [PheS]₂[PheT]₂, [MetG]₂, [IysS]₂, [HisS]₂, [GlyS]₂[GlyQ]₂). (B) Estimate of the number of ribosomes required to synthesize 10⁹ peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr⁻¹. Grey lines correspond to the estimated complex abundance calculated at different growth rates. See Appendix Extending Estimates to a Continuum of Growth Rates for a more detail description of this calculation.

755 average ribosome copy number per cell. The addition factor f_a refers to the fraction of actively
 756 translating ribosomes, and allows us to account for the possibility of nonfunctional, immature
 757 ribosomes or active sequestration of ribosomes, mediated by the secondary-messenger molecule
 758 alarmones, guanosine pentaphosphate [(p)ppGpp] at slow growth (Dennis et al., 2004; Dai et al.,
 759 2016). Knowing the number of peptide bonds formed per cell permits us to compute the translation-
 760 limited growth rate as

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

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

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

771 This is plotted as a function of ribosomal fraction Φ_R in **Figure 10(B)**, where we take $L_R = 7459 \text{ AA}$,
 772 corresponding to the length in amino acids for all ribosomal subunits of the 50S and 30S complex
 773 (BNID: 101175), and $f_a = 1$. In **Figure 10(C)** we use the recent measurements of f_a from Dai et al.
 774 (2016) to estimate the active fraction of ribosomal protein across the proteomic data sets and
 775 number of other recent measurements. We see that cells are consistently skirting the limit in
 776 growth rate set by **Equation 3** as nutrient conditions vary.

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

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

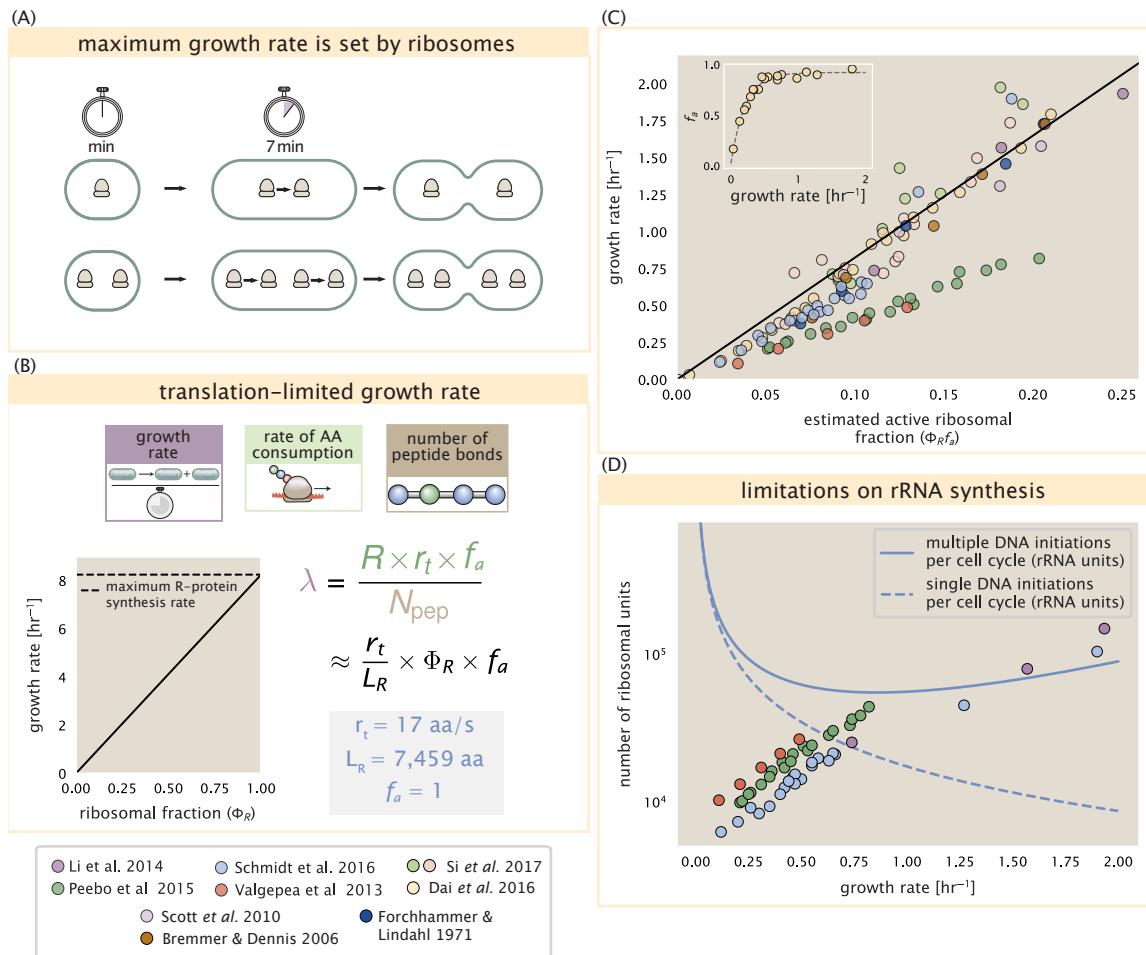


Figure 10. Translation-limited growth rate. (A) An inherent maximum growth rate is set by the time to synthesize all ribosomal subunits. This growth rate is given by r_t/L_R , where r_t is the elongation rate and L_R is the total number of amino acids that make up the entire ribosomal complex. This rate is independent of the number of ribosomes and instead is limited by the time required to double an individual ribosome. (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 corresponding to the rRNA copy number is calculated by multiplying the number of rRNA operons by the estimated number of (# ori) at each growth rate. The quantity (# ori) was calculated using Equation 4 and the measurements from [Si et al. \(2017\)](#) 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.

799 Relationship Between Cell Size and Growth Rate

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

807 Cells grow at a near-maximal rate dictated by their total ribosomal mass fraction Φ_R , at least at
 808 moderate growth rates above 0.5 hr⁻¹ (where f_a is close to 1, and r_t is near its maximal rate). Here,
 809 growth rate can be increased only by increasing Φ_R , though the simple addition of more ribosomes
 810 is likely constrained by aspects physical constrains like macromolecular crowding (*Delarue et al.,*
 811 *2018; Soler-Bistué et al., 2020*). As *E. coli* grows faster, large swaths of its proteome increase in
 812 absolute abundance. It is now well-documented that *E. coli* cells add a constant volume per origin
 813 of replication (termed a "unit cell" or "initiation mass"), which is robust to a remarkable array of
 814 cellular perturbations (*Si et al., 2017*). To consider this dependency in the context of the proteomic
 815 data, we used measurements from *Si et al. (2017)* (*Figure 11(A)*) to estimate the average number
 816 of origins per cell $\langle \# \text{ ori} \rangle$ at different growth rates. $\langle \# \text{ ori} \rangle$ is set by how often replication must be
 817 initiated per cell doubling under steady-state growth. This can be quantified as

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

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

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

832 Nutrient-Mediated Regulation of Proteomic Composition and Growth Rate

833 As we have seen, cell size, total proteomic content, and the number of ribosomes are all intercon-
 834 nected and influence the achievable growth rate. The drastic change in these parameters across
 835 different growth conditions suggests a hypothesis that each parameter is being tuned to better
 836 match the cell's biosynthetic capacity to the specific environment. Take, as another illustration
 837 of this, the recent experimental work by *Dai et al. (2016)*. In one set of experiments the authors
 838 considered growth in cells whose primary glucose transport system was disrupted ($\Delta ptsG$). Un-
 839 surprisingly, the growth rate was reduced, and was measured at about two-fold slower than their
 840 wild-type line. This change, however, was not simply the result of now-limiting carbon uptake.
 841 Instead, cells accommodated the perturbation by also reducing their ribosomal mass fraction by a
 842 factor of two, which is still in line with **Equation 3** under translation-limited growth. In this final
 843 section, we explore the interconnection between cell size, ribosome content, and growth rate by
 844 formulating a minimal model of growth rate control. We use it to quantitatively show how tuning
 845 these parameters help cells maximize their growth rate for a particular environment.

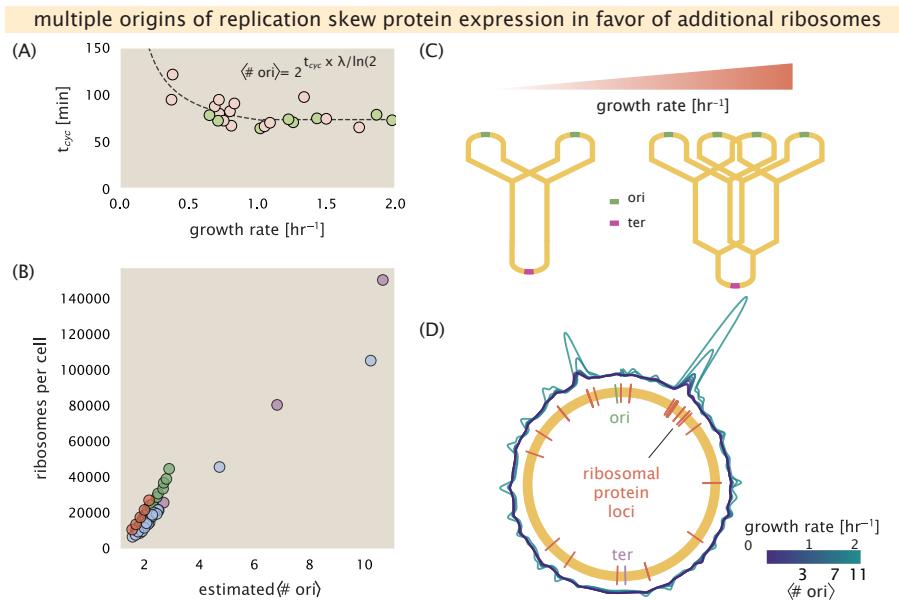


Figure 11. Cells increase absolute ribosome abundance with $\langle \# \text{ori} \rangle$. (A) Experimental data from Si *et al.* (2017). Dashed line shows fit to the data, which were used to estimate $\langle \# \text{ori} \rangle$. t_{cyc} was assumed to vary in proportion to τ for doubling times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

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

853 Furthermore, (p)ppGpp can inhibit the initiation of DNA replication by mediating a change
 854 in transcriptional activity and the supercoiling state of the origin of replication (Kraemer *et al.*,
 855 2019). These observations all raise the possibility that it is through (p)ppGpp that cells mediate the
 856 growth-rate dependent changes in $\langle \# \text{ori} \rangle$, cell size, and ribosomal abundance and activity (Zhu
 857 and Dai, 2019; Büke *et al.*, 2020). Indeed, recent work in a (p)ppGpp deficient strain of *E. coli* found
 858 that cells exhibited $\langle \# \text{ori} \rangle$ and cell sizes that were more consistent with a fast growth state where
 859 (p)ppGpp levels are normally low (Fernández-Coll *et al.*, 2020).

860 Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid 861 Scarcity

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

866 The rate of elongation r_e will depend on how quickly the ribosomes can match codons with
 867 their correct amino-acyl tRNA, along with the subsequent steps of peptide bond formation and
 868 translocation. We therefore coarse-grain the steps of elongation to two time-scales, 1) the time

869 required to find and bind each correct amino-acyl tRNA, and 2) the remaining steps in peptide
 870 elongation that will not depend on the amino acid availability. The availability of amino acids will
 871 depend on their cellular concentration, which we treat as a single effective species, $[AA]_{\text{eff}}$. Under
 872 this model, other molecular players required for translation like elongation factors and GTP are
 873 considered in sufficient abundance, which appear to be valid assumptions given our analysis of the
 874 proteomic data and energy production thus far. The time to translate each codon is given by the
 875 inverse of the elongation rate r_t , which can be written as,

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

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

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

881 where the elongation rate r_t is now given by

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

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

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

885 where the time from 0 to t is an arbitrary length of time, and the square brackets indicate concen-
 886 trations per unit time. Integrating [Equation 8](#) yields.

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

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

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

889 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 an
 890 expression for $[AA]_{\text{eff}}$ in hand, we can now solve [Equation 7](#) for r_t which is a quadratic function with
 891 a physically-meaningful root of

$$r_t = \frac{t(r_{AA} + r_t^{\max}Rf_a) + K_DV - \sqrt{(r_{AA}t + r_t^{\max}Rf_at + K_DV)^2 - 4(Rf_at)(r_t^{\max}r_{AA}t)}}{2Rf_at}. \quad (11)$$

892 In [Figure 12\(B\)](#), we illustrate how the elongation rate depends on the ribosomal copy number.
 893 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
 894 [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
 895 copy numbers, the observed elongation rate is dependent primarily on the ratio of K_D/Vr_{AA} [as
 896 $r_t^{\max} \times R \times f_a \ll r_{AA}$, point (1) in [Figure 12\(B\)](#)]. As the ribosome copy number is increased such
 897 that the amino acid supply rate and consumption rate are nearly equal [point (2) in [Figure 12\(B\)](#)],
 898 the observed elongation rate begins to decrease sharply. When the ribosome copy number is
 899 increased even further, consumption at the maximum elongation rate exceeds the supply rate,
 900 yielding a significantly reduced elongation rate [point (3) in [Figure 12B](#)]. While the elongation rate
 901 will always be dominated by the amino acid supply rate at sufficiently low ribosome copy numbers,
 902 the elongation rate at larger ribosome abundances can be increased by tuning f_a such that not all
 903 ribosomes are elongating, reducing the total consumption rate.

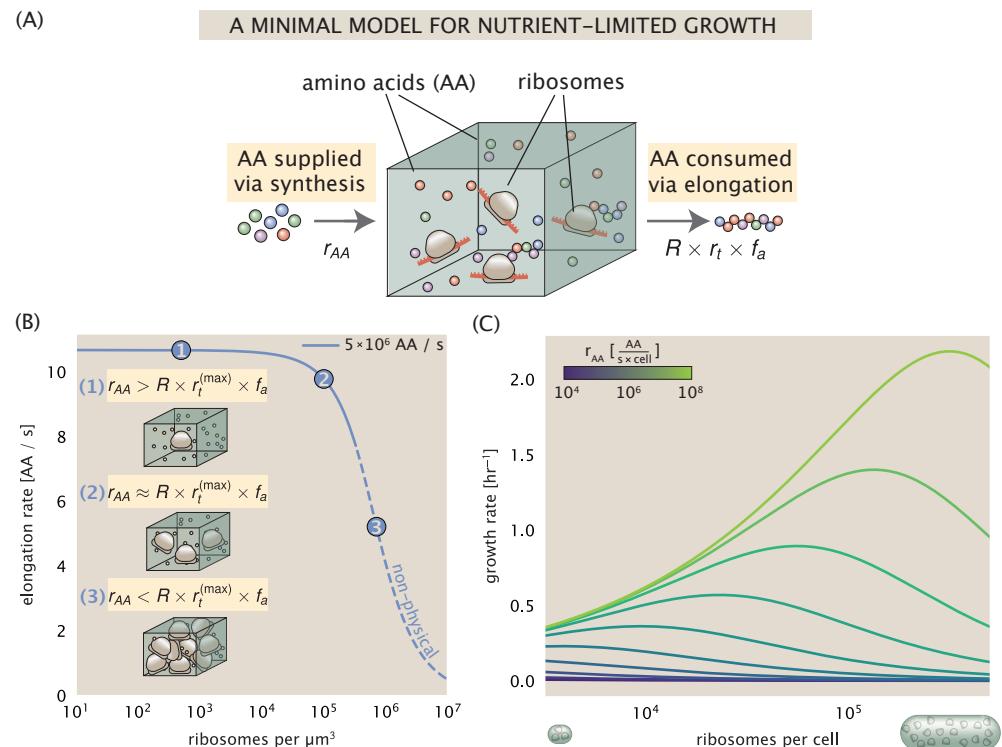


Figure 12. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times f_a$, 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 in a unit volume μm^3 . The three points correspond to three regimes of ribosome copy numbers and are shown schematically on the left-hand side. The region of the curve shown as dashed lines represents a non-physical copy number, but is shown for illustrative purposes. This curve was generated using the parameters $r_{AA} = 5 \times 10^6 \text{ AA} / \text{s}$, $K_D = 5 \text{ mM}$, and $r_t^{(\max)} = 17.1 \text{ AA} / \text{s}$. (C) The cellular growth rate is plotted as a function of total cellular ribosome copy number for different cellular amino acid supply rates, with blue and green curves corresponding to low and high supply rates, respectively. As the ribosome copy number is increased, so too is the cell volume and total protein abundance. We direct the reader to the Supplemental Information for discussion on the inference of the realtionship between cell volume, number of peptide bonds, and ribosome copy number.

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

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

918 Also of note is the growth rate profiles shown for low amino acid supply rates [purple and
 919 blue lines in **Figure 12(C)**], representing growth in nutrient-poor media. In these conditions, there
 920 no longer exists a peak in growth, at least in the range of physiologically-relevant ribosome copy
 921 numbers. Instead, cells limit their pool of actively translating ribosomes by decreasing f_a (**Dai**
 922 *et al.*, 2016), which would help maintain the pool of available amino acids $[AA]_{\text{eff}}$ and increase the
 923 achievable elongation rate. This observation is in agreement with the central premise of the cellular
 924 resource allocation principle proposed by **Scott et al. (2010)**; **Klumpp et al. (2009)**; **Klumpp and Hwa**
 925 (**2014**) and **Hui et al. (2015)**.

926 Discussion

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

942 In general, the copy numbers of the complexes involved in these processes were reasonable
 943 agreement with our order-of-magnitude estimates. Since many of these estimates represent soft
 944 lower-bound quantities, this suggests that cells do not express proteins grossly in excess of what
 945 is needed for a particular growth rate. Several exceptions, however, also highlight the dichotomy
 946 between a proteome that appears to "optimize" expression according to growth rate and one that
 947 must be able to quickly adapt to environments of different nutritional quality. Take, for example, the
 948 expression of carbon transporters. Shown in **Figure 2(B)**, we find that cells always express a similar
 949 number of glucose transporters irrespective of growth condition. At the same time, it is interesting
 950 to note that many of the alternative carbon transporters are still expressed in low but non-zero
 951 numbers (≈ 10 -100 copies per cell) across growth conditions. This may relate to the regulatory

953 configuration for many of these operons, which require the presence of a metabolite signal in
 954 order for alternative carbon utilization operons to be induced (*Monod, 1949; Laxhuber et al., 2020*).
 955 Furthermore, upon induction, these transporters are expressed and present in abundances in close
 956 agreement with a simple estimate.

957 Of the processes illustrated in *Figure 1*, we arrive at a ribosome-centric view of cellular growth
 958 rate control. This is in some sense unsurprising given the long-held observation that *E. coli* and
 959 many other organisms vary their ribosomal abundance as a function of growth conditions and
 960 growth rate *Scott et al. (2010); Metzl-Raz et al. (2017)*. However, through our dialogue with the
 961 proteomic data, two additional key points emerge. The first relates to our question of what process
 962 sets the absolute speed limit of bacterial growth. While a cell can parallelize many of its processes
 963 simply by increasing the abundance of specific proteins or firing multiple rounds of DNA replication,
 964 this is not so for synthesis of ribosomes (*Figure 10(A)*). The translation time for each ribosome [\approx
 965 6 min, *Dill et al. (2011)*] places an inherent limit on the growth rate that can only be surpassed
 966 if the cell were to increase their polypeptide elongation rate, or if they could reduce the total
 967 protein and rRNA mass of the ribosome. The second point relates to the long-observed correlations
 968 between growth rate and cell size (*Schaechter et al., 1958; Si et al., 2017*), and between growth
 969 rate and ribosomal mass fraction. While both trends have sparked tremendous curiosity and
 970 driven substantial amounts of research in their own regards, these relationships are themselves
 971 intertwined. In particular, it is the need for cells to increase their absolute number of ribosomes
 972 under conditions of rapid growth that require cells to also grow in size. Further experiments are
 973 needed to test the validity of this hypothesis. In particular, we believe that the change in growth
 974 rate in response to translation-inhibitory drugs (such as chloramphenicol) could be quantitatively
 975 predicted, given one had precision measurement of the relevant parameters, including the fraction
 976 of actively translating ribosomes f_a and changes in the metabolic capacity of the cell (i.e. the
 977 parameter r_{AA} in our minimal model) for a particular growth condition.

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

992 This work is by no means meant to be an exhaustive dissection of bacterial growth rate control,
 993 and there are many aspects of the bacterial proteome and growth that we neglected to consider.
 994 For example, other recent work (*Liebermeister et al., 2014; Hui et al., 2015; Schmidt et al., 2016*)
 995 has explored how the proteome is structured and how that structure depends on growth rate. In
 996 the work of *Hui et al. (2015)*, the authors coarse-grained the proteome into six discrete categories
 997 being related to either translation, catabolism, anabolism, and others related to signaling and
 998 core metabolism. The relative mass fraction of the proteome occupied by each sector could be
 999 modulated by external application of drugs or simply by changing the nutritional content of the
 1000 medium. While we have explored how the quantities of individual complexes are related to cell
 1001 growth, we acknowledge that higher-order interactions between groups of complexes or metabolic
 1002 networks at a systems-level may reveal additional insights into how these growth-rate dependences
 1003 are mechanistically achieved. Furthermore, while we anticipate the conclusions summarized here

1004 are applicable to a wide collection of bacteria with similar lifestyles as *E. coli*, other bacteria and
1005 archaea may have evolved other strategies that were not considered. Further experiments with the
1006 level of rigor now possible in *E. coli* will need to be performed in a variety of microbial organisms to
1007 learn more about how regulation of proteomic composition and growth rate control has evolved
1008 over the past 3.5 billion years.

1009 **Acknowledgements**

1010 We thank Matthias Heinemann, Alexander Schmidt, and Gene-Wei Li for additional input regarding
1011 their data. We also thank members of the Phillips, Theriot, Kondev, and Garcia labs for useful
1012 discussions. R.P. is supported by La Fondation Pierre-Gilles de Gennes, the Rosen Center at Caltech,
1013 and the NIH 1R35 GM118043 (MIRA). J.A.T. is supported by the Howard Hughes Medical Institute,
1014 and NIH Grant R37-AI036929. N.M.B is a HHMI Fellow of The Jane Coffin Childs Memorial Fund.

1015 **Competing Interests**

1016 The authors declare no competing interests.

1017 Appendix for: Fundamental limits on 1018 the rate of bacterial cell division

1019 **Nathan M. Belliveau^{†, 1}, Griffin Chure^{†, 2}, Christina L. Hueschen³, Hernan G.**
1020 **Garcia⁴, Jane Kondev⁵, Daniel S. Fisher⁶, Julie A. Theriot^{1, 7, *}, Rob Phillips^{8, 9, *}**

*For correspondence:

¹⁰²¹ These authors contributed equally
¹⁰²² to this work

¹⁰²³ ¹Department of Biology, University of Washington, Seattle, WA, USA; ²Department of
Applied Physics, California Institute of Technology, Pasadena, CA, USA; ³Department of
Chemical Engineering, Stanford University, Stanford, CA, USA; ⁴Department of Molecular
Cell Biology and Department of Physics, University of California Berkeley, Berkeley, CA,
USA; ⁵Department of Physics, Brandeis University, Waltham, MA, USA; ⁶Department of
Applied Physics, Stanford University, Stanford, CA, USA; ⁷Allen Institute for Cell Science,
Seattle, WA, USA; ⁸Division of Biology and Biological Engineering, California Institute of
Technology, Pasadena, CA, USA; ⁹Department of Physics, California Institute of
Technology, Pasadena, CA, USA; *Co-corresponding authors. Address correspondence to
¹⁰²⁹ phillips@pboc.caltech.edu and jtheriot@uw.edu
¹⁰³⁰

1031 **Contents**

1032	Introduction	1
1033	Uptake of Nutrients	2
1034	Nitrogen Transport	4
1035	Carbon Transport	4
1036	Phosphorus and Sulfur Transport	6
1037	Limits on Transporter Expression	8
1038	Energy Production	8
1039	ATP Synthesis	8
1040	Generating the Proton Electrochemical Gradient	9
1041	Energy Production in a Crowded Membrane.	9
1042	Synthesis of the Cell Envelope	11
1043	Lipid Synthesis	11
1044	Peptidoglycan Synthesis	13
1045	Function of the Central Dogma	15
1046	DNA	15
1047	dNTP synthesis	16
1048	DNA Replication	16
1049	RNA Synthesis	18
1050	rRNA	18
1051	mRNA	19
1052	tRNA	19
1053	RNA Polymerase and σ -factor Abundance	20
1054	Translation and Ribosomal Synthesis	21
1055	tRNA Synthetases	21
1056	Protein Synthesis	22
1057	Translation and Ribosomal Synthesis as a Growth-Rate Limiting Step	22
1058	Relationship Between Cell Size and Growth Rate	26
1059	Nutrient-Mediated Regulation of Proteomic Composition and Growth Rate	26
1060	Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid Scarcity	27
1061	Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and Metabolic Capacity.	30
1063	Discussion	30
1064	Acknowledgements	32
1065	Competing Interests	32
1066	Experimental Details Behind Proteomic Data	36
1067	Fluorescence based measurements	36
1068	Ribosomal profiling measurements	36
1069	Mass spectrometry measurements	37
1070	Summary of Proteomic Data	37
1071	Estimation of Cell Size and Surface Area	38

1072	Estimation of Total Protein Content per Cell	39
1073	Estimating Volume and Number of Amino Acids from Ribosome Copy Number	40
1074	Additional Considerations of Schmidt <i>et al.</i> Data Set	40
1075	Effect of cell volume on reported absolute protein abundances	44
1076	Relaxing assumption of constant protein concentration across growth conditions	44
1077	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	45
1078	Calculation of Complex Abundance	45
1079	Extending Estimates to a Continuum of Growth Rates	46
1080	Estimation of the total cell mass	46
1081	Complex Abundance Scaling With Cell Volume	48
1082	A Relation for Complex Abundance Scaling With Surface Area	49
1083	Number of Lipids	49
1084	Number of Murein Monomers	49
1085	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	49
1086	Calculation of active ribosomal fraction.	50
1087	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	50

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).

1088 **Experimental Details Behind Proteomic Data**

1089 Here we provide a brief summary of the experiments behind each proteomic data set. The purpose
 1090 of this section is to identify how the authors arrived at absolute protein abundances. In the
 1091 following section (Section Summary of Proteomic Data) we will then provide a summary of the
 1092 final protein abundance measurements that were used throughout the main text. Table 1 provides
 1093 an overview of the publications we considered. These are predominately mass spectrometry-
 1094 based, with the exception of the work from Li *et al.* (2014) which used ribosomal profiling, and the
 1095 fluorescence-based counting done in Taniguchi *et al.* (2010).

1096 **Fluorescence based measurements**

1097 In the work of Taniguchi *et al.* (2010), the authors used a chromosomal YFP fusion library where
 1098 individual strains have a specific gene tagged with a YFP-coding sequence. 1018 of their 1400
 1099 attempted strains were used in the work. A fluorescence microscope was used to collect cellular
 1100 YFP intensities across all these strains. Through automated image analysis, the authors normalized
 1101 intensity measurements by cell size to account for the change in size and expression variability
 1102 across the cell cycle. Following correction of YFP intensities for cellular autofluorescence, final
 1103 absolute protein levels were determined by a calibration curve with single-molecule fluorescence
 1104 intensities. This calibration experiment was performed separately using a purified YFP solution.

1105 **Ribosomal profiling measurements**

1106 The work of Li *et al.* (2014) takes a sequencing based approach to estimate protein abundance.
 1107 Ribosomal profiling, which refers to the deep sequencing of ribosome-protected mRNA fragments,
 1108 can provide a quantitative measurement of the protein synthesis rate. As long as the protein
 1109 life-time is long relative to the cell doubling time, it is possible to estimate absolute protein copy
 1110 numbers. The absolute protein synthesis rate has units of proteins per generation, and for stable
 1111 proteins will also correspond to the protein copy number per cell.

1112 In the experiments, ribosome-protected mRNA is extracted from cell lysate and selected on
 1113 a denaturing polyacrylamide gel for deep sequencing (15–45 nt long fragments collected and
 1114 sequenced by using an Illumina HiSeq 2000 in Li *et al.* (2014)). Counts of ribosome footprints from
 1115 the sequencing data were then corrected empirically for position-dependent biases in ribosomal
 1116 density across each gene, as well as dependencies on specific sequences including the Shine-
 1117 Dalgarno sequence. These data-corrected ribosome densities represent relative protein synthesis
 1118 rates. Absolute protein synthesis rates are obtained by multiplying the relative rates by the total

1119 cellular protein per cell. The total protein per unit volume was determined with the Lowry method
 1120 to quantify total protein, calibrated against bovine serum albumin (BSA). By counting colony-forming
 1121 units following serial dilution of their cell cultures, they then calculated the total protein per cell.

1122 Mass spectrometry measurements

1123 Perhaps not surprisingly, the data is predominantly mass spectrometry based. This is largely due to
 1124 tremendous improvements in the sensitivity of mass spectrometers, as well as improvements in
 1125 sample preparation and data analysis pipelines. It is now a relatively routine task to extract protein
 1126 from a cell and quantify the majority of proteins present by shotgun proteomics. In general, this
 1127 involves lysing cells, enzymatically digesting the proteins into short peptide fragments, and then
 1128 introducing them into the mass spectrometer (e.g. with liquid chromatography and electrospray
 1129 ionization), which itself can have multiple rounds of detection and further fragmentation of the
 1130 peptides.

1131 Most quantitative experiments rely on labeling protein with stable isotopes, which allow multiple
 1132 samples to be measured together by the mass spectrometer. By measuring samples of known total
 1133 protein abundance simultaneously (i.e. one sample of interest, and one reference), it is possible to
 1134 determine relative protein abundances. Absolute protein abundances can be estimated following
 1135 the same approach used above for ribosomal profiling, which is to multiply each relative abundance
 1136 measurement by the total cellular protein per cell. This is the approach taken by *Valgepea et al. (2013)*
 1137 and *Peebo et al. (2015)*, with relative protein abundances determined based on the relative
 1138 peptide intensities (label free quantification 'LFQ' intensities). For the data of *Valgepea et al. (2013)*,
 1139 total protein per cell was determined by measuring total protein by the Lowry method, and counting
 1140 colony-forming units following serial dilution. For the data from *Peebo et al. (2015)*, the authors did
 1141 not determine cell quantities and instead report the cellular protein abundances in protein per unit
 1142 volume by assuming a mass density of 1.1 g/ml, with a 30% dry mass fraction.

1143 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide
 1144 fragments of known abundance. These can serve as a direct way to calibrate mass spectrometry
 1145 signal intensities to absolute mass. This is the approach taken by *Schmidt et al. (2016)*. In addition
 1146 to a set of shotgun proteomic measurements to determine proteome-wide relative abundances,
 1147 the authors also performed absolute quantification of 41 proteins covering over four orders of
 1148 magnitude in cellular abundance. Here, a synthetic peptide was generated for each of the proteins,
 1149 doped into each protein sample, and used these to determine absolute protein abundances of the
 1150 41 proteins. These absolute measurements, determined for every growth condition, were then
 1151 used as a calibration curve to convert proteomic-wide relative abundances into absolute protein
 1152 abundance per cell. A more extensive discussion of the *Schmidt et al. (2016)* data set can be found
 1153 in Section Additional Considerations of Schmidt et al. Data Set.

1154 Summary of Proteomic Data

1155 In the work of the main text we only used the data from *Valgepea et al. (2013)*; *Li et al. (2014)*;
 1156 *Peebo et al. (2015)*; *Schmidt et al. (2016)*. As shown in *Figure 13(A)*, the reported total protein
 1157 abundances in the work of *Taniguchi et al. (2010)* and *Soufi et al. (2015)* differed quite substantially
 1158 from the other work. For the work of *Taniguchi et al. (2010)* this is in part due to a lower coverage in
 1159 total proteomic mass quantified, though we also noticed that most proteins appear undercounted
 1160 when compared to the other data.

1161 *Figure 13(B)* summarizes the total protein mass for each data point in our final compiled data
 1162 set. We note that protein abundances were all scaled so they followed a common growth rate-
 1163 dependent change in total protein mass. While our inclination initially was to leave reported copy
 1164 numbers untouched, a notable discrepancy in the scaling total protein per cell between *Schmidt*
 1165 *et al. (2016)* and the other data sets forced us to dig deeper into those measurements (compare
 1166 *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 13(A)*). The particular trend in *Schmidt et al.*
 1167 *(2016)* appears to be due to assumptions of cell size and we provide a more extensive discussion

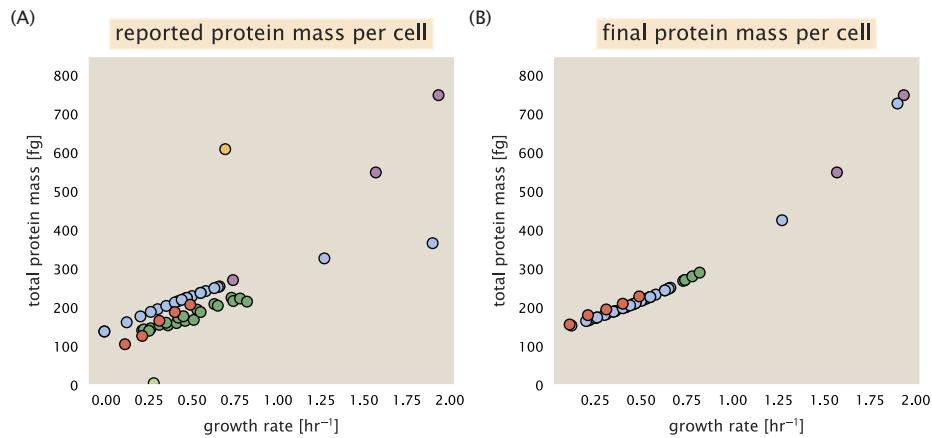


Figure 13. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as original reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al.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

1186 Estimation of Cell Size and Surface Area

1187 Since most of the proteomic data sets lack cell size (i.e. volume) measurements, we chose instead
 1188 to use a common estimate of size for any analysis requiring cell size or surface area. Since each
 1189 of the data sets used either K-12 MG1655 or its derivative, BW25113 (from the lab of Barry L.
 1190 Wanner; the parent strain of the Keio collection (*Datsenko and Wanner, 2000*; *Baba et al., 2006*)),
 1191 we fit the MG1655 cell size data from the supplemental material of *Si et al. (2017, 2019)* using the
 1192 `optimize.curve_fit` function from the Scipy python package (*Virtanen et al., 2020*).

1193 The average size measurements from each of their experiments are shown in Figure **Figure 15**,
 1194 with cell length and width shown in (A) and (B), respectively. The length data was well described by
 1195 the exponential function $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, while the width data was well described by $0.64 e^{0.24 \cdot \lambda}$.

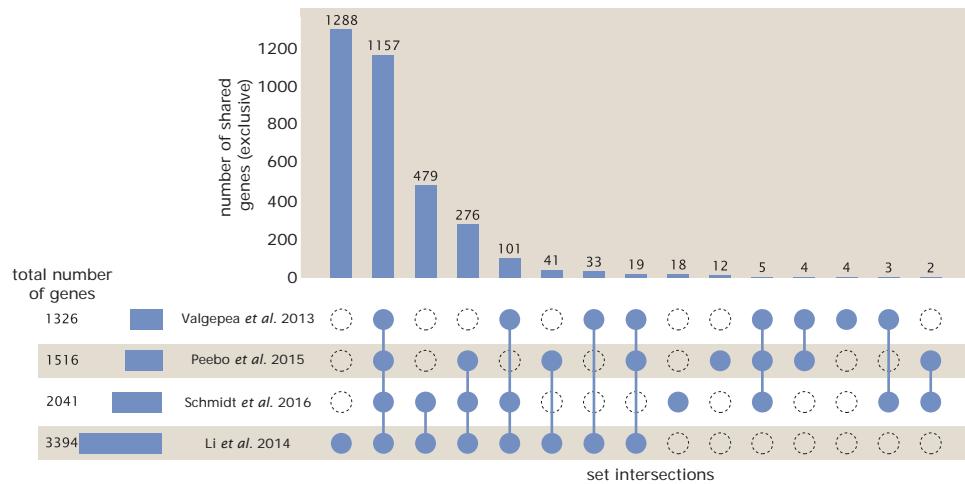


Figure 14. Comparison of proteomic coverage across different data sets. An UpSet diagram (Lex et al., 2014) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of Valgepea et al. (2013); Li et al. (2014); Schmidt et al. (2016); Peebo et al. (2015).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

1196 μm. In order to estimate cell size we take the cell as a cylinders with two hemispherical ends (Si
 1197 et al., 2017; Basan et al., 2015). Specifically, cell size is estimated from,

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

1198 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
 1199 the cell surface area is given by,

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

1200 where η is the aspect ratio ($\eta = l/w$) (Ojkic et al., 2019).

1201 Estimation of Total Protein Content per Cell

1202 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the
 1203 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
 1204 protein content from the total dry mass of the cell. Here we begin by noting that for almost the
 1205 entire range of growth rates considered here, protein, DNA, and RNA were reported to account
 1206 for at least 90 % of the dry mass (Basan et al. (2015)). The authors also found that the total dry
 1207 mass concentration was roughly constant across growth conditions. Under such a scenario, we can
 1208 calculate the total dry mass concentration for protein, DNA, and RNA, which is given by $1.1 \text{ g/ml} \times$
 1209 $30\% \times 90\%$ or about $[M_p] = 300 \text{ fg per fl}$. Multiplying this by our prediction of cell size gives the
 1210 total dry mass per cell.

1211 However, even if dry mass concentration is relatively constant across growth conditions, it
 1212 is not obvious how protein concentration might vary due to the substantial increase in rRNA at
 1213 faster growth rates (Dai et al. (2016)). This is a well-documented result that arises from an increase
 1214 in ribosomal abundance at faster growth rates (Scott et al. (2010)). To proceed therefore rely on
 1215 experimental measurements of total DNA content per cell that also come from Basan et al., and RNA
 1216 to protein ratios that were measured in Dai et al. (and cover the entire range of growth conditions
 1217 considered here). These are reproduced in Figure Figure 16(A) and (B), respectively.

1218 Assuming that the protein, DNA, and RNA account for 90 % of the total dry mass, the protein
 1219 mass can then determined by first subtracting the experimentally measured DNA mass, and then

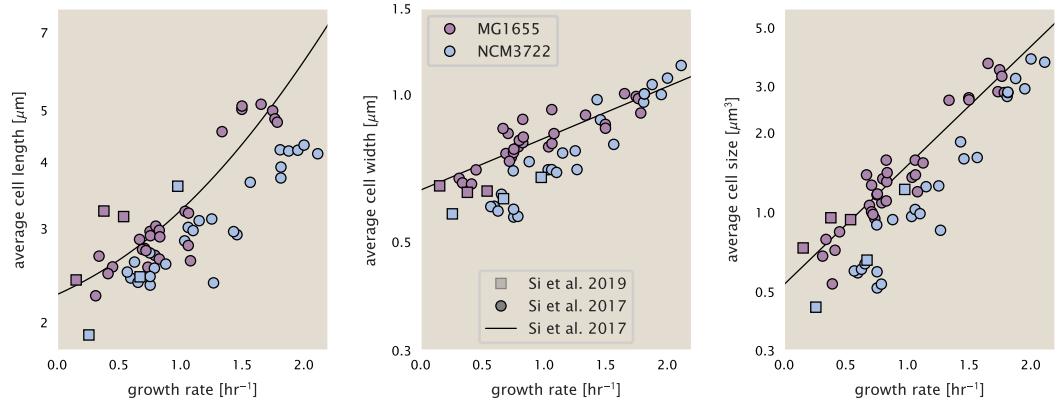


Figure 15. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300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

1220 using the experimental estimate of the RNA to protein ratio. The total protein per cell is will be
1221 related to the summed RNA and protein mass by,

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

1222 (RP_{ratio}) refers to the RNA to protein ratio as measured by Dai et al.. In Figure **Figure 16(C)** we plot
1223 the estimated cellular concentrations for protein, DNA, and RNA from these calculations, and in
1224 Figure **Figure 16(D)** we plot their total expected mass per cell. This later quantity is the growth
1225 rate-dependent total protein mass that was used to estimate total protein abundance across all
1226 data sets (and summarized in **Figure 13(B)**).

1227 Estimating Volume and Number of Amino Acids from Ribosome Copy Number

1228 Towards the end of the main text, we examine a coarse-grained model of nutrient-limited growth. A
1229 key point in our analysis was to consider how elongation rate r_i and growth rate λ vary with respect
1230 to the experimentally observed changes in cell size, total number of peptide bonds per cell N_{pep} ,
1231 and ribosomal content. In order to do maintain parameters in line with the experimental data, but
1232 otherwise allow us to explore the model, we performed a phenomenological fit of N_{pep} and V as
1233 a function of the measured ribosomal copy number R . As has been described in the preceding
1234 sections of this supplement, we estimate cell volume for each growth condition using the size
1235 measurements from **Si et al. (2017, 2019)**, and N_{pep} is approximated by taking the total protein
1236 mass and dividing this number by the average mass of an amino acid, 110 Da (BNID: 104877).

1237 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression
1238 of the log transform of these parameters as a function of the log transform of the ribosome copy
1239 number. Using optimization by minimization, we estimated the best-fit values of the intercept and
1240 slope for each regression. **Figure 17** shows the result of each regression as a dashed line.

1241 Additional Considerations of Schmidt et al. Data Set

1242 While the data set from **Schmidt et al. (2016)** remains a heroic effort that our labs continue to
1243 return to as a resource, there were steps taken in their calculation of protein copy number that
1244 we felt needed further consideration. In particular, the authors made an assumption of constant
1245 cellular protein concentration across all growth conditions and used measurements of cell volume

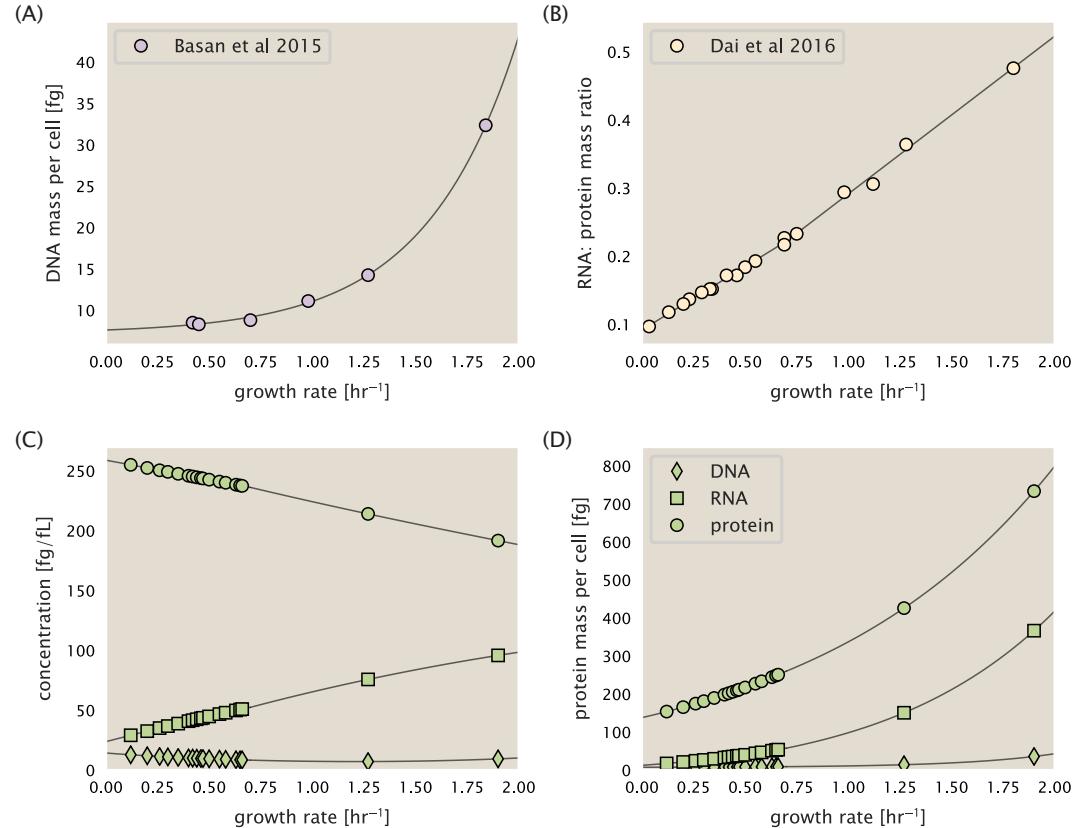


Figure 16. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23\lambda} + 7.2$ fg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

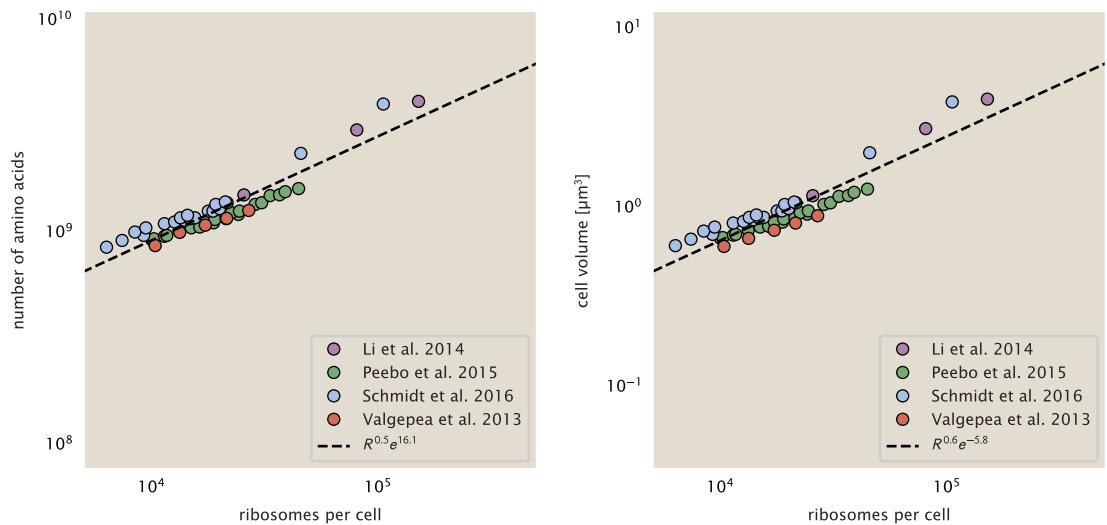


Figure 17. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows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

1246 that appear inconsistent with an expected exponential scaling of cell size with growth rate that is
 1247 well-documented in *E. coli* (*Schaechter et al. (1958); Taheri-Araghi et al. (2015); Si et al. (2017)*).

1248 We begin by looking at their cell volume measurements, which are shown in blue in Figure
 1249 **Figure 18**. As a comparison, we also plot cell sizes reported in three other recent papers: measure-
 1250 ments from Taheri-Araghi et al. and Si et al. come from the lab of Suckjoon Jun, while those from
 1251 Basan et al. come from the lab of Terence Hwa. Each set of measurements used microscopy and
 1252 cell segmentation to determine the length and width, and then calculated cell size by treating the
 1253 cell as a cylinder with two hemispherical ends, as we considered in the previous section. While there
 1254 is notable discrepancy between the two research groups, which are both using strain NCM3722,
 1255 Basan et al. found that this came specifically from uncertainty in determining the cell width. This is
 1256 prone to inaccuracy given the small cell size and optical resolution limits (further described in their
 1257 supplemental text). Perhaps the more concerning point is that while each of these alternative mea-
 1258 surements show an exponential increase in cell size at faster growth rates, the measurements used
 1259 by Schmidt et al. appear to plateau. This resulted in an analogous trend in their final reported total
 1260 cellular protein per cell as shown in Figure **Figure 19** (purple data points), and is in disagreement
 1261 with other measurements of total protein at these growth rates (*Basan et al., 2015*).

1262 Since it is not obvious how measurements of cell size influenced their reported protein abun-
 1263 dances, in the following subsections we begin by considering this calculation. We then consider
 1264 three different approaches to estimate the growth-rate dependent total protein mass to compare
 1265 with those values reported from *Schmidt et al. (2016)*. The results of this are summarized in **Fig-**
 1266 **ure 18(B)**, with the original values from both *Schmidt et al. (2016)* and *Li et al. (2014)* shown in
 1267 **Figure 18(A)** for reference. For most growth conditions, we find that total protein per cell is still in
 1268 reasonable agreement. However, for the fastest growth conditions, with glycerol + supplemented
 1269 amino acids, and LB media, all estimates are substantially higher than those originally reported.
 1270 This is the main reason why we chose to readjusted protein abundance as shown in **Figure 13(B)**
 1271 (with the calculation described in section Estimation of Total Protein Content per Cell).

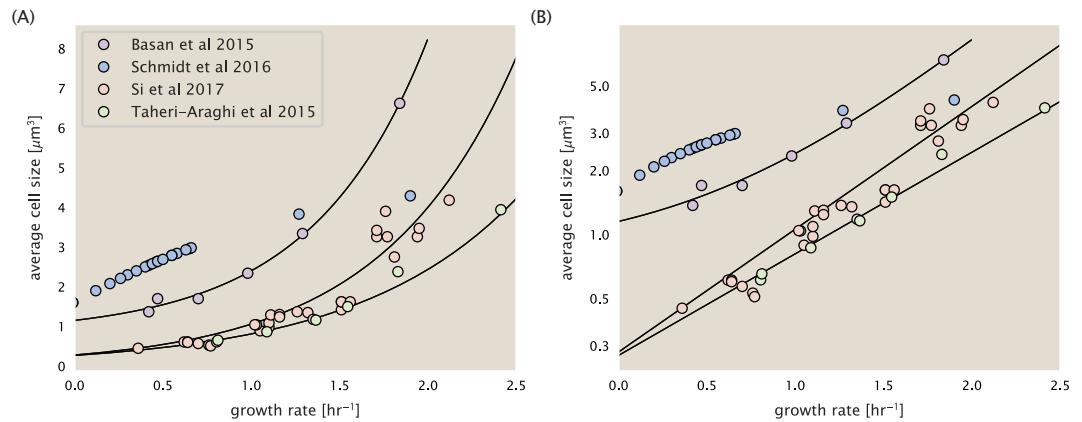


Figure 18. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are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

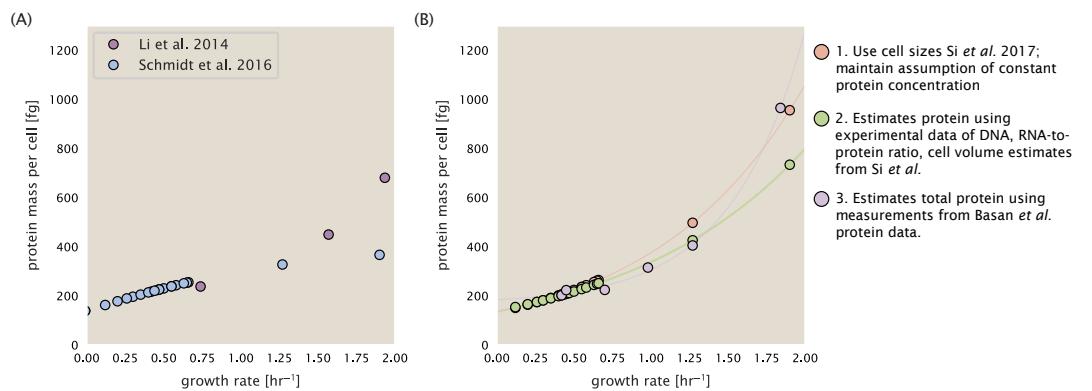


Figure 19. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt *et al.* (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. *light red*: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

1272 **Effect of cell volume on reported absolute protein abundances**

1273 As noted in section Experimental Details Behind Proteomic Data, the authors calculated proteome-
 1274 wide protein abundances by first determining absolute abundances of 41 pre-selected proteins,
 1275 which relied on adding synthetic heavy reference peptides into their protein samples at known
 1276 abundance. This absolute quantitation was performed in replicate for each growth condition.
 1277 Separately, the authors also performed a more conventional mass spectrometry measurement
 1278 for samples from each growth condition, which attempted to maximize the number of quantified
 1279 proteins but only provided relative abundances based on peptide intensities. Finally, using their 41
 1280 proteins with absolute abundances already determined, they then created calibration curves with
 1281 which to relate their relative intensity to absolute protein abundance for each growth condition. This
 1282 allowed them to estimate absolute protein abundance for all proteins detected in their proteome-
 1283 wide data set. Combined with their flow cytometry cell counts, they were then able to determine
 1284 absolute abundance of each protein detected on a per cell basis.

1285 While this approach provided absolute abundances, another necessary step to arrive at total
 1286 cellular protein was to account for any protein loss during their various protein extraction steps.
 1287 Here the authors attempted to determine total protein separately using a BCA protein assay. In
 1288 personal communications, it was noted that determining reasonable total protein abundances
 1289 by BCA across their array of growth conditions was particularly troublesome. Instead, they noted
 1290 confidence in their total protein measurements for cells grown in M9 minimal media + glucose
 1291 and used this as a reference point with which to estimate the total protein for all other growth
 1292 conditions.

1293 For cells grown in M9 minimal media + glucose an average total mass of $M_p = 240$ fg per
 1294 cell was measured. Using their reported cell volume, reported as $V_{orig} = 2.84$ fl, a cellular protein
 1295 concentration of $[M_p]_{orig} = M_p/V_{orig} = 85$ fg/fl. Now, taking the assumption that cellular protein
 1296 concentration is relatively independent of growth rate, they could then estimate the total protein
 1297 mass for all other growth conditions from,

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

1298 where $M_{P,i}$ represents the total protein mass per cell and V_i is the cell volume for each growth
 1299 condition i as measured in Volkmer and Heinemann, 2011. Here the thinking is that the values
 1300 of $M_{P,i}$ reflects the total cellular protein for growth condition i , where any discrepancy from their
 1301 absolute protein abundance is assumed to be due to protein loss during sample preparation. The
 1302 protein abundances from their absolute abundance measurements noted above were therefore
 1303 scaled to their estimates and are shown in Figure 19 (purple data points).

1304 If we instead consider the cell volumes predicted in the work of Si *et al.*, we again need to take
 1305 growth in M9 minimal media + glucose as a reference with known total mass, but we can follow a
 1306 similar approach to estimate total protein mass for all other growth conditions. Letting $V_{Si_glu} = 0.6$
 1307 fl be the predicted cell volume, the cellular protein concentration becomes $[M_p]_{Si} = M_p/V_{Si_glu} =$
 1308 400 fg/fl. The new total protein mass per cell can then be calculated from,

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

1309 where $M'_{P,i}$ is the new protein mass prediction, and $V_{Si,i}$ refers to the new volume prediction for each
 1310 condition i . These are shown as red data points in Figure 19(B).

1311 **Relaxing assumption of constant protein concentration across growth conditions**

1312 We next relax the assumption that cellular protein concentration is constant and instead, attempt
 1313 to estimate it using experimental data. Here we use the estimation of total protein mass per cell
 1314 detailed in section Estimation of Total Protein Content per Cell for all data points in the Schmidt
 1315 *et al.* (2016) data set. The green data points in Figure 19(B) show this prediction, and this represents
 1316 the approach used to estimate total protein per cell for all data sets.

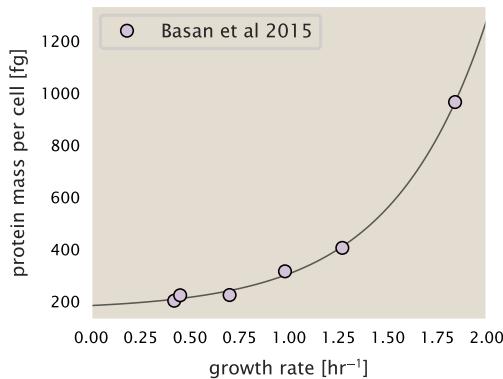


Figure 20. Total cellular protein reported in Basan *et al.* 2015. Measured protein mass as a function of growth rate as reproduced from Basan *et al.* 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given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

1317 Comparison with total protein measurements from Basan *et al.* 2015.

1318 One of the challenges in our estimates in the preceding sections is the need to estimate protein
 1319 concentration and cell volumes. These are inherently difficult to accurately due to the small size
 1320 of *E. coli*. Indeed, for all the additional measurements of cell volume included in Figure **Figure 18**,
 1321 no measurements were performed for cells growing at rates below $0.5\ hr^{-1}$. It therefore remains to
 1322 be determined whether our extrapolated cell volume estimates are appropriate, with the possibility
 1323 that the logarithmic scaling of cell size might break down for slower growth.

1324 In our last approach we therefore attempt to estimate total protein using experimental data
 1325 that required no estimates of concentration or cell volume. Specifically, in the work of Basan *et al.*,
 1326 the authors measured total protein per cell for a broad range of growth rates (reproduced in Figure
 1327 **Figure 20**). These were determined by first measuring bulk protein from cell lysate, measured by
 1328 the colorimetric Biuret method (*You et al. (2013)*), and then abundance per cell was calculated from
 1329 cell counts from either plating cells or a Coulter counter. While it is unclear why Schmidt *et al.* was
 1330 unable to take a similar approach, the results from Basan *et al* appear more consistent with our
 1331 expectation that cell mass will increase exponentially with faster growth rates. In addition, although
 1332 they do not consider growth rates below about $0.5\ hr^{-1}$, it is interesting to note that the protein
 1333 mass per cell appears to plateau to a minimum value at slow growth. In contrast, our estimates
 1334 using cell volume so far have predicted that total protein mass should continue to decrease slightly
 1335 for slower growing cells. By fitting this data to an exponential function dependent on growth rate,
 1336 we could then estimate the total protein per cell for each growth condition considered by *Schmidt*
 1337 *et al. (2016)*. These are plotted as red data points in **Figure 19(B)**.

1338 Calculation of Complex Abundance

1339 All protein data quantified the abundance of individual proteins per cell. However, this work
 1340 requires estimates on the abundance of individual protein *complexes*, rather than the copy number
 1341 of individual proteins. In this section, we outline the approach we used to annotate proteins as
 1342 being part of a macromolecular complex and how we computed their absolute abundances per cell.

1343 Protein complexes, and proteins individually, often have a variety of names, both longform and
 1344 shorthand. As individual proteins can have a variety of different synonyms, we sought to ensure
 1345 that each protein annotated in the data sets used the same synonym. To do use, we relied heavily
 1346 on the EcoCyc Genomic Database (*Keseler et al., 2017*). Each protein in available data sets included
 1347 an annotation of one of the gene name synonyms as well as an accession ID – either a UniProt
 1348 or Blattner "b-number". We programmatically matched up individual accession IDs between the

1349 proteins in different data sets. In cases where accession IDs matched but the gene names were
 1350 different, we manually verified that the gene product was the same between the datasets and chose
 1351 a single synonym. All code used in the data cleaning and unification procedures can be found on
 1352 the associated [GitHub repository] (DOI:XXX) associated with this paper as well as on the associated
 1353 [paper website](#).

1354 With each protein conforming to a single identification scheme, we then needed to identify the
 1355 molecular complexes each protein was a member of. Additionally, we needed to identify how many
 1356 copies of each protein were present in each complex (i.e. the subunit copy number) and compute
 1357 the estimated abundance complex that accounted for fluctuations in subunit stoichiometry. To
 1358 map proteins to complexes, we accessed the EcoCyc *E. coli* database **Keseler et al. (2017)** using
 1359 PathwayTools version 23.0 **Karp et al. (2019)**. With a license for PathWay Tools, we mapped each
 1360 unique protein to its annotated complexes via the BioCyc Python package. As we mapped each
 1361 protein with *all* of its complex annotations, there was redundancy in the dataset. For example,
 1362 ribosomal protein L20 (RplT) is annotated to be a component of the 50S ribosome (EcoCyc complex
 1363 CPLX-03962) as well as a component of the mature 70S ribosome (EcoCyc complex CPLX-03964).

1364 In addition to the annotated complex, we collected information on the stoichiometry of each
 1365 macromolecular complex. For a complex with N_{subunits} protein species, for each protein subunit i
 1366 we first calculate the number of complexes that *could* be formed given the measured protein copy
 1367 numbers per cell,

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

1368 Here, $P_{\text{subunit } i}^{(\text{measured})}$ refers to the measured protein copy number of species i , and m refers to the
 1369 number of monomers present for that protein in the complex. For example, the 70S mature
 1370 ribosome complex has 55 protein components, all of which are present in a single copy except
 1371 L4 (RpL), which is present in 4 copies ($m = 4$). For each ribosomal protein, we then calculate the
 1372 maximum number of complexes that could be formed using **Equation 17**. This example, along with
 1373 example from 5 other macromolecular complexes, can be seen in **Figure 21**.

1374 It is important to note that measurement noise, efficiency of protein extraction, and differences
 1375 in protein stability will mean that the precise value of each calculation will be different for each
 1376 component of a given complex. Thus, to report the total complex abundance, we use the arithmetic
 1377 mean of across all subunits in the complex,

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

1378 in **Figure 21**, we show this mean value as a grey line for a variety of different complexes. Addi-
 1379 tionally, we have built an interactive figure accessible on the [paper website](#) where the validity of
 1380 this approach can be examined for any complex with more than two subunits (thus, excluding
 1381 monomers and dimers).

1382 Extending Estimates to a Continuum of Growth Rates

1383 In the main text, we considered a standard stopwatch of 5000 s to estimate the abundance of
 1384 the various protein complexes considered. In addition to point estimates, we also showed the
 1385 estimate as a function of growth rate as transparent grey curves. In this section, we elaborate
 1386 on this continuum estimate, giving examples of estimates that scale with either cell volume, cell
 1387 surface area, or number of origins of replication.

1388 Estimation of the total cell mass

1389 For many of the processes estimated in the main text we relied on a cellular dry mass of ≈ 300 fg
 1390 from which we computed elemental and protein fractions using knowledge of fractional composition
 1391 of the dry mass. At modest growth rates, such as the 5000 s doubling time used in the main text, this

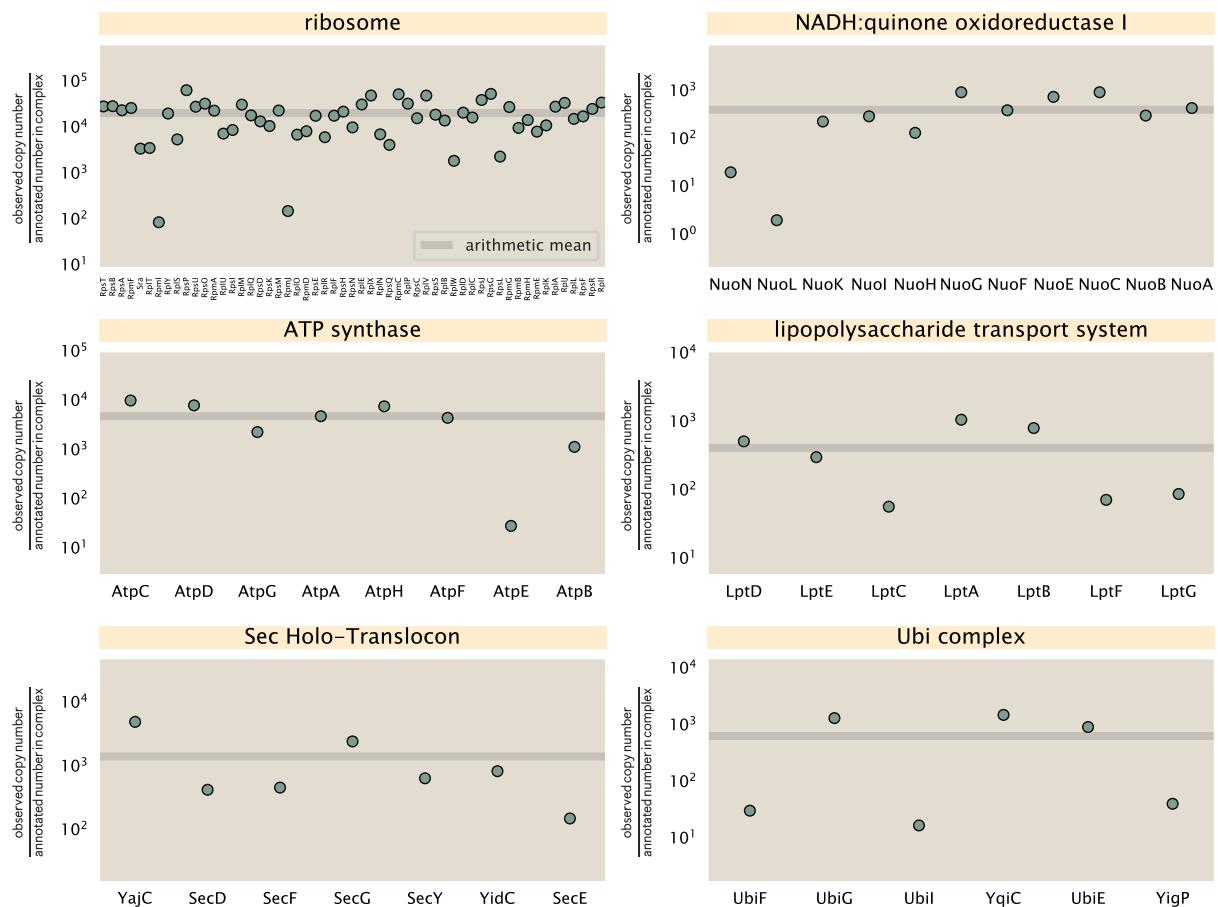


Figure 21. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds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

1392 is a reasonable number to use as the typical cell mass is ≈ 1 pg and *E. coli* cells can approximated as
 1393 70% water by volume. However, as we have shown in the preceding sections, the cell size is highly
 1394 dependent on the growth rate. This means that a dry mass of 300 fg cannot be used reliably across
 1395 all growth rates.

1396 Rather, using the phenomenological description of cell volume scaling exponentially with growth
 1397 rate, and using a rule-of-thumb of a cell buoyant density of ≈ 1.1 pg / fL (BNID: 103875), we can
 1398 calculate the cell dry mass across a range of physiological growth rates as

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

1399 where a and b are constants with units of μm^3 and hr, respectively. The value of these constants can
 1400 be estimated from the careful volume measurements performed by *Si et al.* (2017,?), as considered
 1401 in Appendix Estimation of Cell Size and Surface Area earlier.

1402 Complex Abundance Scaling With Cell Volume

1403 Several of the estimates performed in the main text are implicitly dependent on the cell volume.
 1404 This includes processes such as ATP utilization and, most prominently, the transport of nutrients,
 1405 whose demand will be proportional to the volume of the cell. Of the latter, we estimated the number
 1406 of transporters that would be needed to shuttle enough carbon, phosphorus, and sulfur across
 1407 the membrane to build new cell mass. To do so, we used elemental composition measurements
 1408 combined with a 300 fg cell dry mass to make the point estimate. As we now have a means to
 1409 estimate the total cell mass as a function of volume, we can generalize these estimates across
 1410 growth rates.

1411 Rather than discussing the particular details of each transport system, we will derive this scaling
 1412 expression in very general terms. Consider that we wish to estimate the number of transporters
 1413 for some substance X , which has been measured to be made up some fraction of the dry mass,
 1414 θ_X . If we assume that, irrespective of growth rate, the cell dry mass is relatively constant (*Basan*
 1415 *et al.*, 2015) and $\approx 30\%$ of the total cell mass, we can state that the total mass of substance X as a
 1416 function of growth rate is

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

1417 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in **Equation 19**. To convert
 1418 this to the number of units N_X of substance X in the cell, we can use the formula weight w_X of a
 1419 single unit of X in conjunction with **Equation 20**,

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

1420 To estimate the number of transporters needed, we make the approximation that loss of units
 1421 of X via diffusion through porins or due to the permeability of the membrane is negligible and that
 1422 a single transporter complex can transport substance X at a rate r_X . As this rate r_X is in units of X
 1423 per time per transporter, we must provide a time window over which the transport process can
 1424 occur. This is related to the cell doubling time τ , which can be calculated from the the growth rate λ
 1425 as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized transport scaling relation of

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

1426 This function is used to draw the continuum estimates for the number of transporters seen in
 1427 Figures 2 and 3 as transparent grey curves. Occasionally, this continuum scaling relationship will
 1428 not precisely agree with the point estimate outlined in the main text. This is due to the choice of \approx
 1429 300 fg total dry mass per cell for the point estimate, whereas we considered more precise values
 1430 of cell mass in the continuum estimate. We note, however, that both this scaling relation and the
 1431 point estimates are meant to describe the order-of-magnitude observed, and not the predict the
 1432 exact values of the abundances.

1433 **Equation 22** is a very general relation for processes where the cell volume is the "natural
 1434 variable" of the problem. This means that, as the cell increases in volume, the requirements for
 1435 substance *X* also scale with volume rather than scaling with surface area, for example. So long as
 1436 the rate of the process, the fraction of the dry mass attributable to the substance, and the formula
 1437 mass of the substance is known, **Equation 22** can be used to compute the number of complexes
 1438 needed. For example, to compute the number of ATP synthases per cell, **Equation 22** can be slightly
 1439 modified to the form

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

1440 where we have included the term N_{ATP} to account for the number of ATP equivalents needed per
 1441 amino acid for translation (≈ 4 , BNID: 114971), and w_{AA} is the average mass of an amino acid. The
 1442 grey curves in Figure 4 o the main text were made using this type of expression.

1443 A Relation for Complex Abundance Scaling With Surface Area

1444 In our estimation for the number of complexes needed for lipid synthesis and peptidoglycan
 1445 maturation, we used a particular estimate for the cell surface area ($\approx 5 \mu\text{m}$, BNID: 101792) and
 1446 the fraction of dry mass attributable to peptidoglycan ($\approx 3\%$, BNID: 101936). Both of these values
 1447 come from glucose-fed *E. coli* in balance growth. As we are interested in describing the scaling as a
 1448 function of the growth rate, we must also consider how these values scale with cell surface area,
 1449 which is the natural variable for these types of processes. In the coming paragraphs, we highlight
 1450 how we incorporate a condition-dependent surface area into our calculation of the number of lipids
 1451 and murein monomers that need to be synthesized and crosslinked, respectively.

1452 Number of Lipids

1453 To compute the number of lipids as a function of growth rate, we make the assumption that some
 1454 features, such as the surface area of a single lipid ($A_{\text{lipid}} \approx 0.5 \text{ nm}^2$, BNID: 106993) and the total
 1455 fraction of the membrane composed of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth
 1456 rate. Using these approximations combined with **Equation 13**, and recognizing that each membrane
 1457 is composed of two leaflets, we can compute the number of lipids as a function of growth rate as

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

1458 where η is the length-to-width aspect ratio and V is the cell volume.

1459 Number of Murein Monomers

1460 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we
 1461 used an empirical measurement that $\approx 3\%$ of the dry mass is attributable to peptidoglycan and that
 1462 a single murien monomer is $m_{\text{murein}} \approx 1000 \text{ Da}$. While the latter is independent of growth rate, the
 1463 former is not. As the peptidoglycan exists as a thin shell with a width of $w \approx 10 \text{ nm}$ encapsulating
 1464 the cell, one would expect the number of murein monomers scales with the surface area of this
 1465 shell. In a similar spirit to our calculation of the number of lipids, the total number of murein
 1466 monomers as a function of growth rate can be calculated as

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

1467 where ρ_{pg} is the density of peptidoglycan.

1468 Complex Abundance Scaling With Number of Origins, and rRNA Synthesis

1469 While the majority of our estimates hinge on the total cell volume or surface area, processes related
 1470 to the central dogma, namely DNA replication and synthesis of rRNA, depend on the number of
 1471 chromosomes present in the cell. As discussed in the main text, the ability of *E. coli* to parallelize the

replication of its chromosome by having multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived in *Si et al. (2017)* and reproduced in the main text and Appendix Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$ below, the average number of origins of replication at a given growth rate can be calculated as

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

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

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

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

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

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

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

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

Calculation of active ribosomal fraction.

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

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

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

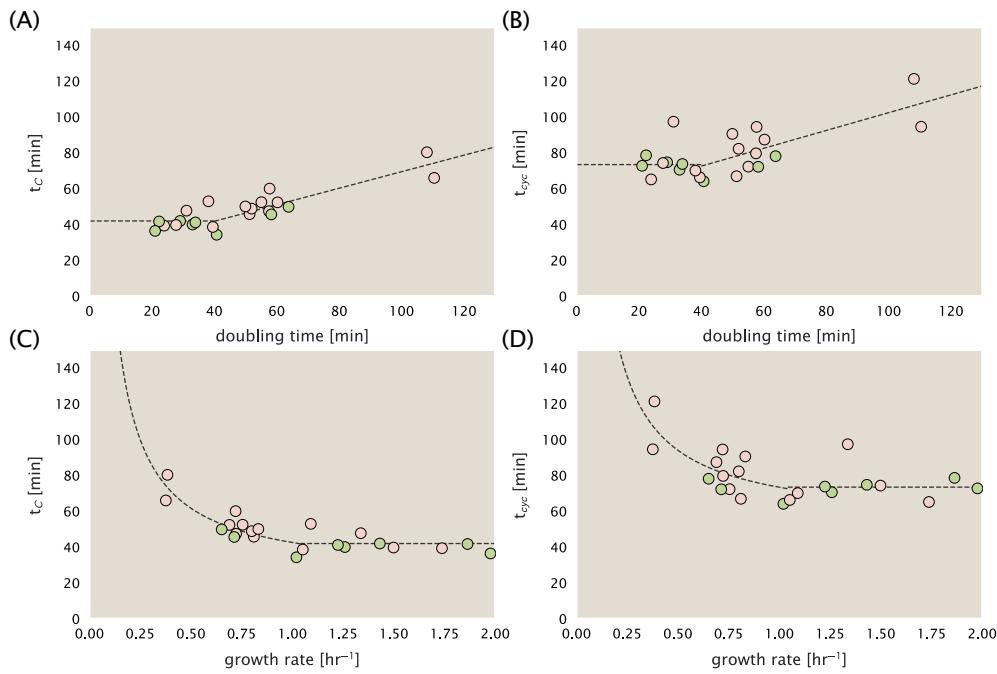


Figure 22. Estimation of $\langle \#ori \rangle / \langle \# ter \rangle$ and $\langle \#ori \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling times (rich media), t_C and t_{cyc} are assumed constant ($t_C = 42$ minutes, $t_{cyc} = 73$ minutes). At the transition, taken to occur at 40 minutes, the dashed line corresponds to an assumed proportional increase in each parameter as a function of the doubling time ($t_C = 0.46\tau + 23.3$ minutes, $t_{cyc} = 0.50\tau + 52.7$ minutes). (C) and (D) plot the same data as in (A) and (B), but as a function of growth rate, given by $\lambda = \ln(2)/\tau$.

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

In Figure 22(A) and (B) we plot the measured t_C and t_{cyc} values versus the doubling time from Si et al. (2017). The authors estimated t_C by marker frequency analysis using qPCR, while $t_{cyc} = t_C + t_D$ were inferred from t_C and τ . In the plots we see that both t_C and t_{cyc} reach a minimum at around 40 and 75 minutes, respectively. For a C period of 40 minutes, this would correspond to a maximum rate of elongation of about 1,000 bp/sec. Since we lacked a specific model to describe how each of these parameters vary with growth condition, we assumed that they were linearly dependent on the doubling time. For each parameter, t_C and t_{cyc} , we split them up into two domains corresponding to poorer nutrient conditions and rich nutrient conditions (cut off at $\tau \approx 40$ minutes where chromosomal replication becomes nearly constant). The fit lines are shown as solid black lines. In Figure 22(C) and (D) we also show t_C and t_{cyc} as a function of growth rate λ along with our piecewise linear fits, which match the plots in the main text.

References

- 1532 **Abelson HT**, Johnson LF, Penman S, Green H. Changes in RNA in Relation to Growth of the Fibroblast: II.
 1533 The Lifetime of mRNA, rRNA, and tRNA in Resting and Growing Cells. *Cell*. 1974 Apr; 1(4):161–165. doi:
 1534 10.1016/0092-8674(74)90107-X.
- 1535
 1536 **Aidellberg G**, Towbin BD, Rothschild D, Dekel E, Bren A, Alon U. Hierarchy of Non-Glucose Sugars in *Escherichia coli*. *BMC Systems Biology*. 2014 Dec; 8(1):133. doi: 10.1186/s12918-014-0133-z.
- 1537
 1538 **Antonenko YN**, Pohl P, Denisov GA. Permeation of Ammonia across Bilayer Lipid Membranes Studied by Ammonium Ion Selective Microelectrodes. *Biophysical Journal*. 1997 May; 72(5):2187–2195.
- 1539
 1540 **Ashburner M**, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene Ontology: tool for the unification of biology. *Nature Genetics*. 2000 May; 25(1):25–29.
- 1541
 1542
 1543 **Ason B**, Bertram JG, Hingorani MM, Beechem JM, O'Donnell M, Goodman MF, Bloom LB. A Model for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*. 2000 Jan; 275(4):3006–3015. doi: 10.1074/jbc.275.4.3006.
- 1544
 1545
 1546
 1547 **Assentoft M**, Kaptan S, Schneider HP, Deitmer JW, de Groot BL, MacAulay N. Aquaporin 4 as a NH₃ Channel. *Journal of Biological Chemistry*. 2016 Sep; 291(36):19184–19195. doi: 10.1074/jbc.M116.740217.
- 1548
 1549 **Baba T**, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*. 2006 Jun; 2(1):2460.
- 1550
 1551
 1552 **Basan M**, Zhu M, Dai X, Warren M, Sévin D, Wang YP, Hwa T. Inflating bacterial cells by increased protein synthesis. *Molecular Systems Biology*. 2015 Oct; 11(10):836.
- 1553
 1554 **Bauer S**, Ziv E. Dense Growth of Aerobic Bacteria in a Bench-Scale Fermentor. *Biotechnology and Bioengineering*. 1976; 18(1):81–94. doi: 10.1002/bit.260180107, eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107.
- 1555
 1556
 1557 **Belliveau NM**, Barnes SL, Ireland WT, Jones DL, Sweredoski MJ, Moradian A, Hess S, Kinney JB, Phillips R. Systematic Approach for Dissecting the Molecular Mechanisms of Transcriptional Regulation in Bacteria. *Proceedings of the National Academy of Sciences*. 2018 May; 115(21):E4796–E4805. doi: 10.1073/pnas.1722055115.
- 1558
 1559
 1560 **Bennett BD**, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology*. 2009 Aug; 5(8):593–599.
- 1561
 1562 **Birnbaum LS**, Kaplan S. Localization of a Portion of the Ribosomal RNA Genes in *Escherichia coli*. *Proceedings of the National Academy of Sciences*. 1971 May; 68(5):925–929. doi: 10.1073/pnas.68.5.925.
- 1563
 1564 **Booth IR**, Mitchell WJ, Hamilton WA. Quantitative Analysis of Proton-Linked Transport Systems. The Lactose Permease of *Escherichia coli*. *Biochemical Journal*. 1979 Sep; 182(3):687–696.
- 1565
 1566 **Bremer H**, Dennis PP. Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus*. 2008 Oct; 3(1).
- 1567
 1568 **Browning DF**, Busby SJW. Local and global regulation of transcription initiation in bacteria. *Nature Reviews Microbiology*. 2016 Aug; 14(10):638–650.
- 1569
 1570 **Büke F**, Grilli J, Lagomarsino MC, Bokinsky G, Tans S. ppGpp is a bacterial cell size regulator. *bioRxiv*. 2020 Jun; 266:2020.06.16.154187.
- 1571
 1572 **Catherwood AC**, Lloyd AJ, Tod JA, Chauhan S, Slade SE, Walkowiak GP, Galley NF, Punekar AS, Smart K, Rea D, Evans ND, Chappell MJ, Roper DI, Dowson CG. Substrate and Stereochemical Control of Peptidoglycan Cross-Linking by Transpeptidation by *Escherichia Coli* PBP1B. *Journal of the American Chemical Society*. 2020 Mar; 142(11):5034–5048. doi: 10.1021/jacs.9b08822.
- 1573
 1574
 1575
 1576 **Cox GB**, Newton NA, Gibson F, Snoswell AM, Hamilton JA. The Function of Ubiquinone in *Escherichia coli*. *Biochemical Journal*. 1970 Apr; 117(3):551–562. doi: 10.1042/bj1170551.
- 1577
 1578 **Dai X**, Zhu M, Warren M, Balakrishnan R, Okano H, Williamson JR, Fredrick K, Hwa T. Slowdown of Translational Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*. 2018 Feb; 9(1):281.
- 1579

- 1580 **Dai X**, Zhu M, Warren M, Balakrishnan R, Patsalo V, Okano H, Williamson JR, Fredrick K, Wang YP, Hwa T. Reduction
1581 of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature Microbiology*. 2016 Dec; 2(2):16231.
- 1583 **Datsenko KA**, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR
1584 products. *Proceedings of the National Academy of Sciences*. 2000 Jun; 97(12):6640–6645.
- 1585 **Delarue M**, Brittingham GP, Pfeffer S, Surovtsev IV, Pinglay S, Kennedy KJ, Schaffer M, Gutierrez JI, Sang D,
1586 Poterewicz G, Chung JK, Plitzko JM, Groves JT, Jacobs-Wagner C, Engel BD, Holt LJ. mTORC1 Controls Phase
1587 Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. *Cell*. 2018 Jul; 174(2):338–
1588 349.e20.
- 1589 **Dennis PP**, Ehrenberg M, Bremer H. Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.
1590 *Microbiology and Molecular Biology Reviews*. 2004 Dec; 68(4):639–668.
- 1591 **Dill KA**, Ghosh K, Schmit JD. Physical Limits of Cells and Proteomes. *Proceedings of the National Academy of
1592 Sciences*. 2011 Nov; 108(44):17876–17882. doi: [10.1073/pnas.1114477108](https://doi.org/10.1073/pnas.1114477108).
- 1593 **Escalante A**, Salinas Cervantes A, Gosset G, Bolívar F. Current Knowledge of the *Escherichia coli* Phosphoenolpyru-
1594 vate–Carbohydrate Phosphotransferase System: Peculiarities of Regulation and Impact on Growth and Prod-
1595 uct Formation. *Applied Microbiology and Biotechnology*. 2012 Jun; 94(6):1483–1494. doi: [10.1007/s00253-012-4101-5](https://doi.org/10.1007/s00253-012-4101-5).
- 1597 **Feist AM**, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BØ. A
1598 Genome-Scale Metabolic Reconstruction for *Escherichia coli* K-12 MG1655 That Accounts for 1260 ORFs and
1599 Thermodynamic Information. *Molecular Systems Biology*. 2007 Jan; 3(1):121. doi: [10.1038/msb4100155](https://doi.org/10.1038/msb4100155).
- 1600 **Fernández-Coll L**, Maciąg-Dorszynska M, Tailor K, Vadia S, Levin PA, Szalewska-Palasz A, Cashel M, Dunn GM.
1601 The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of
1602 Growth Rates. *mBio*. 2020 Apr; 11(2):45.
- 1603 **Fijalkowska IJ**, Schaaper RM, Jonczyk P. DNA Replication Fidelity in *Escherichia coli*: A Multi-DNA Polymerase
1604 Affair. *FEMS Microbiology Reviews*. 2012 Nov; 36(6):1105–1121. doi: [10.1111/j.1574-6976.2012.00338.x](https://doi.org/10.1111/j.1574-6976.2012.00338.x).
- 1605 **Finkelstein IJ**, Greene EC. Molecular Traffic Jams on DNA. *Annual Review of Biophysics*. 2013 May; 42(1):241–263.
- 1606 **Gallagher LA**, Bailey J, Manoil C. Ranking Essential Bacterial Processes by Speed of Mutant Death. *Proceedings
1607 of the National Academy of Sciences*. 2020 Jul; doi: [10.1073/pnas.2001507117](https://doi.org/10.1073/pnas.2001507117).
- 1608 **Gama-Castro S**, Salgado H, Santos-Zavaleta A, Ledezma-Tejeida D, Muñiz-Rascado L, García-Sotelo JS, Alquicira-
1609 Hernández K, Martínez-Flores I, Pannier L, Castro-Mondragón JA, Medina-Rivera A, Solano-Lira H, Bonavides-
1610 Martínez C, Pérez-Rueda E, Alquicira-Hernández S, Porrón-Sotelo L, López-Fuentes A, Hernández-Koutoucheva
1611 A, Moral-Chávez VD, Rinaldi F, et al. RegulonDB Version 9.0: High-Level Integration of Gene Regulation,
1612 Coexpression, Motif Clustering and Beyond. *Nucleic Acids Research*. 2016 Jan; 44(D1):D133–D143. doi:
1613 [10.1093/nar/gkv1156](https://doi.org/10.1093/nar/gkv1156).
- 1614 **Ge J**, Yu G, Ator MA, Stubbe J. Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class I Ribonucleotide
1615 Reductase. *Biochemistry*. 2003 Sep; 42(34):10071–10083. doi: [10.1021/bi034374r](https://doi.org/10.1021/bi034374r).
- 1616 **Godin M**, Delgado FF, Son S, Grover WH, Bryan AK, Tzur A, Jorgensen P, Payer K, Grossman AD, Kirschner
1617 MW, Manalis SR. Using buoyant mass to measure the growth of single cells. *Nature Methods*. 2010 Apr;
1618 7(5):387–390.
- 1619 **Goldman SR**, Nair NU, Wells CD, Nickels BE, Hochschild A. The Primary σ Factor in *Escherichia coli* Can Access the
1620 Transcription Elongation Complex from Solution *In Vivo*. *eLife*. 2015 Sep; 4:e10514. doi: [10.7554/eLife.10514](https://doi.org/10.7554/eLife.10514).
- 1621 **Guo Y**, Li D, Zhang S, Yang Y, Liu JJ, Wang X, Liu C, Milkie DE, Moore RP, Tulu US, Kiehart DP, Hu J, Lippincott-
1622 Schwartz J, Betzig E, Li D. Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale
1623 Resolution on Millisecond Timescales. *Cell*. 2018 Nov; 175(5):1430–1442.e17.
- 1624 **Harris LK**, Theriot JA. Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in
1625 microbiology*. 2018 Oct; 26(10):815–832.
- 1626 **Harris RM**, Webb DC, Howitt SM, Cox GB. Characterization of PitA and PitB from *Escherichia coli*. *Journal of
1627 Bacteriology*. 2001 Sep; 183(17):5008–5014. doi: [10.1128/JB.183.17.5008-5014.2001](https://doi.org/10.1128/JB.183.17.5008-5014.2001).

- 1628 **Hauryliuk V**, Atkinson GC, Murakami KS, Tenson T, Gerdes K. Recent functional insights into the role of (p)ppGpp
1629 in bacterial physiology. *Nature Reviews Microbiology*. 2015 Apr; 13(5):298–309.
- 1630 **van Heeswijk WC**, Westerhoff HV, Boogerd FC. Nitrogen Assimilation in *Escherichia coli*: Putting Molecular
1631 Data into a Systems Perspective. *Microbiology and Molecular Biology Reviews*. 2013 Dec; 77(4):628–695. doi:
1632 [10.1128/MMBR.00025-13](https://doi.org/10.1128/MMBR.00025-13).
- 1633 **Heldal M**, Norland S, Tumyr O. X-Ray Microanalytic Method for Measurement of Dry Matter and Elemental
1634 Content of Individual Bacteria. *Applied and Environmental Microbiology*. 1985 Nov; 50(5):1251–1257.
- 1635 **Helmstetter CE**, Cooper S. DNA synthesis during the division cycle of rapidly growing *Escherichia coli* Br. *Journal*
1636 of *Molecular Biology*. 1968 Feb; 31(3):507–518.
- 1637 **Henkel SG**, Beek AT, Steinsiek S, Stagge S, Bettenbrock K, de Mattos MJT, Sauter T, Sawodny O, Ederer M. Basic
1638 Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying Oxygen Conditions. *PLoS ONE*.
1639 2014 Sep; 9(9):e107640. doi: [10.1371/journal.pone.0107640](https://doi.org/10.1371/journal.pone.0107640).
- 1640 **Hui S**, Silverman JM, Chen SS, Erickson DW, Basan M, Wang J, Hwa T, Williamson JR. Quantitative proteomic
1641 analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*. 2015
1642 Feb; 11(2):e784–e784.
- 1643 **Inglede w WJ**, Poole RK. The Respiratory Chains of *Escherichia coli*. *Microbiological Reviews*. 1984; 48(3):222–271.
1644 doi: [10.1128/MMBR.48.3.222-271.1984](https://doi.org/10.1128/MMBR.48.3.222-271.1984).
- 1645 **Ireland WT**, Beeler SM, Flores-Bautista E, Belliveau NM, Sweredoski MJ, Moradian A, Kinney JB, Phillips R.
1646 Deciphering the Regulatory Genome of *Escherichia coli*, One Hundred Promoters at a Time. *bioRxiv*. 2020 Jan;
1647 doi: [10.1101/2020.01.18.910323](https://doi.org/10.1101/2020.01.18.910323).
- 1648 **Jacob F**, Monod J. Genetic Regulatory Mechanisms in the Synthesis of Proteins. *Journal of Molecular Biology*.
1649 1961 Jun; 3(3):318–356. doi: [10.1016/S0022-2836\(61\)80072-7](https://doi.org/10.1016/S0022-2836(61)80072-7).
- 1650 **Jensen RB**, Wang SC, Shapiro L. A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*.
1651 2001 Sep; 20(17):4952–4963.
- 1652 **Jun S**, Si F, Pugatch R, Scott M. Fundamental Principles in Bacterial Physiology - History, Recent Progress, and
1653 the Future with Focus on Cell Size Control: A Review. *Reports on Progress in Physics*. 2018 May; 81(5):056601.
1654 doi: [10.1088/1361-6633/aaa628](https://doi.org/10.1088/1361-6633/aaa628).
- 1655 **Kapanidis AN**, Margeat E, Laurence TA, Doose S, Ho SO, Mukhopadhyay J, Kortkhonja E, Mekler V, Ebright RH,
1656 Weiss S. Retention of Transcription Initiation Factor Σ70 in Transcription Elongation: Single-Molecule Analysis.
1657 *Molecular Cell*. 2005 Nov; 20(3):347–356. doi: [10.1016/j.molcel.2005.10.012](https://doi.org/10.1016/j.molcel.2005.10.012).
- 1658 **Karp PD**, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, Keseler IM, Krummenacker M, Midford PE,
1659 Ong Q, Ong WK, Paley SM, Subhraveti P. The BioCyc collection of microbial genomes and metabolic pathways.
1660 *Briefings in Bioinformatics*. 2019 Aug; 20(4):1085–1093.
- 1661 **Karr JR**, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW. A
1662 Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*. 2012 Jul; 150(2):389–401. doi:
1663 [10.1016/j.cell.2012.05.044](https://doi.org/10.1016/j.cell.2012.05.044).
- 1664 **Keseler IM**, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, Fulcher C, Gama-Castro
1665 S, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti
1666 P, Velázquez-Ramírez DA, Weaver D, Collado-Vides J, Paulsen I, et al. The EcoCyc database: reflecting new
1667 knowledge about *Escherichia coli*K-12. *Nucleic Acids Research*. 2017 Jan; 45(D1):D543–D550.
- 1668 **Khademi S**, O'Connell J, Remis J, Robles-Colmenares Y, Miercke LJW, Stroud RM. Mechanism of Ammonia
1669 Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*. 2004 Sep; 305(5690):1587–1594. doi:
1670 [10.1126/science.1101952](https://doi.org/10.1126/science.1101952).
- 1671 **Khademian M**, Imlay JA. *Escherichia coli* Cytochrome c Peroxidase Is a Respiratory Oxidase That Enables the
1672 Use of Hydrogen Peroxide as a Terminal Electron Acceptor. *Proceedings of the National Academy of Sciences*.
1673 2017 Aug; 114(33):E6922–E6931. doi: [10.1073/pnas.1701587114](https://doi.org/10.1073/pnas.1701587114).
- 1674 **Klumpp S**, Hwa T. Bacterial Growth: Global Effects on Gene Expression, Growth Feedback and Proteome
1675 Partition. *Current Opinion in Biotechnology*. 2014 Aug; 28:96–102. doi: [10.1016/j.copbio.2014.01.001](https://doi.org/10.1016/j.copbio.2014.01.001).

- 1676 Klumpp S, Zhang Z, Hwa T. Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*. 2009
1677 Dec; 139(7):1366–1375. doi: [10.1016/j.cell.2009.12.001](https://doi.org/10.1016/j.cell.2009.12.001).
- 1678 Kostinski S, Reuveni S. Ribosome Composition Maximizes Cellular Growth Rates in *E. Coli*. *Physical Review Letters*. 2020 Jul; 125(2):028103. doi: [10.1103/PhysRevLett.125.028103](https://doi.org/10.1103/PhysRevLett.125.028103).
- 1680 Kraemer JA, Sanderlin AG, Laub MT. The Stringent Response Inhibits DNA Replication Initiation in *E. coli* by
1681 Modulating Supercoiling of oriC. *mBio*. 2019 Aug; 10(4):822.
- 1682 Lascu I, Gonin P. The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and*
1683 *Biomembranes*. 2000 Jun; 32(3):237–246. doi: [10.1023/A:1005532912212](https://doi.org/10.1023/A:1005532912212).
- 1684 Laxhuber KS, Morrison MJ, Chure G, Belliveau NM, Strandkvist C, Naughton KL, Phillips R. Theoretical investiga-
1685 tion of a genetic switch for metabolic adaptation. *PLOS ONE*. 2020 May; 15(5):e0226453.
- 1686 Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: visualization of intersecting sets. *IEEE Transactions*
1687 *on Visualization and Computer Graphics*. 2014; 20(12):1983–1992.
- 1688 Li GW, Burkhardt D, Gross C, Weissman JS. Quantifying Absolute Protein Synthesis Rates Reveals Principles
1689 Underlying Allocation of Cellular Resources. *Cell*. 2014 Apr; 157(3):624–635. doi: [10.1016/j.cell.2014.02.033](https://doi.org/10.1016/j.cell.2014.02.033).
- 1690 Liebermeister W, Noor E, Flamholz A, Davidi D, Bernhardt J, Milo R. Visual account of protein investment in
1691 cellular functions. *Proceedings of the National Academy of Sciences*. 2014 Jun; 111(23):8488–8493.
- 1692 Liu M, Durfee T, Cabrera JE, Zhao K, Jin DJ, Blattner FR. Global Transcriptional Programs Reveal a Carbon Source
1693 Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*. 2005 Apr; 280(16):15921–15927. doi:
1694 [10.1074/jbc.M41405200](https://doi.org/10.1074/jbc.M41405200).
- 1695 Lu D, Grayson P, Schulten K. Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol
1696 Facilitator GlpF. *Biophysical Journal*. 2003 Nov; 85(5):2977–2987. doi: [10.1016/S0006-3495\(03\)74718-3](https://doi.org/10.1016/S0006-3495(03)74718-3).
- 1697 Lynch M, Marinov GK. The Bioenergetic Costs of a Gene. *Proceedings of the National Academy of Sciences*.
1698 2015 Dec; 112(51):15690–15695. doi: [10.1073/pnas.1514974112](https://doi.org/10.1073/pnas.1514974112).
- 1699 Maaløe O. Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On. Goldberger
1700 R, editor, *Gene Expression*, Springer; 1979.
- 1701 Metzl-Raz E, Kafri M, Yaakov G, Soifer I, Gurvich Y, Barkai N. Principles of cellular resource allocation revealed
1702 by condition-dependent proteome profiling. *eLife*. 2017; 6:e03528.
- 1703 Mikucki JA, Pearson A, Johnston DT, Turchyn AV, Farquhar J, Schrag DP, Anbar AD, Priscu JC, Lee PA. A Contem-
1704 porary Microbially Maintained Subglacial Ferrous "Ocean". *Science*. 2009 Apr; 324(5925):397–400.
- 1705 Milo R, Jorgensen P, Moran U, Weber G, Springer M. BioNumbers—the Database of Key Numbers in Molecular
1706 and Cell Biology. *Nucleic Acids Research*. 2010 Jan; 38(suppl_1):D750–D753. doi: [10.1093/nar/gkp889](https://doi.org/10.1093/nar/gkp889).
- 1707 Monod J. The Phenomenon of Enzymatic Adaptation and Its Bearings on Problems of Genetics and Cellular
1708 Differentiation. *Growth Symposium*. 1947; 9:223–289.
- 1709 Monod J. The Growth of Bacterial Cultures. *Annual Review of Microbiology*. 1949 Oct; 3(1):371–394.
- 1710 Mooney RA, Darst SA, Landick R. Sigma and RNA Polymerase: An On-Again, Off-Again Relationship? *Molecular*
1711 *Cell*. 2005 Nov; 20(3):335–345. doi: [10.1016/j.molcel.2005.10.015](https://doi.org/10.1016/j.molcel.2005.10.015).
- 1712 Mooney RA, Landick R. Tethering Σ 70 to RNA Polymerase Reveals High in Vivo Activity of σ Factors and Σ 70-
1713 Dependent Pausing at Promoter-Distal Locations. *Genes & Development*. 2003 Nov; 17(22):2839–2851. doi:
1714 [10.1101/gad.1142203](https://doi.org/10.1101/gad.1142203).
- 1715 Morgenstein RM, Bratton BP, Nguyen JP, Ouzounov N, Shaevitz JW, Gitai Z. RodZ Links MreB to Cell Wall
1716 Synthesis to Mediate MreB Rotation and Robust Morphogenesis. *Proceedings of the National Academy of*
1717 *Sciences*. 2015 Oct; 112(40):12510–12515. doi: [10.1073/pnas.1509610112](https://doi.org/10.1073/pnas.1509610112).
- 1718 Neidhardt FC, Ingraham JL, Schaechter M. *Physiology of the Bacterial Cell - A Molecular Approach*, vol. 1.
1719 Elsevier; 1991.
- 1720 Ojkic N, Serbanescu D, Banerjee S. Surface-to-volume scaling and aspect ratio preservation in rod-shaped
1721 bacteria. *eLife*. 2019 Aug; 8:642.

- 1722 **Patrick M**, Dennis PP, Ehrenberg M, Bremer H. Free RNA Polymerase in *Escherichia coli*. *Biochimie*. 2015 Dec; 1723 119:80–91. doi: [10.1016/j.biochi.2015.10.015](https://doi.org/10.1016/j.biochi.2015.10.015).
- 1724 **Peebo K**, Valgepea K, Maser A, Nahku R, Adamberg K, Vilu R. Proteome Reallocation in *Escherichia coli* with 1725 Increasing Specific Growth Rate. *Molecular BioSystems*. 2015; 11(4):1184–1193. doi: [10.1039/C4MB00721B](https://doi.org/10.1039/C4MB00721B).
- 1726 **Perdue SA**, Roberts JW. σ^{70} -Dependent Transcription Pausing in *Escherichia coli*. *Journal of Molecular Biology*. 1727 2011 Oct; 412(5):782–792. doi: [10.1016/j.jmb.2011.02.011](https://doi.org/10.1016/j.jmb.2011.02.011).
- 1728 **Phillips R**. Membranes by the Numbers. In: *Physics of Biological Membranes* Cham: Springer, Cham; 2018.p. 1729 73–105.
- 1730 **Ramos S**, Kaback HR. The Relation between the Electrochemical Proton Gradient and Active Transport in 1731 *Escherichia coli* Membrane Vesicles. *Biochemistry*. 1977 Mar; 16(5):854–859. doi: [10.1021/bi00624a007](https://doi.org/10.1021/bi00624a007).
- 1732 **Ranganathan S**, Tee TW, Chowdhury A, Zomorodi AR, Yoon JM, Fu Y, Shanks JV, Maranas CD. An Integrated 1733 Computational and Experimental Study for Overproducing Fatty Acids in *Escherichia Coli*. *Metabolic Engineering*. 1734 2012 Nov; 14(6):687–704. doi: [10.1016/j.ymben.2012.08.008](https://doi.org/10.1016/j.ymben.2012.08.008).
- 1735 **Rogers H**, Perkins H, Ward J. *Microbial Cell Walls and Membranes*. London: Chapman and Hall; 1980.
- 1736 **Roller BRK**, Stoddard SF, Schmidt TM. Exploiting rRNA operon copy number to investigate bacterial reproductive 1737 strategies. *Nature microbiology*. 2016 Sep; 1(11):1–7.
- 1738 **Rosenberg H**, Gerdes RG, Chegwidden K. Two Systems for the Uptake of Phosphate in *Escherichia coli*. *Journal 1739 of Bacteriology*. 1977 Aug; 131(2):505–511.
- 1740 **Rudd SG**, Valerie NCK, Helleday T. Pathways Controlling dNTP Pools to Maintain Genome Stability. *DNA Repair*. 1741 2016 Aug; 44:193–204. doi: [10.1016/j.dnarep.2016.05.032](https://doi.org/10.1016/j.dnarep.2016.05.032).
- 1742 **Ruppe A**, Fox JM. Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*. 2018 Dec; 1743 8(12):11722–11734. doi: [10.1021/acscatal.8b03171](https://doi.org/10.1021/acscatal.8b03171).
- 1744 **Sánchez-Romero MA**, Molina F, Jiménez-Sánchez A. Organization of Ribonucleoside Diphosphate Reductase 1745 during Multifork Chromosome Replication in *Escherichia coli*. *Microbiology*. 2011 Aug; 157(8):2220–2225. doi: 1746 [10.1099/mic.0.049478-0](https://doi.org/10.1099/mic.0.049478-0).
- 1747 **Schaechter M**, Maaløe O, Kjeldgaard NO. Dependency on Medium and Temperature of Cell Size and Chemical 1748 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*. 1958 Dec; 19(3):592–606.
- 1749 **Schmidt A**, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heine- 1750 mann M. The Quantitative and Condition-Dependent *Escherichia coli* Proteome. *Nature Biotechnology*. 2016 1751 Jan; 34(1):104–110. doi: [10.1038/nbt.3418](https://doi.org/10.1038/nbt.3418).
- 1752 **Scott M**, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. Interdependence of cell growth and gene expression: 1753 origins and consequences. *Science*. 2010 Nov; 330(6007):1099–1102.
- 1754 **Sekowska A**, Kung HF, Danchin A. Sulfur Metabolism in *Escherichia coli* and Related Bacteria: Facts and Fiction. 1755 *Journal of Molecular Microbiology and Biotechnology*. 2000; 2(2):34.
- 1756 **Shi H**, Bratton BP, Gitai Z, Huang KC. How to Build a Bacterial Cell: MreB as the Foreman of *E. Coli* Construction. 1757 *Cell*. 2018 Mar; 172(6):1294–1305. doi: [10.1016/j.cell.2018.02.050](https://doi.org/10.1016/j.cell.2018.02.050).
- 1758 **Si F**, Le Treut G, Sauls JT, Vadia S, Levin PA, Jun S. Mechanistic Origin of Cell-Size Control and Homeostasis in 1759 Bacteria. *Current Biology*. 2019 Jun; 29(11):1760–1770.e7. doi: [10.1016/j.cub.2019.04.062](https://doi.org/10.1016/j.cub.2019.04.062).
- 1760 **Si F**, Li D, Cox SE, Sauls JT, Azizi O, Sou C, Schwartz AB, Erickstad MJ, Jun Y, Li X, Jun S. Invariance of Initiation Mass 1761 and Predictability of Cell Size in *Escherichia coli*. *Current Biology*. 2017 May; 27(9):1278–1287.
- 1762 **Sirk A**, Zatyka M, Sadowy E, Hulanicka D. Sulfate and Thiosulfate Transport in *Escherichia coli* K-12: Evidence 1763 for a Functional Overlapping of Sulfate- and Thiosulfate-Binding Proteins. *Journal of Bacteriology*. 1995 Jul; 1764 177(14):4134–4136. doi: [10.1128/jb.177.14.4134-4136.1995](https://doi.org/10.1128/jb.177.14.4134-4136.1995).
- 1765 **Sohlenkamp C**, Geiger O. Bacterial Membrane Lipids: Diversity in Structures and Pathways. *FEMS Microbiology 1766 Reviews*. 2016 Jan; 40(1):133–159. doi: [10.1093/femsre/fuv008](https://doi.org/10.1093/femsre/fuv008).

- 1767 **Soler-Bistué A**, Aguilar-Pierlé S, Garcia-Garcera M, Val ME, Sismeiro O, Varet H, Sieira R, Krin E, Skovgaard O,
1768 Comerci DJ, Eduardo P C Rocha, Mazel D. Macromolecular crowding links ribosomal protein gene dosage to
1769 growth rate in *Vibrio cholerae*. *BMC Biology*. 2020 Dec; 18(1):1–18.
- 1770 **Soufi B**, Krug K, Harst A, Macek B. Characterization of the *E. coli* proteome and its modifications during growth
1771 and ethanol stress. *Frontiers in Microbiology*. 2015 Feb; 6:198.
- 1772 **Stasi R**, Neves HI, Spira B. Phosphate Uptake by the Phosphonate Transport System PhnCDE. *BMC Microbiology*.
1773 2019 Apr; 19. doi: 10.1186/s12866-019-1445-3.
- 1774 **Stevenson BS**, Schmidt TM. Life History Implications of rRNA Gene Copy Number in *Escherichia coli*. *Applied
1775 and Environmental Microbiology*. 2004 Nov; 70(11):6670–6677. doi: 10.1128/AEM.70.11.6670-6677.2004.
- 1776 **Stouthamer AH**. A Theoretical Study on the Amount of ATP Required for Synthesis of Microbial Cell Material.
1777 *Antonie van Leeuwenhoek*. 1973 Dec; 39(1):545–565. doi: 10.1007/BF02578899.
- 1778 **Stouthamer AH**, Bettenhausen CW. A continuous culture study of an ATPase-negative mutant of *Escherichia
1779 coli*. *Archives of Microbiology*. 1977 Jun; 113(3):185–189.
- 1780 **Svenningsen SL**, Kongstad M, Stenum TS, Muñoz-Gómez AJ, Sørensen MA. Transfer RNA Is Highly Unstable
1781 during Early Amino Acid Starvation in *Escherichia coli*. *Nucleic Acids Research*. 2017 Jan; 45(2):793–804. doi:
1782 10.1093/nar/gkw1169.
- 1783 **Szenk M**, Dill KA, de Graff AMR. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1784 Membrane Real Estate Hypothesis. *Cell Systems*. 2017 Aug; 5(2):95–104. doi: 10.1016/j.cels.2017.06.005.
- 1785 **Taheri-Araghi S**, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, Vergassola M, Jun S. Cell-size control and
1786 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*. 2015 Feb; 25(3):385–391.
- 1787 **Taniguchi Y**, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS. Quantifying *E. coli* proteome and
1788 transcriptome with single-molecule sensitivity in single cells. *Science (New York, NY)*. 2010 Jul; 329(5991):533–
1789 538.
- 1790 **Tatusov RL**, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein
1791 functions and evolution. *Nucleic Acids Research*. 2000 Jan; 28(1):33–36.
- 1792 **Taymaz-Nikerel H**, Borujeni AE, Verheijen PJT, Heijnen JJ, van Gulik WM. Genome-Derived Minimal
1793 Metabolic Models for *Escherichia coli* MG1655 with Estimated in Vivo Respiratory ATP Stoichiometry.
1794 *Biotechnology and Bioengineering*. 2010; 107(2):369–381. doi: 10.1002/bit.22802, _eprint:
1795 https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802.
- 1796 **The Gene Ontology Consortium**. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids
1797 Research*. 2018 Nov; 47(D1):D330–D338.
- 1798 **Valgepea K**, Adamberg K, Seiman A, Vilu R. *Escherichia coli* Achieves Faster Growth by Increasing Catalytic and
1799 Translation Rates of Proteins. *Molecular BioSystems*. 2013; 9(9):2344. doi: 10.1039/c3mb70119k.
- 1800 **Virtanen P**, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser
1801 W, Bright J, van der Walt SJ, Brett M, Wilson J, Jarrod Millman K, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson
1802 E, Carey C, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*.
1803 2020; 17:261–272. doi: https://doi.org/10.1038/s41592-019-0686-2.
- 1804 **Volkmer B**, Heinemann M. Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate
1805 Data Conversion for Systems Biology Modeling. *PLOS ONE*. 2011 Jul; 6(7):e23126.
- 1806 **Vollmer W**, Blanot D, De Pedro MA. Peptidoglycan Structure and Architecture. *FEMS Microbiology Reviews*.
1807 2008 Mar; 32(2):149–167. doi: 10.1111/j.1574-6976.2007.00094.x.
- 1808 **Weber J**, Senior AE. ATP Synthesis Driven by Proton Transport in F1FO-ATP Synthase. *FEBS Letters*. 2003;
1809 545(1):61–70. doi: 10.1016/S0014-5793(03)00394-6.
- 1810 **Willsky GR**, Bennett RL, Malamy MH. Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
1811 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*. 1973 Feb; 113(2):529–
1812 539.
- 1813 **You C**, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, Wang YP, Lenz P, Yan D, Hwa T. Coordination of bacterial
1814 proteome with metabolism by cyclic AMP signalling. *Nature*. 2013 Aug; 500(7462):301–306.

- 1815 **Yu X**, Liu T, Zhu F, Khosla C. In Vitro Reconstitution and Steady-State Analysis of the Fatty Acid Synthase from
1816 Escherichia Coli. *Proceedings of the National Academy of Sciences*. 2011 Nov; 108(46):18643–18648. doi:
1817 [10.1073/pnas.1110852108](https://doi.org/10.1073/pnas.1110852108).
- 1818 **Zhang L**, Jiang W, Nan J, Almqvist J, Huang Y. The *Escherichia coli* CysZ Is a pH Dependent Sulfate Transporter That
1819 Can Be Inhibited by Sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2014 Jul; 1838(7):1809–1816.
1820 doi: [10.1016/j.bbamem.2014.03.003](https://doi.org/10.1016/j.bbamem.2014.03.003).
- 1821 **Zhang Z**, Aboulwafa M, Saier MH. Regulation of Crp Gene Expression by the Catabolite Repressor/Activator,
1822 Cra, in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*. 2014; 24(3):135–141. doi:
1823 [10.1159/000362722](https://doi.org/10.1159/000362722).
- 1824 **Zhu M**, Dai X. Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in
1825 *Escherichia coli*. *Nucleic Acids Research*. 2019 Mar; 47(9):4684–4693.