

Fundamental limits on the rate of bacterial cell division

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

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

¹⁴

¹⁵ **Abstract** Recent years have seen a deluge of experiments dissecting the relationship between bacterial
¹⁶ growth rate, cell size, and protein content, quantifying the abundance of proteins across growth conditions with
¹⁷ unprecedented resolution. However, we still lack a rigorous understanding of what sets the scale of these
¹⁸ quantities and when protein abundances should (or should not) depend on growth rate. Here, we seek to
¹⁹ quantitatively understand this relationship across a collection of *Escherichia coli* proteomic data sets covering \approx
²⁰ 4000 proteins and 31 growth conditions. We estimate the basic requirements for steady-state growth by
²¹ considering key processes in nutrient transport, energy generation, cell envelope biogenesis, and the central
²² dogma, from which ribosome biogenesis emerges as a primary determinant of growth rate. We conclude by
²³ exploring a model of ribosomal regulation as a function of the nutrient supply, revealing a mechanism that ties
²⁴ cell size and growth rate to ribosomal content.

²⁵

Introduction

²⁷ The observed range of bacterial growth rates is enormously diverse. In natural environments, some microbial
²⁸ organisms might double only once per year (*Mikucki et al., 2009*) while in comfortable laboratory conditions,
²⁹ growth can be rapid with several divisions per hour (*Schaechter et al., 1958*). This six order-of-magnitude differ-
³⁰ ence in time scales encompasses different microbial species and lifestyles, yet even for a single species such as
³¹ *Escherichia coli*, the growth rate can be modulated over a similar scale by tuning the type and amount of nutrients
³² in the growth medium. This remarkable flexibility in growth rate illustrates the intimate relationship between en-
³³ vironmental conditions and the rates at which cells convert nutrients into new cellular material – a relationship
³⁴ that has remained a major topic of inquiry in bacterial physiology for over a century (*Jun et al., 2018*).

³⁵ As was noted by Jacques Monod, “the study of the growth of bacterial cultures does not constitute a specialized
³⁶ subject or branch of research, it is the basic method of Microbiology.” Those words ring as true today as they did
³⁷ when they were written 70 years ago (*Monod, 1949*). Indeed, the study of bacterial growth has undergone a
³⁸ renaissance. Many of the key questions addressed by the pioneering efforts in the middle of the last century can
³⁹ be revisited by examining them through the lens of the increasingly refined molecular census that is available for
⁴⁰ bacteria such as the microbial workhorse *E. coli*. In this work, we explore an amalgamation of recent proteomic
⁴¹ data sets to explore fundamental limits of bacterial growth.

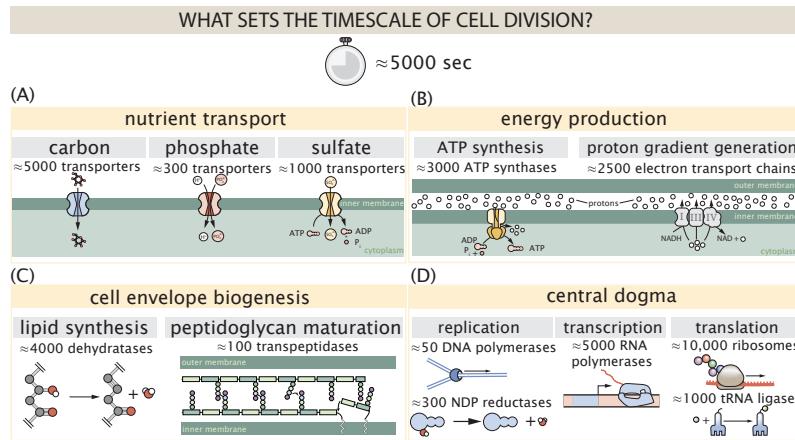


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP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

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

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

72

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

73

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

74

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

75

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

76

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

77

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

110 Uptake of Nutrients

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

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

129 Nitrogen Transport

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

142 Carbon Transport

143 We begin with the most abundant element in *E. coli* by mass, carbon. Using ≈ 0.3 pg as the typical *E. coli* dry
144 mass (BNID: 103904), we estimate that $\sim 10^{10}$ carbon atoms must be brought into the cell in order to double all
145 of the carbon-containing molecules (*Figure 2(A, top)*). Typical laboratory growth conditions provide carbon as a
146 single class of sugar such as glucose, galactose, or xylose to name a few. *E. coli* has evolved myriad mechanisms
147 by which these sugars can be transported across the cell membrane. One such mechanism of transport is via
148 the PTS system which is a highly modular system capable of transporting a diverse range of sugars (*Escalante*
149 *et al., 2012*). The glucose-specific component of this system transports ≈ 200 glucose molecules per second per
150 transporter (BNID: 114686). Making the assumption that this is a typical sugar transport rate, coupled with the
151 need to transport $\sim 10^{10}$ carbon atoms, we arrive at the conclusion that on the order of 1000 transporters must
152 be expressed in order to bring in enough carbon atoms to divide in 5000 s, diagrammed in the top panel of
153 *Figure 2(A)*. This estimate, along with the observed average number of the PTS system carbohydrate transporters
154 present in the proteomic data, is shown in *Figure 2(A)*. While we estimate 1500 transporters are needed with a
155 5000 s division time, we can abstract this calculation to consider any particular growth rate given knowledge of the
156 cell density and volume as a function of growth rate and direct the reader to the Appendix Extending Estimates
157 to a Continuum of Growth Rates for more information. As revealed in *Figure 2(A)*, experimental measurements
158 exceed the estimate by several fold, suggesting that transport of carbon into the cell is not rate limiting for cell

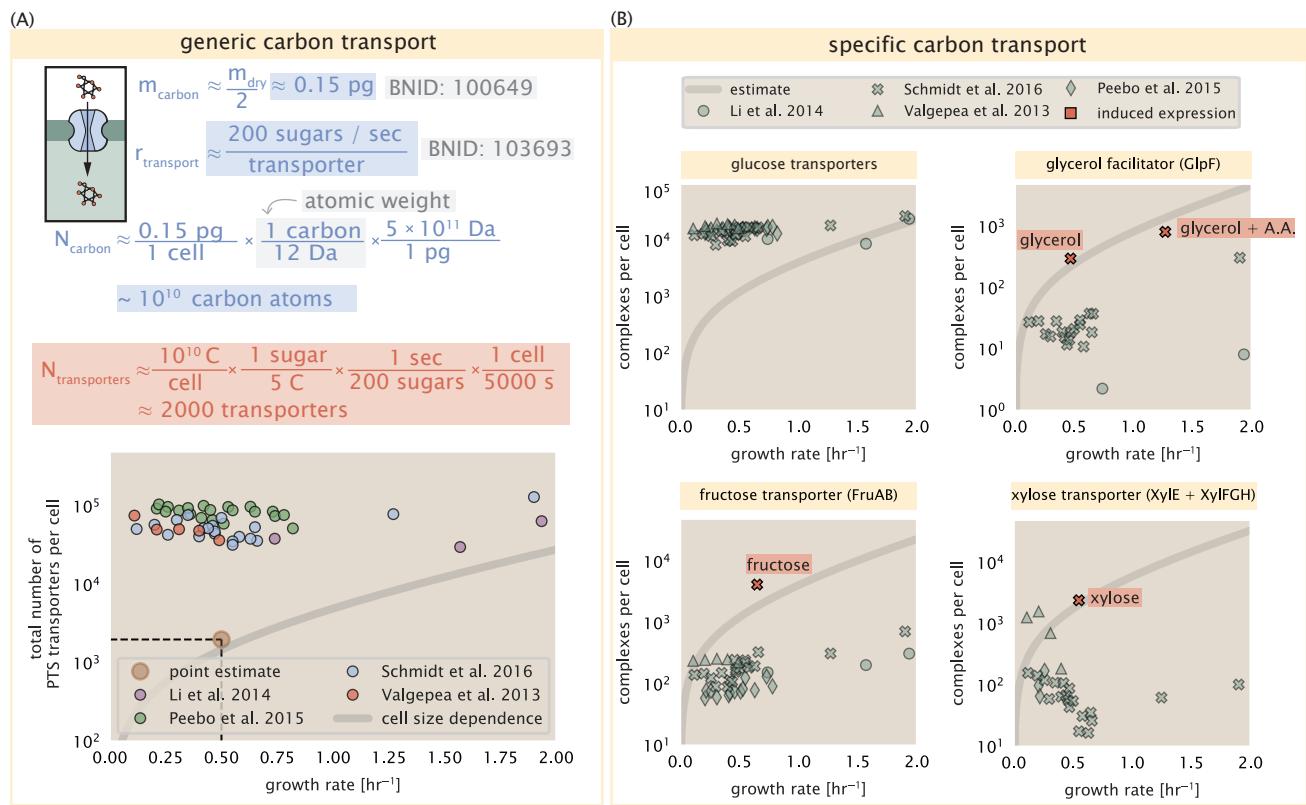


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules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

division. Abstracting this point estimate at 5000 s to a continuum of growth rates (grey line in *Figure 2(A)*) reveals an excess of transporters even at faster growth rates.

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

Many metabolic operons are regulated with dual-input logic gates that are only expressed when glucose concentrations are low (mediated by cyclic-AMP receptor protein CRP) and the concentration of other carbon sources are elevated (Gama-Castro et al., 2016; Zhang et al., 2014b). A famed example of such dual-input regulatory logic

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

187 Phosphorus and Sulfur Transport

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

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

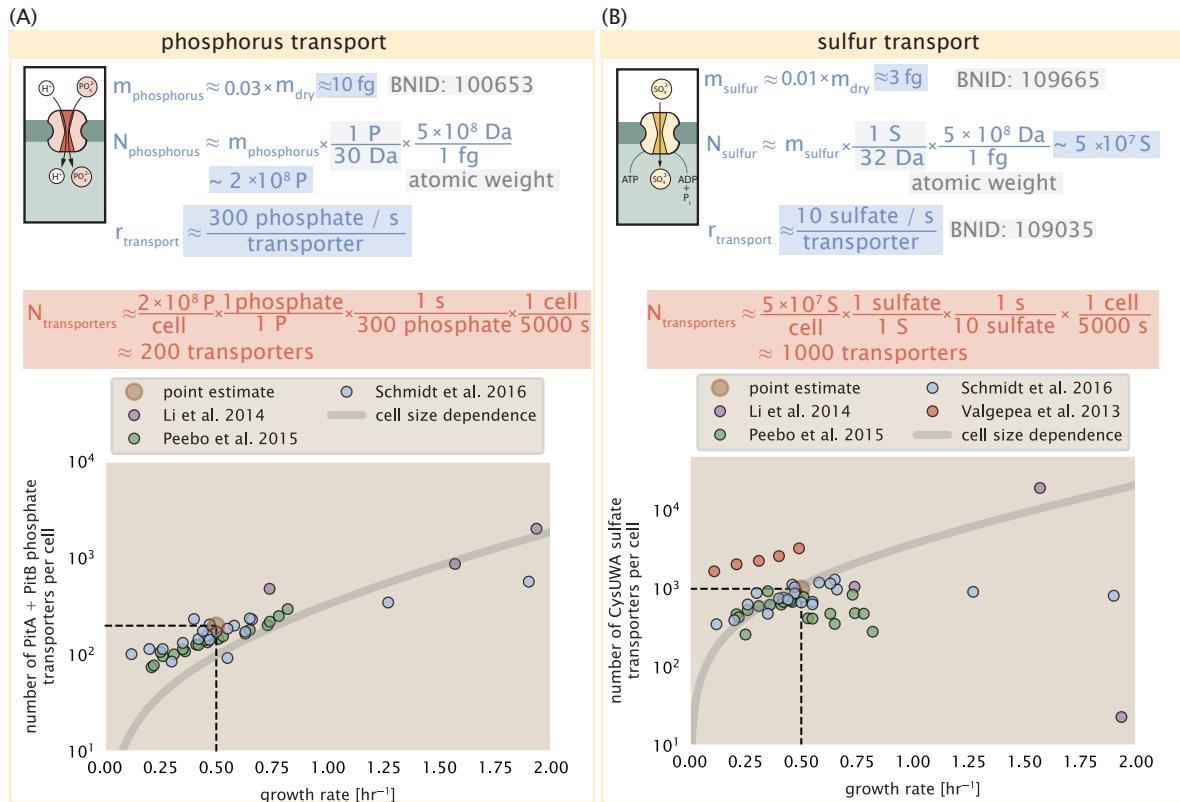


Figure 3. Estimates and measurements of phosphate and sulfate transport systems as a function of growth rate. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell.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

221 Limits on Transporter Expression

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

233 Energy Production

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

239 ATP Synthesis

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

252 To estimate the total number of ATP equivalents consumed during a cell cycle, we will make the approximation
253 that there are $\approx 3 \times 10^6$ proteins per cell with an average protein length of ≈ 300 peptide bonds (BNID: 115702;
254 108986; 104877). Taking these values together, we estimate that the typical *E. coli* cell consumes $\sim 5 \times 10^9$ ATP
255 per cell cycle on protein synthesis alone. Assuming that the ATP synthases are operating at their fastest possible
256 rate, ≈ 3000 ATP synthases are needed to keep up with the energy demands of the cell. This estimate and a
257 comparison with the data are shown in *Figure 4* (A). Despite our assumption of maximal ATP production rate per
258 synthase and approximation of all NTP consuming reactions being the same as ATP, we find that an estimate
259 of a few thousand complete synthases per cell to agree well with the experimental data. Much as we did for
260 the estimates of transporter copy number in the previous section, we can generalize this estimate to consider a
261 continuum of growth rates rather than a point estimate of 5000 s. Given knowledge of how the cell volume scales
262 with growth rate (*Si et al., 2017*), the density of the cytoplasm ($\rho \approx 1 \text{ pg} / \text{fL}$), and the empirical determination that
263 approximately half of the dry mass is protein, we can compute the energy demand as a function of growth rte,
264 indicated by the gray line in *Figure 4*.

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

270 Generating the Proton Electrochemical Gradient

271 In order to produce ATP, the F₁-F₀ ATP synthase itself must consume energy. Rather than burning through its own
272 product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the elec-
273 trochemical potential established across the inner membrane through cellular respiration. This electrochemical
274 gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains
275 as they reduce NADH. In *E. coli*, this potential difference is ≈ -200 mV (BNID: 102120). A simple estimate of the
276 inner membrane as a capacitor with a working voltage of -200 mV reveals that $\approx 2 \times 10^4$ protons must be present
277 in the intermembrane space.

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

284 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of intense biochemical
285 and biophysical study over the past half century (*Ingledeew and Poole, 1984; Khademian and Imlay, 2017; Cox et al.,*
286 *1970; Henkel et al., 2014*). A recent work (*Szenk et al., 2017*) examined the respiratory capacity of the *E. coli* electron
287 transport complexes using structural and biochemical data, revealing that each electron transport chain rapidly
288 pumps protons into the intermembrane space at a rate of ≈ 1500 protons per second (BNID: 114704; 114687).
289 Using our estimate of the number of ATP synthases required per cell [*Figure 4(A)*], coupled with these recent
290 measurements, we estimate that ≈ 1000 electron transport complexes would be necessary to facilitate the $\sim 5 \times 10^6$
291 protons per second diet of the cellular ATP synthases. This estimate (along with a generalization to the entire range
292 of observed growth rates) is in agreement with the number of complexes identified in the proteomic datasets (plot
293 in *Figure 4(B)*). This suggests that every ATP synthase must be accompanied by ≈ 1 functional electron transport
294 chain. Again, if this were a rate limiting process for bacterial growth, one must conclude that it is not possible
295 for the cell to simply increase the production of both the number of electron transport chain complexes as well
296 as ATP synthases. As both of these components only function bound to the inner membrane, we now turn our
297 attention towards the available space in the membrane as well as surface-area-to-volume constraints.

298 Energy Production in a Crowded Membrane.

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

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

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

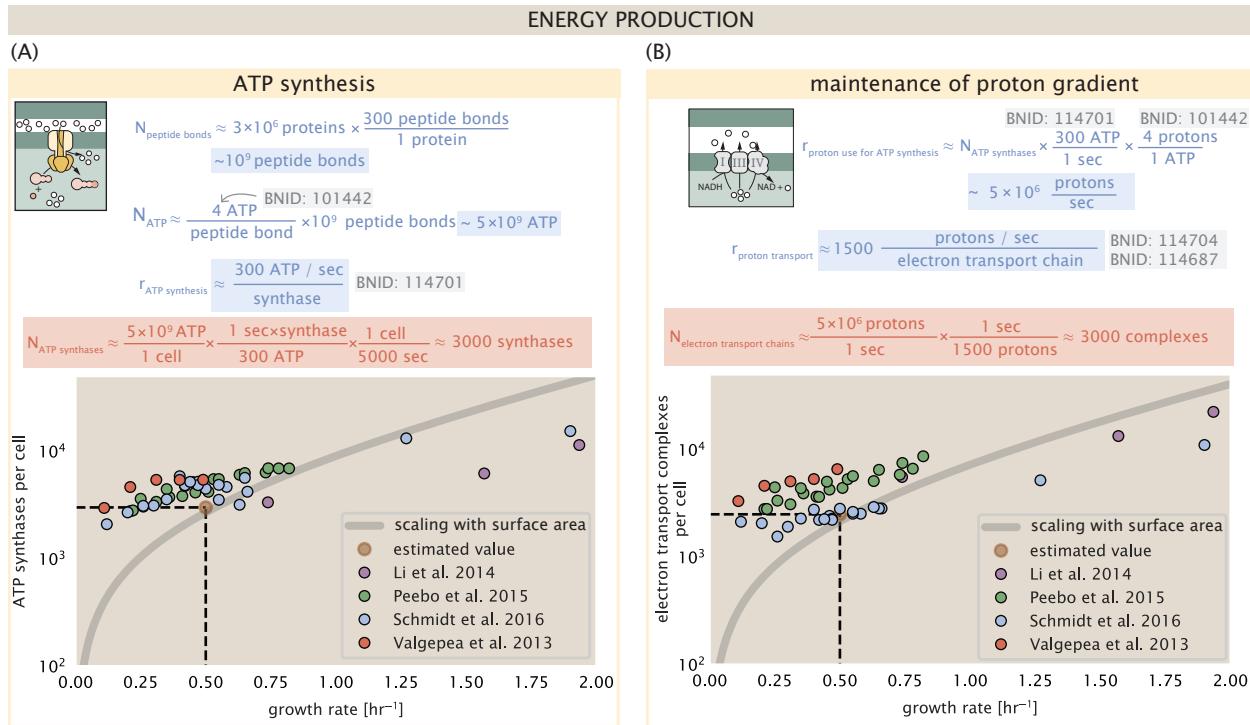


Figure 4. The abundance of F_1 - F_0 ATP synthases and electron transport chain complexes as a function of growth rate.

(A) Estimate of the number of F_1 - F_0 ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes. Points in plot correspond to the mean number of complete F_1 - F_0 ATP synthase complexes that can be formed given proteomic measurements and the subunit stoichiometry $[AtpE]_{10}[AtpF]_2[AtpB][AtpC][AtpH][AtpA]_3[AtpG][AtpD]_3$.

(B) Estimate of the number of electron transport chain complexes needed to maintain a membrane potential of -200 mV given estimate of number of F_1 - F_0 ATP synthases from (A). Points in plot correspond to the average number of complexes identified as being involved in aerobic respiration by the Gene Ontology identifier GO:0019646 that could be formed given proteomic observations. These complexes include cytochromes *bd1* ($[CydA][CydB][CydX][CydH]$), *bdII* ($[AppC][AppB]$), *bo3*, ($[CyoD][CyoA][CyoB][CyoC]$) and NADH:quinone oxioreducase I ($[NuoA][NuoH][NuoJ][NuoK][NuoL][NuoM][NuoN][NuoB][NuoC][NuoE][NuoF][NuoG][NuoI]$) and II ($[Ndh]$). Grey lines in both (A) and (B) correspond to the estimate procedure described, but applied to a continuum of growth rates. We direct the reader to the Supporting Information for a more thorough description of this approach.

area occupied due to the additional proteins of the electron transport chain. This suggests that, even under anaerobic growth, there will be sufficient membrane space for ATP production in general.

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

Synthesis of the Cell Envelope

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

Lipid Synthesis

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

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

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

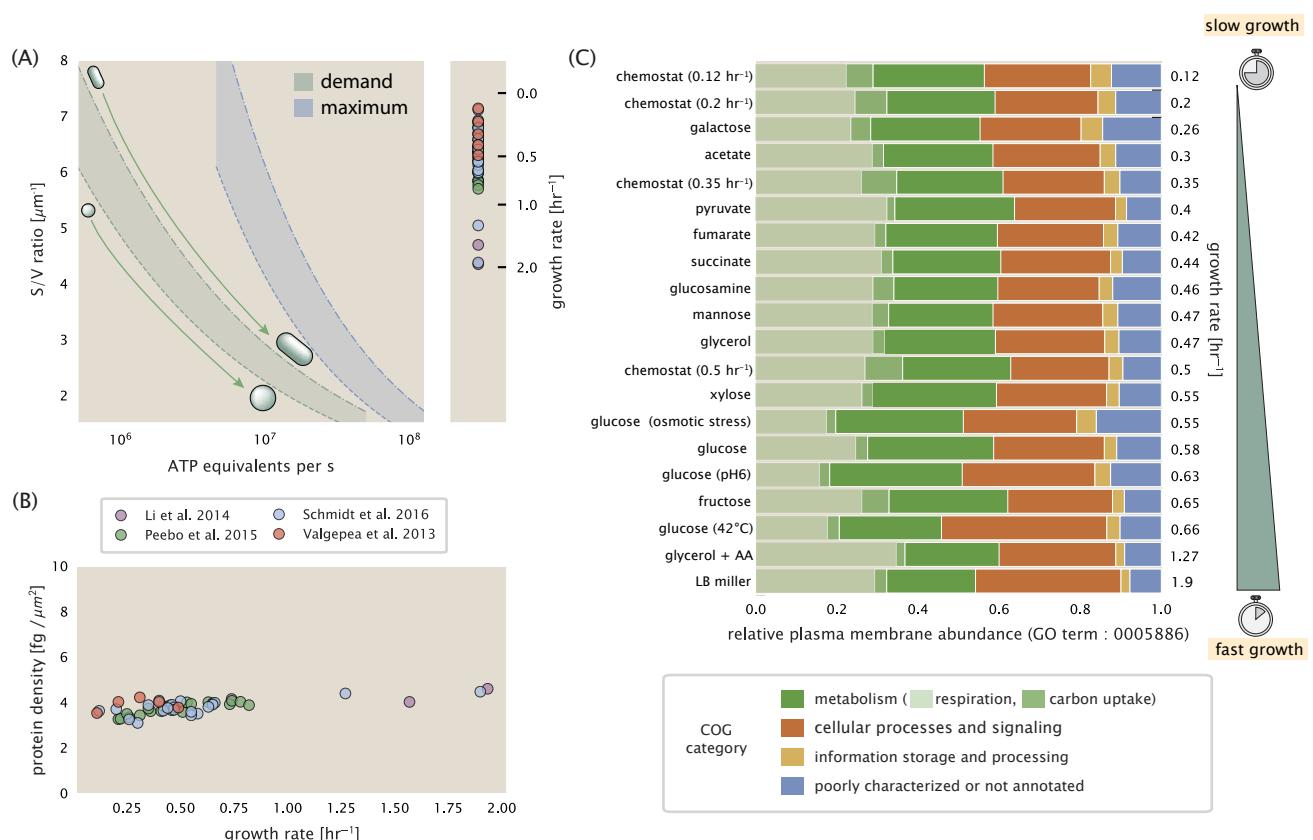


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (—·—) reflecting a rod-shaped bacterium like *E. coli* with a typical aspect ratio (length / width) of 4 (Shi et al., 2018). Approximately 50% of the bacterial inner membrane is assumed to be protein, with the remainder lipid. The right plot shows the measured growth rates, and their estimated S/V ratio using growth rate dependent size measurements from Si et al. (2017) (See Appendix Estimation of Cell Size and Surface Area on calculation). (B) Total protein mass per μm^2 calculated for proteins with inner membrane annotation (GO term: 0005886). (C) Relative protein abundances by mass based on COG annotation. Metabolic proteins are further separated into respiration (F_1 - F_0 ATP synthase, NADH dehydrogenase I, succinate:quinone oxidoreductase, cytochrome bo_3 ubiquinol oxidase, cytochrome bd-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)).

CELL ENVELOPE BIOSYNTHESIS

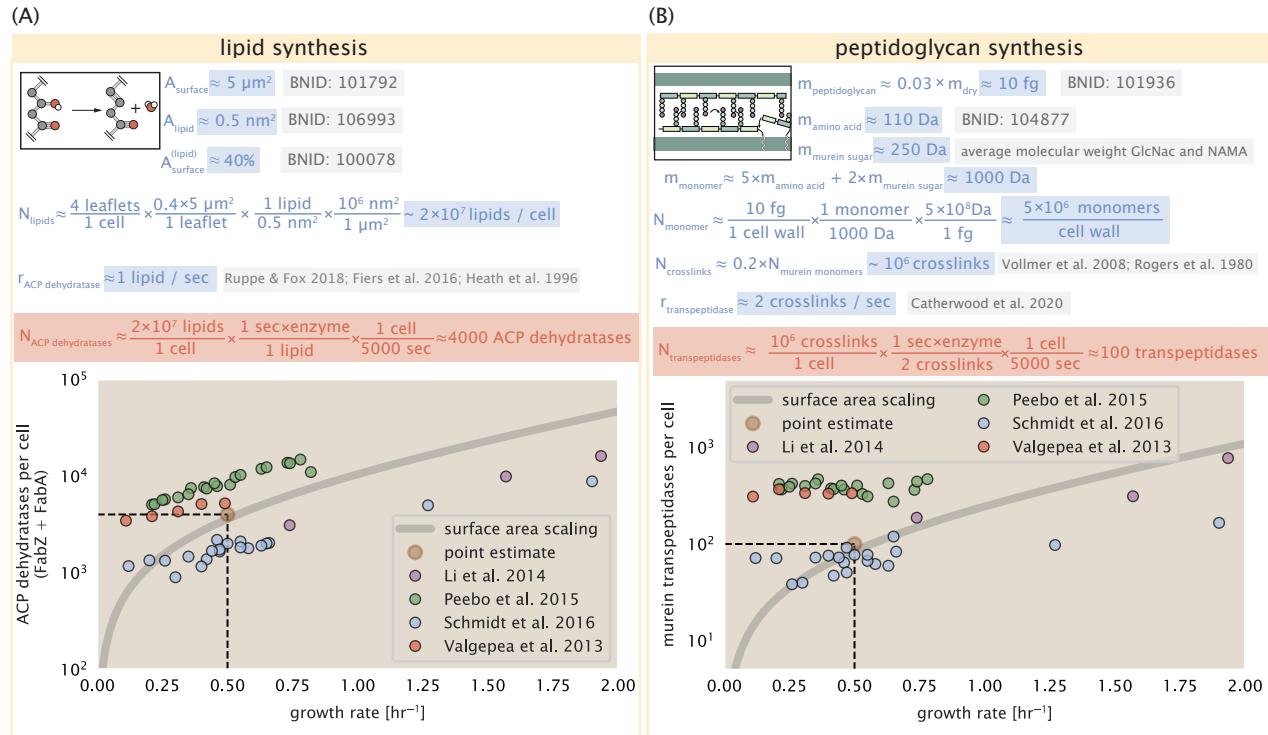


Figure 6. Estimation of the key components involved in cell envelope biosynthesis. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred from experimental measurements via a stochastic kinetic model described in [Ruppe and Fox \(2018\)](#). Bottom panel shows the experimentally observed complex copy numbers using the stoichiometries $[\text{FabZ}]_2$ 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.

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

Peptidoglycan Synthesis

While variation in cell size can vary substantially across growth conditions, bacterial cells demonstrate exquisite control over their cell shape. This is primarily due to the cell wall, a stiff meshwork of polymerized disaccharides interspersed with short peptide crosslinks termed the peptidoglycan. The cell wall is also a vital structural component that counteracts turgor pressure. In *E. coli*, this enormous peptidoglycan molecule is a few nanometers

384 thick and resides within the periplasmic space between the inner and outer membrane. The formation of the
385 peptidoglycan is an intricate process, involving the bacterial actin homolog MreB (*Shi et al., 2018*) along with a
386 variety of membrane-bound and periplasmic enzymes (*Morgenstein et al., 2015*). The coordinated action of these
387 components result in a highly-robust polymerized meshwork that maintains cell shape even in the face of large-
388 scale perturbations and can restore rod-shaped morphology even after digestion of the peptidoglycan (*Harris and*
389 *Theriot, 2018; Shi et al., 2018*).

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

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

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

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

423 Function of the Central Dogma

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

429 DNA

430 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal plasmids
431 up to ~ 100 kbp in length. We will focus our quantitative thinking solely on the chromosome of *E. coli* which harbors
432 ≈ 5000 genes and $\approx 5 \times 10^6$ base pairs. To successfully divide and produce viable progeny, this chromosome must

433 be faithfully replicated and segregated into each nascent cell. We again rely on the near century of literature
434 in molecular biology to provide some insight on the rates and mechanics of the replicative feat as well as the
435 production of the required starting materials, dNTPs.

436 dNTP synthesis

437 We begin our exploration DNA replication by examining the production of the deoxyribonucleotide triphosphates
438 (dNTPs). The four major dNTPs (dATP, dTTP, dCTP, and dGTP) are synthesized *de novo* in separate pathways, re-
439 quiring different building blocks. However, a critical step present in all dNTP synthesis pathways is the conversion
440 from ribonucleotide to deoxyribonucleotide via the removal of the 3' hydroxyl group of the ribose ring (Rudd *et al.*,
441 2016). This reaction is mediated by a class of enzymes termed ribonucleotide reductases, of which *E. coli* possesses
442 two aerobically active complexes (termed I and II) and a single anaerobically active enzyme. Due to their peculiar
443 formation of a radical intermediate, these enzymes have received much biochemical, kinetic, and structural char-
444 acterization. One such work (Ge *et al.*, 2003) performed a detailed *in vitro* measurement of the steady-state kinetic
445 rates of these complexes, revealing a turnover rate of ≈ 10 dNTP per second.

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

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

460 DNA Replication

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

475 It is well known that *E. coli* can parallelize its DNA replication such that multiple chromosomes are being repli-
476 cated at once, with as many as 10 - 12 replication forks at a given time (Bremer and Dennis, 2008; Si *et al.*, 2017).
477 Thus, even in rapidly growing cultures, we expect only a few polymerases (≈ 10) are needed to replicate the chro-
478 mosome per cell doubling. However, as shown in Figure 7(B), DNA polymerase III is nearly an order of
479 magnitude more abundant. This discrepancy can be understood by considering its binding constant to DNA. DNA
480 polymerase III is highly processive, facilitated by a strong affinity of the complex to the DNA. *In vitro* biochemical
481 characterization has quantified the K_D of DNA polymerase III holoenzyme to single-stranded and double-stranded

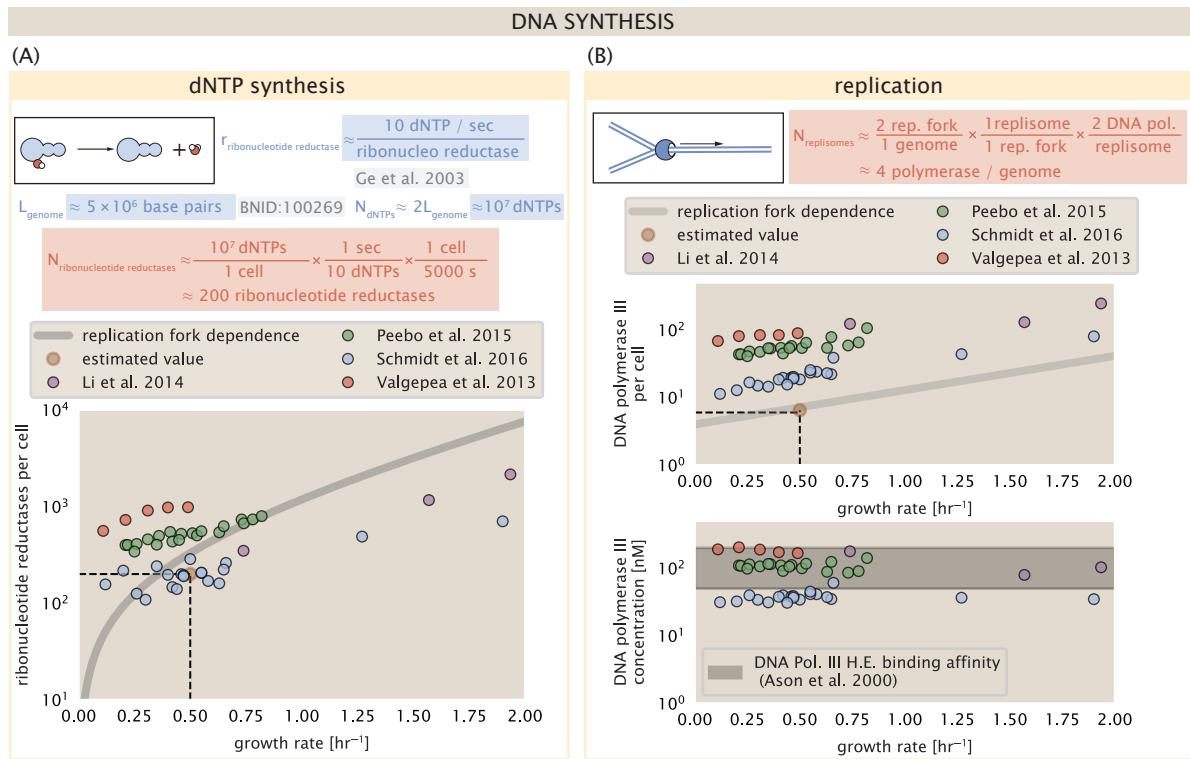


Figure 7. Complex abundance estimates for dNTP synthesis and DNA replication. (A) Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of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

482 DNA to be 50 and 200 nM, respectively (*Ason et al., 2000*). The bottom plot in *Figure 7(B)* shows that the concen-
483 tration of the DNA polymerase III across all data sets and growth conditions is within this range. Thus, while the
484 copy number of the DNA polymerase III is in excess of the strict number required to replicate the genome, its
485 copy number appears to vary such that its concentration is approximately equal to the dissociation constant to
486 the DNA. While the processes regulating the initiation of DNA replication are complex and involve more than just
487 the holoenzyme, these data indicate that the kinetics of replication rather than the explicit copy number of the
488 DNA polymerase III holoenzyme is the more relevant feature of DNA replication to consider. In light of this, the
489 data in *Figure 7(B)* suggests that for bacteria like *E. coli*, DNA replication does no that represent a rate-limiting step
490 in cell division. However, it is worth noting that for bacterium like *C. crescentus* whose chromosomal replication is
491 initiated only once per cell cycle (*Jensen et al., 2001*), the time to double their chromosome likely represents an
492 upper limit to their growth rate.

493 RNA Synthesis

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

502 rRNA

503 We begin with an estimation of the number of RNA polymerases needed to synthesize the rRNA that serve as
504 catalytic and structural elements of the ribosome. Each ribosome contains three rRNA molecules of lengths 120,
505 1542, and 2904 nucleotides (BNID: 108093), meaning each ribosome contains \approx 4500 nucleotides. As the *E. coli*
506 RNA polymerase transcribes DNA to RNA at a rate of \approx 40 nucleotides per second (BNID: 101904), it takes a single
507 RNA polymerase \approx 100 s to synthesize the RNA needed to form a single functional ribosome. Therefore, in a 5000
508 s division time, a single RNA polymerase transcribing rRNA at a time would result in only \approx 50 functional ribosomal
509 rRNA units – far below the observed number of \approx 10^4 ribosomes per cell.

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

518 *E. coli*, like many other bacterium, have evolved a clever mechanism to surpass this kinetic limit for the rate
519 of rRNA production. Rather than having only one copy of each rRNA gene, *E. coli* has seven copies of the operon
520 (BIND: 100352) four of which are localized directly adjacent to the origin of replication (*Birnbaum and Kaplan,*
521 *1971*). As fast growth also implies an increased gene dosage due to parallelized chromosomal replication, the
522 total number of rRNA genes can be on the order of \approx 10 – 70 copies at moderate to fast growth rates (*Stevenson*
523 *and Schmidt, 2004*). Given a 5000 second division time, we can make the lower-bound estimate that the typical
524 cell will have \approx 7 copies of the rRNA operon. Synthesizing one functional rRNA unit per second per rRNA operon,
525 a total of 5×10^4 rRNA units can be synthesized, comfortably above the observed number of ribosomes per cell.

526 How many RNA polymerases are then needed to constantly transcribe 7 copies of the rRNA genes? We ap-
527 proach this estimate by considering the maximum number of RNA polymerases tiled along the rRNA genes with
528 a loading rate of 1 per second and a transcription rate of 40 nucleotides per second. Considering that a RNA
529 polymerase has a physical footprint of approximately 40 nucleotides (BNID: 107873), we can expect \approx 1 RNA poly-
530 merase per 80 nucleotides. With a total length of \approx 4500 nucleotides per operon and 7 operons per cell, the

531 maximum number of RNA polymerases that can be transcribing rRNA at any given time is \approx 500. As we will see in
532 the coming sections, the synthesis of rRNA is the dominant requirement of the RNA polymerase pool.

533 mRNA

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

541 This estimation captures the *steady-state* mRNA copy number, meaning that at any given time, there will exist
542 approximately 3000 unique mRNA molecules. To determine the *total* number of mRNA that need to be synthesized
543 over the cell's lifetime, we must consider degradation of the mRNA. In most bacteria, mRNAs are rather unstable
544 with life times on the order of several minutes (BNID: 104324; 106253; 111927; 111998). For convenience, we
545 assume that the typical mRNA in our cell of interest has a typical lifetime of \approx 300 seconds. Using this value, we
546 can determine the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell.
547 While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional dynamics, we
548 state here that the total mRNA production rate must be on the order of \approx 15 mRNA per second. In *E. coli*, the
549 average protein is \approx 300 amino acids in length (BNID: 108986), meaning that the corresponding mRNA is \approx 900
550 nucleotides which we will further approximate as \approx 1000 nucleotides to account for the non-protein coding regions
551 on the 5' and 3' ends. This means that the cell must have enough RNA polymerase molecules about to sustain a
552 transcription rate of $\approx 1.5 \times 10^4$ nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA
553 at a clip of 40 nucleotides per second, we arrive at a comfortable estimate of \approx 250 RNA polymerase complexes
554 needed to satisfy the mRNA demands of the cell. It is worth noting that this number is approximately half of that
555 required to synthesize enough rRNA, as we saw in the previous section. We find this to be a striking result as these
556 250 RNA polymerase molecules are responsible for the transcription of the \approx 4000 protein coding genes that are
557 not ribosome associated.

558 tRNA

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

571 RNA Polymerase and σ -factor Abundance

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

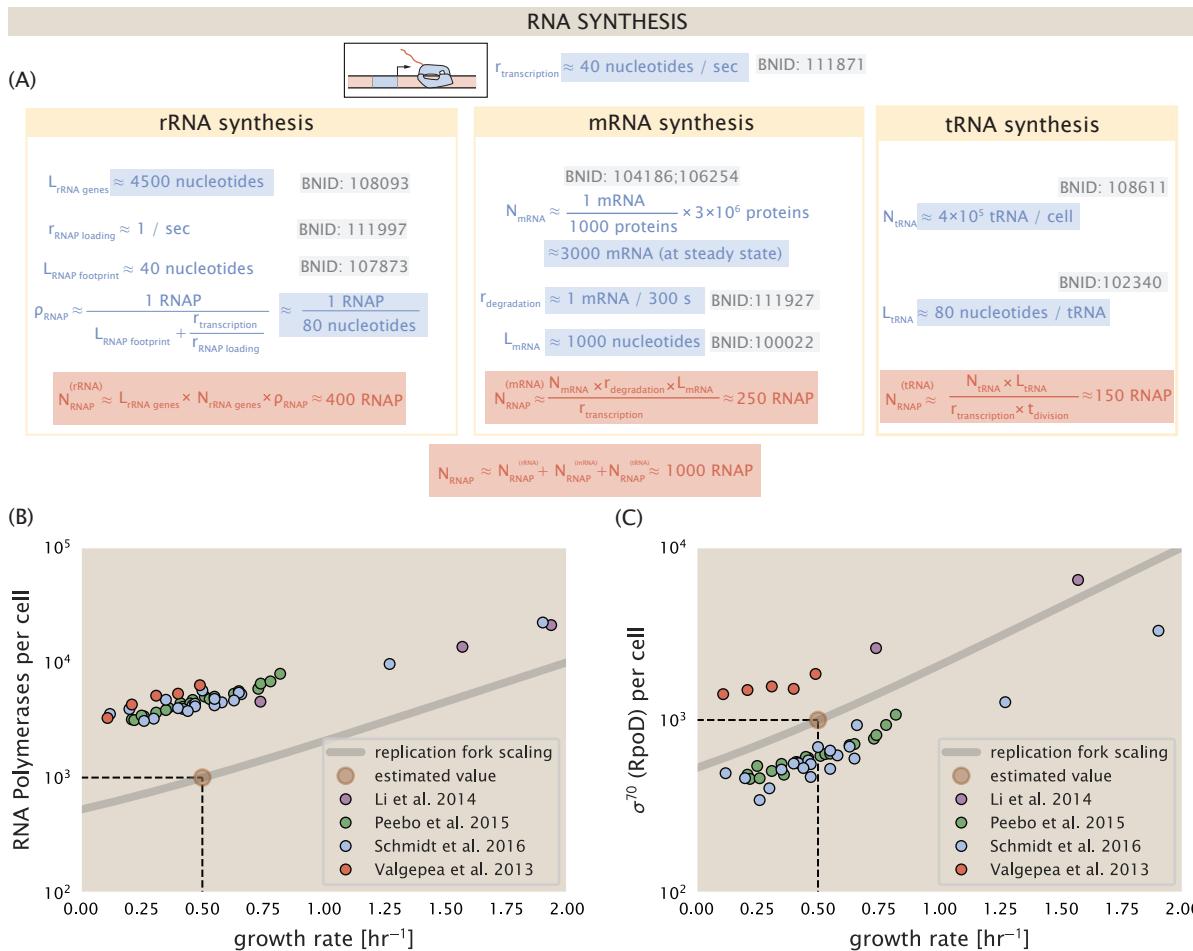


Figure 8. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. (B) The RNA polymerase core enzyme copy number as a function of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

579 polymerase present for each other as they move along the DNA (*Finkelstein and Greene, 2013*).

580 In addition, it is also vital to consider the role of σ -factors which help RNA polymerase identify and bind to trans-
581criptional start sites (*Browning and Busby, 2016*). Here we consider σ^{70} (RpoD) which is the dominant "general-
582 purpose" σ -factor in *E. coli*. While initially thought of as being solely involved in transcriptional initiation, the past
583 two decades of single-molecule work has revealed a more multipurpose role for σ^{70} including facilitating transcrip-
584 tional elongation (*Kapanidis et al., 2005; Goldman et al., 2015; Perdue and Roberts, 2011; Mooney and Landick,
585 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
586 agreement with our estimate of the number of transcriptional complexes needed.

587 These estimates provide insight as to the observed magnitude of both RNA polymerase and the σ -70 factor.
588 As we have done in the previous sections, and described in Appendix Extending Estimates to a Continuum of
589 Growth Rates, we can generalize these estimates across a wide range of growth rates (grey line in **Figure 8(B)**).
590 While there remains some disagreement in the magnitude of the copy number, this estimate appears to very
591 adequately describe the growth rate dependence of these complexes. Furthermore, these findings illustrate that
592 transcription cannot be the rate limiting step in bacterial division. **Figure 8 (A)** reveals that the availability of RNA
593 polymerase is not a limiting factor for cell division as the cell always has an apparent ~ 10 -fold excess than needed.
594 Furthermore, if more transcriptional activity was needed to satisfy the cellular requirements, more σ^{70} -factors
595 could be expressed to utilize a larger fraction of the RNA polymerase pool.

596 Translation and Ribosomal Synthesis

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

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

607 tRNA Synthetases

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

612 How many tRNAs are needed to facilitate this remarkable number of amino acid delivery events to the trans-
613 lating ribosomes? It is important to note that tRNAs are recycled after they've passed through the ribosome and
614 can be recharged with a new amino acid, ready for another round of peptide bond formation. While some *in vitro*
615 data exists on the turnover of tRNA in *E. coli* for different amino acids, we can make a reasonable estimate by
616 comparing the number of amino acids to be polymerized to cell division time. Using our stopwatch of 5000 s and
617 10^9 amino acids, we arrive at a requirement of $\approx 2 \times 10^5$ tRNA molecules to be consumed by the ribosome per
618 second.

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

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

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

640 Protein Synthesis

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

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

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

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

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

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

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

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

672 Alternatively, since N_{pep} is related to the total protein mass through the molecular weight of each protein, we
673 can also consider the growth rate in terms of the fraction of the total proteome mass dedicated to ribosomal

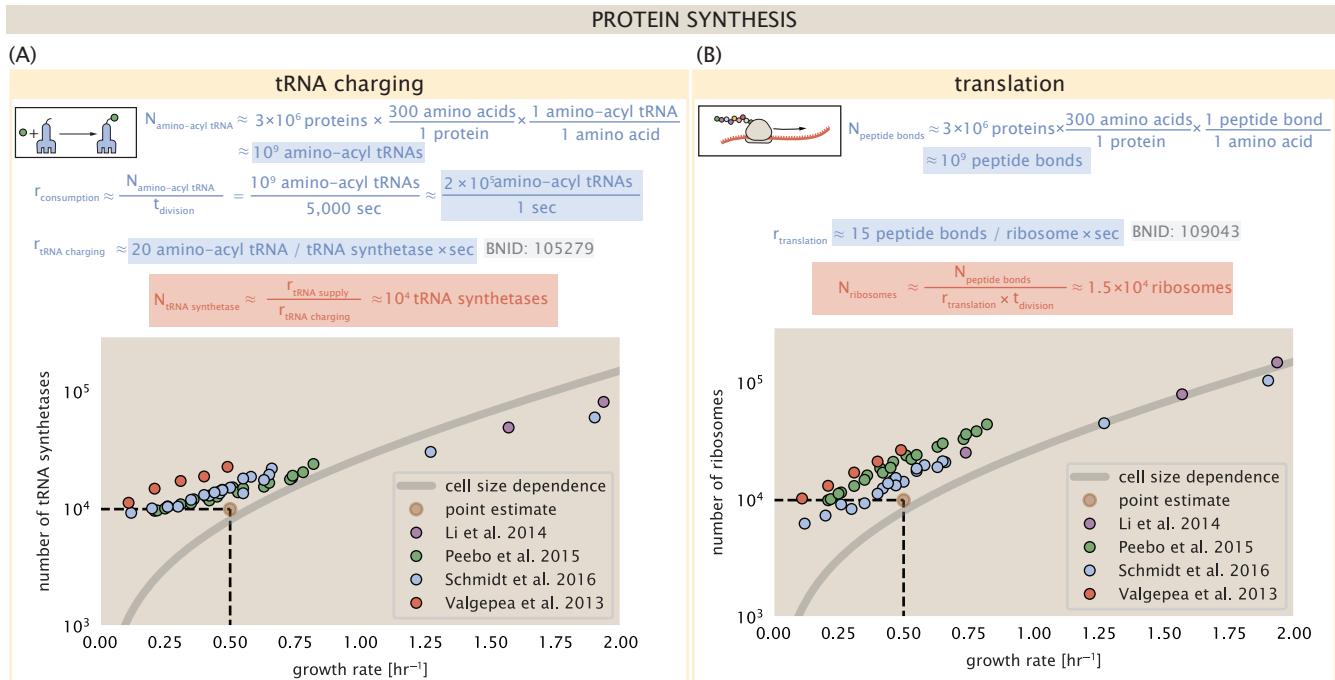


Figure 9. Estimation of the required tRNA synthetases and ribosomes. (A) Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [Glx], [IleS], [LeuS], [ValS], [AlaS]₂, [AsnS]₂, [AspS]₂, [TyrS]₂, [TrpS]₂, [ThrS]₂, [SerS]₂, [ProS]₂, [PheS]₂[PheT]₂, [MetG]₂, [LysS]₂, [HisS]₂, [GlyS]₂[GlyQ]₂). (B) Estimation of the number of ribosomes required to synthesize 10^9 peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr⁻¹. Grey lines correspond to the estimated complex abundance calculated at different growth rates. See Appendix Extending Estimates to a Continuum of Growth Rates for a more detail description of this calculation.

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

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

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

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

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

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

Relationship Between Cell Size and Growth Rate

The relationship between cell size and growth rate has long been of interest in the study of bacterial physiology, particularly following the now six decade-old observation that cell volume appears to increase exponentially with growth rate; known as Schaechter's growth law (*Schaechter et al., 1958; Taheri-Araghi et al., 2015*). However, the mechanism that governs this relationship, and even the question of whether the change in average cell size is truly exponential, has remained under debate (*Harris and Theriot, 2018*). Given the importance of cell size in relating the total protein mass that must be doubled (as well as in setting other parameters like the surface-area-to-volume ratio), we examine the influence size may have in setting the scaling of protein abundance and growth

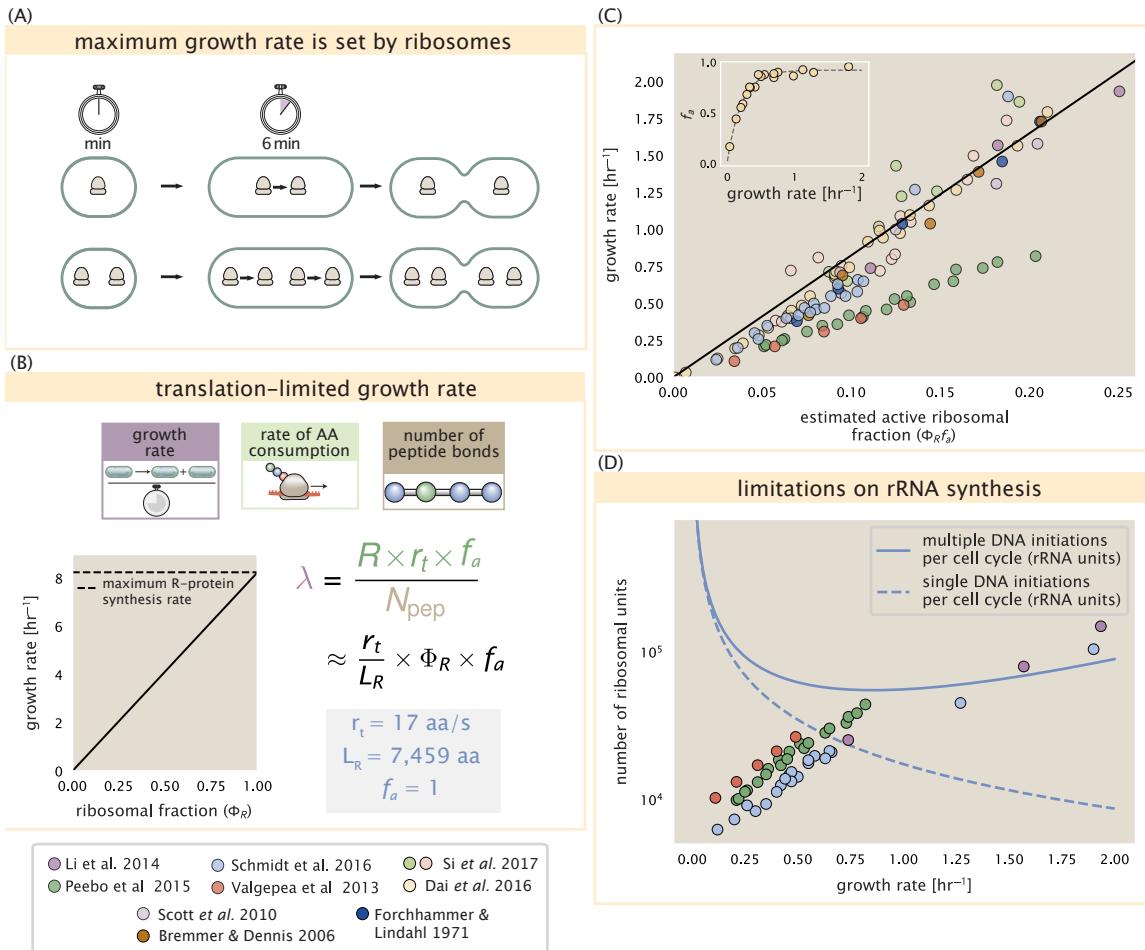


Figure 10. Translation-limited growth rate. (A) An inherent maximum growth rate is set by the time to synthesize all ribosomal subunits. This growth rate is given by r_t/L_R , where r_t is the elongation rate and L_R is the total number of amino acids that make up the entire ribosomal complex. This rate is independent of the number of ribosomes; instead limited serial requirement for each ribosome to double itself. (B) Translation-limited growth as a function of the ribosomal fraction. The solid line is calculated for an elongation rate of 17 aa per second. The dashed line corresponds to the maximum rate of ribosomal protein (R-protein) synthesis considered in part (A). (C) Actively translating ribosomal fraction versus growth rate. The actively translating ribosomal fraction is calculated using the estimated values of f_a from [Dai et al. \(2016\)](#) (shown in inset; see Appendix Calculation of active ribosomal fraction for additional detail). (D) Maximum number of rRNA units that can be synthesized as a function of growth rate. Solid curve corresponds to the rRNA copy number is calculated by multiplying the number of rRNA operons by the estimated number of (# ori) at each growth rate. The quantity (# ori) was calculated using Equation 4 and the measurements from [Si et al. \(2017\)](#) that are plotted in [Figure 11\(A\)](#). The dashed line shows the maximal number of functional rRNA units produced from a single chromosomal initiation per cell cycle.

722 rate across the proteomic datasets.

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

732 The average number of origins (# ori) is set by how often replication must be initiated per cell doubling under
733 steady-state growth. This can be quantified as

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

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

739 Why does *E. coli* add a constant volume per $\langle \# \text{ ori} \rangle$? To consider how this trend pertains to growth, we must
740 consider how the proteome size and composition changes with respect to growth rate. In **Figure 11(D)**, we ana-
741 lyze the position-dependent protein expression across the chromosome for each of the growth conditions from
742 *Schmidt et al. (2016)*. Here, we have calculated a running Gaussian average of protein copy number (20 kbp st.
743 dev. averaging window) based on each gene's transcriptional start site, which were then median-subtracted to
744 account for the differences in total protein abundance with each growth condition. Importantly, we find that
745 the major deviations in protein copy number are largely restricted to regions of ribosomal protein genes, with
746 substantially higher deviations observed for cells with high $\langle \# \text{ ori} \rangle$ (teal), as compared to those with low $\langle \# \text{ ori} \rangle$
747 (purple). This is particularly apparent for genes closer to the origin, where the majority of ribosomal proteins are
748 found. This suggests that in addition to the linear scaling between protein abundance and $\langle \# \text{ ori} \rangle$, the relative
749 ribosomal abundance is tuned in proportion to $\langle \# \text{ ori} \rangle$. Given the increased rRNA gene dosage required at faster
750 growth rates, additional rounds of DNA replication have the effect of skewing DNA dosage in favor of additional
751 ribosomal synthesis. Since growth rate depends specifically on the ribosomal fraction Φ_R , this result suggests that
752 cells are changing their size as a way to vary the absolute number of ribosomes per cell and tune Φ_R according to
753 better match available nutrient conditions.

754 **Alarmone-Mediated Regulation Controls the Rate of Protein Synthesis**

755 As we have seen, cell size, total proteomic content, and the number of ribosomes are all interconnected and
756 influence the achievable growth rate. The drastic change in these parameters across different growth conditions
757 also suggests that cells are tuning them to better match their biosynthetic capacity to the specific environment.
758 Take, as another illustration of this, the recent experimental work by *Dai et al. (2016)*. In one set of experiments the
759 authors considered growth in cells whose primary glucose transport system was disrupted ($\Delta ptsG$). Unsurprisingly,
760 the growth rate was reduced, and was measured at about two-fold slower than their wild-type line. This change,
761 however, was not simply the result of now-limiting carbon uptake. Instead, cells accommodated this perturbation
762 by also reducing their ribosomal mass fraction by a factor of two, which is still in line with **Equation 3** under
763 translation-limited growth. In this final, we explore the interconnection between cell size, ribosome content, and
764 growth rate by formulating a minimal model of growth rate control. We use it to quantitatively show how tuning
765 these parameters help cells maximize their growth rate.

766 To react to changes in nutrient conditions, bacteria rely on the synthesis or degradation of secondary-messenger
767 molecules like (p)ppGpp, which cause global changes in transcriptional and translational activity. In *E. coli*, amino
768 acid starvation causes the accumulation of de-acylated tRNAs at the ribosome's A-site and leads to a strong in-
769 crease in (p)ppGpp synthesis activity by the enzyme RelA (*Hauryliuk et al., 2015*). Cells also accumulate (p)ppGpp

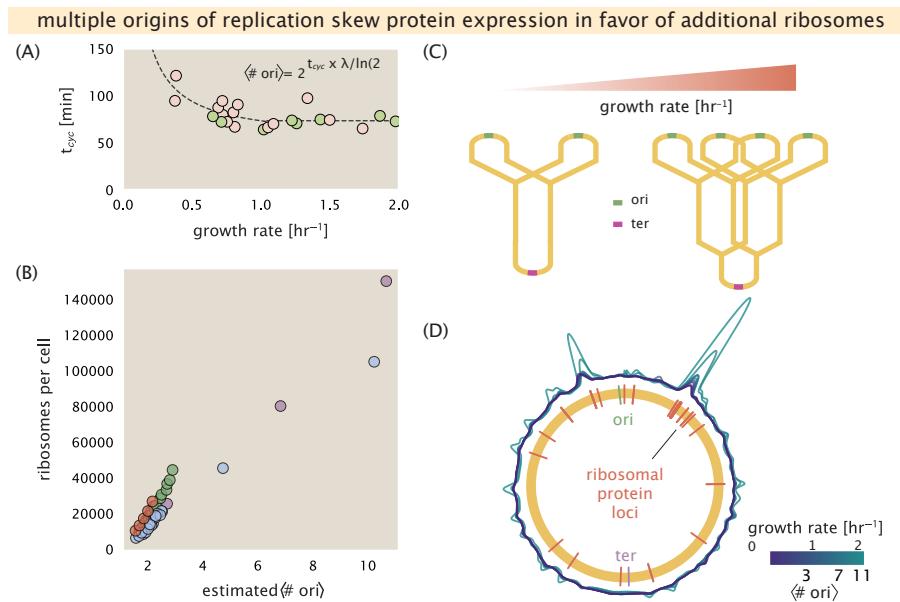


Figure 11. Cells increase absolute ribosome abundance with $\langle \# \text{ori} \rangle$. (A) Experimental data from Si *et al.* (2017). Dashed line shows fit to the data, which were used to estimate $\langle \# \text{ori} \rangle$. t_{cyc} was assumed to vary in proportion to τ for doubling times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

770 during steady-state growth in poorer growth conditions, which leads to a decrease in the fraction of actively translating ribosomes, f_a (with $f_a \approx 0.5$ at a growth rate of $\approx 0.3 \text{ hr}^{-1}$).

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

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

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

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

$$793 \frac{1}{r_t} = \frac{1}{k_{on}\alpha[AA]_{\text{eff}}} + \frac{1}{r_t^{\max}}. \quad (5)$$

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

$$798 K_D = \frac{r_t^{\max}}{\alpha k_{on}}, \quad (6)$$

799 where the elongation rate r_t is now given by

$$800 r_t = \frac{r_t^{\max}}{1 + K_D/[AA]_{\text{eff}}}. \quad (7)$$

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

$$804 \int_0^t \frac{d[AA]_{\text{eff}}}{dt} dt = \int_0^t ([r_{AA}] - [r_t \times R \times f_a]) dt, \quad (8)$$

805 where the time from 0 to t is an arbitrary length of time, and the square brackets indicate concentrations per unit
806 time. Integrating *Equation 8* yields.

$$807 [AA]_{\text{eff}} = t([r_{AA}] - [r_t \times R \times f_a]). \quad (9)$$

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

$$809 [AA]_{\text{eff}} = \frac{t(r_{AA} - r_t \times R \times f_a)}{V}, \quad (10)$$

804 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 expression for
805 $[AA]_{\text{eff}}$ in hand, we can now solve **Equation 7** for r_t , which is a quadratic function with a physically-meaningful root
806 of

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

807 In **Figure 12(B)**, we illustrate how the elongation rate depends on the ribosomal copy number. Here, we have
808 considered a unit volume $V = 1\mu\text{m}^3$, a unit time $t = 1\text{ s}$, a $K_D = 5\text{ mM}$ (inferred from **Bennett et al. (2009)**),
809 $f_a = 1$, and an arbitrarily chosen $r_{AA} = 5 \times 10^6 \text{ AA} \cdot \text{s}^{-1} \cdot \mu\text{m}^{-3}$. At low ribosome copy numbers, the observed
810 elongation rate is dependent primarily on the ratio of K_D/Vr_{AA} [as $r_t^{\max} \times R \times f_a \ll r_{AA}$, point (1) in **Figure 12(B)**].
811 As the ribosome copy number is increased such that the amino acid supply rate and consumption rate are nearly
812 equal [point (2) in **Figure 12(B)**], the observed elongation rate begins to decrease sharply. When the ribosome
813 copy number is increased even further, consumption at the maximum elongation rate exceeds the supply rate,
814 yielding a significantly reduced elongation rate [point (3) in **Figure 12B**]. While the elongation rate will always be
815 dominated by the amino acid supply rate at sufficiently low ribosome copy numbers, the elongation rate at larger
816 ribosome abundances can be increased by tuning f_a such that not all ribosomes are elongating, reducing the total
817 consumption rate.

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

822 Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and Metabolic 823 Capacity.

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

828 Earlier, we considered ribosome biosynthesis as the growth-rate determining cellular process in **Equation 2** by
829 stating that the cellular growth rate λ was related to the ribosome abundance, elongation rate, active ribosome
830 fraction, and the total number of peptide bonds to be formed, N_{pep} . We return to this limit in light of our expression
831 for a condition-dependent elongation rate r_t given by **Equation 11**. **Figure 12(C)** shows how the observed growth
832 rate depends on the rate of amino acid supply r_{AA} as a function of the cellular ribosome copy number. A feature
833 immediately apparent is the presence of a maximal growth rate whose dependence on R (and consequently, the
834 cell volume) increases with increasing r_{AA} . Importantly, however, there is an optimum set of R , N_{pep} , and V that
835 are strictly dependent on the value of r_{AA} . Increasing the ribosomal concentration beyond the cell's metabolic
836 capacity has the adverse consequence of depleting the supply of amino acids and a concomitant decrease in the
837 elongation rate r_t , [**Figure 12(B)**].

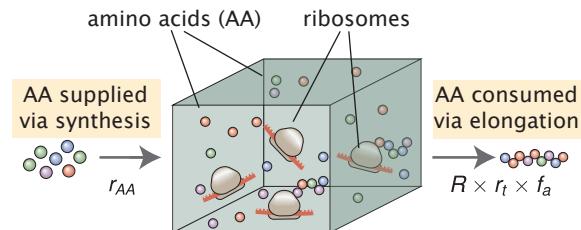
838 Also of note is the growth rate profiles shown for low amino acid supply rates [purple and blue lines in **Fig-**
839 **ure 12(C)**], representing growth in nutrient-poor media. In these conditions, there no longer exists a peak in
840 growth, at least in the range of physiologically-relevant ribosome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

845 Discussion

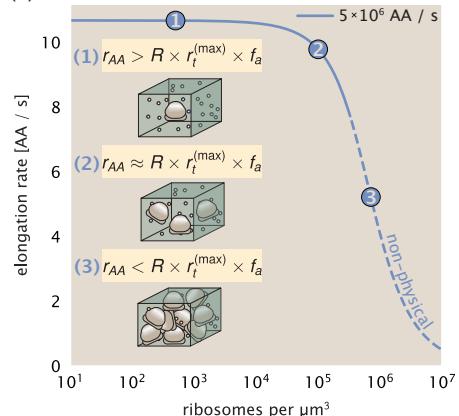
846 Continued experimental and technological improvements have led to a treasure trove of quantitative biological
847 data (**Hui et al., 2015; Schmidt et al., 2016; Si et al., 2017; Gallagher et al., 2020; Peebo et al., 2015; Valgepea et al.,**
848 **2013**), and an ever advancing molecular view and mechanistic understanding of the constituents that support
849 bacterial growth (**Taheri-Araghi et al., 2015; Morgenstein et al., 2015; Si et al., 2019; Karr et al., 2012; Kostinski and**

(A)

A MINIMAL MODEL FOR NUTRIENT-LIMITED GROWTH



(B)



(C)

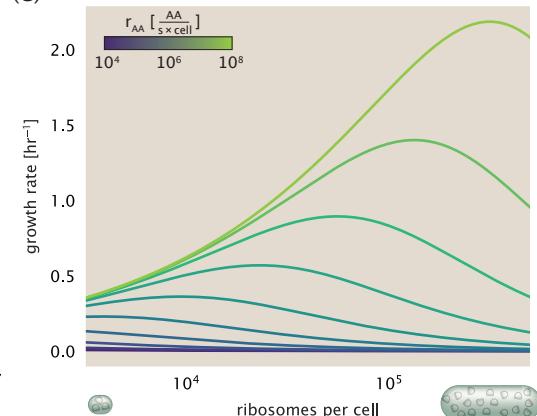


Figure 12. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times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

850 **Reuveni, 2020**). In this work we have compiled what we believe to be the state-of-the-art knowledge on proteomic
851 copy number across a broad range of growth conditions in *E. coli*. We have made this data accessible through a
852 [GitHub repository](#), and an interactive figure that allows exploration of specific protein and protein complex copy
853 numbers. Through a series of order-of-magnitude estimates that traverse key steps in the bacterial cell cycle, this
854 proteomic data has been a resource to guide our understanding of two key questions: what biological processes
855 limit the absolute speed limit of bacterial growth, and how do cells alter their molecular constituents as a function
856 of changes in growth rate or nutrient availability? While not exhaustive, our series of estimates provide insight
857 on the scales of macromolecular complex abundance across four classes of cellular processes – the transport of
858 nutrients, the production of energy, the synthesis of the membrane and cell wall, and the numerous steps of the
859 central dogma.

860 In general, the copy numbers of the complexes involved in these processes were reasonable agreement with
861 our order-of-magnitude estimates. Since many of these estimates represent soft lower-bound quantities, this
862 suggests that cells do not express proteins grossly in excess of what is needed for a particular growth rate. Several
863 exceptions, however, also highlight the dichotomy between a proteome that appears to "optimize" expression
864 according to growth rate and one that must be able to quickly adapt to environments of different nutritional quality.
865 Take, for example, the expression of carbon transporters. Shown in **Figure 2(B)**, we find that cells always express
866 a similar number of glucose transporters irrespective of growth condition. At the same time, it is interesting to
867 note that many of the alternative carbon transporters are still expressed in low but non-zero numbers (≈ 10 -
868 100 copies per cell) across growth conditions. This may relate to the regulatory configuration for many of these
869 operons, which require the presence of a metabolite signal in order for alternative carbon utilization operons to
870 be induced (**Monod, 1949; Laxhuber et al., 2020**). Furthermore, upon induction, these transporters are expressed
871 and present in abundances in close agreement with a simple estimate.

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

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

901 growth conditions associated with the proteomic data and further consideration will be needed for anaerobic
902 growth.

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

918 Acknowledgements

919 We thank Matthias Heinemann, Alexander Schmidt, and Gene-Wei Li for additional input regarding their data. We
920 also thank members of the Phillips, Theriot, Kondev, and Garcia labs for useful discussions. R.P. is supported by
921 La Fondation Pierre-Gilles de Gennes, the Rosen Center at Caltech, and the NIH 1R35 GM118043 (MIRA). N.M.B is
922 a HHMI Fellow of The Jane Coffin Childs Memorial Fund.

923 Competing Interests

924 The authors declare no competing interests.

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

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

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

938	Contents	
939	Introduction	1
940	Uptake of Nutrients	4
941	Nitrogen Transport	4
942	Carbon Transport	4
943	Phosphorus and Sulfur Transport	6
944	Limits on Transporter Expression	8
945	Energy Production	8
946	ATP Synthesis	8
947	Generating the Proton Electrochemical Gradient	9
948	Energy Production in a Crowded Membrane.	9
949	Synthesis of the Cell Envelope	11
950	Lipid Synthesis	11
951	Peptidoglycan Synthesis	13
952	Function of the Central Dogma	14
953	DNA	14
954	dNTP synthesis	15
955	DNA Replication	15
956	RNA Synthesis	17
957	rRNA	17
958	mRNA	18
959	tRNA	18
960	RNA Polymerase and σ -factor Abundance	18
961	Translation and Ribosomal Synthesis	20
962	tRNA Synthetases	20
963	Protein Synthesis	21
964	Translation and Ribosomal Synthesis as a Growth-Rate Limiting Step	21
965	Relationship Between Cell Size and Growth Rate	23
966	Alarmone-Mediated Regulation Controls the Rate of Protein Synthesis	25
967	Ribosomal Elongation Rate and Cellular Growth Rate are Linked by Amino Acid Scarcity	27
968	Optimal Growth Rate, Ribosomal Content, and Cell Size Depend on Nutrient Availability and Metabolic Capacity.	28
970	Discussion	28
971	Acknowledgements	31
972	Competing Interests	31
973	Experimental Details Behind Proteomic Data	35
974	Fluorescence based measurements	35
975	Ribosomal profiling measurements	35
976	Mass spectrometry measurements	36
977	Summary of Proteomic Data	36
978	Estimation of Cell Size and Surface Area	37

979	Estimation of Total Protein Content per Cell	39
980	Estimating Volume and Number of Amino Acids from Ribosome Copy Number	39
981	Additional Considerations of Schmidt <i>et al.</i> Data Set	39
982	Effect of cell volume on reported absolute protein abundances	41
983	Relaxing assumption of constant protein concentration across growth conditions	43
984	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	43
985	Calculation of Complex Abundance	44
986	Extending Estimates to a Continuum of Growth Rates	46
987	Estimation of the total cell mass	46
988	Complex Abundance Scaling With Cell Volume	46
989	A Relation for Complex Abundance Scaling With Surface Area	47
990	Number of Lipids	47
991	Number of Murein Monomers	47
992	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	48
993	Calculation of active ribosomal fraction.	48
994	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	48

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).

995 Experimental Details Behind Proteomic Data

996 Here we provide a brief summary of the experiments behind each proteomic data set. The purpose of this section
997 is to identify how the authors arrived at absolute protein abundances. In the following section (Section Summary
998 of Proteomic Data) we will then provide a summary of the final protein abundance measurements that were
999 used throughout the main text. Table 1 provides an overview of the publications we considered. These are pre-
1000 dominately mass spectrometry-based, with the exception of the work from Li *et al.* (2014) which used ribosomal
1001 profiling, and the fluorescence-based counting done in Taniguchi *et al.* (2010).

1002 Fluorescence based measurements

1003 In the work of Taniguchi *et al.* (2010), the authors used a chromosomal YFP fusion library where individual strains
1004 have a specific gene tagged with a YFP-coding sequence. 1018 of their 1400 attempted strains were used in the
1005 work. A fluorescence microscope was used to collect cellular YFP intensities across all these strains. Through au-
1006 tomated image analysis, the authors normalized intensity measurements by cell size to account for the change
1007 in size and expression variability across the cell cycle. Following correction of YFP intensities for cellular autoflu-
1008 orescence, final absolute protein levels were determined by a calibration curve with single-molecule fluorescence
1009 intensities. This calibration experiment was performed separately using a purified YFP solution.

1010 Ribosomal profiling measurements

1011 The work of Li *et al.* (2014) takes a sequencing based approach to estimate protein abundance. Ribosomal pro-
1012 filing, which refers to the deep sequencing of ribosome-protected mRNA fragments, can provide a quantitative
1013 measurement of the protein synthesis rate. As long as the protein life-time is long relative to the cell doubling
1014 time, it is possible to estimate absolute protein copy numbers. The absolute protein synthesis rate has units of
1015 proteins per generation, and for stable proteins will also correspond to the protein copy number per cell.

1016 In the experiments, ribosome-protected mRNA is extracted from cell lysate and selected on a denaturing poly-
1017 acrylamide gel for deep sequencing (15–45 nt long fragments collected and sequenced by using an Illumina HiSeq
1018 2000 in Li *et al.* (2014)). Counts of ribosome footprints from the sequencing data were then corrected empiri-
1019 cally for position-dependent biases in ribosomal density across each gene, as well as dependencies on specific
1020 sequences including the Shine-Dalgarno sequence. These data-corrected ribosome densities represent relative
1021 protein synthesis rates. Absolute protein synthesis rates are obtained by multiplying the relative rates by the total
1022 cellular protein per cell. The total protein per unit volume was determined with the Lowry method to quantify
1023 total protein, calibrated against bovine serum albumin (BSA). By counting colony-forming units following serial
1024 dilution of their cell cultures, they then calculated the total protein per cell.

1025 **Mass spectrometry measurements**

1026 Perhaps not surprisingly, the data is predominantly mass spectrometry based. This is largely due to tremendous
1027 improvements in the sensitivity of mass spectrometers, as well as improvements in sample preparation and data
1028 analysis pipelines. It is now a relatively routine task to extract protein from a cell and quantify the majority of
1029 proteins present by shotgun proteomics. In general, this involves lysing cells, enzymatically digesting the proteins
1030 into short peptide fragments, and then introducing them into the mass spectrometer (e.g. with liquid chromatog-
1031 raphy and electrospray ionization), which itself can have multiple rounds of detection and further fragmentation
1032 of the peptides.

1033 Most quantitative experiments rely on labeling protein with stable isotopes, which allow multiple samples to
1034 be measured together by the mass spectrometer. By measuring samples of known total protein abundance simul-
1035 taneously (i.e. one sample of interest, and one reference), it is possible to determine relative protein abundances.
1036 Absolute protein abundances can be estimated following the same approach used above for ribosomal profil-
1037 ing, which is to multiply each relative abundance measurement by the total cellular protein per cell. This is the
1038 approach taken by *Valgepea et al. (2013)* and *Peebo et al. (2015)*, with relative protein abundances determined
1039 based on the relative peptide intensities (label free quantification 'LFQ' intensities). For the data of *Valgepea et al.*
1040 (*2013*), total protein per cell was determined by measuring total protein by the Lowry method, and counting colony-
1041 forming units following serial dilution. For the data from *Peebo et al. (2015)*, the authors did not determine cell
1042 quantities and instead report the cellular protein abundances in protein per unit volume by assuming a mass
1043 density of 1.1 g/ml, with a 30% dry mass fraction.

1044 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide fragments of known
1045 abundance. These can serve as a direct way to calibrate mass spectrometry signal intensities to absolute mass.
1046 This is the approach taken by *Schmidt et al. (2016)*. In addition to a set of shotgun proteomic measurements to
1047 determine proteome-wide relative abundances, the authors also performed absolute quantification of 41 proteins
1048 covering over four orders of magnitude in cellular abundance. Here, a synthetic peptide was generated for each of
1049 the proteins, doped into each protein sample, and used these to determine absolute protein abundances of the 41
1050 proteins. These absolute measurements, determined for every growth condition, were then used as a calibration
1051 curve to convert proteomic-wide relative abundances into absolute protein abundance per cell. A more extensive
1052 discussion of the *Schmidt et al. (2016)* data set can be found in Section Additional Considerations of Schmidt et al.
1053 Data Set.

1054 **Summary of Proteomic Data**

1055 In the work of the main text we only used the data from *Valgepea et al. (2013)*; *Li et al. (2014)*; *Peebo et al. (2015)*;
1056 *Schmidt et al. (2016)*. As shown in *Figure 13(A)*, the reported total protein abundances in the work of *Taniguchi*
1057 *et al. (2010)* and *Soufi et al. (2015)* differed quite substantially from the other work. For the work of *Taniguchi*
1058 *et al. (2010)* this is in part due to a lower coverage in total proteomic mass quantified, though we also noticed that
1059 most proteins appear undercounted when compared to the other data.

1060 *Figure 13(B)* summarizes the total protein mass for each data point in our final compiled data set. We note that
1061 protein abundances were all scaled so they followed a common growth rate-dependent change in total protein
1062 mass. While our inclination initially was to leave reported copy numbers untouched, a notable discrepancy in the
1063 scaling total protein per cell between *Schmidt et al. (2016)* and the other data sets forced us to dig deeper into
1064 those measurements (compare *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 13(A)*). The particular trend
1065 in *Schmidt et al. (2016)* appears to be due to assumptions of cell size and we provide a more extensive discussion
1066 and analysis of that data set in section Additional Considerations of Schmidt et al. Data Set. As a compromise, and
1067 in an effort to treat all data equally, we instead scaled all protein abundance values to a data-driven estimate of
1068 total protein per cell. Here we used cell size measurements from *Si et al. (2017, 2019)*, and an estimate of total
1069 protein content through expected dry mass. Total protein per cell was estimated using available data on total
1070 DNA, RNA, and protein from *Basan et al. (2015)*; *Dai et al. (2016)*, which account for the majority of dry mass in the
1071 cell. We consider these details in sections Estimation of Cell Size and Surface Area and Estimation of Total Protein
1072 Content per Cell that follows.

1073 Lastly, in *Figure 14* we show the total proteomic coverage and overlap of proteins quantified across each data

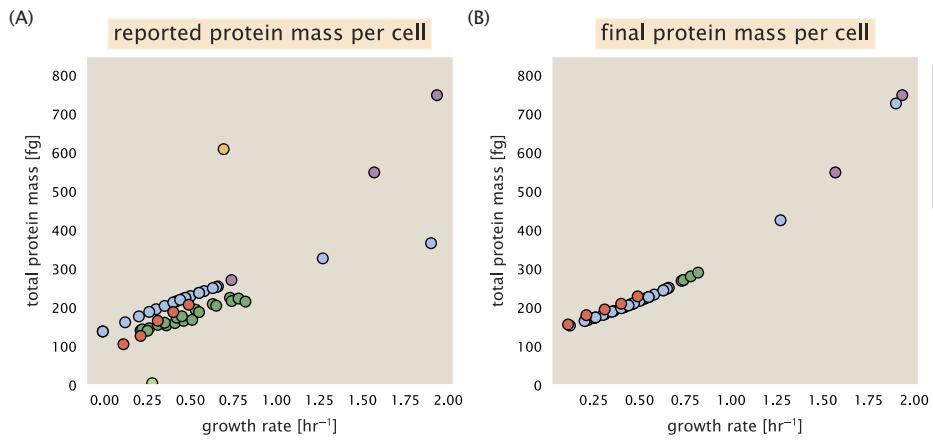


Figure 13. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as original reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al. (2016)*. Note that the data from *Peebo et al. (2015)* only reported protein abundances per unit volume and total protein per cell was found by multiplying these by the growth-rate dependent cell size as determined by *Si et al. (2017)*. (B) Adjusted total protein abundances across the proteomic data sets are summarized. Protein abundances were adjusted so that all data shared a common set of growth-rate dependent total protein per cell and cellular protein concentration following the cell size expectations of *Si et al. (2017)* (see section on Estimation of Cell Size and Surface Area for further details).

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).

1082 Estimation of Cell Size and Surface Area

1083 Since most of the proteomic data sets lack cell size (i.e. volume) measurements, we chose instead to use a common
 1084 estimate of size for any analysis requiring cell size or surface area. Since each of the data sets used either K-12
 1085 MG1655 or its derivative, BW25113 (from the lab of Barry L. Wanner; the parent strain of the Keio collection
 1086 (*Datsenko and Wanner, 2000*; *Baba et al., 2006*)), we fit the MG1655 cell size data from the supplemental material
 1087 of *Si et al. (2017, 2019)* using the optimize.curve_fit function from the Scipy python package (*Virtanen et al., 2020*).

1088 The average size measurements from each of their experiments are shown in Figure *Figure 15*, with cell length
 1089 and width shown in (A) and (B), respectively. The length data was well described by the exponential function $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, while the width data was well described by $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. In order to estimate cell size we take the
 1090 cell as a cylinders with two hemispherical ends (*Si et al., 2017*; *Basan et al., 2015*). Specifically, cell size is estimated
 1091 from,

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

1093 where r is half the cell width. A best fit to the data is described by $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$. Calculation of the cell surface
 1094 area is given by,

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

1095 where η is the aspect ratio ($\eta = l/w$) (*Ojkic et al., 2019*).

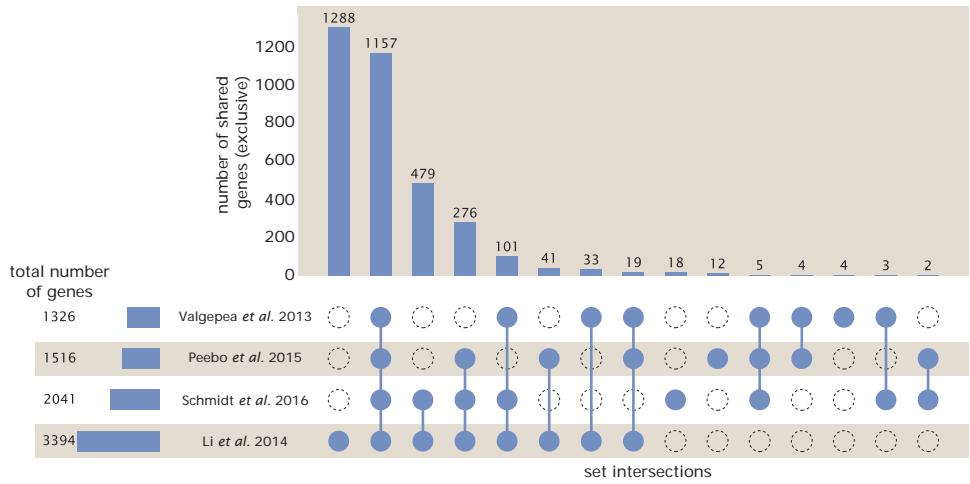


Figure 14. Comparison of proteomic coverage across different data sets. An UpSet diagram (*Lex et al., 2014*) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of *Valgepea et al. (2013)*; *Li et al. (2014)*; *Schmidt et al. (2016)*; *Peebo et al. (2015)*. The total number of genes reported in each individual data set are noted on the bottom left, while their overlap is summarized in the bar chart. Each column here refers to the intersection of a set of data sets identified by blue circles.

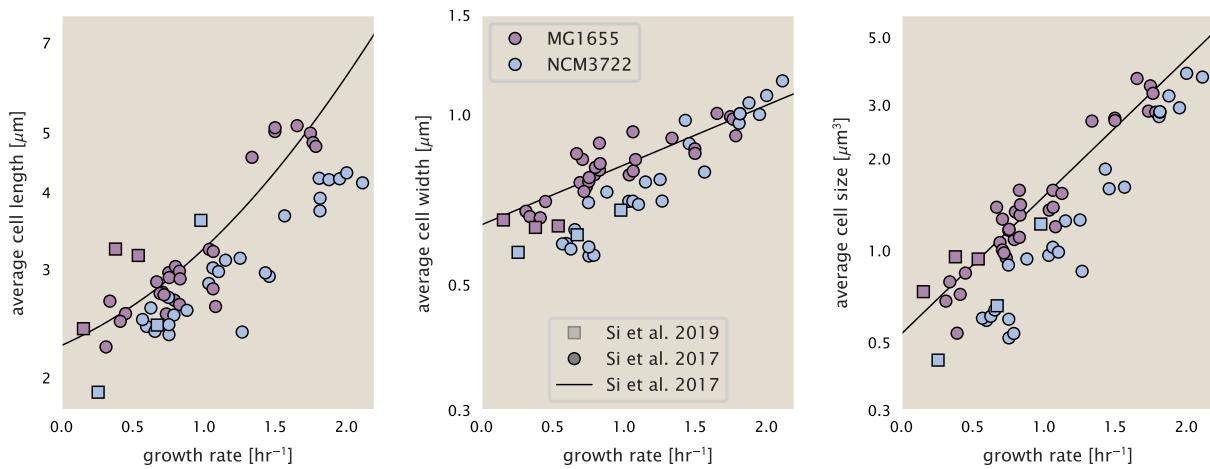


Figure 15. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 images and 5,000-30,000 for each set of samples; which likely means about 1,000-5,000 measurements per mean value reported here since they considered about 6 conditions at a time). Fits were made to the MG1655 strain data; length: $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, 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 $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$.

1096 Estimation of Total Protein Content per Cell

1097 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell size from
1098 the fit shown in Figure **Figure 15(C)** ($0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$). We then estimate the total protein content from the total
1099 dry mass of the cell. Here we begin by noting that for almost the entire range of growth rates considered here,
1100 protein, DNA, and RNA were reported to account for at least 90 % of the dry mass (*Basan et al. (2015)*). The authors
1101 also found that the total dry mass concentration was roughly constant across growth conditions. Under such a
1102 scenario, we can calculate the total dry mass concentration for protein, DNA, and RNA, which is given by 1.1 g/ml
1103 $\times 30\% \times 90\%$ or about $[M_p] = 300 \text{ fg per fl}$. Multiplying this by our prediction of cell size gives the total dry mass
1104 per cell.

1105 However, even if dry mass concentration is relatively constant across growth conditions, it is not obvious how
1106 protein concentration might vary due to the substantial increase in rRNA at faster growth rates (*Dai et al. (2016)*).
1107 This is a well-documented result that arises from an increase in ribosomal abundance at faster growth rates (*Scott*
1108 *et al. (2010)*). To proceed therefore rely on experimental measurements of total DNA content per cell that also
1109 come from Basan *et al.*, and RNA to protein ratios that were measured in Dai *et al.* (and cover the entire range of
1110 growth conditions considered here). These are reproduced in Figure **Figure 16(A)** and (B), respectively.

1111 Assuming that the protein, DNA, and RNA account for 90 % of the total dry mass, the protein mass can then de-
1112 termined by first subtracting the experimentally measured DNA mass, and then using the experimental estimate
1113 of the RNA to protein ratio. The total protein per cell is will be related to the summed RNA and protein mass by,

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

1114 (RP_{ratio} refers to the RNA to protein ratio as measured by Dai *et al.*). In Figure **Figure 16(C)** we plot the estimated
1115 cellular concentrations for protein, DNA, and RNA from these calculations, and in Figure **Figure 16(D)** we plot their
1116 total expected mass per cell. This later quantity is the growth rate-dependent total protein mass that was used to
1117 estimate total protein abundance across all data sets (and summarized in **Figure 13(B)**).

1118 Estimating Volume and Number of Amino Acids from Ribosome Copy Number

1119 Towards the end of the main text, we examine a coarse-grained model of nutrient-limited growth. A key point
1120 in our analysis was to consider how elongation rate r_e and growth rate λ vary with respect to the experimentally
1121 observed changes in cell size, total number of peptide bonds per cell N_{pep} , and ribosomal content. In order to
1122 do maintain parameters in line with the experimental data, but otherwise allow us to explore the model, we
1123 performed a phenomenological fit of N_{pep} and V as a function of the measured ribosomal copy number R . As has
1124 been described in the preceding sections of this supplement, we estimate cell volume for each growth condition
1125 using the size measurements from *Si et al. (2017, 2019)*, and N_{pep} is approximated by taking the total protein mass
1126 and dividing this number by the average mass of an amino acid, 110 Da (BNID: 104877).

1127 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression of the log trans-
1128 form of these parameters as a function of the log transform of the ribosome copy number. Using optimization by
1129 minimization, we estimated the best-fit values of the intercept and slope for each regression. ?? shows the result
1130 of each regression as a dashed line, given in terms of the linear-scaled data.

1131 Additional Considerations of Schmidt *et al.* Data Set

1132 While the data set from *Schmidt et al. (2016)* remains a heroic effort that our labs continue to return to as a
1133 resource, there were steps taken in their calculation of protein copy number that we felt needed further consider-
1134 ation. In particular, the authors made an assumption of constant cellular protein concentration across all growth
1135 conditions and used measurements of cell volume that appear inconsistent with an expected exponential scaling
1136 of cell size with growth rate that is well-documented in *E. coli* (*Schaechter et al. (1958); Taheri-Araghi et al. (2015);*
1137 *Si et al. (2017)*).

1138 We begin by looking at their cell volume measurements, which are shown in blue in Figure **Figure 18**. As a
1139 comparison, we also plot cell sizes reported in three other recent papers: measurements from Taheri-Araghi *et*
1140 *al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from Basan *et al.* come from the lab of Terence
1141 Hwa. Each set of measurements used microscopy and cell segmentation to determine the length and width, and

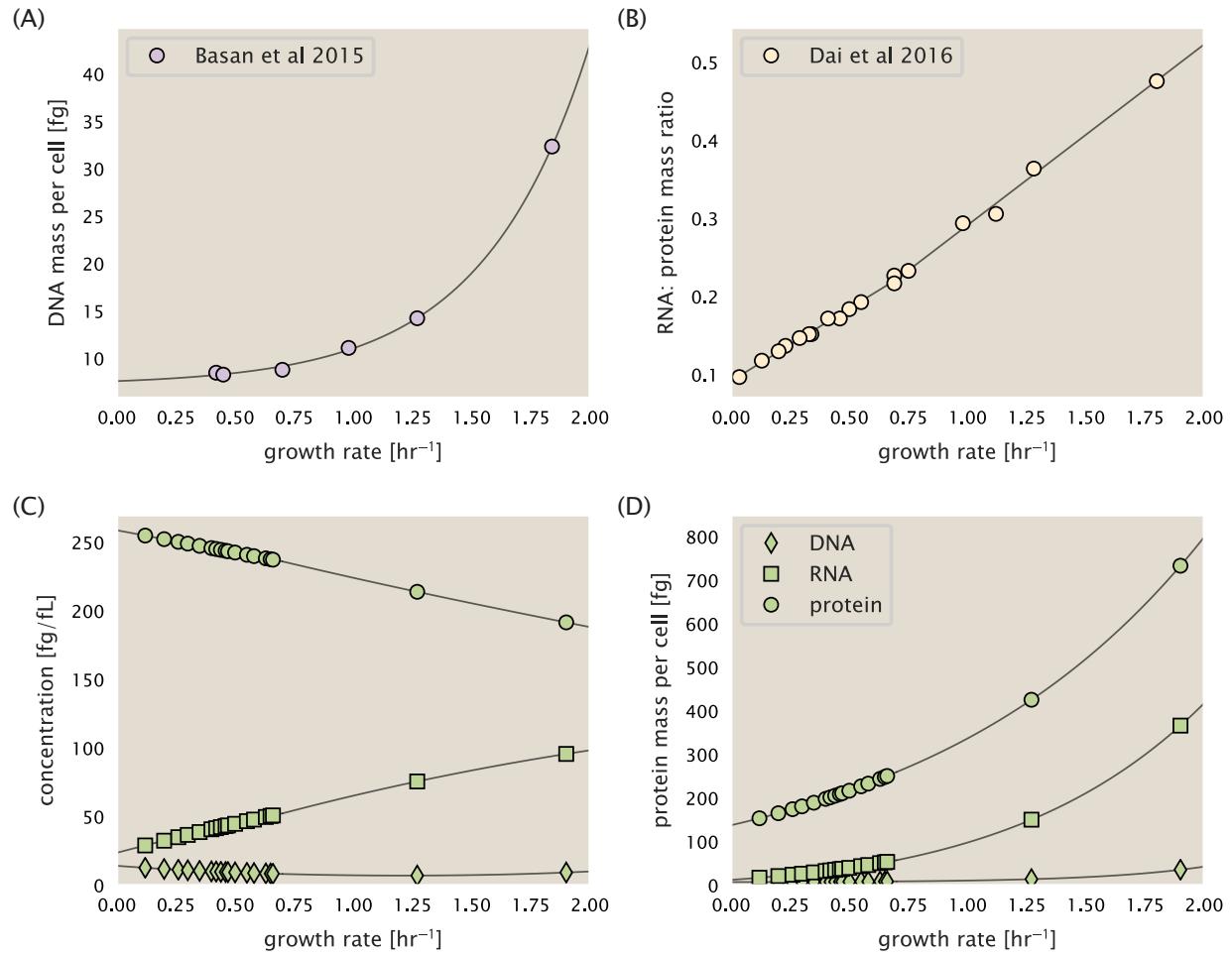


Figure 16. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23 \cdot \lambda} + 7.2$ fg per cell, where λ is the growth rate in hr⁻¹). (B) RNA to protein measurements as a function of growth rate. The data was fit to two lines: for growth rates below 0.7 hr⁻¹, the RNA/protein ratio is $0.18 \cdot \lambda + 0.093$, while for growth rates faster than 0.7 hr⁻¹ the RNA/protein ratio is given by $0.25 \cdot \lambda + 0.035$. For (A) and (B) cells are grown under varying levels of nutrient limitation, with cells grown in minimal media with different carbon sources for the slowest growth conditions, and rich-defined media for fast growth rates. (C) Predictions of cellular protein, DNA, and RNA concentration. (D) Total cellular mass predicted for protein, DNA, and RNA using the cell size predictions from Si *et al.*. Symbols (diamond: DNA, square: RNA, circle: protein) show estimated values of mass concentration and mass per cell for the specific growth rates in Schmidt *et al.* (2016).

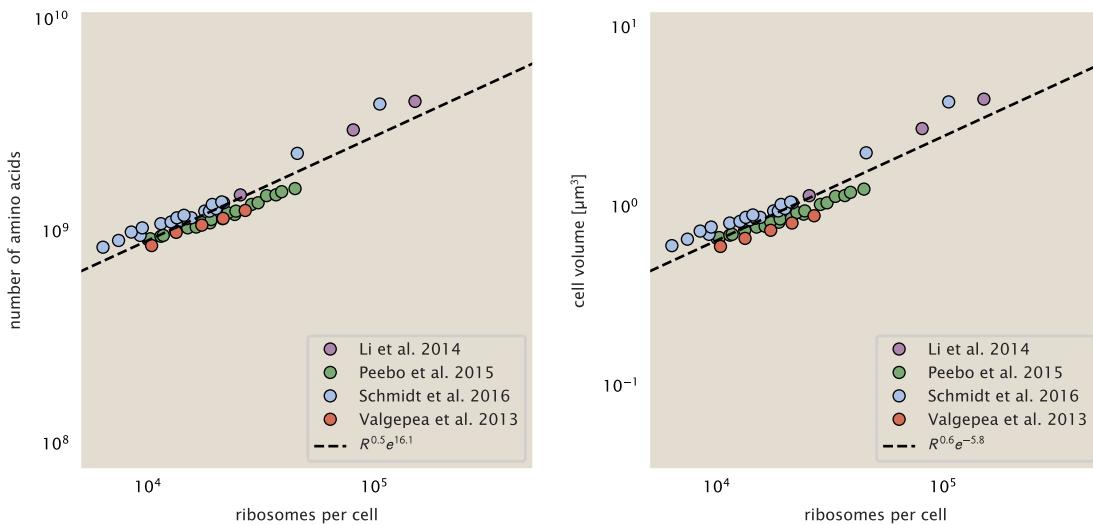


Figure 17. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows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

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

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

Effect of cell volume on reported absolute protein abundances

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

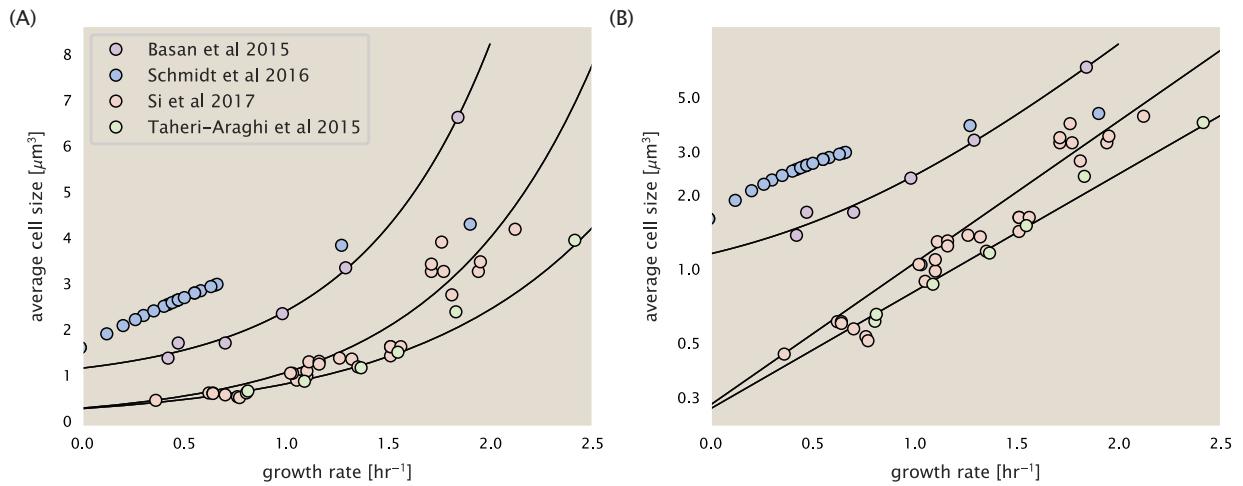


Figure 18. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 ([Volkmer and Heinemann \(2011\)](#)) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa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

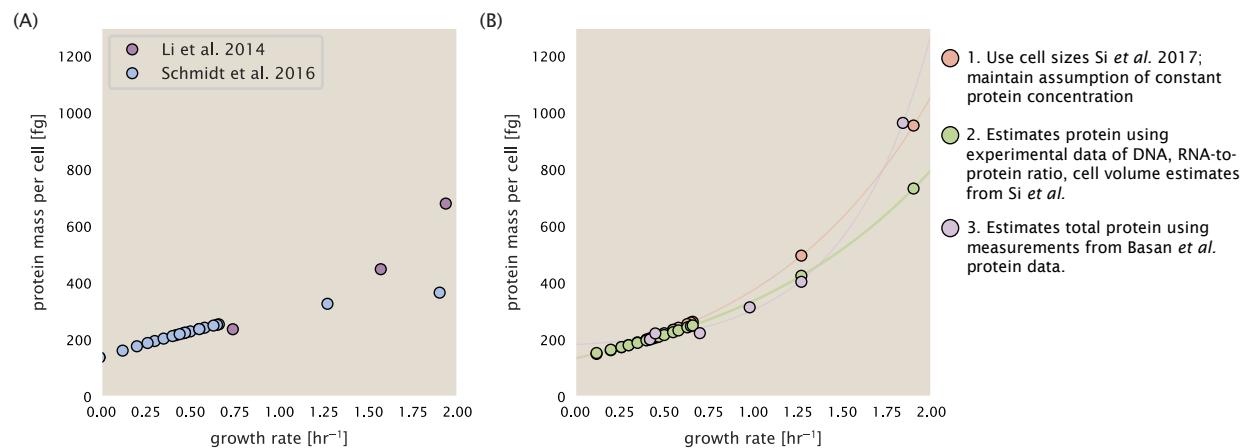


Figure 19. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt *et al.* (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. *light red*: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

their flow cytometry cell counts, they were then able to determine absolute abundance of each protein detected on a per cell basis.

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

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

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

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

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

$$M'_{P_i} = [M_p]_{Si} \cdot V_{Si_i} \quad (16)$$

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

Relaxing assumption of constant protein concentration across growth conditions

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

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

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

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

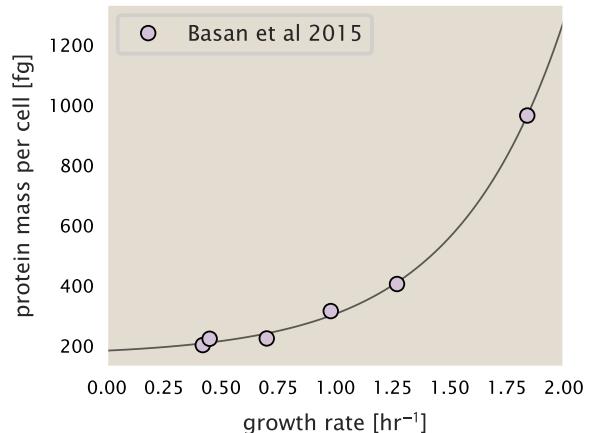


Figure 20. Total cellular protein reported in Basan et al. 2015. Measured protein mass as a function of growth rate as reproduced from Basan *et al.* 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given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

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

Calculation of Complex Abundance

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

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

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

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

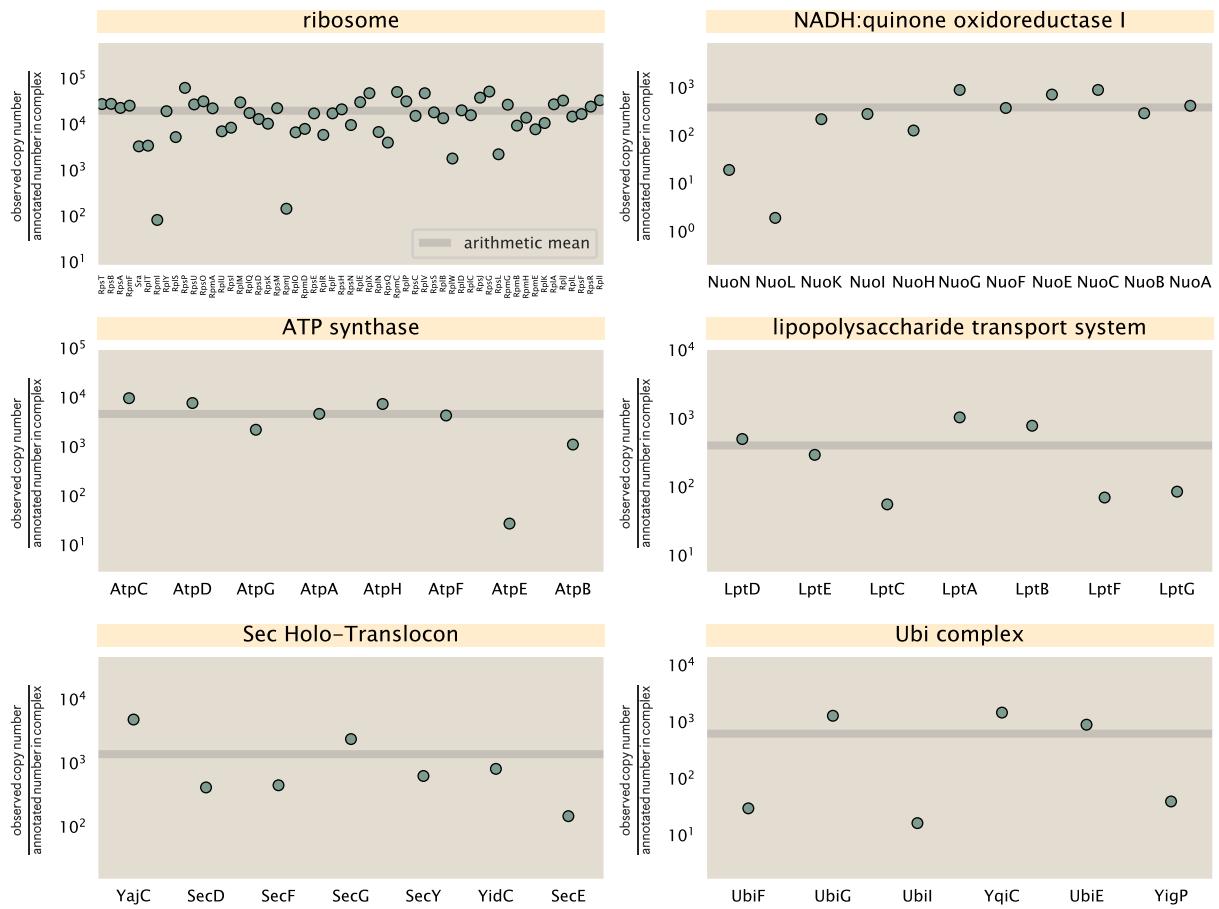


Figure 21. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds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

1245 complexes that could be formed given the measured protein copy numbers per cell,

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

1246 Here, $P_{\text{subunit } i}^{(\text{measured})}$ refers to the measured protein copy number of species i , and m refers to the number of monomers
 1247 present for that protein in the complex. For example, the 70S mature ribosome complex has 55 protein compo-
 1248 nents, all of which are present in a single copy except L4 (RplL), which is present in 4 copies ($m = 4$). For each
 1249 ribosomal protein, we then calculate the maximum number of complexes that could be formed using **Equation 17**.
 1250 This example, along with example from 5 other macromolecular complexes, can be seen in **Figure 21**.

1251 It is important to note that measurement noise, efficiency of protein extraction, and differences in protein
1252 stability will mean that the precise value of each calculation will be different for each component of a given complex.
1253 Thus, to report the total complex abundance, we use the arithmetic mean of across all subunits in the complex,

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

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

1257 Extending Estimates to a Continuum of Growth Rates

1258 In the main text, we considered a standard stopwatch of 5000 s to estimate the abundance of the various protein
1259 complexes considered. In addition to point estimates, we also showed the estimate as a function of growth rate
1260 as transparent grey curves. In this section, we elaborate on this continuum estimate, giving examples of estimates
1261 that scale with either cell volume, cell surface area, or number of origins of replication.

1262 Estimation of the total cell mass

1263 For many of the processes estimated in the main text we relied on a cellular dry mass of ≈ 300 fg from which we
1264 computed elemental and protein fractions using knowledge of fractional composition of the dry mass. At modest
1265 growth rates, such as the 5000 s doubling time used in the main text, this is a reasonable number to use as the
1266 typical cell mass is ≈ 1 pg and *E. coli* cells can approximated as 70% water by volume. However, as we have shown
1267 in the preceding sections, the cell size is highly dependent on the growth rate. This means that a dry mass of 300
1268 fg cannot be used reliably across all growth rates.

1269 Rather, using the phenomenological description of cell volume scaling exponentially with growth rate, and
1270 using a rule-of-thumb of a cell buoyant density of ≈ 1.1 pg / fL (BNID: 103875), we can calculate the cell dry mass
1271 across a range of physiological growth rates as

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

1272 where a and b are constants with units of μm^3 and hr, respectively. The value of these constants can be estimated
1273 from the careful volume measurements performed by *Si et al. (2017,?)*, as considered in Appendix Estimation of
1274 Cell Size and Surface Area earlier.

1275 Complex Abundance Scaling With Cell Volume

1276 Several of the estimates performed in the main text are implicitly dependent on the cell volume. This includes
1277 processes such as ATP utilization and, most prominently, the transport of nutrients, whose demand will be pro-
1278 portional to the volume of the cell. Of the latter, we estimated the number of transporters that would be needed
1279 to shuttle enough carbon, phosphorus, and sulfur across the membrane to build new cell mass. To do so, we
1280 used elemental composition measurements combined with a 300 fg cell dry mass to make the point estimate. As
1281 we now have a means to estimate the total cell mass as a function of volume, we can generalize these estimates
1282 across growth rates.

1283 Rather than discussing the particular details of each transport system, we will derive this scaling expression in
1284 very general terms. Consider that we wish to estimate the number of transporters for some substance X , which
1285 has been measured to be made up some fraction of the dry mass, θ_X . If we assume that, irrespective of growth
1286 rate, the cell dry mass is relatively constant (*Basan et al., 2015*) and $\approx 30\%$ of the total cell mass, we can state that
1287 the total mass of substance X as a function of growth rate is

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

1288 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in *Equation 19*. To convert this to the
1289 number of units N_X of substance X in the cell, we can use the formula weight w_X of a single unit of X in conjunction
1290 with *Equation 20*,

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

1291 To estimate the number of transporters needed, we make the approximation that loss of units of X via diffusion
1292 through porins or due to the permeability of the membrane is negligible and that a single transporter complex
1293 can transport substance X at a rate r_X . As this rate r_X is in units of X per time per transporter, we must provide
1294 a time window over which the transport process can occur. This is related to the cell doubling time τ , which can
1295 be calculated from the the growth rate λ as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized
1296 transport scaling relation of

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

1297 This function is used to draw the continuum estimates for the number of transporters seen in Figures 2 and
 1298 3 as transparent grey curves. Occasionally, this continuum scaling relationship will not precisely agree with the
 1299 point estimate outlined in the main text. This is due to the choice of ≈ 300 fg total dry mass per cell for the point
 1300 estimate, whereas we considered more precise values of cell mass in the continuum estimate. We note, however,
 1301 that both this scaling relation and the point estimates are meant to describe the order-of-magnitude observed,
 1302 and not the predict the exact values of the abundances.

1303 **Equation 22** is a very general relation for processes where the cell volume is the "natural variable" of the
 1304 problem. This means that, as the cell increases in volume, the requirements for substance X also scale with
 1305 volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the
 1306 dry mass attributable to the substance, and the formula mass of the substance is known, **Equation 22** can be used
 1307 to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell,
 1308 **Equation 22** can be slightly modified to the form

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

1309 where we have included the term N_{ATP} to account for the number of ATP equivalents needed per amino acid for
 1310 translation (≈ 4 , BNID: 114971), and w_{AA} is the average mass of an amino acid. The grey curves in Figure 4 o the
 1311 main text were made using this type of expression.

1312 A Relation for Complex Abundance Scaling With Surface Area

1313 In our estimation for the number of complexes needed for lipid synthesis and peptidoglycan maturation, we used
 1314 a particular estimate for the cell surface area ($\approx 5 \mu\text{m}$, BNID: 101792) and the fraction of dry mass attributable to
 1315 peptidoglycan ($\approx 3\%$, BNID: 101936). Both of these values come from glucose-fed *E. coli* in balance growth. As we
 1316 are interested in describing the scaling as a function of the growth rate, we must also consider how these values
 1317 scale with cell surface area, which is the natural variable for these types of processes. In the coming paragraphs,
 1318 we highlight how we incorporate a condition-dependent surface area into our calculation of the number of lipids
 1319 and murein monomers that need to be synthesized and crosslinked, respectively.

1320 Number of Lipids

1321 To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as
 1322 the surface area of a single lipid ($A_{\text{lipid}} \approx 0.5 \text{ nm}^2$, BNID: 106993) and the total fraction of the membrane composed
 1323 of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth rate. Using these approximations combined with
 1324 **Equation 13**, and recognizing that each membrane is composed of two leaflets, we can compute the number of
 1325 lipids as a function of growth rate as

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

1326 where η is the length-to-width aspect ratio and V is the cell volume.

1327 Number of Murein Monomers

1328 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we used an empirical
 1329 measurement that $\approx 3\%$ of the dry mass is attributable to peptidoglycan and that a single murien monomer is
 1330 $m_{\text{murein}} \approx 1000 \text{ Da}$. While the latter is independent of growth rate, the former is not. As the peptidoglycan exists as
 1331 a thin shell with a width of $w \approx 10 \text{ nm}$ encapsulating the cell, one would expect the number of murein monomers
 1332 scales with the surface area of this shell. In a similar spirit to our calculation of the number of lipids, the total
 1333 number of murein monomers as a function of growth rate can be calculated as

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

1334 where ρ_{pg} is the density of peptidoglycan.

1335 **Complex Abundance Scaling With Number of Origins, and rRNA Synthesis**

1336 While the majority of our estimates hinge on the total cell volume or surface area, processes related to the central
1337 dogma, namely DNA replication and synthesis of rRNA, depend on the number of chromosomes present in the
1338 cell. As discussed in the main text, the ability of *E. coli* to parallelize the replication of its chromosome by having
1339 multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived
1340 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
1341 average number of origins of replication at a given growth rate can be calculated as

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

1342 where t_{cyc} is the total time of replication and division. We can make the approximation that $t_{\text{cyc}} \approx 70$ min, which is
1343 the time it takes two replisomes to copy an entire chromosome.

1344 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of replication. Thus,
1345 at a given growth rate λ , the average dosage of rRNA operons per cell D_{rRNA} is

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

1346 This makes the approximation that *all* rRNA operons are localized around the origin. In reality, the operons
1347 are some distance away from the origin, making **Equation 27** an approximation (*Dennis et al., 2004*).

1348 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
1349 approximation, **Equation 26** illustrates that is not precisely true, even at slow growth rates. In estimating the
1350 number of RNA polymerases as a function of growth rate, we consider that regardless of the number of rRNA
1351 operons, they are all sufficiently loaded with RNA polymerase such that each operon produces one rRNA per
1352 second. Thus, the total number of RNA polymerase as a function of the growth rate can be calculated as

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

1353 where L_{operon} is the total length of an rRNA operon (≈ 4500 bp) and $\rho_{\text{RNA polymerase}}$ is packing density of RNA poly-
1354 merase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

1355 **Calculation of active ribosomal fraction.**

1356 In the main text we used the active ribosomal fraction f_a that was reported in the work of *Dai et al. (2016)* to
1357 estimate the active ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We lacked any specific model to
1358 consider how f_a should vary with growth rate, and instead find that the data is well-approximated by fitting to an
1359 exponential curve ($f_a = -0.889 e^{4.6 \cdot \lambda} + 0.922$; dashed line in inset of **Figure 10(C)**). We use this function to estimate
1360 f_a for each of the data points shown in **Figure 10(C)**.

1361 **Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.**

1362 *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
1363 protein makes up a majority of the cell's dry mass, the change in cell size is also a reflection of the changes in
1364 proteomic composition and total abundance across growth conditions. Given the potential constraints on rRNA
1365 synthesis and changes in ribosomal copy number with $\langle \#ori \rangle$, it becomes important to also consider how pro-
1366 tein copy numbers vary with the state of chromosomal replication. This is particularly true when trying to make
1367 sense of the changes in ribosomal fraction and growth-rate dependent changes in proteomic composition at a
1368 mechanistic level. As considered in the main text, it is becoming increasingly apparent that regulation through
1369 the secondary messengers (p)ppGpp may act to limit DNA replication and also reduce ribosomal activity in poorer
1370 nutrient conditions. In this context, both $\langle \#ori \rangle$, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio become important parameters
1371 to consider and keep tract of. An increase in $\langle \#ori \rangle / \langle \#ter \rangle$ ratio in particular, causes a relatively higher gene
1372 dosage in rRNA and r-protein genes due to skew in genes near the origin, where the majority of these are located

1373 In the main text we estimated the change in $\langle \#ori \rangle$ with growth rate using the nutrient-limited wild-type cell
1374 data from *Si et al. (2017)*. We consider their measurements of DNA replication time (t_C , 'C' period of cell division),
1375 total cell cycle time (t_{cyc} , 'C' + 'D' period of cell division), and doubling time τ from wild-type *E. coli* growing across
1376 a range of growth conditions. Here we show how we estimate this parameter, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio

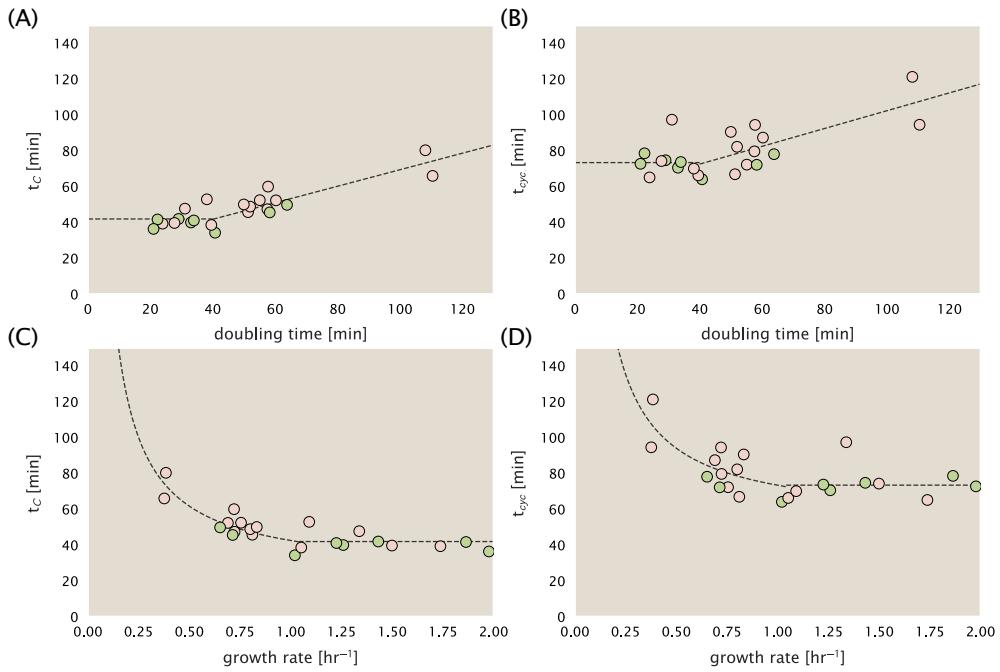


Figure 22. Estimation of $\langle \#ori \rangle / \langle \# ter \rangle$ and $\langle \#ori \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling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

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

In Figure 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.

1392 **References**

- 1393 Abelson, H., Johnson, L., Penman, S., and Green, H. (1974). Changes in RNA in relation to growth of the fibroblast: II. The lifetime
1394 of mRNA, rRNA, and tRNA in resting and growing cells. *Cell*, 1(4):161–165.
- 1395 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia*
1396 *coli*. *BMC Systems Biology*, 8(1):133.
- 1397 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by ammo-
1398 nium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1399 Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris,
1400 M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock,
1401 G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1402 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model
1403 for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding
1404 Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- 1405 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH_3 Channel.
1406 *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1407 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006).
1408 Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*,
1409 2(1):2460.
- 1410 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein
1411 synthesis. *Molecular Systems Biology*, 11(10):836.
- 1412 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*,
1413 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 1414 Belliveau, N. M., Barnes, S. L., Ireland, W. T., Jones, D. L., Sweredoski, M. J., Moradian, A., Hess, S., Kinney, J. B., and Phillips, R.
1415 (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proceedings
1416 of the National Academy of Sciences*, 115(21):E4796–E4805.
- 1417 Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009). Absolute metabolite concentra-
1418 tions and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology*, 5(8):593–599.
- 1419 Birnbaum, L. S. and Kaplan, S. (1971). Localization of a Portion of the Ribosomal RNA Genes in *Escherichia coli*. *Proceedings of the
1420 National Academy of Sciences*, 68(5):925–929.
- 1421 Booth, I. R., Mitchell, W. J., and Hamilton, W. A. (1979). Quantitative analysis of proton-linked transport systems. The lactose
1422 permease of *Escherichia coli*. *Biochemical Journal*, 182(3):687–696.
- 1423 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Expon-
1424 ential Growth Rates. *EcoSal Plus*, 3(1).
- 1425 Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews
1426 Microbiology*, 14(10):638–650.
- 1427 Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*,
1428 266:2020.06.16.154187.
- 1429 Catherwood, A. C., Lloyd, A. J., Tod, J. A., Chauhan, S., Slade, S. E., Walkowiak, G. P., Galley, N. F., Punekar, A. S., Smart, K., Rea, D.,
1430 Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan
1431 Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- 1432 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in *Escherichia coli*.
1433 *Biochemical Journal*, 117(3):551–562.
- 1434 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Transla-
1435 tional Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- 1436 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T.
1437 (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature
1438 Microbiology*, 2(2):16231.

- 1439 Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.
1440 *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1441 Delarue, M., Brittingham, G. P., Pfeffer, S., Surovtsev, I. V., Pinglay, S., Kennedy, K. J., Schaffer, M., Gutierrez, J. I., Sang, D.,
1442 Poterewicz, G., Chung, J. K., Plitzko, J. M., Groves, J. T., Jacobs-Wagner, C., Engel, B. D., and Holt, L. J. (2018). mTORC1 Con-
1443 trols Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. *Cell*, 174(2):338–349.e20.
- 1444 Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.
1445 *Microbiology and Molecular Biology Reviews*, 68(4):639–668.
- 1446 Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of
1447 Sciences*, 108(44):17876–17882.
- 1448 Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyru-
1449 vate–carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation. *Ap-
1450 plied Microbiology and Biotechnology*, 94(6):1483–1494.
- 1451 Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., and Palsson,
1452 B. Ø. (2007). A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and
1453 thermodynamic information. *Molecular Systems Biology*, 3(1):121.
- 1454 Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M., and Dunny, G. M.
1455 (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of Growth
1456 Rates. *mBio*, 11(2):45.
- 1457 Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair.
1458 *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1459 Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–263.
- 1460 Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of
1461 the National Academy of Sciences*.
- 1462 Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-
1463 Hernández, K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-Lira, H., Bonavides-Martínez,
1464 C., Pérez-Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes, A., Hernández-Koutoucheva, A., Moral-Chávez,
1465 V. D., Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB version 9.0: High-level integration of gene regulation, coexpression,
1466 motif clustering and beyond. *Nucleic Acids Research*, 44(D1):D133–D143.
- 1467 Ge, J., Yu, G., Ator, M. A., and Stubbe, J. (2003). Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class I Ribonucleotide
1468 Reductase. *Biochemistry*, 42(34):10071–10083.
- 1469 Godin, M., Delgado, F. F., Son, S., Grover, W. H., Bryan, A. K., Tzur, A., Jorgensen, P., Payer, K., Grossman, A. D., Kirschner, M. W.,
1470 and Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390.
- 1471 Goldman, S. R., Nair, N. U., Wells, C. D., Nickels, B. E., and Hochschild, A. (2015). The primary σ factor in *Escherichia coli* can access
1472 the transcription elongation complex from solution *in vivo*. *eLife*, 4:e10514.
- 1473 Guo, Y., Li, D., Zhang, S., Yang, Y., Liu, J.-J., Wang, X., Liu, C., Milkie, D. E., Moore, R. P., Tulu, U. S., Kiehart, D. P., Hu, J., Lippincott-
1474 Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolu-
1475 tion on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- 1476 Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in
1477 microbiology*, 26(10):815–832.
- 1478 Harris, R. M., Webb, D. C., Howitt, S. M., and Cox, G. B. (2001). Characterization of PitA and PitB from *Escherichia coli*. *Journal of
1479 Bacteriology*, 183(17):5008–5014.
- 1480 Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of
1481 (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- 1482 Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content
1483 of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1484 Helmstetter, C. E. and Cooper, S. (1968). DNA synthesis during the division cycle of rapidly growing *Escherichia coli* Br. *Journal
1485 of Molecular Biology*, 31(3):507–518.

- 1486 Henkel, S. G., Beek, A. T., Steinsiek, S., Stagge, S., Bettenbrock, K., de Mattos, M. J. T., Sauter, T., Sawodny, O., and Ederer, M.
1487 (2014). Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying Oxygen Conditions. *PLoS ONE*,
1488 9(9):e107640.
- 1489 Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative
1490 proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2):e784-
1491 e784.
- 1492 Ingledew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*, 48(3):222-271.
- 1493 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B., and Phillips, R. (2020).
1494 Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a time. *bioRxiv*.
- 1495 Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3(3):318-
1496 356.
- 1497 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*,
1498 20(17):4952-4963.
- 1499 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the
1500 future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1501 Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonja, E., Mekler, V., Ebright, R. H.,
1502 and Weiss, S. (2005). Retention of Transcription Initiation Factor Σ 70 in Transcription Elongation: Single-Molecule Analysis.
1503 *Molecular Cell*, 20(3):347-356.
- 1504 Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E.,
1505 Ong, Q., Ong, W. K., Paley, S. M., and Subhraveti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways.
1506 *Briefings in Bioinformatics*, 20(4):1085-1093.
- 1507 Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., Assad-Garcia, N., Glass, J. I., and Covert, M. W.
1508 (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389-401.
- 1509 Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S.,
1510 Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-
1511 Ramírez, D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge
1512 about *Escherichia coli*K-12. *Nucleic Acids Research*, 45(D1):D543-D550.
- 1513 Khademi, S., O'Connell, J., Remis, J., Robles-Colmenares, Y., Miercke, L. J. W., and Stroud, R. M. (2004). Mechanism of Ammonia
1514 Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*, 305(5690):1587-1594.
- 1515 Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that enables the use of
1516 hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of Sciences*, 114(33):E6922-E6931.
- 1517 Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition.
1518 *Current Opinion in Biotechnology*, 28:96-102.
- 1519 Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*,
1520 139(7):1366-1375.
- 1521 Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*,
1522 125(2):028103.
- 1523 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation in *E. coli*by
1524 Modulating Supercoiling of oriC. *mBio*, 10(4):822.
- 1525 Lascur, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and Biomem-
1526 branes*, 32(3):237-246.
- 1527 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical
1528 investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1529 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting sets. *IEEE Transactions
1530 on Visualization and Computer Graphics*, 20(12):1983-1992.
- 1531 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles
1532 underlying allocation of cellular resources. *Cell*, 157(3):624-635.

- 1533 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in
1534 cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1535 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon
1536 Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- 1537 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facili-
1538 tator GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1539 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*,
1540 112(51):15690–15695.
- 1541 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On*. Gene Expression.
1542 Springer.
- 1543 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed
1544 by condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1545 Mikucki, J. A., Pearson, A., Johnston, D. T., Turchyn, A. V., Farquhar, J., Schrag, D. P., Anbar, A. D., Priscu, J. C., and Lee, P. A. (2009).
1546 A Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.
- 1547 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular
1548 and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1549 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation.
1550 *Growth Symposium*, 9:223–289.
- 1551 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1552 Mooney, R. A., Darst, S. A., and Landick, R. (2005). Sigma and RNA Polymerase: An On-Again, Off-Again Relationship? *Molecular
1553 Cell*, 20(3):335–345.
- 1554 Mooney, R. A. and Landick, R. (2003). Tethering $\Sigma 70$ to RNA polymerase reveals high *in vivo* activity of σ factors and $\Sigma 70$ -
1555 dependent pausing at promoter-distal locations. *Genes & Development*, 17(22):2839–2851.
- 1556 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall
1557 synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–
1558 12515.
- 1559 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 1560 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria.
1561 *eLife*, 8:642.
- 1562 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- 1563 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with
1564 increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 1565 Perdue, S. A. and Roberts, J. W. (2011). σ^{70} -dependent Transcription Pausing in *Escherichia coli*. *Journal of Molecular Biology*,
1566 412(5):782–792.
- 1567 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1568 Ramos, S. and Kaback, H. R. (1977). The relation between the electrochemical proton gradient and active transport in *Escherichia
1569 coli* membrane vesicles. *Biochemistry*, 16(5):854–859.
- 1570 Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorrodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., and Maranas, C. D. (2012). An
1571 integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering*,
1572 14(6):687–704.
- 1573 Rogers, H., Perkins, H., and Ward, J. (1980). *Microbial Cell Walls and Membranes*. Chapman and Hall, London.
- 1574 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproduc-
1575 tive strategies. *Nature microbiology*, 1(11):1–7.
- 1576 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of
1577 Bacteriology*, 131(2):505–511.

- 1578 Rudd, S. G., Valerie, N. C. K., and Helleday, T. (2016). Pathways controlling dNTP pools to maintain genome stability. *DNA Repair*,
1579 44:193–204.
- 1580 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–
1581 11734.
- 1582 Sánchez-Romero, M. A., Molina, F., and Jiménez-Sánchez, A. (2011). Organization of ribonucleoside diphosphate reductase
1583 during multifork chromosome replication in *Escherichia coli*. *Microbiology*, 157(8):2220–2225.
- 1584 Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical
1585 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- 1586 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann,
1587 M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 1588 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression:
1589 origins and consequences. *Science*, 330(6007):1099–1102.
- 1590 Sekowska, A., Kung, H.-F., and Danchin, A. (2000). Sulfur Metabolism in *Escherichia coli* and Related Bacteria: Facts and Fiction.
1591 *Journal of Molecular Microbiology and Biotechnology*, 2(2):34.
- 1592 Shi, H., Bratton, B. P., Gitai, Z., and Huang, K. C. (2018). How to Build a Bacterial Cell: MreB as the Foreman of *E. coli* Construction.
1593 *Cell*, 172(6):1294–1305.
- 1594 Si, F., Le Treut, G., Sauls, J. T., Vadia, S., Levin, P. A., and Jun, S. (2019). Mechanistic Origin of Cell-Size Control and Homeostasis in
1595 Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 1596 Si, F., Li, D., Cox, S. E., Sauls, J. T., Azizi, O., Sou, C., Schwartz, A. B., Erickstad, M. J., Jun, Y., Li, X., and Jun, S. (2017). Invariance of
1597 Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- 1598 Sirko, A., Zatyka, M., Sadowy, E., and Hulanicka, D. (1995). Sulfate and thiosulfate transport in *Escherichia coli* K-12: Evidence for
1599 a functional overlapping of sulfate- and thiosulfate-binding proteins. *Journal of Bacteriology*, 177(14):4134–4136.
- 1600 Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology
1601 Reviews*, 40(1):133–159.
- 1602 Soler-Bistué, A., Aguilar-Pierlé, S., García-Garcera, M., Val, M.-E., Sismeiro, O., Varet, H., Sieira, R., Krin, E., Skovgaard, O., Comerci,
1603 D. J., Eduardo P. C. Rocha, and Mazel, D. (2020). Macromolecular crowding links ribosomal protein gene dosage to growth
1604 rate in *Vibrio cholerae*. *BMC Biology*, 18(1):1–18.
- 1605 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth
1606 and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1607 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*,
1608 19.
- 1609 Stevenson, B. S. and Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in *Escherichia coli*. *Applied and
1610 Environmental Microbiology*, 70(11):6670–6677.
- 1611 Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van
1612 Leeuwenhoek*, 39(1):545–565.
- 1613 Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*.
1614 *Archives of Microbiology*, 113(3):185–189.
- 1615 Svenningsen, S. L., Kongstad, M., Stenum, T. S. n., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017). Transfer RNA is highly
1616 unstable during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Research*, 45(2):793–804.
- 1617 Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1618 Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 1619 Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Cell-size control
1620 and homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 1621 Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A., and Xie, X. S. (2010). Quantifying *E. coli* proteome and
1622 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538.

- 1623 Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of
1624 protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- 1625 Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic
1626 models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*,
1627 107(2):369–381. _eprint: <https://onlinelibrary.wiley.com/doi/10.1002/bit.22802>.
- 1628 The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*,
1629 47(D1):D330–D338.
- 1630 Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and
1631 translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 1632 van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen Assimilation in *Escherichia coli*: Putting Molecular
1633 Data into a Systems Perspective. *Microbiology and Molecular Biology Reviews*, 77(4):628–695.
- 1634 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,
1635 Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson,
1636 E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero,
1637 E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0:
1638 Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1639 Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate
1640 Data Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.
- 1641 Vollmer, W., Blanot, D., and De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*,
1642 32(2):149–167.
- 1643 Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1FO-ATP synthase. *FEBS Letters*, 545(1):61–70.
- 1644 Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
1645 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- 1646 You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C. W., Wang, Y.-P., Lenz, P., Yan, D., and Hwa, T. (2013). Coordination
1647 of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- 1648 Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from
1649 *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 108(46):18643–18648.
- 1650 Zhang, L., Jiang, W., Nan, J., Almqvist, J., and Huang, Y. (2014a). The *Escherichia coli* CysZ is a pH dependent sulfate transporter
1651 that can be inhibited by sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(7):1809–1816.
- 1652 Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014b). Regulation of crp gene expression by the catabolite repressor/activator, cra,
1653 in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1654 Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in
1655 *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.