

Fundamental limits on the rate of bacterial cell division

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

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

14

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

25

Introduction

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

As noted by Jacques Monod, “the study of the growth of bacterial cultures does not constitute a specialized subject or branch of research, it is the basic method of Microbiology.” Those words ring as true today as they did when they were written 70 years ago (*Monod, 1949*) and the study of bacterial growth has recently undergone a renaissance. Many of the key questions addressed by the pioneering efforts in the middle of the last century can be revisited by examining them through the lens of the increasingly refined molecular census that is available for bacteria such as the microbial workhorse *E. coli*.

Several of the evergreen questions about bacterial growth that were originally raised by microbiologists in

42 the middle of the 20th century can now be reframed in light of this newly available data. For example, what
43 biological processes are the primary determinants for how quickly bacterial cells can grow and reproduce? How
44 do cells alter the absolute numbers and relative ratios of their molecular constituents as a function of changes in
45 growth rate or nutrient availability? In this paper, we address these two questions from two distinct angles. First,
46 as a result of an array of high-quality proteome-wide measurements of *E. coli* under diverse growth conditions,
47 we have a census that allows us to explore how the number of key molecular players change as a function of
48 growth rate. Here, we have compiled a combination of data sets collected over the past decade using either mass
49 spectrometry (*Schmidt et al., 2016; Peebo et al., 2015; Valgepea et al., 2013*) or ribosomal profiling (*Li et al., 2014*)
50 of *E. coli* across 36 unique growth rates (see Appendix Experimental Details Behind Proteomic Data for further
51 discussion of the data). Second, by compiling molecular turnover rate measurements for many of the fundamental
52 processes associated with bacterial growth, we make quantitative estimates of key cellular processes (schematized
53 in **Figure 1**) to determine whether our current understanding of the dynamics of these processes are sufficient to
54 explain the magnitude of the observed protein copy numbers across conditions. The census, combined with these
55 estimates, provide a window into the question of whether the rates of central processes such as energy generation
56 or DNA synthesis are regulated systematically as a function of cell growth rate by altering protein copy number.

57 Throughout this work, we consider an archetypal growth rate of $\approx 0.5 \text{ hr}^{-1}$ corresponding to a doubling time of
58 ≈ 5000 seconds, as the the data sets heavily sample this regime. While we formulate point estimates for the protein
59 abundances at this division time, we also consider how these values will vary at other growth rates due to changes
60 in cell volume, surface area, and chromosome copy number (*Taheri-Araghi et al., 2015*). Broadly, we find that for
61 the majority of these estimates, the protein copy numbers appear tuned for the task of cell doubling at a given
62 growth rate. Thus, our understanding of the kinetics of myriad biological processes is sufficient to quantitatively
63 explain the observed abundances of these proteins.

64 From these estimates, it emerges that translation, particularly the synthesis of ribosomal proteins, is a plausible
65 candidate that limits the rate of cell division. We reach this conclusion by considering that ribosome synthesis is
66 1) a rate limiting step for the *fastest* bacterial division, and 2) a major determinant of bacterial growth across the
67 nutrient conditions we have considered under steady state, exponential growth. This enables us to suggest that the
68 long-observed correlation between growth rate and cell size (*Schaechter et al., 1958; Si et al., 2017*) can be simply
69 attributed to the increased absolute number of ribosomes per cell under conditions supporting extremely rapid
70 growth. To better understand how the observed alterations in absolute protein abundances, and in particular,
71 changes in ribosome copy number, influence growth rate across different nutrient conditions we consider a
72 minimal model of cellular growth. Our conclusions from these analyses provide important insight into how *E.*
73 *coli* optimizes growth across conditions of differing nutrient availability and identifies fundamental constraints in
74 bacterial growth more broadly.

114 Nutrient Transport

115 We begin by considering the critical transport processes diagrammed in **Figure 1(A)**. In order to build new cellular
116 mass, the molecular and elemental building blocks must be scavenged from the environment in different forms.
117 Carbon, for example, is acquired via the transport of carbohydrates and sugar alcohols with some carbon sources
118 receiving preferential treatment in their consumption (*Monod, 1947*). Phosphorus, sulfur, and nitrogen, on the
119 other hand, are harvested primarily in the forms of inorganic salts, namely phosphate, sulfate, and ammonia (*Jun*
120 *et al., 2018; Assentoft et al., 2016; Stasi et al., 2019; Antonenko et al., 1997; Rosenberg et al., 1977; Willsky et al.,*
121 *1973*). All of these compounds have different permeabilities across the cell membrane (*Phillips, 2018*) and most
122 require some energetic investment either via ATP hydrolysis or through the proton electrochemical gradient to
123 bring the material across the hydrophobic cell membrane.

124 The elemental composition of *E. coli* has received much quantitative attention over the past half century
125 (*Neidhardt et al., 1991; Taymaz-Nikerel et al., 2010; Heldal et al., 1985; Bauer and Ziv, 1976*), providing us with
126 a starting point for estimating how many atoms of each element must be scavenged from the environment. A
127 synthesis of these studies presents an approximate dry mass composition of $\approx 45\%$ carbon (BioNumber ID: 100649,
128 see **Box 1**), $\approx 15\%$ nitrogen (BNID: 106666), $\approx 3\%$ phosphorus (BNID: 100653), and 1% sulfur (BNID: 100655) with
129 remainder being attributable to oxygen, hydrogen, and various metals. We use this stoichiometric breakdown to
130 estimate the abundance and growth rate dependence of a variety of transporters responsible for carbon uptake,

76

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

77

This work relies heavily on "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.

78

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.

79

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.

80

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.

81

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

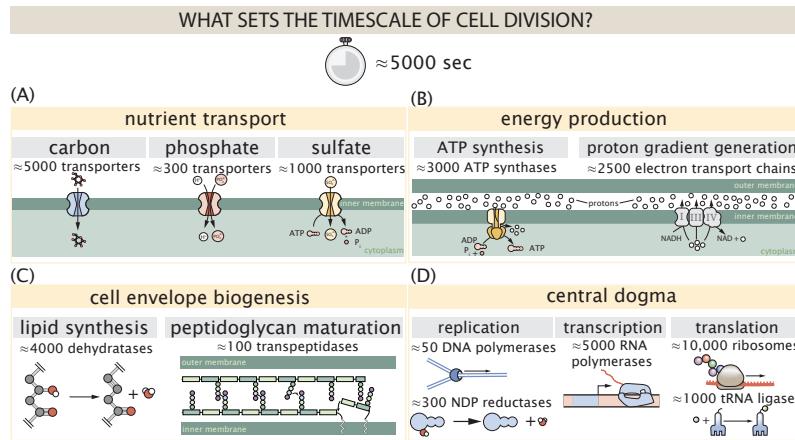


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP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

131 and provide more extensive investigation of the other critical elements – phosphorus, sulfur, and nitrogen – in the
 132 Appendix ??.

133 Using $\approx 0.3 \text{ pg}$ as the typical *E. coli* dry mass (BNID: 103904) at a growth rate of $\approx 0.5 \text{ hr}^{-1}$ coupled with an
 134 approximation that $\approx 50\%$ of this is carbon, we estimate that $\sim 10^{10}$ carbon atoms must be brought into the cell
 135 in order to double all of the carbon-containing molecules (*Figure 2(A, top)*). Typical laboratory growth conditions
 136 provide carbon as a single class of sugar (such as glucose, galactose, or xylose) for which *E. coli* has evolved myriad
 137 mechanisms by which they can be transported across the cell membrane. One such mechanism of transport is
 138 via the PTS system which is a highly modular system capable of transporting a diverse range of sugars (*Escalante et al., 2012*). The glucose-specific component of this system transports ≈ 200 glucose molecules (≈ 1200 carbon
 139 atoms) per second per transporter (BNID: 114686). Making the assumption that this is a typical sugar transport
 140 rate, coupled with the need to transport $\sim 10^{10}$ carbon atoms, we then expect on the order of ≈ 1000 transporters
 141 must be expressed in order to bring in enough carbon atoms, diagrammed in the top panel of *Figure 2(A)*.
 142

143 As revealed in *Figure 2(A)*, experimental measurements exceed the estimate by several fold, implying that
 144 the cell is capable of transporting more carbon atoms than needed. While we estimate ≈ 1000 transporters are
 145 needed with a 5000 s division time, we can abstract this calculation to consider any particular growth rate given
 146 knowledge of the cell density and volume as a function of growth rate and direct the reader to the Appendix
 147 Extending Estimates to a Continuum of Growth Rates for more information. This abstraction, shown as a grey line
 148 in *Figure 2(A)*, reveals an excess of transporters even at faster growth rates. This contrasts with our observations
 149 for uptake of phosphorus and sulfur, which align well with our expectations across different growth conditions
 150 (*Figure 2-Figure Supplement 1* and discussed further in Appendix ??)

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

161 Many metabolic operons are regulated with dual-input logic gates that are only expressed when glucose

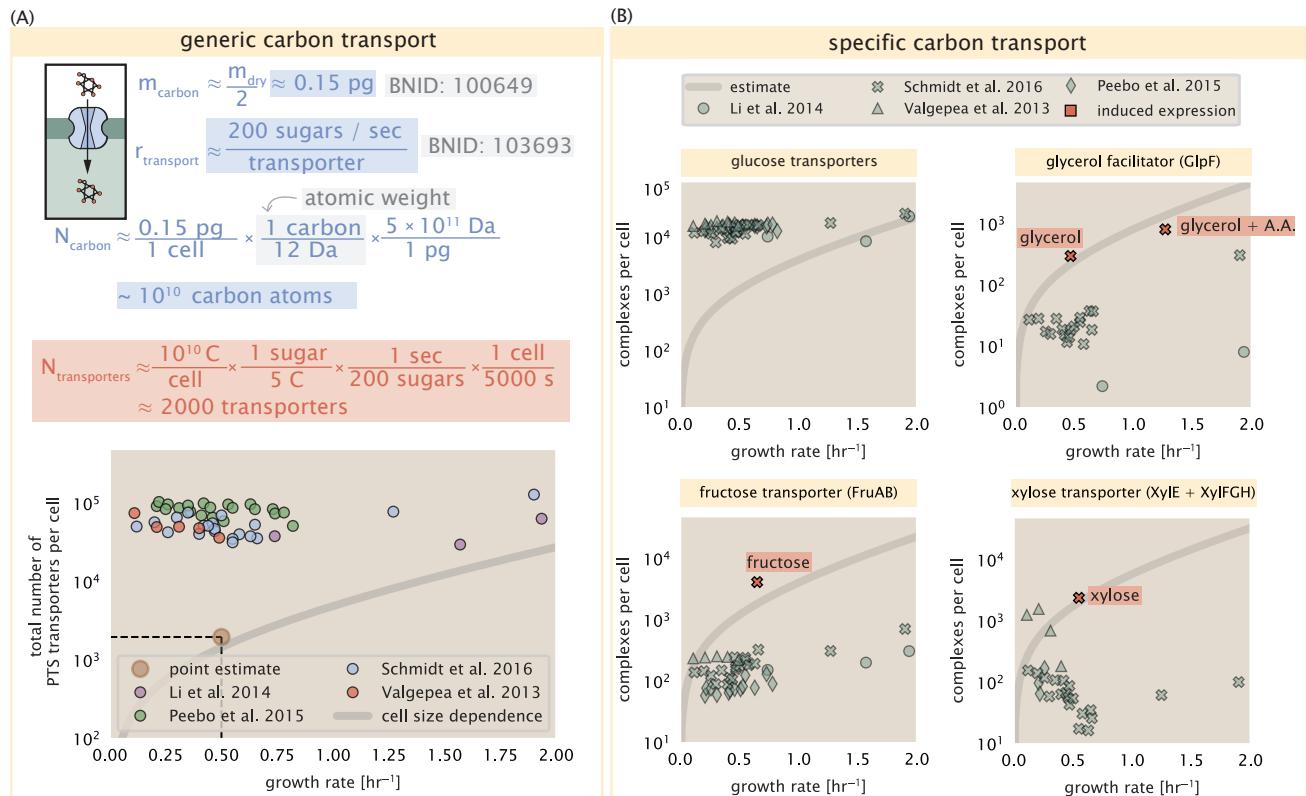


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules at a rate of ≈ 200 per second. Bottom plot shows the estimated number of transporters needed at a growth rate of ≈ 0.5 per hr (light-brown point and dashed lines). Colored points correspond to the mean number of complexes involved in carbohydrate import (complexes annotated with the Gene Ontology terms GO:0009401 and GO:0098704) for different growth conditions across different published datasets. (B) The abundance of various specific carbon transport systems plotted as a function of the population growth rate. The rates of substrate transport to compute the continuum growth rate estimate (grey line) were 200 glucose- s^{-1} (BNID: 103693), 2000 glycerol- s^{-1} (Li et al., 2003), 200 fructose- s^{-1} (assumed to be similar to PtsI, BNID: 103693), and 50 xylose- s^{-1} (assumed to be comparable to LacY, BNID: 103159). Red points and highlighted text indicate conditions in which the only source of carbon in the growth medium induces expression of the transport system. Grey line in (A) and (B) represents the estimated number of transporters per cell at a continuum of growth rates.

Figure 2-Figure supplement 1. Estimates and observed abundances of phosphate and sulfate transporters.

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., 2014; Gama-Castro et al., 2016; Belliveau et al., 2018; Ireland et al., 2020*). Points colored in red in *Figure 2(B)* (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. The low but non-zero abundances may reflect the specific regulatory logic involved, requiring that cells are able to transport some minimal amount of an alternative carbon source in order to induce expression of these alternative carbon-source systems.

172 Limits on Transporter Expression

If acquisition of nutrients was a limiting process in cell division under the typical growth conditions explored here, could expression simply be increased to accommodate faster growth? A way to approach this question is to compute the amount of space in the bacterial membrane that could be occupied by nutrient transporters. Considering a rule-of-thumb for the surface area of *E. coli* of about $5 \mu\text{m}^2$ (BNID: 101792), we expect an areal density for 1000 transporters to be approximately 200 transporters/ μm^2 . For a typical transporter occupying about $50 \text{ nm}^2/\text{dimer}$, this amounts to about only $\approx 1\%$ of the total inner membrane area (*Szenk et al., 2017*). Additionally, bacterial cell membranes typically have densities of 10^5 proteins/ μm^2 (*Phillips, 2018*), implying that the cell could accommodate more transporters if any one species was limiting. As we will see, processes like cell wall synthesis and energy production also require space on the membrane and this places additional limitations on cell size and surface area that we will consider further in the coming sections.

183 Cell Envelope Biogenesis

We next turn to the processes that build and maintain the cell envelope. In contrast to nutrient transporters, which support the synthesis of biomolecules throughout the cell and therefore need to scale with the cell size, here we must consider the synthesis of components that will need to scale with the surface area of the cell, namely the synthesis of lipids and peptidoglycan. *E. coli* is a rod-shaped bacterium with a remarkably robust length-to-width aspect ratio of $\approx 4:1$ (*Harris and Theriot, 2018; Ojkic et al., 2019*). At modest growth rates, the total cell surface area is $\approx 5 \mu\text{m}^2$ (BNID: 101792, *Milo et al. (2010)*). Assuming this surface area is approximately the same between the inner and outer membranes of *E. coli*, and the fact that each membrane is itself a lipid bilayer, cells have a total membrane surface area of $\approx 20 \mu\text{m}^2$. In Appendix ?? we describe the calculation of cell surface area as a function of cell size.

193 Lipid Synthesis

The dense packing of the membrane with proteins means that the cell membranes are not composed entirely of lipid molecules, with 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 can estimate $\sim 2 \times 10^7$ lipids per cell, which is 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*). 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*). Thus, given this rate and the need to synthesize $\approx 2 \times 10^7$ lipids over 5000 seconds, one can estimate that a typical cell requires ≈ 4000 ACP dehydratases. This is in reasonable agreement with the experimentally observed copy numbers of FabZ and FabA (*Figure 3(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 3(A)*), which captures the observed growth rate dependent expression of these two enzymes.

210 **Peptidoglycan Synthesis**

211 Bacterial cells demonstrate exquisite control over their cell shape. This is primarily due to the cell wall, a stiff, several
212 nanometer thick meshwork of polymerized disaccharides. The formation of the peptidoglycan is an intricate
213 process involving many macromolecular players (*Shi et al., 2018; Morgenstein et al., 2015*), whose coordinated
214 action maintains cell shape and integrity even in the face of large-scale perturbations (*Harris and Theriot, 2018;*
215 *Shi et al., 2018*). The peptidoglycan alone comprises $\approx 3\%$ of the cellular dry mass (BNID: 101936), making it
216 the most massive molecule in *E. coli*. The polymerized unit of the peptidoglycan is a N-acetylglucosamine and
217 N-acetylmuramic acid disaccharide, of which the former is functionalized with a short pentapeptide. With a mass
218 of ≈ 1000 Da, this unit, which we refer to as a murein monomer, it is polymerized to form long strands in the
219 periplasm which are then attached to each other via their peptide linkers. Together, these quantities provide an
220 estimate of $\approx 5 \times 10^6$ murein monomers per cell.

221 The crosslinking of the pentapeptides between adjacent glycan strands is responsible for maintaining the
222 structural integrity of the cell wall and, in principle, each murein monomer can be involved in such a crosslink.
223 In some microbes, such as in gram-positive bacterium *Staphylococcus aureus*, the extent of crosslinking can be
224 large with $> 90\%$ of pentapeptides forming a connection between glycan strands. In *E. coli*, however, a much
225 smaller proportion ($\approx 20\%$) of the peptides are crosslinked, resulting in a weaker and more porous cell wall *Vollmer*
226 *et al. (2008); Rogers et al. (1980)*. The formation of these crosslinks occurs primarily during the polymerization
227 of the murein monomers and is facilitated by a family of enzymes called transpeptidases. The four primary
228 transpeptidases of *E. coli* have only recently been quantitatively characterized *in vivo* via liquid chromatography
229 mass spectrometry (*Catherwood et al., 2020*), which revealed a notably slow kinetic turnover rate of ≈ 2 crosslinking
230 reactions formed per second per enzyme.

231 Assembling these quantities permits us to make an estimate that on the order of ≈ 100 transpeptidases per cell
232 are needed for complete maturation of the peptidoglycan, given a division time of ≈ 5000 seconds; a value that
233 is comparable to experimental observations (*Figure 3(B)*). Expanding this estimate to account for the changing
234 volume of the peptidoglycan as a function of growth rate (grey line in *Figure 3(B)*) also qualitatively captures the
235 observed dependence in the data, though systematic disagreements between the different data sets makes the
236 comparison more difficult.

237 **Limits on Cell Wall Biogenesis**

238 While the processes we have considered represent only a small portion of proteins devoted to cell wall biogenesis,
239 we find it unlikely that they limit cellular growth in general. The relative amount of mass required for lipid and
240 peptidoglycan components decrease at faster growth rates due to a decrease in their surface area to volume (S/V)
241 ratio (*Ojkic et al., 2019*). In addition, despite the slow catalytic rate of FabZ and FabA in lipid synthesis, experimental
242 data and recent computational model has shown that the rate of fatty-acid synthesis can be drastically increased
243 by increasing the concentration of FabZ (*Yu et al., 2011; Ruppe and Fox, 2018*). With a proteome size of $\approx 3 \times 10^6$
244 proteins, a hypothetical 10-fold increase in expression from 4000 to 40,000 ACP dehydratases would result in a \approx
245 1% increase in the size of the proteome.

246 **Energy Production**

247 Cells consume and generate energy predominantly in the form of NTPs in order to grow. The high-energy
248 phosphodiester bonds of (primarily) ATP power a variety of cellular processes that drive biological systems away
249 from thermodynamic equilibrium. We next turn to the synthesis of ATP as a potential process that may limit growth,
250 which also requires us to consider the maintenance of the electrochemical proton gradient which powers it.

251 **ATP Synthesis**

252 Hydrolysis of the terminal phosphodiester bond of ATP into ADP (or alternatively GTP and GDP) and an inorganic
253 phosphate provides the kinetic driving force in a wide array of biochemical reactions. One such reaction is the
254 formation of peptide bonds during translation, which requires ≈ 2 ATPs for the charging of an amino acid to
255 the tRNA and ≈ 2 GTP for the formation of each peptide bond. Assuming the ATP costs associated with error
256 correction and post-translational modifications of proteins are negligible, we can make the approximation that each
257 peptide bond has a net cost of ≈ 4 ATP (BNID: 101442, *Milo et al. (2010)*). Formation of GTP from ATP is achieved

CELL ENVELOPE BIOSYNTHESIS

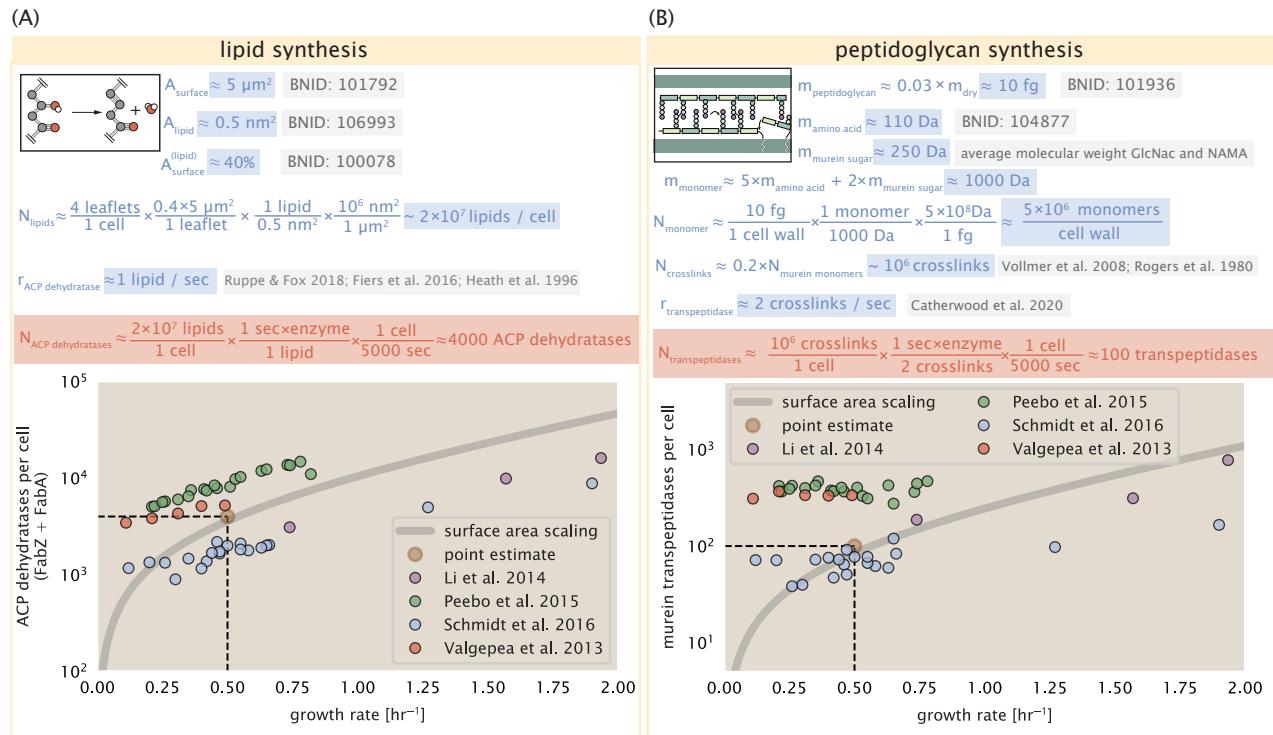


Figure 3. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred from experimental measurements via a stochastic kinetic model described in [Ruppe and Fox \(2018\)](#). Bottom panel shows the experimentally observed complex copy numbers using the stoichiometries $[\text{FabA}]_2$ and $[\text{FabZ}]_2$. (B) An estimate for the number of peptidoglycan transpeptidases needed to complete maturation of the peptidoglycan. The mass of the murein monomer was estimated by approximating each amino acid in the pentapeptide chain as having a mass of 110 Da and each sugar in the disaccharide having a mass of ≈ 250 Da. The *in vivo* rate of transpeptidation $r_{\text{E. coli}}$ 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.

258 via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy investment
259 (*Lascu and Gonin, 2000*) and therefore consider all NTP requirements of the cell to be functionally equivalent to
260 being exclusively ATP. In total, the energetic costs of peptide bond formation consume $\approx 80\%$ of the cells ATP
261 budget (BNID: 107782; 106158; 101637; 111918, *Lynch and Marinov (2015)*; *Stouthamer (1973)*). The pool of ATP is
262 produced by the F₁-F₀ ATP synthase – a membrane-bound rotary motor which under ideal conditions can yield \approx
263 300 ATP per second (BNID: 114701; *Weber and Senior (2003)*).

264 To estimate the total number of ATP equivalents consumed during a cell cycle, we will make the approximation
265 that there are $\approx 3 \times 10^6$ proteins per cell with an average protein length of ≈ 300 peptide bonds (BNID: 115702;
266 108986; 104877). Taking these values together, we find that the typical *E. coli* cell consumes $\sim 5 \times 10^9$ ATP per
267 cell cycle on protein synthesis alone. Assuming that each ATP synthases operates at its maximal speed (300 ATP
268 per second per synthase), ≈ 3000 ATP synthases are needed to keep up with the energy demands of the cell.
269 This estimate and a is comparable with the experimental observations, shown in *Figure 4* (A). We note that this
270 estimate assumes all ATP is synthesized via ATP synthase and neglects synthesis via fermentative metabolism.
271 This assumption may explain why at the fastest growth rates ($\approx 2 \text{ hr}^{-1}$), our continuum estimate predicts more
272 synthase than is experimentally observed (gray line in *Figure 4*). At rapid growth rates, *E. coli* enters a type of
273 overflow metabolism where fermentative metabolism becomes pronounced *Szenk et al. (2017)*.

274 Generating the Proton Electrochemical Gradient

275 In order to produce ATP, the F₁-F₀ ATP synthase itself must consume energy. Rather than burning through its
276 own product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the
277 electrochemical potential established across the inner membrane through cellular respiration. This electrochemical
278 gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains as
279 they reduce NADH. In *E. coli*, this potential difference is ≈ -200 mV (BNID: 102120). A simple estimate of the inner
280 membrane as a capacitor with a working voltage of -200 mV reveals that $\approx 2 \times 10^4$ protons must be present in the
281 intermembrane space. However, the constant rotation of the ATP synthases would abolish this potential difference
282 in a few milliseconds if it were not being actively maintained.

283 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of intense biochemical
284 and biophysical study (*Ingledeew and Poole, 1984*; *Khademian and Imlay, 2017*; *Cox et al., 1970*; *Henkel et al., 2014*).
285 A recent work (*Szenk et al., 2017*) examined the respiratory capacity of the *E. coli* electron transport complexes
286 using structural and biochemical data, revealing that each electron transport chain rapidly pumps protons into the
287 intermembrane space at a rate of ≈ 1500 protons per second (BNID: 114704; 114687). Using our estimate of the
288 number of ATP synthases required per cell [*Figure 4(A)*], coupled with these recent measurements, we estimate
289 that ≈ 1000 electron transport complexes would be necessary to facilitate the $\sim 5 \times 10^6$ protons per second diet of
290 the cellular ATP synthases. This estimate is in agreement with the number of complexes identified in the proteomic
291 datasets (plot in *Figure 4(B)*). This suggests that every ATP synthase must be accompanied by ≈ 1 functional electron
292 transport chain.

293 Biosynthesis in a Crowded Membrane

294 Our estimates thus far have focused on biochemistry at the periphery of the cell and have generally been
295 concordant with the abundances predicted by our estimates. However, as surface area and cell volume do not
296 scale identically, it is worth considering what physical limits may be imposed for transport or energy production
297 given the S/V ratio, which as we've noted will decrease at faster growth rates.

298 In our estimate of ATP production above we found that a cell demands about 5×10^9 ATP per cell cycle or 10^6
299 ATP/s. With a cell volume of roughly 1 fL (BNID: 100004), this corresponds to about 2×10^{10} ATP per fL of cell volume,
300 in line with previous estimates (*Stouthamer and Bettenhausen, 1977*; *Szenk et al., 2017*). In *Figure 5* (A) we plot
301 this ATP demand as a function of the S/V ratio in green, where we have considered a range of cell shapes from
302 spherical to rod-shaped with an aspect ratio (length/width) equal to 4. In order to consider the maximum ATP that
303 could be produced, we consider the amount of ATP that can be generated by a membrane filled with ATP synthase
304 and electron transport complexes, which provides a maximal production of about 3 ATP / (nm²·s) (*Szenk et al.,
305 2017*). This is shown in blue in *Figure 5(A)*, which shows that at least for the growth rates observed (right column
306 in plot), the energy demand is roughly an order of magnitude less. Interestingly, *Szenk et al. (2017)* also found

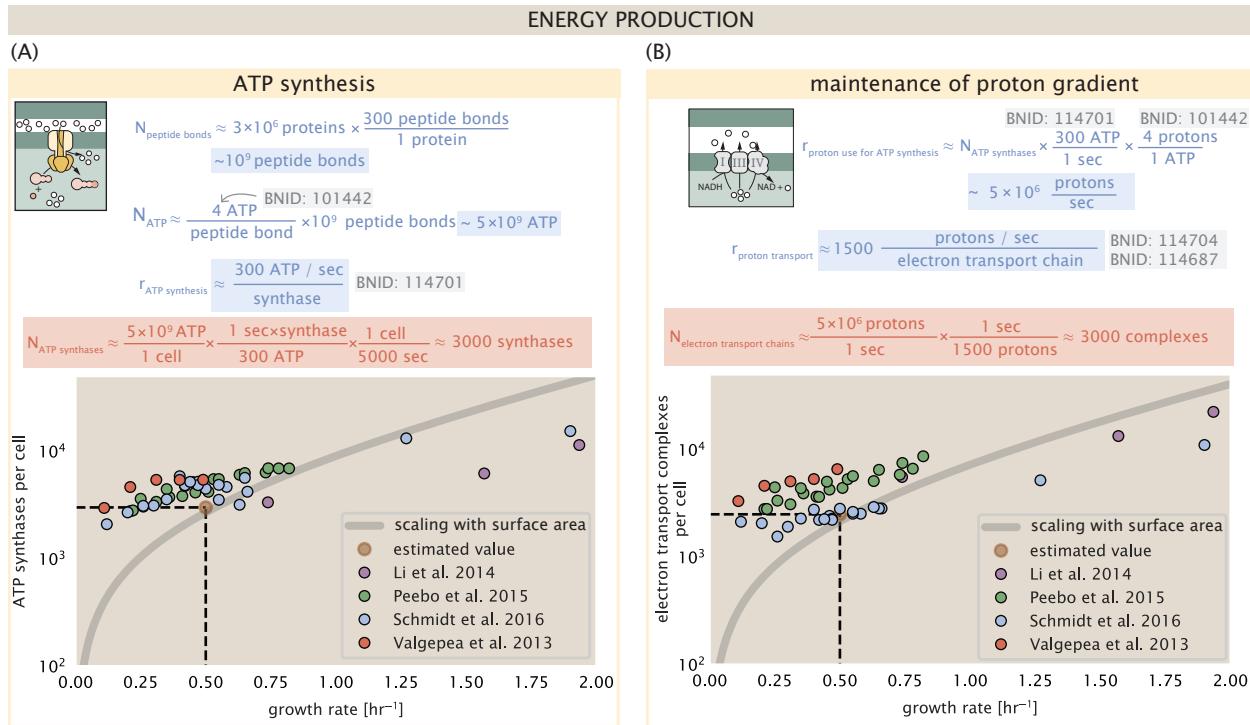


Figure 4. The abundance of F_1 - F_0 ATP synthases and electron transport chain complexes as a function of growth rate.

(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 oxioreductase I ($[NuoA][NuoH][NuoJ][NuoK][NuoL][NuoM][NuoN][NuoB][NuoC][NuoE][NuoF][NuoG][NuoI]$) and II ($[Ndh]$). Grey lines in both (A) and (B) correspond to the estimate procedure described, but applied to a continuum of growth rates. We direct the reader to the Supporting Information for a more thorough description of this approach.

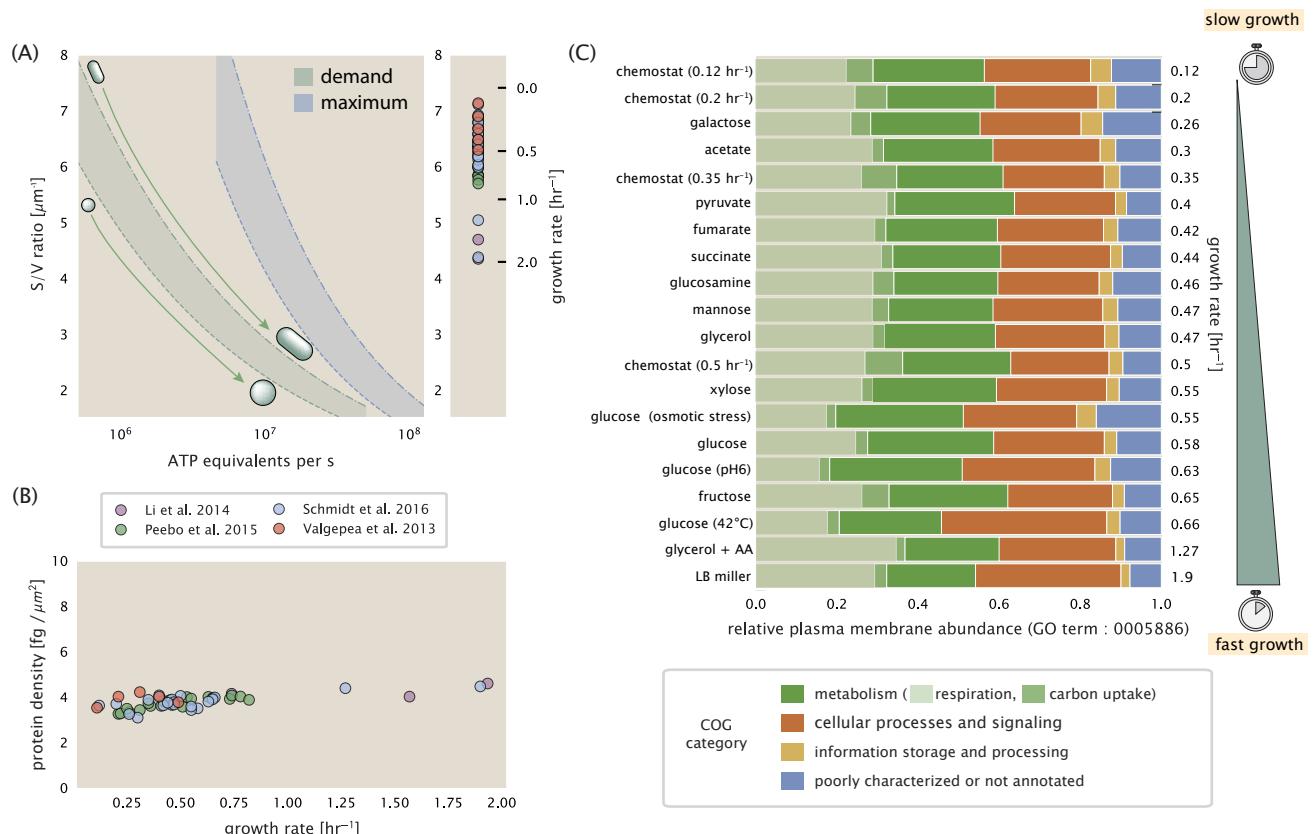


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (---)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

307 that ATP production by respiration is less efficient than by fermentation per membrane area occupied due to the
 308 additional proteins of the electron transport chain. This suggests that, even under anaerobic growth, there will be
 309 sufficient membrane space for ATP production.

310 The analysis highlights the diminishing capacity to provide resources as the cell increases in size. However,
 311 the maximum energy production in *Figure 5(A)* does represent a somewhat unachievable limit since the inner
 312 membrane must also include other proteins including those required for lipid and membrane synthesis. To
 313 better understand the overall proteomic makeup of the inner membrane, we therefore used Gene Ontology (GO)
 314 annotations (*Ashburner et al., 2000; The Gene Ontology Consortium, 2018*) to identify all proteins embedded or
 315 peripheral to the inner membrane (GO term: 0005886). Those associated but not membrane-bound include
 316 proteins like MreB and FtsZ and must nonetheless be considered as a vital component occupying space on the
 317 membrane. In *Figure 5(B)*, we find that the total protein mass per μm^2 is nearly constant across growth rates.
 318 Interestingly, when we consider the distribution of proteins grouped by their Clusters of Orthologous Groups (COG)
 319 (*Tatusov et al., 2000*), the relative abundance for those in metabolism (including ATP synthesis via respiration) is
 320 also relatively constant across growth rates, suggesting that no one process (energy production, nutrient uptake,
 321 etc.) is particularly dominating even at fast growth rates *Figure 5(C)*.

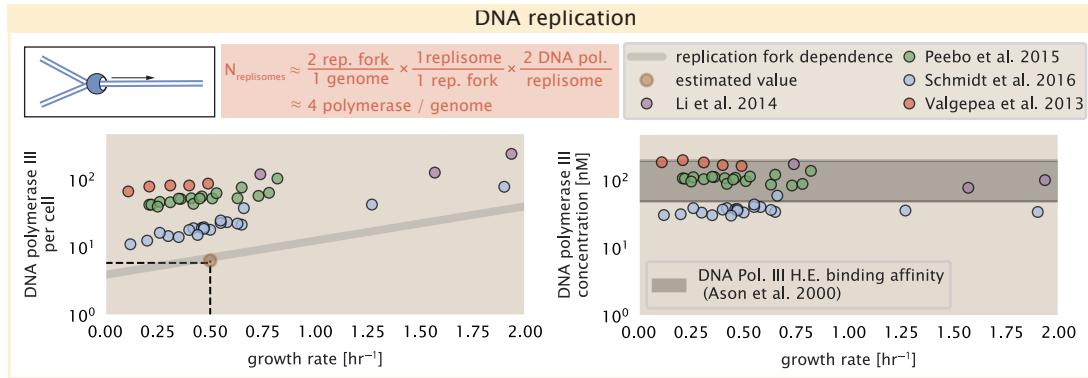


Figure 6. Complex abundance estimates for dNTP synthesis and DNA replication. An estimate for the minimum number of DNA polymerase holoenzyme complexes needed to facilitate replication of a single genome. Points in the left-hand plot correspond to the total number of DNA polymerase III holoenzyme complexes ($[DnaE]_3[DnaQ]_3[HolE]_3[DnaX]_5[HolB][HolA][DnaN]_4[HolC]_4[HolD]_4$) per cell. Right-hand plot shows the effective concentration of DNA polymerase III holoenzyme (See Appendix Estimation of Cell Size and Surface Area for calculate of cell volumes). Grey lines in left-hand panel show the estimated number of complexes needed as a function of growth, the details of which are described in the Supplemental Information.

Figure 6–Figure supplement 1. Estimate and observations of the abundance of ribonucleotide reductase, a key component in dNTP synthesis.

322 Processes of the Central Dogma

323 Up to this point, we have considered a variety of transport and biosynthetic processes that are critical to acquiring
 324 and generating new cell mass. While there are of course many other metabolic processes we could consider, we
 325 now turn our focus to some of the most central processes which *must* be undertaken irrespective of the growth
 326 conditions – those of the central dogma.

327 DNA Replication

328 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal plasmids
 329 up to ~ 100 kbp in length. While we consider the starting material dNTPs in **Figure 6–Figure Supplement 1** and
 330 discussed further in Appendix ??, here we focus our quantitative thinking on the chromosome of *E. coli* which
 331 harbors ≈ 5000 genes and $\approx 5 \times 10^6$ base pairs. To successfully divide and produce viable progeny, this chromosome
 332 must be faithfully replicated and segregated into each nascent cell.

333 Replication is initiated at a single region of the chromosome termed the *oriC* locus at which a pair of DNA
 334 polymerase III begin their high-fidelity replication of the genome in opposite directions (Fijalkowska et al., 2012).
 335 *In vitro* measurements have shown that DNA Polymerase III copies DNA at a rate of ≈ 600 nucleotides per second
 336 (BNID: 104120). Therefore, to replicate a single chromosome, two replisomes (containing two DNA polymerase III
 337 each) moving at their maximal rate would copy the entire genome in ≈ 4000 s. Thus, with a division time of 5000 s,
 338 there is sufficient time for a pair of replisomes complexes to replicate the entire genome.

339 In rapidly growing cultures, bacteria like *E. coli* initiate as many as 10 - 12 replication forks at a given time (Bremer
 340 and Dennis, 2008; Si et al., 2017), we expect only a few DNA polymerases (≈ 10) are needed. However, as shown in
 341 **Figure 6** DNA polymerase III is nearly an order of magnitude more abundant. This discrepancy can be understood
 342 by considering its binding constant to DNA. *In vitro* characterization has quantified the K_d of DNA polymerase III
 343 holoenzyme to single-stranded and double-stranded DNA to be 50 and 200 nM, respectively (Ason et al., 2000).
 344 The right-hand plot in **Figure 6** shows that the concentration of DNA polymerase III across all data sets is within this
 345 range. Thus, its copy number appears to vary such that its concentration is approximately equal to the dissociation
 346 constant to the DNA. While the processes regulating the initiation of DNA replication are complex and involve more
 347 than just the holoenzyme, these data indicate that the kinetics of replication rather than the explicit copy number
 348 of the DNA polymerase III holoenzyme is the more relevant feature of DNA replication to consider. In light of this,
 349 the data in **Figure 6** suggests that for bacteria like *E. coli*, DNA replication does not represent a rate-limiting step in
 350 cell division. However, it is worth noting that for bacterium like *C. crescentus* whose chromosomal replication is

351 initiated only once per cell cycle (*Jensen et al., 2001*), the time to double their chromosome indeed represents an
352 upper limit to their growth rate.

353 RNA Synthesis

354 We now turn our attention to the next stage of the central dogma – the transcription of DNA to form RNA. We
355 consider three major groupings of RNA, namely the RNA associated with ribosomes (rRNA), the RNA encoding the
356 amino-acid sequence of proteins (mRNA), and the RNA which links codon sequence to amino-acid identity during
357 translation (tRNA).

358 rRNA serves as the catalytic and structural framework along with myriad ribosomal proteins as part of a
359 complete ribosomal complex. Each ribosome contains three rRNA molecules of lengths 120, 1542, and 2904
360 nucleotides (BNID: 108093), meaning each ribosome contains \approx 4500 nucleotides overall. *In vivo* measurements
361 of the kinetics of rRNA transcription have revealed that RNA polymerases are loaded onto the promoter of an
362 rRNA gene at a rate of \approx 1 per second (BNID: 111997, 102362). If RNA polymerases are constantly loaded at this
363 rate, then we can assume that \approx 1 functional rRNA unit is synthesized per second per rRNA operon. While *E. coli*
364 possesses 7 of these operons per chromosome, the fact that chromosome replication can be parallelized means
365 that the average dosage of rRNA genes can be substantially higher at fast growth rates. At a growth rate of \approx 0.5
366 hr $^{-1}$, however, the average cell has \approx 1 copy of its chromosome and therefore approximately \approx 7 copies of the
367 rRNA operons, therefore producing \approx 7 rRNA units per second. With a 5000 second division time, this means the
368 cell is able to generate around 3×10^4 functional rRNA units, comparable within an order of magnitude to the
369 number of ribosomes per cell.

370 How many RNA polymerases are then needed to constantly transcribe the required rRNA? If one polymerase is
371 loaded per second, and the transcription rate is \approx 40 nucleotides per second, then the typical spacing between
372 polymerases will be \approx 40 nucleotides. However, we must note that the polymerase itself has a footprint of \approx 40
373 nucleotides (BNID: 107873), meaning that one could expect to find one RNA polymerase per 80 nucleotide stretch
374 of an rRNA gene. With a total length of \approx 4500 nucleotides per operon and 7 operons per cell, the number of RNA
375 polymerases transcribing rRNA at any given time is then \approx 500 per cell.

376 As outlined in *Figure 7*, and discussed further the Appendix ??, synthesis of mRNA and tRNA together require
377 on the order of \approx 400 RNAP. Thus, in total, one would expect the typical cell to require \approx 1000 RNAP to satisfy its
378 transcriptional demands. As is revealed in *Figure 7(B)*, this estimate is about an order of magnitude below the
379 observed number of RNA polymerase complexes per cell (\approx 5000 - 7000). The difference between the estimated
380 number of RNA polymerase needed for transcription and these observations, however, are consistent with recent
381 literature revealing that \approx 80 % of RNA polymerases in *E. coli* are not transcriptionally active (*Patrick et al., 2015*).

382 Our estimates also neglect other mechanistic features of transcription and transcriptional initiation more
383 broadly. For example, we acknowledge that some fraction of the RNAP pool is nonspecifically bound to DNA
384 during its search for promoters from which to begin transcription. Furthermore, we ignore the obstacles that RNA
385 polymerase and DNA polymerase present to each other as they move along the DNA (*Finkelstein and Greene,*
386 *2013*). Finally, we neglect the fact that RNA polymerase also require σ -factors for promoter recognition and
387 transcription initiation (*Browning and Busby, 2016*). In *Figure 7-Figure Supplement 1*, we show that the predicted
388 RNA polymerase copy number indeed is more comparable with the abundance of σ -70 (RpoD), the primary sigma
389 factor in *E. coli*. There therefore remains more to be investigated as to what sets the observed abundance of RNA
390 polymerase in these proteomic data sets.

391 Protein Synthesis

392 The last process of the central dogma – the translation of RNA into function protein – is the final subject in our
393 dialogue between back-of-the-envelope estimates and comparison with proteomic data. Here we consider the
394 number of ribosomes needed to replicate the cellular proteome. While the rate at which ribosomes translates is
395 well known to have a growth rate dependence *Dai et al. (2018)*, here we make the approximation that translation
396 occurs at a rate of \approx 15 amino acids per second per ribosome (BNID: 100233). Under this approximation and
397 assuming a division time of 5000 s, we can easily arrive at an estimate of \approx 10^4 ribosomes needed to replicate
398 the entire protein mass (*Figure 8*). This point estimate and the corresponding estimate across a continuum of
399 growth rates proves to be notably comparable to the experimental observations, shown in *Figure 8(B)*. While the

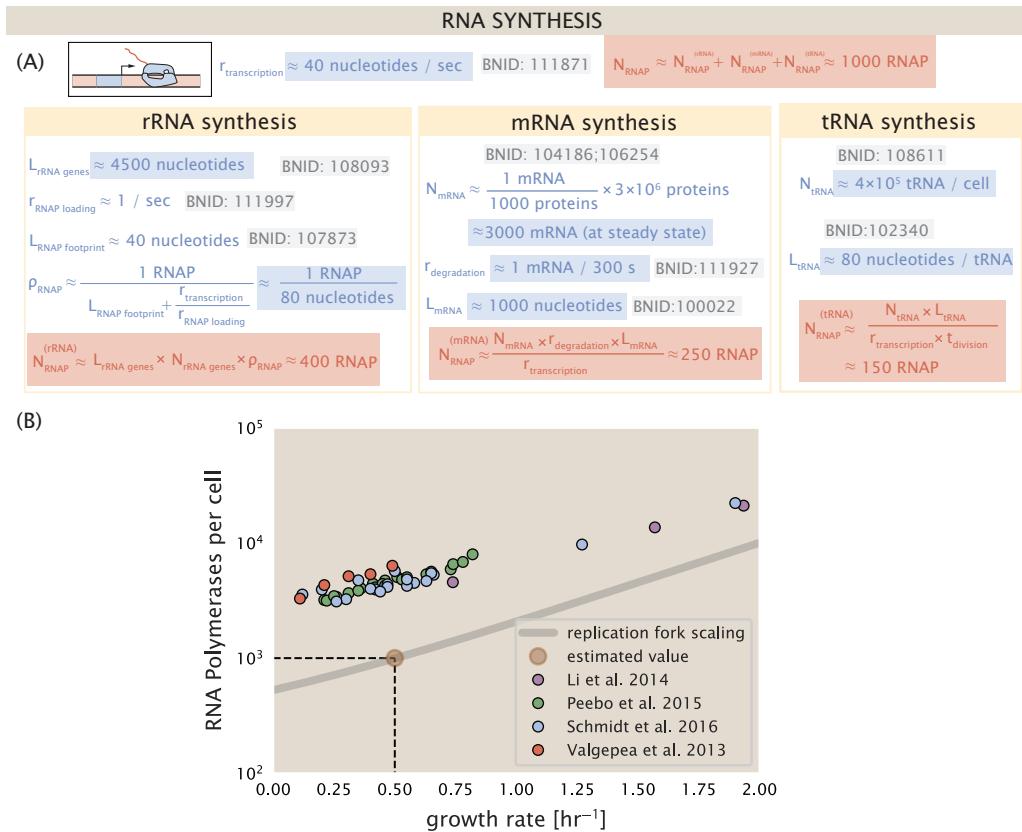


Figure 7. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively.(B) The RNA polymerase core enzyme copy number as a function of growth rate. Colored points correspond to the average number RNA polymerase core enzymes that could be formed given a subunit stoichiometry of $[RpoA]_2[RpoC][RpoB]$.

Figure 7-Figure supplement 1. Abundance and growth rate dependence of σ -70.

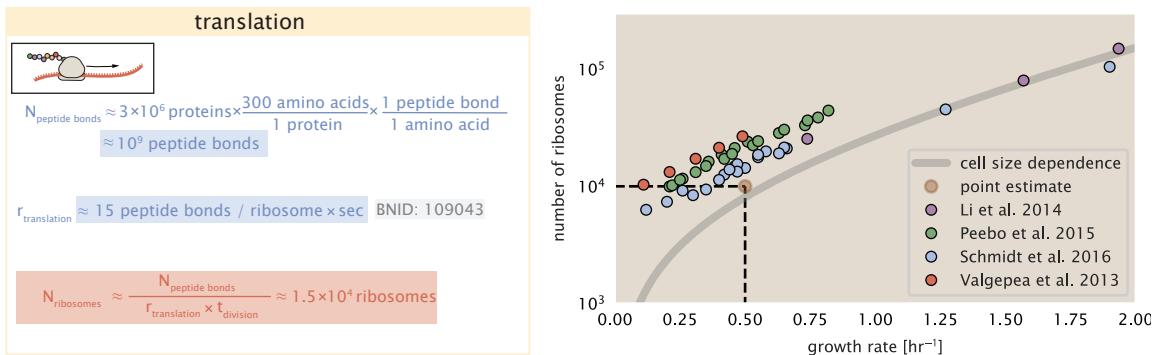


Figure 8. Estimation of the required number of ribosomes. Estimation of the number of ribosomes required to synthesize 10^9 peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr^{-1} . Grey lines correspond to the estimated complex abundance calculated at different growth rates.

Figure 8–Figure supplement 1. Estimate and observed abundance and growth rate dependence of tRNA ligases.

400 ribosome is responsible for the formation of peptide bonds, we do not diminish the importance of the charging of
 401 tRNAs with the appropriate amino acid, a process which occurs with remarkable accuracy. In the Appendix and in
 402 **Figure 8–Figure Supplement 1**, we consider the process of ligating tRNAs to their corresponding amino acid and
 403 again find notable comparability with the data.

404 Translation and Ribosomal Synthesis as a Rate-Limiting Step

405 Having now completed our circuit through key processes of cellular growth, we find that on the whole our
 406 estimates have been reasonably successful in predicting the scale of absolute protein copy number as well as
 407 the observed dependence on the growth rate. This suggests that cells modulate their proteomic composition
 408 and absolute protein abundance to better match their growth rate requirements, without any one process
 409 representing a particular bottleneck. In our effort to identify key limitations on growth, however, there are two
 410 notable observations that we highlight here. The first is a recurring theme throughout our estimates, which is
 411 that any inherent biochemical rate limitation can be overcome by expressing more proteins. We can view this as a
 412 parallelization of each biosynthesis task, and this provides some basis to explain why bacteria tend to increase
 413 their protein content (and cell size) as they grow faster [cites]. The second is that the synthesis of ribosomal
 414 proteins presents a special case where parallelization is not possible (**Figure 9(A)**), suggesting the possibility that it
 415 may represent a rate-limiting step in maximizing growth rate. In the remaining sections we show how ribosome
 416 synthesis indeed limits the achievable growth rate, and we consider how *E. coli* deals with this constraint through
 417 control its ribosome copy number and proteomic composition. We then use a minimal model of growth rate
 418 control to gain insight into how the observed changes in proteomic composition and absolute abundance enable
 419 cells to better maximize their growth rate across different nutrient conditions.

420 Maximum Growth Rate is Determined by the Ribosomal Mass Fraction

421 To gain some intuition into how ribosomes influence bacterial growth, we again consider the total number of
 422 peptide bonds that must be synthesized, which we denote as N_{pep} . With cells growing exponentially in time (**Godin**
 423 **et al., 2010**), the rate of cellular growth will be related to the rate of protein synthesis by

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

424 where λ is the cell growth rate in s^{-1} , r_t is the maximum elongation rate in $\text{AA} \cdot \text{s}^{-1}$, and R is the average ribosome
 425 copy number per cell. The addition factor f_a refers to the fraction of actively translating ribosomes, and allows us to
 426 account for the possibility of nonfunctional, immature ribosomes or active sequestration of ribosomes, mediated
 427 by the secondary-messenger molecule alarmones, guanosine pentaphosphate [(p)ppGpp] at slow growth (**Dennis**
 428 **et al., 2004; Dai et al., 2016**).

429 Alternatively, since we are most interested in the role of ribosomal synthesis on growth rate, we instead consider
 430 this in terms of the ribosomal mass fraction, denoted by Φ_R . N_{pep} is related to the total protein mass through

431 the molecular weight of each protein, we can also consider the growth rate in terms of the fraction of the total
 432 proteome mass dedicated to ribosomal proteins. By making the approximation that an average amino acid has a
 433 molecular weight of 110 Da (BNID: 104877), the total protein mass m_{protein} is related to N_{pep} by $(m_{\text{protein}}/110 \text{ Da}) \times N_A$,
 434 where N_A is Avogadro's number. Similarly, R is related to the ribosomal protein mass by $R \approx (m_R/800 \text{ Da}) \times N_A$,
 435 where 800 Da reflects the summed molecular weight of all ribosomal subunits. This allows us to approximate
 436 $R/N_{\text{pep}} \approx \Phi_R/L_R$, where Φ_R is the ribosomal mass fraction m_{protein}/m_R , and L_R the ratio of 800 kDa / 110 Da per
 437 amino acid or, alternatively, the total length in amino acids that make up a ribosome. The translation-limited
 438 growth rate can then be written in the form

$$\lambda_{\text{translation-limited}} \approx \frac{r_t}{L_R} \Phi_R f_a. \quad (2)$$

439 This is plotted as a function of the ribosomal fraction Φ_R in **Figure 9(B)**, where we take $L_R \approx 7500 \text{ AA}$, corresponding
 440 to the length in amino acids for all ribosomal subunits of the 50S and 30S complex (BNID: 101175), and $f_a = 1$. To
 441 allow us to consider the the proteomic data, we use recent measurements of f_a from **Dai et al. (2016)** to estimate
 442 the active fraction of ribosomal protein across each proteomic data set (**Figure 9(C)**). We find that cells in general
 443 appear to skirt this limit in growth rate as nutrient conditions vary. There is a notable discrepancy in the data from
 444 **Peebo et al. (2015); Valgepea et al. (2013)**, where cells appear to grow substantially slower given their estimated
 445 ribosomal fraction. We have also collected a number of recent measurements of ribosomal fraction and find them
 446 most consistent with the measurements from **Li et al. (2014); Schmidt et al. (2016)** (**Figure 9-Figure Supplement 1**).

447 The growth rate defined by **Equation 2** reflects mass-balance under steady state growth and has long provided
 448 a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function of growth rate (**Maaløe,**
 449 **1979; Scott et al., 2010**). The maximum rate, when $\Phi_R = 1$, could only be achieved if a cell contained only
 450 ribosomes. This corresponds to the synthesis time of all ribosomal subunits, $L_R/r_t \approx 7 \text{ minutes}$ (**Dill et al., 2011**)
 451 and interestingly, is independent of the absolute number of ribosomes. To return to our earlier comments on
 452 parallelization, it is this step that is rate-limiting, with each ribosome being required to produce a second ribosome.
 453 Unless elongation rate increased, or cells could trim their total ribosomal protein mass, this dependency limits both
 454 the maximum growth rate (when $\Phi_R = 1$), and also the achievable growth rate under more moderate values of Φ_R .

455 rRNA Synthesis Presents a Potential Bottleneck during Rapid Growth

456 *E. coli* rarely exhibits growth rates above 2 hr^{-1} (**Bremer and Dennis, 2008**), which is still well-below the synthesis
 457 rate of a single ribosome, and below the growth rates reported for several other bacteria (**Roller et al., 2016**). Here
 458 we need to also consider ribosomal synthesis from the perspective of limiting rRNA synthesis, which as we have
 459 found earlier, will depend on the number of rRNA operons to transcribe rRNA.

460 Due to multiple rounds of chromosomal replication per cell doubling, the effective number of rRNA operons
 461 increases with growth rate and will vary in proportion to the average number of origins per cell, $\langle \# \text{ ori} \rangle$. This
 462 parameter is set by how often replication must be initiated per cell doubling in order to maintain steady state
 463 growth, and quantified by

$$\langle \# \text{ ori} \rangle = 2^{\tau_{\text{cyc}}/\tau} = 2^{\tau_{\text{cyc}} \lambda / \ln(2)}. \quad (3)$$

464 Here, τ_{cyc} is the cell cycle time (referring to the time from replication initiation to cell division), and τ is the cell
 465 doubling time. We used the experimental measurements of τ_{cyc} and τ from **Si et al. (2017)** [supp fig] to calculate $\langle \#$
 466 $\text{ori} \rangle$ with **Equation 3** as a function of growth rates. While rRNA operons are predominantly located near to origin of
 467 replication, 2 of the 7 operons are not (BNID: 100352) and we describe our complete estimation of the number of
 468 rRNA operons in Appendix [ref appendix section].

469 Returning to our rule-of-thumb of 1 functional rRNA unit per second per operon, we estimate the maximum
 470 number of ribosomes that could be made as a function of growth rate (**Figure 9(C)**, blue curve). Although we expect
 471 this estimate to drastically overestimate rRNA abundance at slower growth rates ($\lambda < 0.5 \text{ hr}^{-1}$), it provides a useful
 472 reference alongside the proteomic measurements. For growth rates above about 1 hr^{-1} , we find that cells will
 473 need to transcribe rRNA near their maximal rate. As a counter example, if *E. coli* did not initiate multiple rounds of
 474 replication, they would be unable to make enough rRNA for the observed number of ribosomes (dashed blue curve
 475 in **Figure 9(C)**). The convergence between the maximum rRNA production and measured ribosome copy number
 476 suggests rRNA synthesis may begin to present a bottleneck at the fastest growth rates due to the still-limited copies
 477 of rRNA genes.

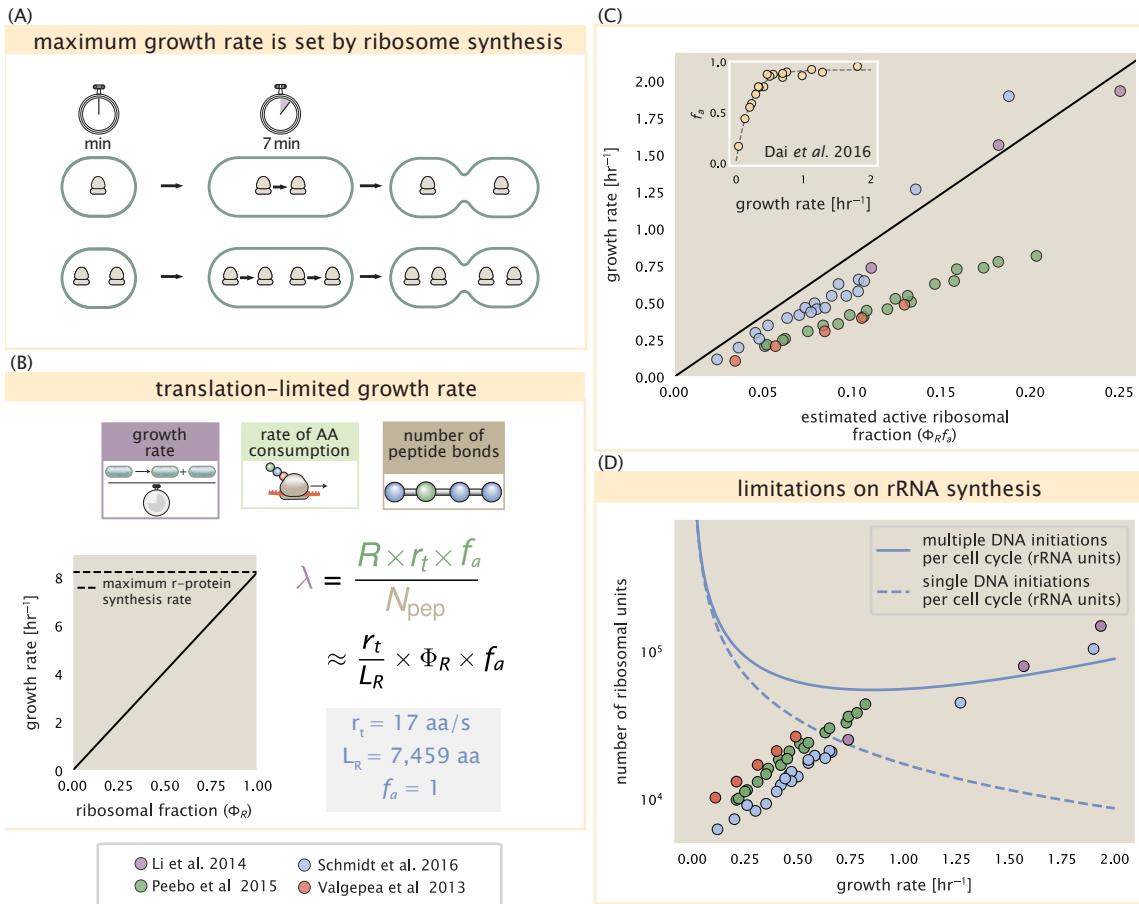


Figure 9. Translation-limited growth rate. (A) An inherent maximum growth rate is set by the time to synthesize all ribosomal subunits. This growth rate is given by r_t/L_R , where r_t is the elongation rate and L_R is the total number of amino acids that make up the entire ribosomal complex. This rate is independent of the number of ribosomes and instead is limited by the time required to double an individual ribosome. (B) Translation-limited growth as a function of the ribosomal fraction. The solid line is calculated for an elongation rate of 17 aa per second. The dashed line corresponds to the maximum rate of ribosomal protein (R-protein) synthesis considered in part (A). (C) Actively translating ribosomal fraction versus growth rate. The actively translating ribosomal fraction is calculated using the estimated values of f_a from [Dai et al. \(2016\)](#) (shown in inset; see Appendix Calculation of active ribosomal fraction for additional detail). (D) Maximum number of rRNA units that can be synthesized as a function of growth rate. Solid curve corresponding to the rRNA copy number is calculated by multiplying the number of rRNA operons by the estimated number of (# ori) at each growth rate. The quantity (# ori) was calculated using Equation 4 and the measurements from [Si et al. \(2017\)](#) that are plotted in [Figure 10\(A\)](#). The dashed line shows the maximal number of functional rRNA units produced from a single chromosomal initiation per cell cycle.

Figure 9–Figure supplement 1. Comparison of $\Phi_R f_a$ with literature and estimation of (# ori).

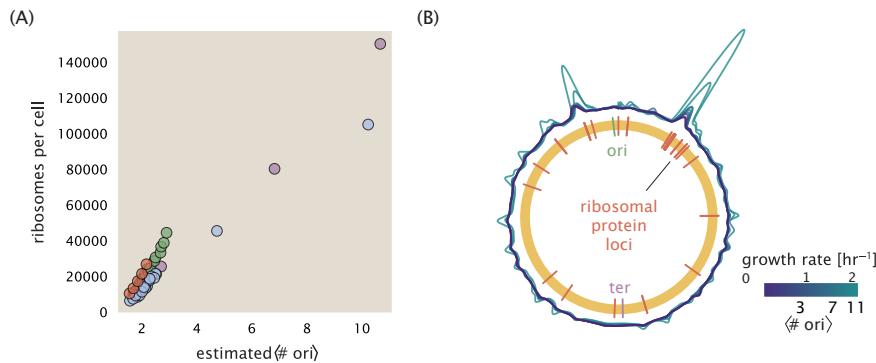


Figure 10. Cells increase both absolute ribosome abundance and Φ_R with (# ori). (A) Plot of the ribosome copy number estimated from the proteomic data against the estimated (# ori) (see Appendix Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$ for additional details). (B) A running Gaussian average (20 kbp st. dev.) of protein copy number is calculated for each growth condition considered by (Schmidt et al., 2016) based on each gene's transcriptional start site. Since total protein abundance increases with growth rate, protein copy numbers are median-subtracted to allow comparison between growth conditions. (# ori) are estimated using the data in (A) and Equation 3.

478 Rapid Growth Requires *E. coli* to Increase Both Cell Size and Ribosomal Mass Fraction

479 In Figure 9(C) we find that above about 0.75 hr^{-1} , the growth rate is determined by the ribosomal mass fraction Φ_R ,
 480 since f_a is close to 1, and r_i is near its maximal rate [cite and refer to figure/ supplemental]. While Φ_R will need to
 481 increase in order for cells to grow faster, the fractional dependence in Equation 2 gives little insight into how this is
 482 actually achieved in the cell and we consider this further here.

483 It is now well-documented that *E. coli* cells add a constant volume per origin of replication, which is robust to
 484 a remarkable array of cellular perturbations (Si et al., 2017). We find that ribosomal copy number also scales its
 485 ribosome copy number in proportion to (# ori) Figure 10(A). Importantly, however, it will only be due to an increase
 486 in Φ_R at these moderate to fast growth rates that cells can achieve an increase in their growth rate. Indeed, we
 487 find that the deviations in protein expression with (# ori) are largely restricted to regions of ribosomal protein
 488 genes Figure 10(B). Here we have calculated the position-dependent protein expression across the chromosome
 489 by a running Gaussian average of protein copy number (20 kbp st. dev. averaging window) based on each gene's
 490 transcriptional start site. These were median-subtracted to account for the change in total protein abundance
 491 with (# ori). This result suggests that Φ_R is also being tuned in proportion to (# ori) under nutrient-limited growth,
 492 and in particular, it is through this additional dependence on Φ_R that *E. coli* exhibits an exponential increase in cell
 493 volume with growth rate.

494 A Minimal Model of Nutrient-Mediated Growth Rate Control

495 While the preceding subsections highlight a dominant role for ribosomes in setting the growth rate, our analysis on
 496 the whole emphasizes that the total proteomic content must also change in response to variable growth conditions
 497 and growth rate. In this final section we use a minimal model of growth rate control to better understand how the
 498 interconnection between ribosomal abundance and total protein influences the observed growth rate.

499 Here we propose that cells modulate their protein abundance in direct response to the availability of nutrients
 500 in their environment. As noted earlier, bacteria can modulate ribosomal activity through the secondary-messenger
 501 molecules like (p)ppGpp in poorer nutrient conditions (Figure 9(C) - inset; Dai et al. (2016)). Importantly, these
 502 secondary-messengers also cause global changes in transcriptional and translational activity (Hauryliuk et al.,
 503 2015; Zhu and Dai, 2019; Büke et al., 2020). In *E. coli*, amino acid starvation leads to the accumulation of de-acetylated
 504 tRNAs at the ribosome's A-site and a strong increase in (p)ppGpp synthesis activity by the enzyme RelA (Hauryliuk
 505 et al., 2015). There is also increasing evidence that (p)ppGpp inhibits the initiation of DNA replication (Kraemer
 506 et al., 2019), resulting in a lower (# ori) and smaller cell size in poorer growth conditions (Fernández-Coll et al.,
 507 2020).

508 To consider this quantitatively, we assume that cells modulate their proteome (N_{pep} , R , Φ_R) to better maximize

their rate of peptide elongation r_t . The elongation rate r_t will depend on how quickly the ribosomes can match codons with their correct amino-acyl tRNA, along with the subsequent steps of peptide bond formation and translocation. This ultimately depends on the cellular concentration amino acids, which we treat as a single effective species, $[AA]_{\text{eff}}$. In our model, we determine the the rate of peptide elongation r_t , and achievable growth rate as simply depending on the supply of amino acids (and, therefore, also amino-acyl tRNAs), through a parameter r_{AA} in units of AA per second, and the rate of amino acid consumption by protein synthesis ($r_t \times R \times f_a$). This is shown schematically in **Figure 11(A)** and derived in Appendix XX. Given our observation that protein synthesis and energy production are not limiting, we assume that other molecular players required by ribosomes like elongation factors and GTP are available in sufficient abundance.

In **Figure 11(B)**, we illustrate how the elongation rate will depend on the ribosomal copy number. Here, we have considered an arbitrarily chosen $r_{AA} = 5 \times 10^6 \text{ AA} \cdot \text{s}^{-1} \cdot \mu\text{m}^{-3}$ and $f_a = 1$ for a cell volume $V = 1 \mu\text{m}^3$ and $N_{\text{pep}} = []$. At low ribosome copy numbers, the observed elongation rate is dependent primarily on $[AA]_{\text{eff}}$ through r_{AA} [as $r_t^{\max} \times R \times f_a \ll r_{AA}$, point (1) in **Figure 11(B)**]. As the ribosome copy number is increased such that the amino acid supply rate and consumption rate are nearly equal [point (2) in **Figure 11(B)**], the observed elongation rate begins to decrease sharply. When the ribosome copy number is increased even further, consumption at the maximum elongation rate exceeds the supply rate, yielding a significantly reduced elongation rate [point (3) in **Figure 11B**]. While the elongation rate will always be dominated by the amino acid supply rate at sufficiently low ribosome copy numbers, the elongation rate at larger ribosome abundances can be increased by tuning f_a such that not all ribosomes are elongating, reducing the total consumption rate.

Optimal Ribosomal Content and Cell Size Depend on Nutrient Availability and Metabolic Capacity
To relate elongation rate to growth rate, we constrain the set of parameters based on our available proteomic measurements; namely, we restrict the values of R , N_{pep} , and cell size to those associated with the amalgamated proteomic data (described in Appendix Estimation of Total Protein Content per Cell). We then consider how changes in the nutrient conditions, through the parameter r_{AA} , influence the maximum growth rate as determined by ???. **Figure 11(C)** shows how the observed growth rate depends on the rate of amino acid supply r_{AA} as a function of the cellular ribosome copy number. A feature immediately apparent is the presence of a maximal growth rate whose dependence on R (and consequently, the cell size) increases with increasing r_{AA} . Importantly, however, there is an optimum set of R , N_{pep} , and V that are strictly dependent on the value of r_{AA} . Increasing the ribosomal concentration beyond the cell's metabolic capacity has the adverse consequence of depleting the supply of amino acids and a concomitant decrease in the elongation rate r_t [**Figure 11(B)**].

Also of note is the growth rate profiles shown for low amino acid supply rates [purple and blue lines in **Figure 11(C)**], representing growth in nutrient-poor media. In these conditions, there no longer exists a peak in growth, at least in the range of physiologically-relevant ribosome copy numbers. Instead, cells limit their pool of actively translating ribosomes by decreasing f_a (Dai et al., 2016), which would help maintain the pool of available amino acids $[AA]_{\text{eff}}$ and increase the achievable elongation rate. This observation is in agreement with the central premise of the cellular resource allocation principle proposed by Scott et al. (2010); Klumpp et al. (2009); Klumpp and Hwa (2014) and Hui et al. (2015).

Discussion

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

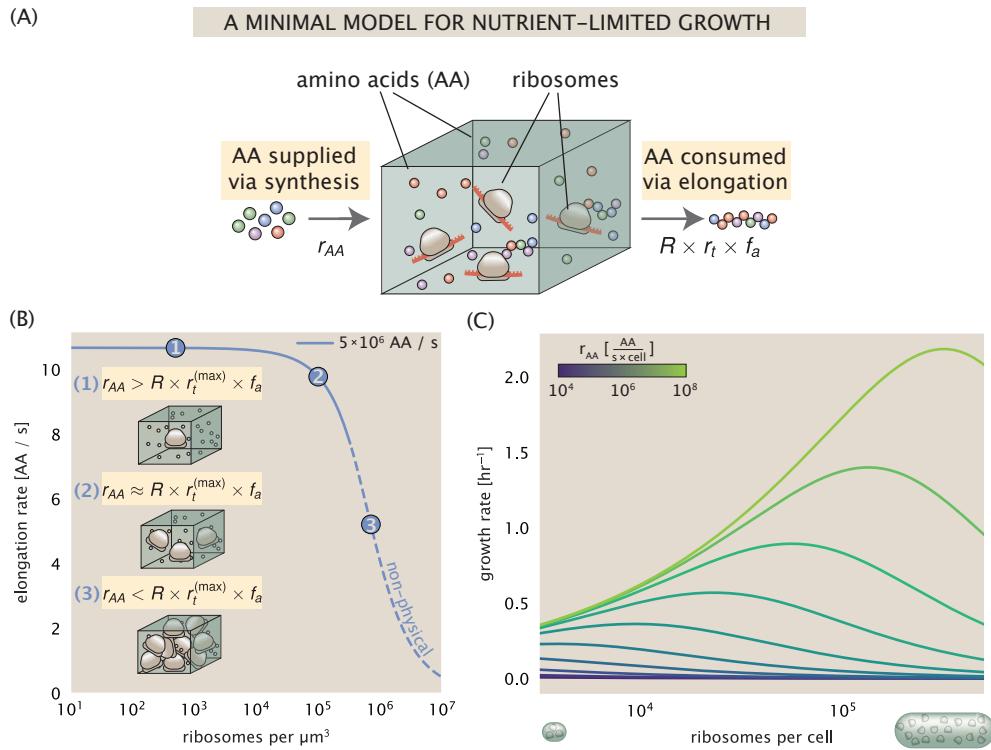


Figure 11. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times f_a$, where r_t is the elongation rate, R is the ribosome copy number in the unit volume, and f_a is the fraction of those ribosomes actively translating. (B) The observed elongation rate is plotted as a function of ribosomes in a unit volume μm^3 . The three points correspond to three regimes of ribosome copy numbers and are shown schematically on the left-hand side. The region of the curve shown as dashed lines represents a non-physical copy number, but is shown for illustrative purposes. This curve was generated using the parameters $r_{AA} = 5 \times 10^6 \text{ AA} / \text{s}$, $K_D = 5 \text{ mM}$, and $r_t^{(\max)} = 17.1 \text{ AA} / \text{s}$. (C) The cellular growth rate is plotted as a function of total cellular ribosome copy number for different cellular amino acid supply rates, with blue and green curves corresponding to low and high supply rates, respectively. As the ribosome copy number is increased, so too is the cell volume and total protein abundance. We direct the reader to the Supplemental Information for discussion on the inference of the realtionship between cell volume, number of peptide bonds, and ribosome copy number.

558 provide insight on the scales of macromolecular complex abundance across four classes of cellular processes – the
559 transport of nutrients, the production of energy, the synthesis of the membrane and cell wall, and the numerous
560 steps of the central dogma.

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

573 Of the processes illustrated in *Figure 1*, we arrive at a ribosome-centric view of cellular growth rate control.
574 This is in some sense unsurprising given the long-held observation that *E. coli* and many other organisms vary
575 their ribosomal abundance as a function of growth conditions and growth rate *Scott et al. (2010); Metzl-Raz et al.*
576 (*2017*). However, through our dialogue with the proteomic data, two additional key points emerge. The first relates
577 to our question of what process sets the absolute speed limit of bacterial growth. While a cell can parallelize
578 many of its processes simply by increasing the abundance of specific proteins or firing multiple rounds of DNA
579 replication, this is not so for synthesis of ribosomes (*Figure 9(A)*). The translation time for each ribosome [\approx 6 min,
580 *Dill et al. (2011)*] places an inherent limit on the growth rate that can only be surpassed if the cell were to increase
581 their polypeptide elongation rate, or if they could reduce the total protein and rRNA mass of the ribosome. The
582 second point relates to the long-observed correlations between growth rate and cell size (*Schaechter et al., 1958; Si*
583 *et al., 2017*), and between growth rate and ribosomal mass fraction. While both trends have sparked tremendous
584 curiosity and driven substantial amounts of research in their own regards, these relationships are themselves
585 intertwined. In particular, it is the need for cells to increase their absolute number of ribosomes under conditions
586 of rapid growth that require cells to also grow in size. Further experiments are needed to test the validity of this
587 hypothesis. In particular, we believe that the change in growth rate in response to translation-inhibitory drugs
588 (such as chloramphenicol) could be quantitatively predicted, given one had precision measurement of the relevant
589 parameters, including the fraction of actively translating ribosomes f_a and changes in the metabolic capacity of the
590 cell (i.e. the parameter r_{AA} in our minimal model) for a particular growth condition.

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

604 This work is by no means meant to be an exhaustive dissection of bacterial growth rate control, and there are
605 many aspects of the bacterial proteome and growth that we neglected to consider. For example, other recent work
606 (*Liebermeister et al., 2014; Hui et al., 2015; Schmidt et al., 2016*) has explored how the proteome is structured
607 and how that structure depends on growth rate. In the work of *Hui et al. (2015)*, the authors coarse-grained the
608 proteome into six discrete categories being related to either translation, catabolism, anabolism, and others related

609 to signaling and core metabolism. The relative mass fraction of the proteome occupied by each sector could be
610 modulated by external application of drugs or simply by changing the nutritional content of the medium. While
611 we have explored how the quantities of individual complexes are related to cell growth, we acknowledge that
612 higher-order interactions between groups of complexes or metabolic networks at a systems-level may reveal
613 additional insights into how these growth-rate dependences are mechanistically achieved. Furthermore, while we
614 anticipate the conclusions summarized here are applicable to a wide collection of bacteria with similar lifestyles as *E.*
615 *coli*, other bacteria and archaea may have evolved other strategies that were not considered. Further experiments
616 with the level of rigor now possible in *E. coli* will need to be performed in a variety of microbial organisms to learn
617 more about how regulation of proteomic composition and growth rate control has evolved over the past 3.5 billion
618 years.

619 Methods

620 Data Analysis and Availability

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

626 Acknowledgements

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

632 Competing Interests

633 The authors declare no competing interests.

634 Appendix for: Fundamental limits on the 635 rate of bacterial cell division

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

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

647	Contents	
648	Introduction	1
649	Nutrient Transport	2
650	Limits on Transporter Expression	6
651	Cell Envelope Biogenesis	6
652	Lipid Synthesis	6
653	Peptidoglycan Synthesis	7
654	Limits on Cell Wall Biogenesis	7
655	Energy Production	7
656	ATP Synthesis	7
657	Generating the Proton Electrochemical Gradient	9
658	Biosynthesis in a Crowded Membrane	9
659	Processes of the Central Dogma	12
660	DNA Replication	12
661	RNA Synthesis	13
662	Protein Synthesis	13
663	Translation and Ribosomal Synthesis as a Rate-Limiting Step	15
664	Maximum Growth Rate is Determined by the Ribosomal Mass Fraction	15
665	rRNA Synthesis Presents a Potential Bottleneck during Rapid Growth	16
666	Rapid Growth Requires <i>E. coli</i> to Increase Both Cell Size and Ribosomal Mass Fraction	18
667	A Minimal Model of Nutrient-Mediated Growth Rate Control	18
668	Optimal Ribosomal Content and Cell Size Depend on Nutrient Availability and Metabolic Capacity	19
669	Discussion	19
670	Methods	22
671	Data Analysis and Availability	22
672	Acknowledgements	22
673	Competing Interests	22
674	Experimental Details Behind Proteomic Data	26
675	Fluorescence based measurements	26
676	Ribosomal profiling measurements	26
677	Mass spectrometry measurements	27
678	Summary of Proteomic Data	27
679	Estimation of Cell Size and Surface Area	28
680	Estimation of Total Protein Content per Cell	30
681	Estimating Volume and Number of Amino Acids from Ribosome Copy Number	30
682	Additional Considerations of Schmidt <i>et al.</i> Data Set	30
683	Effect of cell volume on reported absolute protein abundances	32
684	Relaxing assumption of constant protein concentration across growth conditions	34
685	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	34

686	Calculation of Complex Abundance	35
687	Extending Estimates to a Continuum of Growth Rates	37
688	Estimation of the total cell mass	37
689	Complex Abundance Scaling With Cell Volume	37
690	A Relation for Complex Abundance Scaling With Surface Area	38
691	Number of Lipids	38
692	Number of Murein Monomers	38
693	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	39
694	Calculation of active ribosomal fraction.	39
695	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	39

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).

696 Experimental Details Behind Proteomic Data

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

703 Fluorescence based measurements

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

711 Ribosomal profiling measurements

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

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

726 **Mass spectrometry measurements**

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

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

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

755 **Summary of Proteomic Data**

756 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)*;
757 *Schmidt et al. (2016)*. As shown in *Figure 12(A)*, the reported total protein abundances in the work of *Taniguchi*
758 *et al. (2010)* and *Soufi et al. (2015)* differed quite substantially from the other work. For the work of *Taniguchi et al.*
759 (*2010*) this is in part due to a lower coverage in total proteomic mass quantified, though we also noticed that most
760 proteins appear undercounted when compared to the other data.

761 *Figure 12(B)* summarizes the total protein mass for each data point in our final compiled data set. We note that
762 protein abundances were all scaled so they followed a common growth rate-dependent change in total protein
763 mass. While our inclination initially was to leave reported copy numbers untouched, a notable discrepancy in the
764 scaling total protein per cell between *Schmidt et al. (2016)* and the other data sets forced us to dig deeper into
765 those measurements (compare *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 12(A)*). The particular trend
766 in *Schmidt et al. (2016)* appears to be due to assumptions of cell size and we provide a more extensive discussion
767 and analysis of that data set in section Additional Considerations of Schmidt et al. Data Set. As a compromise, and
768 in an effort to treat all data equally, we instead scaled all protein abundance values to a data-driven estimate of
769 total protein per cell. Here we used cell size measurements from *Si et al. (2017, 2019)*, and an estimate of total
770 protein content through expected dry mass. Total protein per cell was estimated using available data on total
771 DNA, RNA, and protein from *Basan et al. (2015)*; *Dai et al. (2016)*, which account for the majority of dry mass in the
772 cell. We consider these details in sections Estimation of Cell Size and Surface Area and Estimation of Total Protein
773 Content per Cell that follows.

774 Lastly, in *Figure 13* we show the total proteomic coverage and overlap of proteins quantified across each data

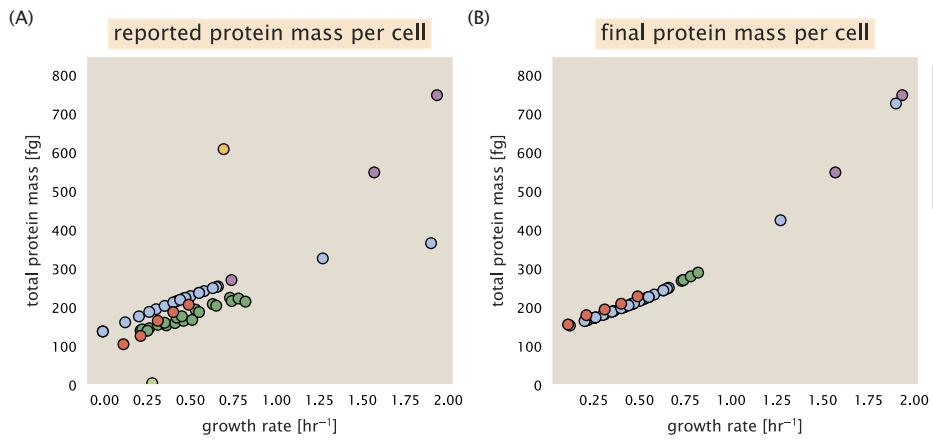


Figure 12. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as original reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al. (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).

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

783 Estimation of Cell Size and Surface Area

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

789 The average size measurements from each of their experiments are shown in Figure **Figure 14**, with cell length
 790 and width shown in (A) and (B), respectively. The length data was well described by the exponential function 0.5
 791 $e^{1.09 \cdot \lambda} + 1.76$ μm , while the width data was well described by $0.64 e^{0.24 \cdot \lambda}$ μm . In order to estimate cell size we take the
 792 cell as a cylinders with two hemispherical ends (*Si et al., 2017*; *Basan et al., 2015*). Specifically, cell size is estimated
 793 from,

$$V = \pi \cdot r^2 \cdot (l - 2r/3), \quad (4)$$

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

$$S = \eta \cdot \pi \left(\frac{\eta \cdot \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V^{2/3}, \quad (5)$$

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

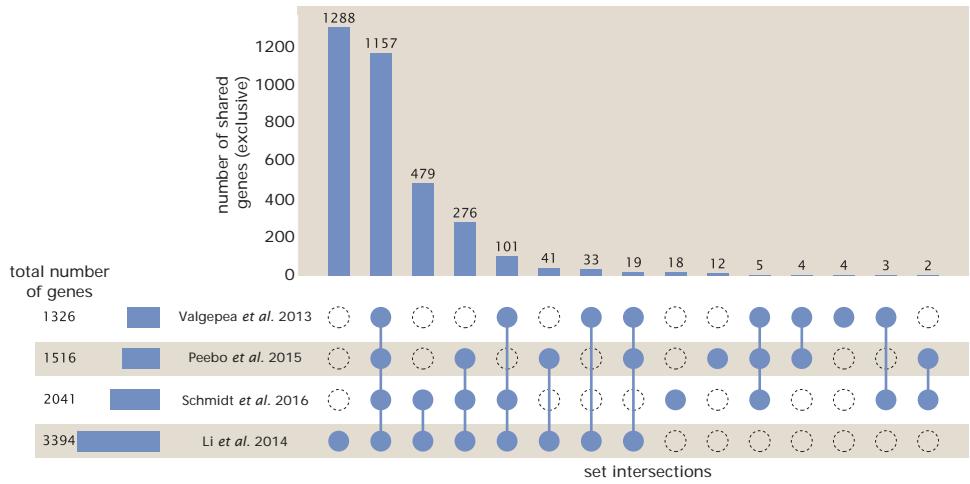


Figure 13. Comparison of proteomic coverage across different data sets. An UpSet diagram (*Lex et al., 2014*) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of *Valgepea et al. (2013)*; *Li et al. (2014)*; *Schmidt et al. (2016)*; *Peebo et al. (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.

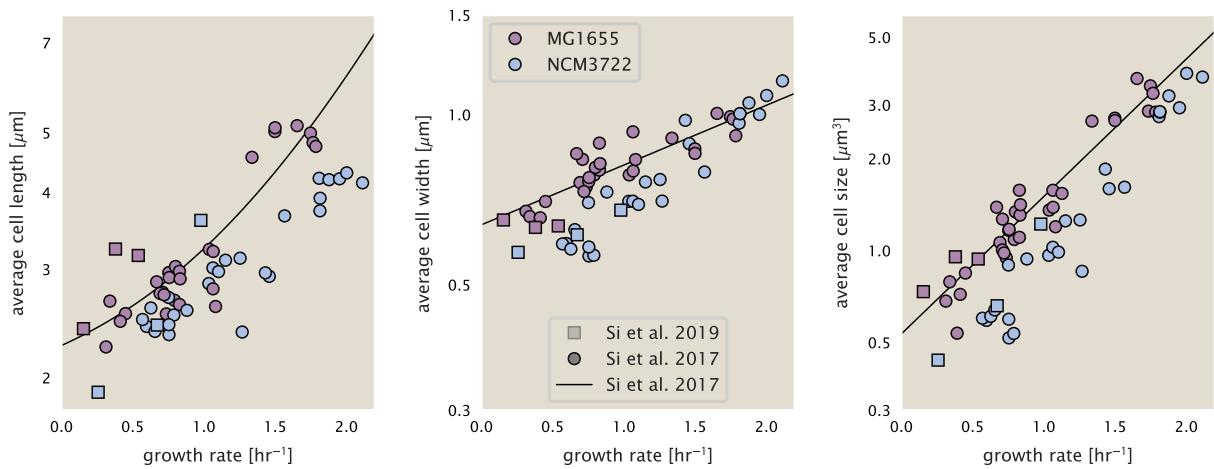


Figure 14. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 images and 5,000-30,000 for each set of samples; which likely means about 1,000-5,000 measurements per mean value reported here since they considered about 6 conditions at a time). Fits were made to the MG1655 strain data; length: $0.5 e^{1.09 \cdot i} + 1.76 \mu\text{m}$, width: $0.64 e^{0.24 \cdot i} \mu\text{m}$. (C) Cell size, V , was calculated as cylinders with two hemispherical ends (Equation 4). The MG1655 strain data gave a best fit of $0.533 e^{1.037 \cdot i} \mu\text{m}^3$.

797 Estimation of Total Protein Content per Cell

798 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell size from the
799 fit shown in Figure **Figure 14(C)** ($0.533 e^{1.037\lambda} \mu\text{m}^3$). We then estimate the total protein content from the total dry
800 mass of the cell. Here we begin by noting that for almost the entire range of growth rates considered here, protein,
801 DNA, and RNA were reported to account for at least 90 % of the dry mass (*Basan et al. (2015)*). The authors also
802 found that the total dry mass concentration was roughly constant across growth conditions. Under such a scenario,
803 we can calculate the total dry mass concentration for protein, DNA, and RNA, which is given by $1.1 \text{ g/ml} \times 30\% \times 90\%$
804 % or about $[M_p] = 300 \text{ fg per fl}$. Multiplying this by our prediction of cell size gives the total dry mass per cell.

805 However, even if dry mass concentration is relatively constant across growth conditions, it is not obvious how
806 protein concentration might vary due to the substantial increase in rRNA at faster growth rates (*Dai et al. (2016)*).
807 This is a well-documented result that arises from an increase in ribosomal abundance at faster growth rates (*Scott*
808 *et al. (2010)*). To proceed therefore rely on experimental measurements of total DNA content per cell that also
809 come from Basan *et al.*, and RNA to protein ratios that were measured in Dai *et al.* (and cover the entire range of
810 growth conditions considered here). These are reproduced in Figure **Figure 15(A)** and (B), respectively.

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

$$M_p = \frac{[M_p + M_{RNA}]}{1 + (RP_{ratio})}. \quad (6)$$

814 (RP_{ratio} refers to the RNA to protein ratio as measured by Dai *et al.*. In Figure **Figure 15(C)** we plot the estimated
815 cellular concentrations for protein, DNA, and RNA from these calculations, and in Figure **Figure 15(D)** we plot their
816 total expected mass per cell. This later quantity is the growth rate-dependent total protein mass that was used to
817 estimate total protein abundance across all data sets (and summarized in **Figure 12(B)**).

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

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

827 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression of the log
828 transform of these parameters as a function of the log transform of the ribosome copy number. Using optimization
829 by minimization, we estimated the best-fit values of the intercept and slope for each regression. **Figure 16** shows
830 the result of each regression as a dashed line.

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

832 While the data set from *Schmidt et al. (2016)* remains a heroic effort that our labs continue to return to as a resource,
833 there were steps taken in their calculation of protein copy number that we felt needed further consideration. In
834 particular, the authors made an assumption of constant cellular protein concentration across all growth conditions
835 and used measurements of cell volume that appear inconsistent with an expected exponential scaling of cell size
836 with growth rate that is well-documented in *E. coli* (*Schaechter et al. (1958); Taheri-Araghi et al. (2015); Si et al.*
837 *(2017)*).

838 We begin by looking at their cell volume measurements, which are shown in blue in Figure **Figure 17**. As a
839 comparison, we also plot cell sizes reported in three other recent papers: measurements from Taheri-Araghi *et*
840 *al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from Basan *et al.* come from the lab of Terence
841 Hwa. Each set of measurements used microscopy and cell segmentation to determine the length and width, and
842 then calculated cell size by treating the cell is a cylinder with two hemispherical ends, as we considered in the

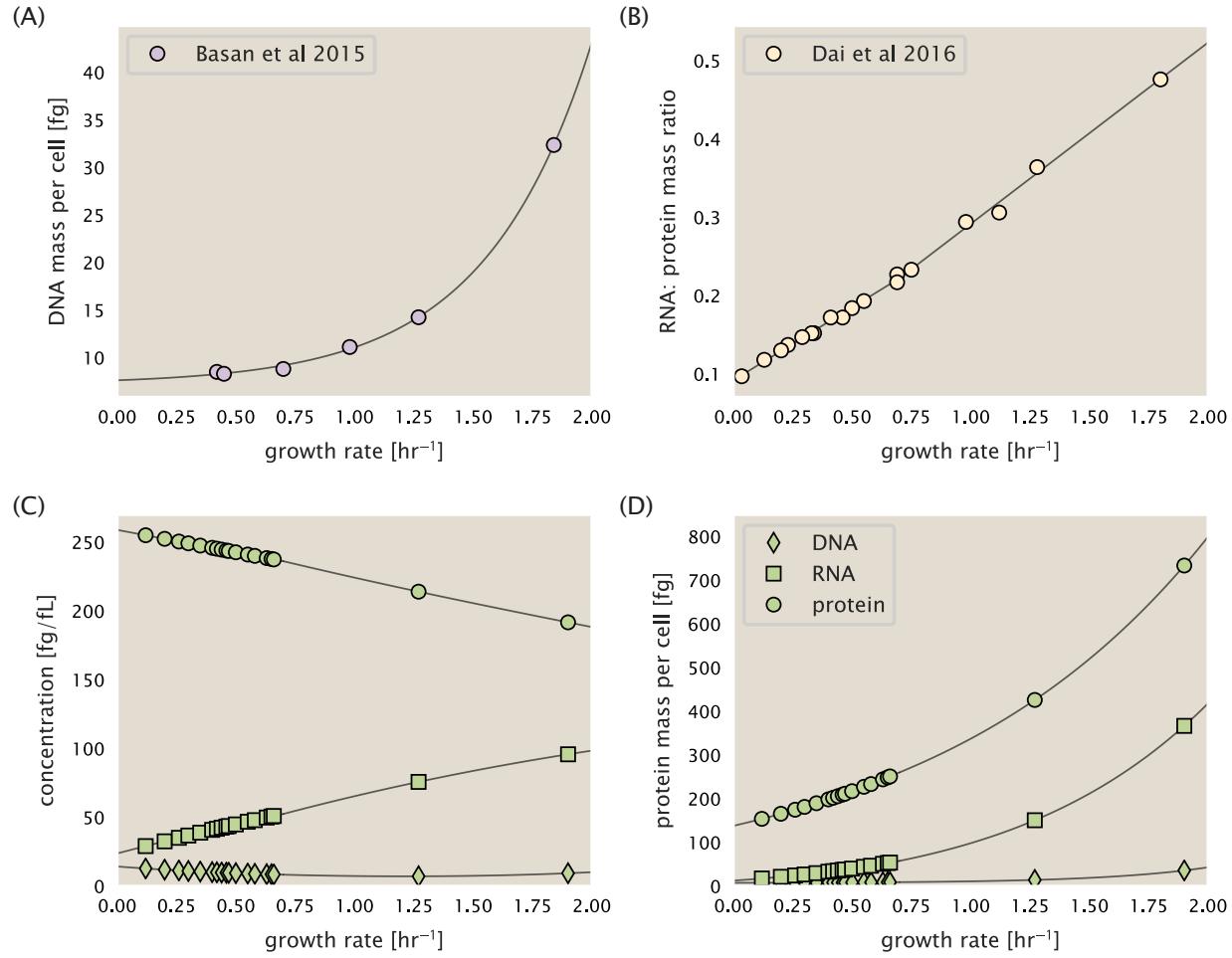


Figure 15. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23 \cdot \lambda} + 7.2$ 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).

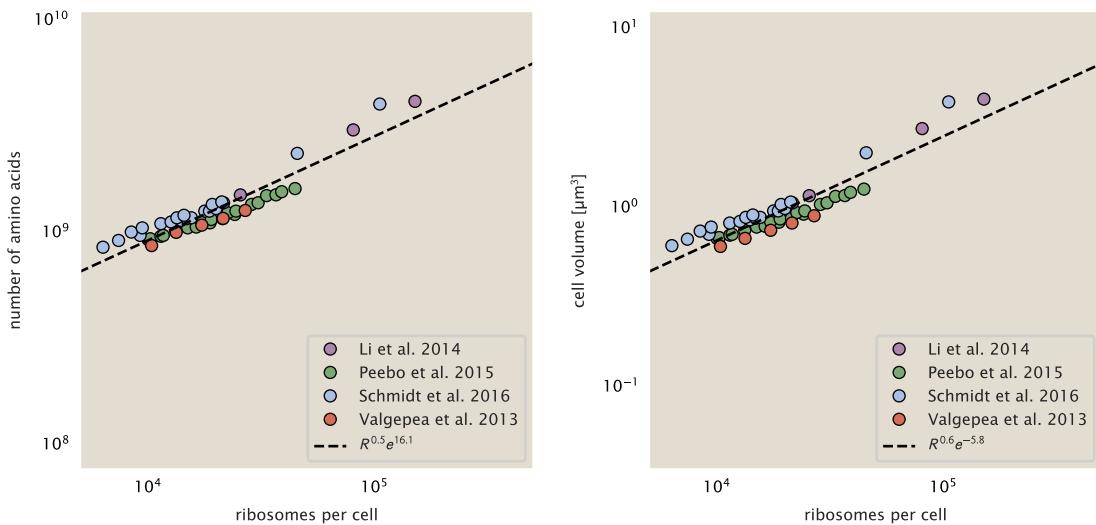


Figure 16. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows the result of the fit, with the functional form of the equation given in the legend with R representing the ribosome copy number.

843 previous section. While there is notable discrepancy between the two research groups, which are both using strain
 844 NCM3722, Basan *et al.* found that this came specifically from uncertainty in determining the cell width. This is prone
 845 to inaccuracy given the small cell size and optical resolution limits (further described in their supplemental text).
 846 Perhaps the more concerning point is that while each of these alternative measurements show an exponential
 847 increase in cell size at faster growth rates, the measurements used by Schmidt *et al.* appear to plateau. This
 848 resulted in an analogous trend in their final reported total cellular protein per cell as shown in Figure **Figure 18**
 849 (purple data points), and is in disagreement with other measurements of total protein at these growth rates (**Basan**
 850 *et al.*, 2015).

851 Since it is not obvious how measurements of cell size influenced their reported protein abundances, in the
 852 following subsections we begin by considering this calculation. We then consider three different approaches to
 853 estimate the growth-rate dependent total protein mass to compare with those values reported from **Schmidt et al.**
 854 (2016). The results of this are summarized in **Figure 17(B)**, with the original values from both **Schmidt et al. (2016)**
 855 and **Li et al. (2014)** shown in **Figure 17(A)** for reference. For most growth conditions, we find that total protein per
 856 cell is still in reasonable agreement. However, for the fastest growth conditions, with glycerol + supplemented
 857 amino acids, and LB media, all estimates are substantially higher than those originally reported. This is the main
 858 reason why we chose to readjusted protein abundance as shown in **Figure 12(B)** (with the calculation described in
 859 section Estimation of Total Protein Content per Cell).

860 Effect of cell volume on reported absolute protein abundances

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

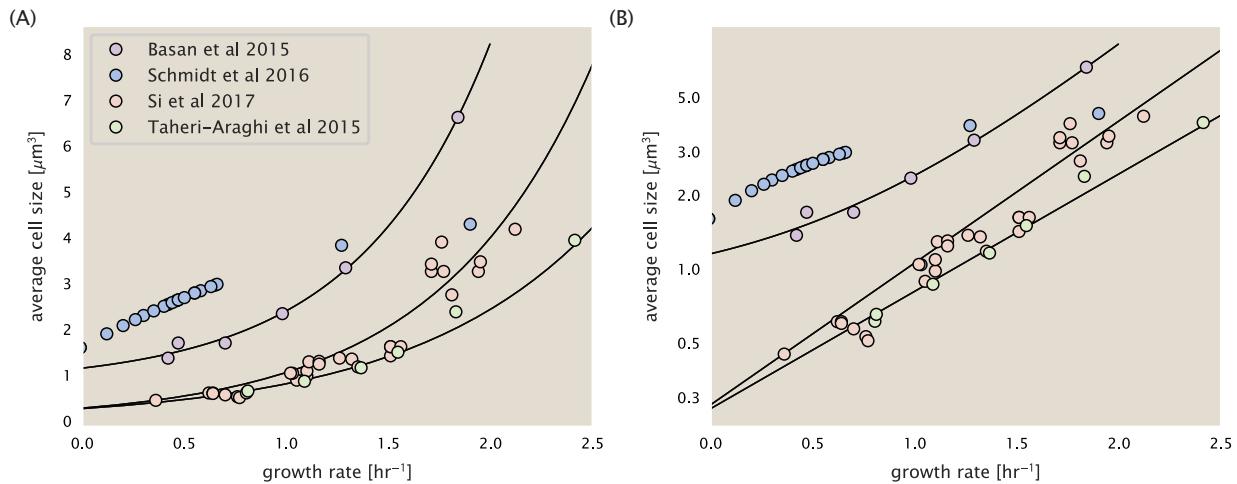


Figure 17. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are 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*.

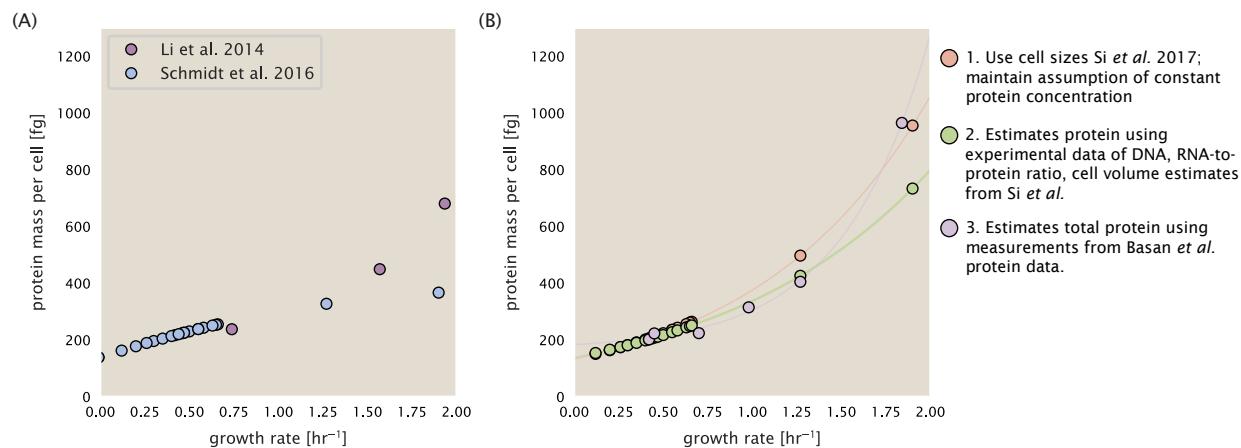


Figure 18. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt et al. (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. light red: Rescaling 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.

871 on a per cell basis.

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

879 For cells grown in M9 minimal media + glucose an average total mass of $M_p = 240$ fg per cell was measured.
880 Using their reported cell volume, reported as $V_{orig} = 2.84$ fl, a cellular protein concentration of $[M_p]_{orig} = M_p/V_{orig} =$
881 85 fg/fl. Now, taking the assumption that cellular protein concentration is relatively independent of growth rate,
882 they could then estimate the total protein mass for all other growth conditions from,

$$M_{P_i} = [M_p]_{orig} \cdot V_i \quad (7)$$

883 where M_{P_i} represents the total protein mass per cell and V_i is the cell volume for each growth condition i as
884 measured in Volkmer and Heinemann, 2011. Here the thinking is that the values of M_{P_i} reflects the total cellular
885 protein for growth condition i , where any discrepancy from their absolute protein abundance is assumed to be due
886 to protein loss during sample preparation. The protein abundances from their absolute abundance measurements
887 noted above were therefore scaled to their estimates and are shown in Figure **Figure 18** (purple data points).

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

$$M'_{P_i} = [M_p]_{Si} \cdot V_{Si_i} \quad (8)$$

893 where M'_{P_i} is the new protein mass prediction, and V_{Si_i} refers to the new volume prediction for each condition i ,
894 These are shown as red data points in Figure **Figure 18(B)**.

895 Relaxing assumption of constant protein concentration across growth conditions

896 We next relax the assumption that cellular protein concentration is constant and instead, attempt to estimate it
897 using experimental data. Here we use the estimation of total protein mass per cell detailed in section **Estimation**
898 of Total Protein Content per Cell for all data points in the **Schmidt *et al.* (2016)** data set. The green data points in
899 **Figure 18(B)** show this prediction, and this represents the approach used to estimate total protein per cell for all
900 data sets.

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

902 One of the challenges in our estimates in the preceding sections is the need to estimate protein concentration
903 and cell volumes. These are inherently difficult to accurately due to the small size of *E. coli*. Indeed, for all the
904 additional measurements of cell volume included in Figure **Figure 17**, no measurements were performed for cells
905 growing at rates below 0.5 hr^{-1} . It therefore remains to be determined whether our extrapolated cell volume
906 estimates are appropriate, with the possibility that the logarithmic scaling of cell size might break down for slower
907 growth.

908 In our last approach we therefore attempt to estimate total protein using experimental data that required no
909 estimates of concentration or cell volume. Specifically, in the work of Basan *et al.*, the authors measured total
910 protein per cell for a broad range of growth rates (reproduced in Figure **Figure 19**). These were determined by
911 first measuring bulk protein from cell lysate, measured by the colorimetric Biuret method (**You *et al.* (2013)**), and
912 then abundance per cell was calculated from cell counts from either plating cells or a Coulter counter. While it
913 is unclear why Schmidt *et al.* was unable to take a similar approach, the results from Basan *et al* appear more
914 consistent with our expectation that cell mass will increase exponentially with faster growth rates. In addition,

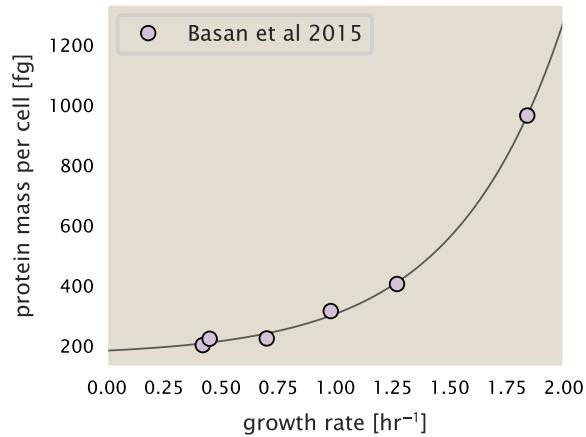


Figure 19. Total cellular protein reported in Basan et al. 2015. Measured protein mass as a function of growth rate as reproduced from Basan et al. 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given by $14.65 e^{2.180 \cdot \lambda} + 172$ fg per cell, where λ is the growth rate in hr^{-1} .

915 although they do not consider growth rates below about 0.5 hr^{-1} , it is interesting to note that the protein mass per
 916 cell appears to plateau to a minimum value at slow growth. In contrast, our estimates using cell volume so far have
 917 predicted that total protein mass should continue to decrease slightly for slower growing cells. By fitting this data
 918 to an exponential function dependent on growth rate, we could then estimate the total protein per cell for each
 919 growth condition considered by Schmidt et al. (2016). These are plotted as red data points in Figure 18(B).

920 Calculation of Complex Abundance

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

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

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

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

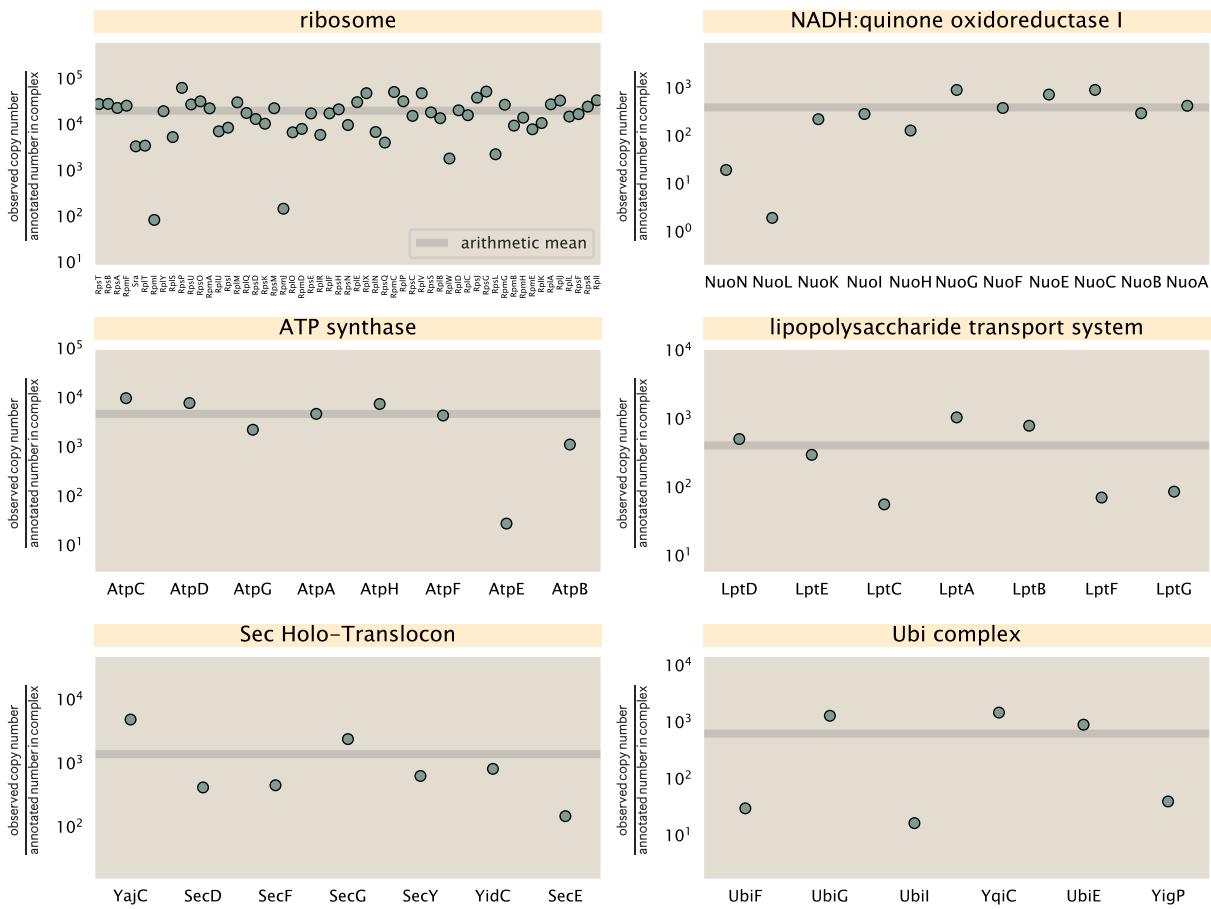


Figure 20. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds to the arithmetic mean across all subunits. These data correspond to measurements from Schmidt et al. (2016) in a glucose-supplemented minimal medium.

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

$$N_{\text{complex}}(\text{subunit } i) = \frac{P_{\text{subunit } i}^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (9)$$

946 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
 947 present for that protein in the complex. For example, the 70S mature ribosome complex has 55 protein components,
 948 all of which are present in a single copy except L4 (RplL), which is present in 4 copies ($m = 4$). For each ribosomal
 949 protein, we then calculate the maximum number of complexes that could be formed using Equation 9. This
 950 example, along with example from 5 other macromolecular complexes, can be seen in Figure 20.

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

$$\langle N_{\text{complex}} \rangle = \frac{1}{N_{\text{subunits}}} \sum_i^{N_{\text{subunits}}} \frac{P_i^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (10)$$

954 in Figure 20, we show this mean value as a grey line for a variety of different complexes. Additionally, we have built
 955 an interactive figure accessible on the [paper website](#) where the validity of this approach can be examined for any
 956 complex with more than two subunits (thus, excluding monomers and dimers).

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

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

962 **Estimation of the total cell mass**

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

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

$$m_{\text{cell}} \approx \rho V(\lambda) \approx \rho a e^{\lambda * b} \quad (11)$$

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

975 **Complex Abundance Scaling With Cell Volume**

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

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

$$m_X \approx 0.3 \times \rho V(\lambda) \theta_X, \quad (12)$$

988 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in **Equation 11**. To convert this to the
989 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
990 with **Equation 12**,

$$N_X \approx \frac{m_X}{w_X}. \quad (13)$$

991 To estimate the number of transporters needed, we make the approximation that loss of units of X via diffusion
992 through porins or due to the permeability of the membrane is negligible and that a single transporter complex
993 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
994 a time window over which the transport process can occur. This is related to the cell doubling time τ , which can
995 be calculated from the the growth rate λ as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized
996 transport scaling relation of

$$N_{\text{transporters}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_X}{w_X r_X \tau}. \quad (14)$$

997 This function is used to draw the continuum estimates for the number of transporters seen in Figures 2 and
 998 3 as transparent grey curves. Occasionally, this continuum scaling relationship will not precisely agree with the
 999 point estimate outlined in the main text. This is due to the choice of ≈ 300 fg total dry mass per cell for the point
 1000 estimate, whereas we considered more precise values of cell mass in the continuum estimate. We note, however,
 1001 that both this scaling relation and the point estimates are meant to describe the order-of-magnitude observed,
 1002 and not the predict the exact values of the abundances.

1003 **Equation 14** is a very general relation for processes where the cell volume is the "natural variable" of the
 1004 problem. This means that, as the cell increases in volume, the requirements for substance X also scale with
 1005 volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the
 1006 dry mass attributable to the substance, and the formula mass of the substance is known, **Equation 14** can be used
 1007 to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell,
 1008 **Equation 14** can be slightly modified to the form

$$N_{\text{ATP synthases}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_{\text{protein}} N_{\text{ATP}}}{w_{AA} r_{\text{ATP}} \tau}, \quad (15)$$

1009 where we have included the term N_{ATP} to account for the number of ATP equivalents needed per amino acid for
 1010 translation (≈ 4 , BNID: 114971), and w_{AA} is the average mass of an amino acid. The grey curves in Figure 4 o the
 1011 main text were made using this type of expression.

1012 A Relation for Complex Abundance Scaling With Surface Area

1013 In our estimation for the number of complexes needed for lipid synthesis and peptidoglycan maturation, we used
 1014 a particular estimate for the cell surface area ($\approx 5 \mu\text{m}$, BNID: 101792) and the fraction of dry mass attributable to
 1015 peptidoglycan ($\approx 3\%$, BNID: 101936). Both of these values come from glucose-fed *E. coli* in balance growth. As we
 1016 are interested in describing the scaling as a function of the growth rate, we must also consider how these values
 1017 scale with cell surface area, which is the natural variable for these types of processes. In the coming paragraphs,
 1018 we highlight how we incorporate a condition-dependent surface area into our calculation of the number of lipids
 1019 and murein monomers that need to be synthesized and crosslinked, respectively.

1020 Number of Lipids

1021 To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as
 1022 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
 1023 of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth rate. Using these approximations combined with
 1024 **Equation 5**, and recognizing that each membrane is composed of two leaflets, we can compute the number of
 1025 lipids as a function of growth rate as

$$N_{\text{lipids}}(\lambda) \approx \frac{4 \text{ leaflets} \times 0.4 \times \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{A_{\text{lipid}}} \quad (16)$$

1026 where η is the length-to-width aspect ratio and V is the cell volume.

1027 Number of Murein Monomers

1028 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we used an empirical
 1029 measurement that $\approx 3\%$ of the dry mass is attributable to peptidoglycan and that a single murein monomer is
 1030 $m_{\text{murein}} \approx 1000 \text{ Da}$. While the latter is independent of growth rate, the former is not. As the peptidoglycan exists as
 1031 a thin shell with a width of $w \approx 10 \text{ nm}$ encapsulating the cell, one would expect the number of murein monomers
 1032 scales with the surface area of this shell. In a similar spirit to our calculation of the number of lipids, the total
 1033 number of murein monomers as a function of growth rate can be calculated as

$$N_{\text{murein monomers}}(\lambda) \approx \frac{\rho_{\text{pg}} w \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{m_{\text{murein}}}, \quad (17)$$

1034 where ρ_{pg} is the density of peptidoglycan.

1035 **Complex Abundance Scaling With Number of Origins, and rRNA Synthesis**

1036 While the majority of our estimates hinge on the total cell volume or surface area, processes related to the central
1037 dogma, namely DNA replication and synthesis of rRNA, depend on the number of chromosomes present in the
1038 cell. As discussed in the main text, the ability of *E. coli* to parallelize the replication of its chromosome by having
1039 multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived in
1040 *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
1041 average number of origins of replication at a given growth rate can be calculated as

$$\langle \#ori \rangle \approx 2^{t_{cyc} \lambda / \ln 2} \quad (18)$$

1042 where t_{cyc} is the total time of replication and division. We can make the approximation that $t_{cyc} \approx 70$ min, which is
1043 the time it takes two replisomes to copy an entire chromosome.

1044 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of replication. Thus,
1045 at a given growth rate λ , the average dosage of rRNA operons per cell D_{rRNA} is

$$D_{rRNA}(\lambda) \approx N_{rRNA \text{ operons}} \times 2^{t_{cyc} \lambda / \ln 2}. \quad (19)$$

1046 This makes the approximation that *all* rRNA operons are localized around the origin. In reality, the operons are
1047 some distance away from the origin, making **Equation 19** an approximation (*Dennis et al., 2004*).

1048 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
1049 approximation, **Equation 18** illustrates that is not precisely true, even at slow growth rates. In estimating the
1050 number of RNA polymerases as a function of growth rate, we consider that regardless of the number of rRNA
1051 operons, they are all sufficiently loaded with RNA polymerase such that each operon produces one rRNA per
1052 second. Thus, the total number of RNA polymerase as a function of the growth rate can be calculated as

$$N_{\text{RNA polymerase}}(\lambda) \approx L_{\text{operon}} D_{rRNA} \rho_{\text{RNA polymerase}}, \quad (20)$$

1053 where L_{operon} is the total length of an rRNA operon (≈ 4500 bp) and $\rho_{\text{RNA polymerase}}$ is packing density of RNA
1054 polymerase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

1055 **Calculation of active ribosomal fraction.**

1056 In the main text we used the active ribosomal fraction f_a that was reported in the work of *Dai et al. (2016)* to
1057 estimate the active ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We lacked any specific model to
1058 consider how f_a should vary with growth rate, and instead find that the data is well-approximated by fitting to an
1059 exponential curve ($f_a = -0.889 e^{4.6 \cdot \lambda} + 0.922$; dashed line in inset of **Figure 9(C)**). We use this function to estimate f_a
1060 for each of the data points shown in **Figure 9(C)**.

1061 **Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.**

1062 *E. coli* shows robust scaling of cell size with the average the number of origins $\langle \#ori \rangle$ per cell (*Si et al., 2017*).
1063 Since protein makes up a majority of the cell's dry mass, the change in cell size is also a reflection of the changes
1064 in proteomic composition and total abundance across growth conditions. Given the potential constraints on
1065 rRNA synthesis and changes in ribosomal copy number with $\langle \#ori \rangle$, it becomes important to also consider how
1066 protein copy numbers vary with the state of chromosomal replication. This is particularly true when trying to make
1067 sense of the changes in ribosomal fraction and growth-rate dependent changes in proteomic composition at a
1068 mechanistic level. As considered in the main text, it is becoming increasingly apparent that regulation through the
1069 secondary messengers (p)ppGpp may act to limit DNA replication and also reduce ribosomal activity in poorer
1070 nutrient conditions. In this context, both $\langle \#ori \rangle$, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio become important parameters to
1071 consider and keep tract of. An increase in $\langle \#ori \rangle / \langle \#ter \rangle$ ratio in particular, causes a relatively higher gene dosage
1072 in rRNA and r-protein genes due to skew in genes near the origin, where the majority of these are located

1073 In the main text we estimated the change in $\langle \#ori \rangle$ with growth rate using the nutrient-limited wild-type cell
1074 data from *Si et al. (2017)*. We consider their measurements of DNA replication time (t_C , 'C' period of cell division),
1075 total cell cycle time (t_{cyc} , 'C' + 'D' period of cell division), and doubling time τ from wild-type *E. coli* growing across
1076 a range of growth conditions. Here we show how we estimate this parameter, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio

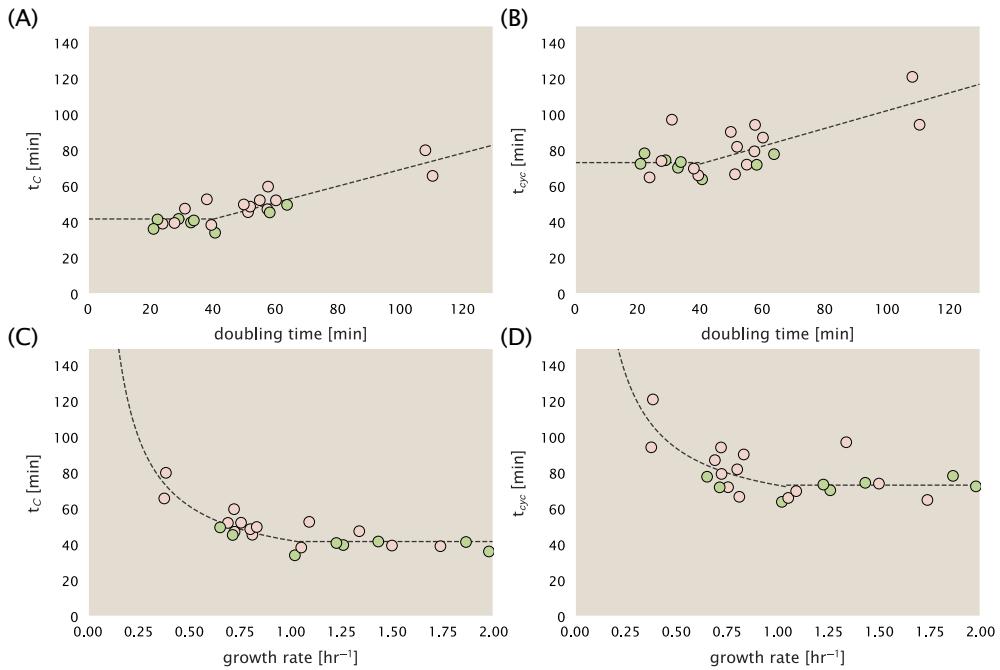


Figure 21. Estimation of $\langle \#ori \rangle / \langle \# ter \rangle$ and $\langle \#ori \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling 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 21(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 21(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.

1092 **References**

- 1093 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia*
1094 *coli*. *BMC Systems Biology*, 8(1):133.
- 1095 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by
1096 ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1097 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,
1098 M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G.
1099 (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1100 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model
1101 for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding
1102 Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- 1103 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH₃ Channel.
1104 *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1105 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006).
1106 Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*,
1107 2(1):2460.
- 1108 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein
1109 synthesis. *Molecular Systems Biology*, 11(10):836.
- 1110 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*,
1111 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 1112 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.
1113 (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proceedings of
1114 the National Academy of Sciences*, 115(21):E4796–E4805.
- 1115 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential
1116 Growth Rates. *EcoSal Plus*, 3(1).
- 1117 Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews
1118 Microbiology*, 14(10):638–650.
- 1119 Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*,
1120 266:2020.06.16.154187.
- 1121 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.,
1122 Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan
1123 Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- 1124 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in *Escherichia coli*.
1125 *Biochemical Journal*, 117(3):551–562.
- 1126 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Translational
1127 Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- 1128 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T.
1129 (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature
1130 Microbiology*, 2(2):16231.
- 1131 Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.
1132 *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1133 Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.
1134 *Microbiology and Molecular Biology Reviews*, 68(4):639–668.
- 1135 Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*,
1136 108(44):17876–17882.
- 1137 Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyru-
1138 vate–carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation.
1139 *Applied Microbiology and Biotechnology*, 94(6):1483–1494.

- 1140 Fernández-Coll, L., Maciag-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M., and Dunny, G. M.
1141 (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of Growth
1142 Rates. *mBio*, 11(2):45.
- 1143 Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair.
1144 *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1145 Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–263.
- 1146 Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of
1147 the National Academy of Sciences*.
- 1148 Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-Hernández,
1149 K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-Lira, H., Bonavides-Martínez, C., Pérez-
1150 Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes, A., Hernández-Koutoucheva, A., Moral-Chávez, V. D.,
1151 Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB version 9.0: High-level integration of gene regulation, coexpression, motif
1152 clustering and beyond. *Nucleic Acids Research*, 44(D1):D133–D143.
- 1153 Godin, M., Delgado, F. F., Son, S., Grover, W. H., Bryan, A. K., Tzur, A., Jorgensen, P., Payer, K., Grossman, A. D., Kirschner, M. W., and
1154 Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390.
- 1155 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-
1156 Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution
1157 on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- 1158 Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in
1159 microbiology*, 26(10):815–832.
- 1160 Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of
1161 (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- 1162 Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content
1163 of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1164 Henkel, S. G., Beek, A. T., Steinsiek, S., Stagge, S., Bettenbrock, K., de Mattos, M. J. T., Sauter, T., Sawodny, O., and Ederer, M. (2014).
1165 Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying Oxygen Conditions. *PLoS ONE*, 9(9):e107640.
- 1166 Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative proteomic
1167 analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2):e784–e784.
- 1168 Ingledew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*, 48(3):222–271.
- 1169 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B., and Phillips, R. (2020).
1170 Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a time. *bioRxiv*.
- 1171 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*,
1172 20(17):4952–4963.
- 1173 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the
1174 future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1175 Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E., Ong,
1176 Q., Ong, W. K., Paley, S. M., and Subhraveti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways.
1177 *Briefings in Bioinformatics*, 20(4):1085–1093.
- 1178 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.
1179 (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–401.
- 1180 Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari,
1181 A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-Ramírez,
1182 D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge about
1183 *Escherichia coli*K-12. *Nucleic Acids Research*, 45(D1):D543–D550.
- 1184 Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that enables the use of
1185 hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of Sciences*, 114(33):E6922–E6931.
- 1186 Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition.
1187 *Current Opinion in Biotechnology*, 28:96–102.

- 1188 Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139(7):1366–
1189 1375.
- 1190 Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*,
1191 125(2):028103.
- 1192 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation in *E. coli* by
1193 Modulating Supercoiling of oriC. *mBio*, 10(4):822.
- 1194 Lascu, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and*
1195 *Biomembranes*, 32(3):237–246.
- 1196 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical
1197 investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1198 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting sets. *IEEE Transactions*
1199 *on Visualization and Computer Graphics*, 20(12):1983–1992.
- 1200 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles
1201 underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1202 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in
1203 cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1204 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon
1205 Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- 1206 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facilitator
1207 GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1208 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*,
1209 112(51):15690–15695.
- 1210 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On*. Gene Expression. Springer.
- 1211 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by
1212 condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1213 Mikucki, J. A., Pearson, A., Johnston, D. T., Turchyn, A. V., Farquhar, J., Schrag, D. P., Anbar, A. D., Priscu, J. C., and Lee, P. A. (2009). A
1214 Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.
- 1215 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular
1216 and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1217 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation.
1218 *Growth Symposium*, 9:223–289.
- 1219 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1220 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall
1221 synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–
1222 12515.
- 1223 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 1224 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria.
1225 *eLife*, 8:642.
- 1226 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- 1227 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with
1228 increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 1229 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1230 Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., and Maranas, C. D. (2012). An
1231 integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering*,
1232 14(6):687–704.

- 1233 Rogers, H., Perkins, H., and Ward, J. (1980). *Microbial Cell Walls and Membranes*. Chapman and Hall, London.
- 1234 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive
1235 strategies. *Nature microbiology*, 1(11):1-7.
- 1236 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of*
1237 *Bacteriology*, 131(2):505-511.
- 1238 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722-
1239 11734.
- 1240 Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical
1241 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592-606.
- 1242 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann,
1243 M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104-110.
- 1244 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression:
1245 origins and consequences. *Science*, 330(6007):1099-1102.
- 1246 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.
1247 *Cell*, 172(6):1294-1305.
- 1248 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
1249 Bacteria. *Current Biology*, 29(11):1760-1770.e7.
- 1250 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
1251 Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278-1287.
- 1252 Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews*,
1253 40(1):133-159.
- 1254 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth
1255 and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1256 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*, 19.
- 1257 Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van
1258 Leeuwenhoek*, 39(1):545-565.
- 1259 Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*.
1260 *Archives of Microbiology*, 113(3):185-189.
- 1261 Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1262 Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95-104.
- 1263 Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Cell-size control and
1264 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385-391.
- 1265 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
1266 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533-538.
- 1267 Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of
1268 protein functions and evolution. *Nucleic Acids Research*, 28(1):33-36.
- 1269 Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic
1270 models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*,
1271 107(2):369-381. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1272 The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*,
1273 47(D1):D330-D338.
- 1274 Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and
1275 translation rates of proteins. *Molecular BioSystems*, 9(9):2344.

- 1276 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,
1277 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,
1278 E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero,
1279 E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0:
1280 Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1281 Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data
1282 Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.
- 1283 Vollmer, W., Blanot, D., and De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*,
1284 32(2):149–167.
- 1285 Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Letters*, 545(1):61–70.
- 1286 Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
1287 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- 1288 You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C. W., Wang, Y.-P., Lenz, P., Yan, D., and Hwa, T. (2013). Coordination of
1289 bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- 1290 Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from
1291 *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 108(46):18643–18648.
- 1292 Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014). Regulation of crp gene expression by the catabolite repressor/activator, cra, in
1293 *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1294 Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in
1295 *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.

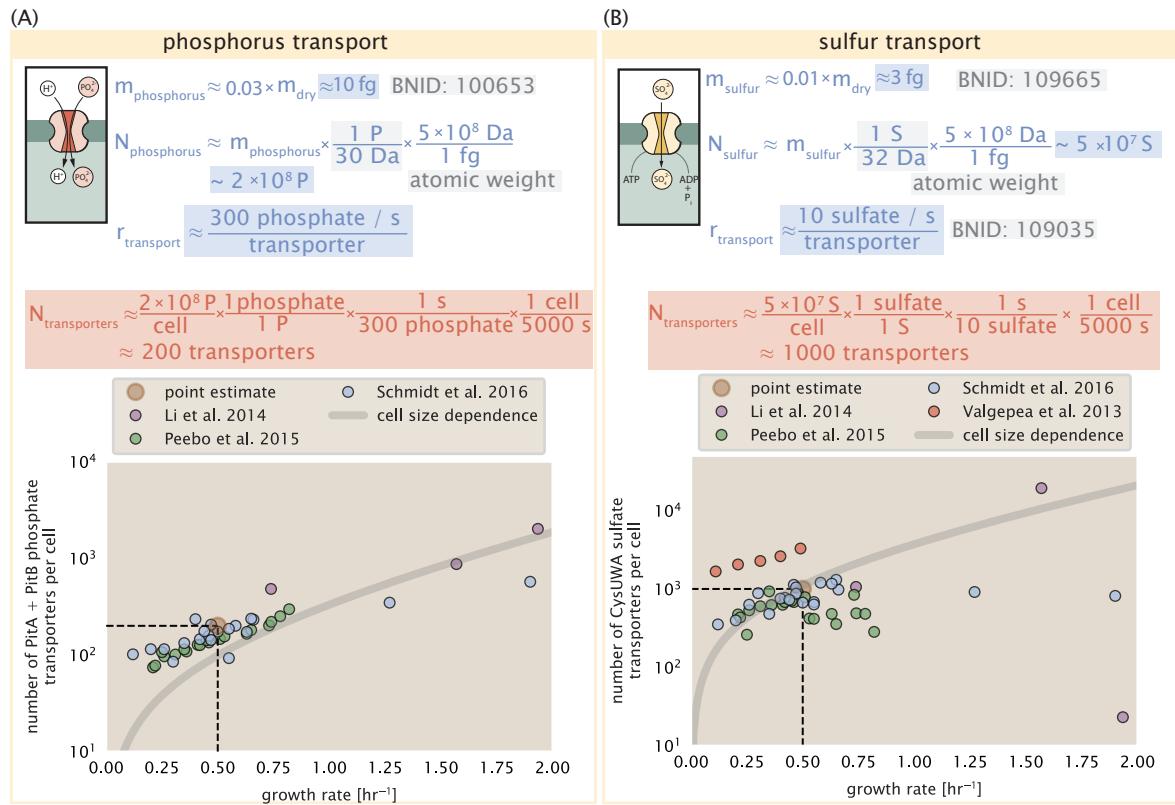


Figure 2-Figure supplement 1. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (B) Estimate of the number of CysUWA complexes necessary to maintain a 1% sulfur *E. coli* dry mass. Points in plot correspond to average number of CysUWA transporter complexes that can be formed given the transporter stoichiometry [CysA]₂[CysU][CysW][Sbp/CysP]. Grey line in (A) and (B) represents the estimated number of transporters per cell at a continuum of growth rates.

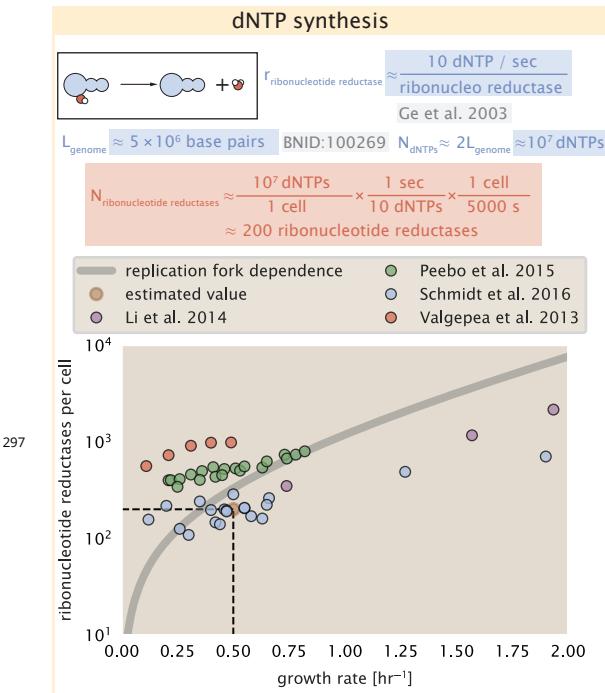


Figure 6–Figure supplement 1. Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ($[\text{NrdA}]_2[\text{NrdB}]_2$) and ribonucleotide reductase II ($[\text{NrdE}]_2[\text{NrdF}]_2$) complexes. Grey lines in top panel show the estimated number of complexes needed as a function of growth, the details of which are described in the Appendix.

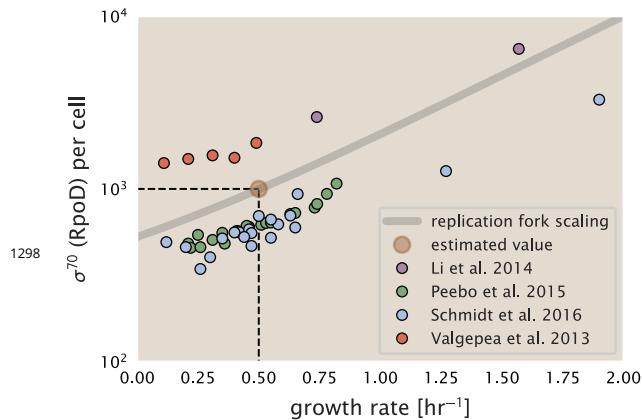


Figure 7–Figure supplement 1. The abundance of σ^{70} as a function of growth rate. Estimated value for the number of RNAP is shown as a translucent brown point and grey line.

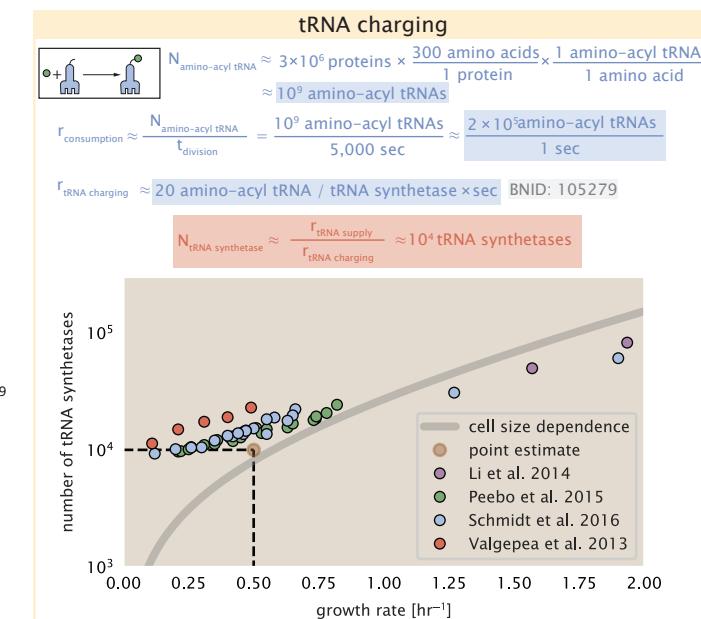


Figure 8-Figure supplement 1. Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [GltX], [IleS], [LeuS], [ValS], [AlaS]₂, [AsnS]₂, [AspS]₂, [TyrS]₂, [TrpS]₂, [ThrS]₂, [SerS]₂, [ProS]₂, [PheS]₂[PheT]₂, [MetG]₂, [lysS]₂, [HisS]₂, [GlyS]₂[GlyQ]₂).

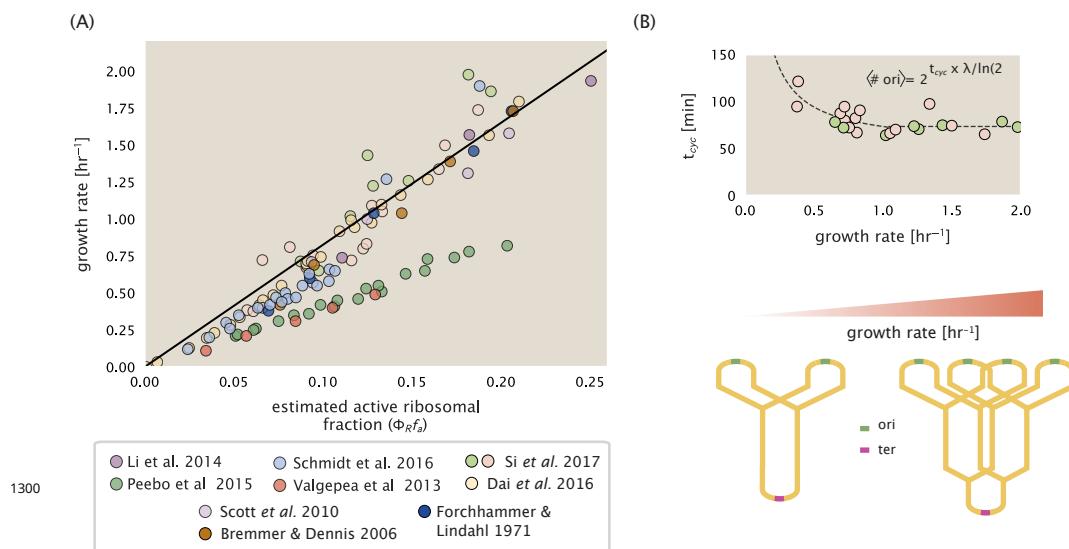


Figure 9-Figure supplement 1. (A) Actively translating ribosomal fraction versus growth rate. The actively translating ribosomal fraction is calculated using the estimated values of f_a from [Dai et al. \(2016\)](#) (shown in inset; see Appendix Calculation of active ribosomal fraction for additional detail). Additional measurements in addition to the proteomic measurements are based on measurements of cellular RNA to protein ratio, with $\Phi_R \approx$ the cellular RNA to protein ratio divided by 2.1 ([Dai et al., 2016](#)). (B) Experimental measurements of the cell doubling time τ and cell cycle time t_{cyc} from [Si et al. \(2017\)](#). Dashed line shows fit to the data, which were used to estimate $\langle \# \text{ori} \rangle$. t_{cyc} was assumed to vary in proportion to τ for doubling times greater than 40 minutes, and reach a minimum value of 73 minutes. See Appendix Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$ for additional details exact estimation of rRNA copy number. Red data points correspond to measurements in strain MG1655, while light green points are for strain NCM3722. Schematic shows the expected increase in replication forks (or number of ori regions) as *E. coli* cells grow faster.