

# Fundamental limits on the rate of bacterial cell division

**Nathan M. Belliveau<sup>†, 1</sup>, Griffin Chure<sup>†, 2</sup>, Christina L. Hueschen<sup>3</sup>, Hernan G. Garcia<sup>4</sup>, Jane Kondev<sup>5</sup>, Daniel S. Fisher<sup>6</sup>, Julie A. Theriot<sup>1, 7, \*</sup>, Rob Phillips<sup>8, 9, \*</sup>**

<sup>1</sup>Department of Biology, University of Washington, Seattle, WA, USA; <sup>2</sup>Department of Applied Physics, California Institute of Technology, Pasadena, CA, USA; <sup>3</sup>Department of Chemical Engineering, Stanford University, Stanford, CA, USA; <sup>4</sup>Department of Molecular Cell Biology and Department of Physics, University of California Berkeley, Berkeley, CA, USA; <sup>5</sup>Department of Physics, Brandeis University, Waltham, MA, USA; <sup>6</sup>Department of Applied Physics, Stanford University, Stanford, CA, USA; <sup>7</sup>Allen Institute for Cell Science, Seattle, WA, USA; <sup>8</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA; <sup>9</sup>Department of Physics, California Institute of Technology, Pasadena, CA, USA; \*Co-corresponding authors. Address correspondence to phillips@pboc.caltech.edu and jtheriot@uw.edu; <sup>†</sup>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 36 growth rates. 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 these estimates, ribosome biogenesis emerges as a primary determinant of growth rate. We expand on this assessment 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 may 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 comparably 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*).

Jacques Monod once remarked that “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*). The quantitative power of this “method” in 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 and physiology that were originally raised by

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

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

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

## 116 Nutrient Transport

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

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

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

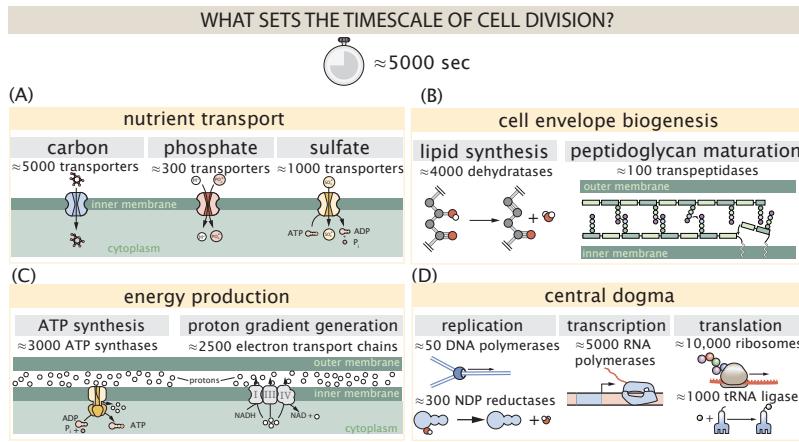
78 This work relies heavily on "back-of-the-envelope" estimates to understand the growth-rate dependent  
79 abundances of molecular complexes. This moniker arises from the limitation that any estimate should be  
80 able to fit on the back of a postage envelope. As such, we must draw a set of rules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

82 **The rule of "one, few, and ten".** The philosophy behind order-of-magnitude estimates is to provide a estimate  
83 of the appropriate scale, not a prediction with many significant digits. We therefore define three different  
84 scales of precision in making estimates. The scale of "one" is reserved for values that range between 1 and 2.  
85 For example, If a particular process has been experimentally measured to transport 1.87 protons for a process  
86 to occur, we approximate this process to require 2 protons per event. The scale of "few" is reserved for values  
87 ranging between 3 and 7. For example, we will often use Avogadro's number to compute the number of  
88 molecules in a cell given a concentration and a volume. Rather than using Avogadro's number as  $6.02214 \times 10^{23}$ ,  
89 we will approximate it as  $5 \times 10^{23}$ . Finally, the scale of "ten" is reserved for values which we know within an  
90 order of magnitude. If a particular protein complex is present at 883 copies per cell, we say that it is present  
91 in approximately  $10^3$  copies per cell. These different scales will be used to arrive at simple estimates that  
92 report the expected scale of the observed data. Therefore, the estimates presented here should not be viewed  
93 as hard-and-fast predictions of precise copy numbers, but as approximate lower (or upper) bounds for the  
94 number of complexes that may be needed to satisfy some cellular requirement.

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

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

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



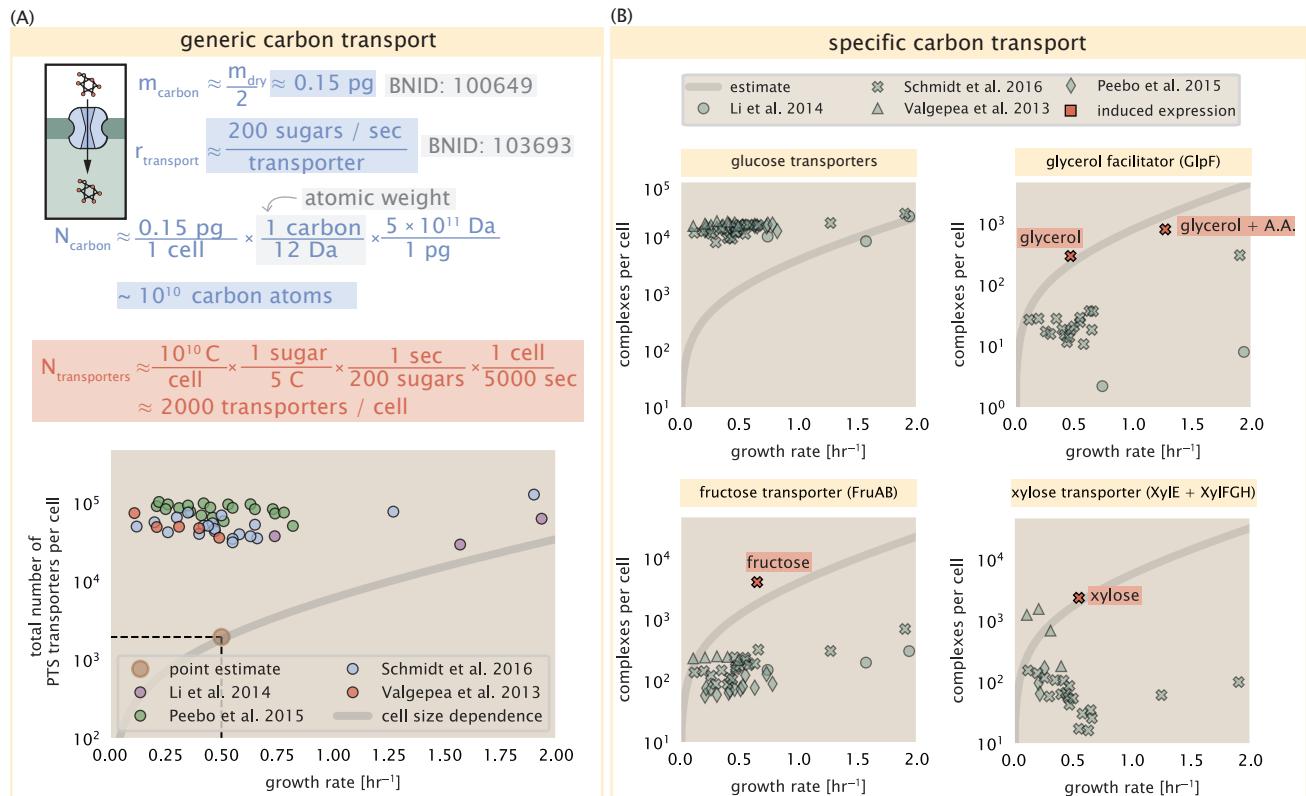
**Figure 1. Transport and synthesis processes necessary for cell division.** We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) cell envelope biogenesis, (C) energy production (namely, ATP synthesis), 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  $0.5 \text{ hr}^{-1}$ , or a cell doubling time of  $\approx 5000$  s.

132 estimate the abundance and growth rate dependence of a variety of transporters responsible for carbon uptake,  
 133 and provide more extensive investigation of the other critical elements – phosphorus, sulfur, and nitrogen – in the  
 134 Appendix ??.

135 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  
 136 approximation that  $\approx 50\%$  of this is carbon, we estimate that  $\sim 10^{10}$  carbon atoms must be brought into the cell  
 137 in order to double all of the carbon-containing molecules (Figure 2(A, top)). Typical laboratory growth conditions  
 138 provide carbon as a single class of sugar (such as glucose, galactose, or xylose) often transported across the cell  
 139 membrane by a specific transporter. One such mechanism of transport is via the PTS system which is a highly  
 140 modular system capable of transporting a diverse range of sugars (Escalante *et al.*, 2012). The glucose-specific  
 141 component of this system transports  $\approx 200$  glucose molecules ( $\approx 1200$  carbon atoms) per second per transporter  
 142 (BNID: 114686). Making the assumption that this is a typical sugar transport rate, coupled with the need to transport  
 143  $\sim 10^{10}$  carbon atoms, we then expect on the order of  $\approx 1000$  transporters must be expressed in order to bring in  
 144 enough carbon atoms (Figure 2(A), top).

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

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



**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  $\approx 5$  C, and each transporter conducts sugar molecules at a rate of  $\approx 200$  per second. Bottom plot shows the estimated number of transporters needed at a growth rate of  $\approx 0.5$  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 differ between these transporter species. To compute the continuum growth rate estimate (grey line), we used the following transport rates for each transporter species: 200 glucose·s<sup>-1</sup> (BNID: 103693), 2000 glycerol·s<sup>-1</sup> (*Lu et al., 2003*), 200 fructose·s<sup>-1</sup> (assumed to be similar to PtsI, BNID: 103693), and 50 xylose·s<sup>-1</sup> (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 lines in (A) and (B) represent 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.

Many metabolic operons are regulated with dual-input logic gates that are only expressed when glucose concentrations are low 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 in terms of number of carbon atoms per molecule and the rate of transport for the particular transporter species. 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.

### Limits on Transporter Expression

If acquisition of nutrients was a limiting process in cell division under the typical growth conditions explored here, the growth rate could be theoretically increased simply by expressing more transporters, but is this feasible at a physiological level? 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/  $\mu\text{m}^2$ . 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/ $\mu\text{m}^2$  (*Phillips, 2018*), implying that the cell could accommodate more membrane and this places additional limitations on cell size and surface area that we will consider further in the coming sections.

## Cell Envelope Biogenesis

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 the 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. In the following paragraphs we will estimate the number of protein complexes needed to produce this membrane surface area as well as the complexes involved in assembling the peptidoglycan scaffold it encapsulates.

### 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  $0.5 \text{ nm}^2$  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  $\approx 1$  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  $\approx 4000$  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

211 the change in membrane surface area as a function of the growth rate (grey line in **Figure 3(A)**), which captures the  
212 observed growth rate dependent expression of these two enzymes.

### 213 Peptidoglycan Synthesis

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

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

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

### 239 Limits on Cell Wall Biogenesis

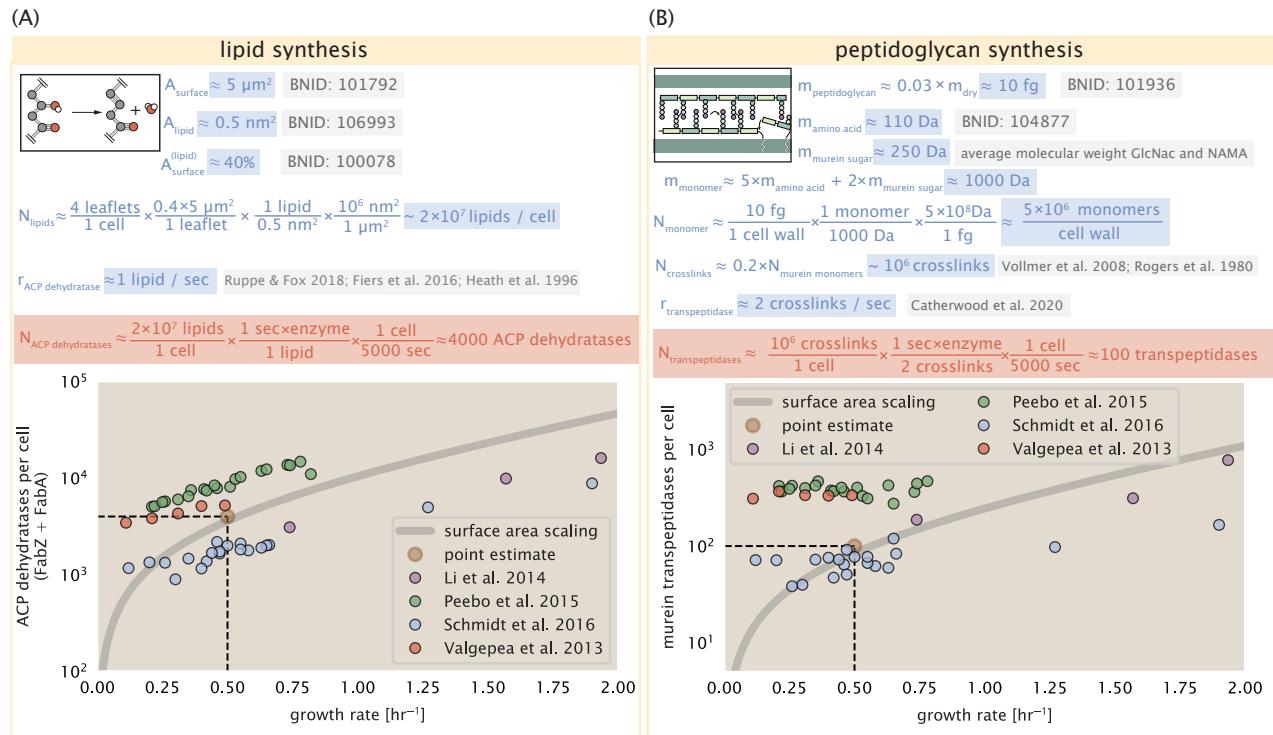
240 While the processes we have considered represent only a small portion of proteins devoted to cell envelope  
241 biogenesis, we find it unlikely that they limit cellular growth in general. The relative amount of mass required for  
242 lipid and peptidoglycan components decrease at faster growth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

248 Considering the assembly of the peptidoglycan, our estimate considers only the transpeptidase enzymes that  
249 are involved lateral and longitudinal elongation of the peptidoglycan. This neglects the presence of other transpep-  
250 tidases that are present in the periplasm and also involved in remodeling and maturation of the peptidoglycan.  
251 It is therefore possible that if this was setting the speed limit for cell division, the simple expression of more  
252 transpeptidases may be sufficient to maintain the structural integrity of the cell wall.

### 253 Energy Production

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

## 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  $\approx 250$  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 **ATP Synthesis**

259 Hydrolysis of the terminal phosphodiester bond of ATP into ADP (or alternatively GTP and GDP) and an inorganic  
260 phosphate provides the thermodynamic driving force in a wide array of biochemical reactions. One such reaction  
261 is the formation of peptide bonds during translation, which requires  $\approx 2$  ATPs for the charging of an amino acid  
262 to the tRNA and  $\approx 2$  GTP for the formation of each peptide bond. Assuming the ATP costs associated with error  
263 correction and post-translational modifications of proteins are negligible, we can make the approximation that each  
264 peptide bond has a net cost of  $\approx 4$  ATP (BNID: 101442, *Milo et al. (2010)*). Formation of GTP from ATP is achieved  
265 via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy investment  
266 (*Lascu and Gonin, 2000*) and therefore consider all NTP requirements of the cell to be functionally equivalent to  
267 being exclusively ATP. In total, the energetic costs of peptide bond formation consume  $\approx 80\%$  of the cells ATP  
268 budget (BNID: 107782; 106158; 101637; 111918, *Lynch and Marinov (2015); Stouthamer (1973)*). The pool of ATP is  
269 produced by the F<sub>1</sub>-F<sub>0</sub> ATP synthase – a membrane-bound rotary motor which under ideal conditions can yield  $\approx$   
270 300 ATP per second (BNID: 114701; *Weber and Senior (2003)*).

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

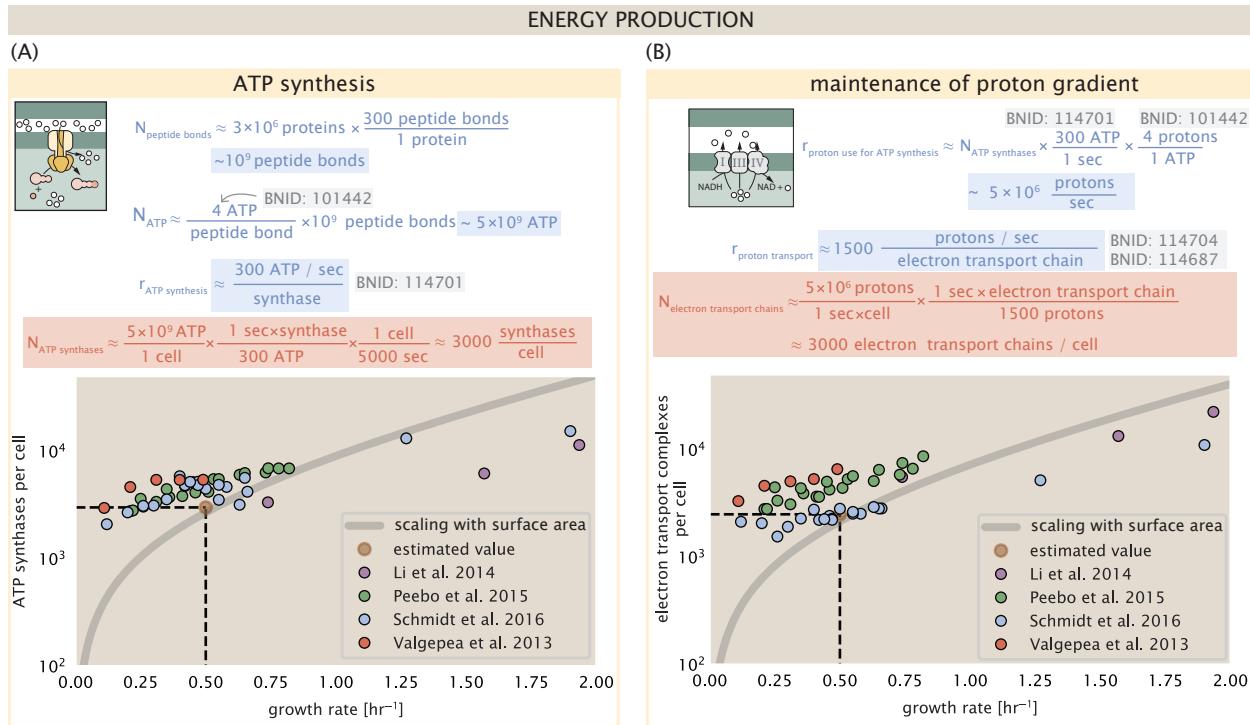
282 **Generating the Proton Electrochemical Gradient**

283 In order to produce ATP, the F<sub>1</sub>-F<sub>0</sub> ATP synthase itself must consume energy. Rather than burning through its  
284 own product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the  
285 electrochemical potential established across the inner membrane through cellular respiration. This electrochemical  
286 gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains as  
287 they reduce NADH. In *E. coli*, this potential difference is  $\approx -200$  mV (BNID: 102120). A simple estimate of the inner  
288 membrane as a capacitor with a working voltage of -200 mV reveals that  $\approx 2 \times 10^4$  protons must be present in the  
289 intermembrane space. However, each rotation of an ATP synthase shuttles  $\approx 4$  protons into the cytosol (BNID:  
290 103390). With a few thousand ATP synthases producing ATP at their maximal rate, the potential difference would  
291 be rapidly abolished in a few milliseconds if it were not being actively maintained.

292 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of intense biochemical  
293 and biophysical study (*Ingledeew and Poole, 1984; Khademian and Imlay, 2017; Cox et al., 1970; Henkel et al., 2014*).  
294 A recent work (*Szenk et al., 2017*) examined the respiratory capacity of the *E. coli* electron transport complexes  
295 using structural and biochemical data, revealing that each electron transport chain rapidly pumps protons into the  
296 intermembrane space at a rate of  $\approx 1500$  protons per second (BNID: 114704; 114687). Using our estimate of the  
297 number of ATP synthases required per cell [*Figure 4(A)*], coupled with these recent measurements, we estimate  
298 that  $\approx 3000$  electron transport complexes would be necessary to facilitate the  $\sim 5 \times 10^6$  protons per second diet of  
299 the cellular ATP synthases. This estimate is in agreement with the number of complexes identified in the proteomic  
300 datasets (plot in *Figure 4(B)*). This suggests that every ATP synthase must be accompanied by  $\approx 1$  functional electron  
301 transport chain.

302 **Limits on Biosynthesis in a Crowded Membrane**

303 Our estimates thus far have focused on biochemistry at the periphery of the cell and have generally been  
304 concordant with the abundances predicted by our estimates. However, as surface area and volume do not scale  
305 identically, it is necessary to consider the physical limits for transport and energy production given the S/V ratio,  
306 which as we've noted will decrease at faster growth rates.



**Figure 4. The abundance of F<sub>1</sub>-F<sub>0</sub> ATP synthases and electron transport chain complexes as a function of growth rate.**

(A) Estimate of the number of F<sub>1</sub>-F<sub>0</sub> 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<sub>1</sub>-F<sub>0</sub> ATP synthase complexes that can be formed given proteomic measurements and the subunit stoichiometry [AtpE]<sub>10</sub>[AtpF]<sub>2</sub>[AtpB][AtpC][AtpH][AtpA]<sub>3</sub>[AtpG][AtpD]<sub>3</sub>. (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<sub>1</sub>-F<sub>0</sub> 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]), *bo<sub>3</sub>*,([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.

307 In our estimate of ATP production above we found that a cell demands about  $5 \times 10^9$  ATP per cell cycle or  $10^6$   
308 ATP/s. With a cell volume of roughly 1 fL (BNID: 100004), this corresponds to about  $2 \times 10^{10}$  ATP per fL of cell volume,  
309 in line with previous estimates (*Stouthamer and Bettenhausen, 1977; Szenk et al., 2017*). In **Figure 5** (A) we plot  
310 this ATP demand as a function of the S/V ratio in green, where we have considered a range of cell shapes from  
311 spherical to rod-shaped with an aspect ratio (length/width) equal to 4. In order to consider the maximum ATP that  
312 could be produced, we consider the amount of ATP that can be generated by a membrane filled with ATP synthase  
313 and electron transport complexes, which provides a maximal production of about 3 ATP / (nm<sup>2</sup>·s) (*Szenk et al.,*  
314 *2017*). This is shown in blue in **Figure 5**(A), which shows that at least for the growth rates observed (right column  
315 in plot), the energy demand is roughly an order of magnitude less. Interestingly, *Szenk et al. (2017)* also found  
316 that ATP production by respiration is less efficient than by fermentation per membrane area occupied due to the  
317 additional proteins of the electron transport chain. This suggests that, even under anaerobic growth, there will be  
318 sufficient membrane space for ATP production.

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

### 331 Processes of the Central Dogma

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

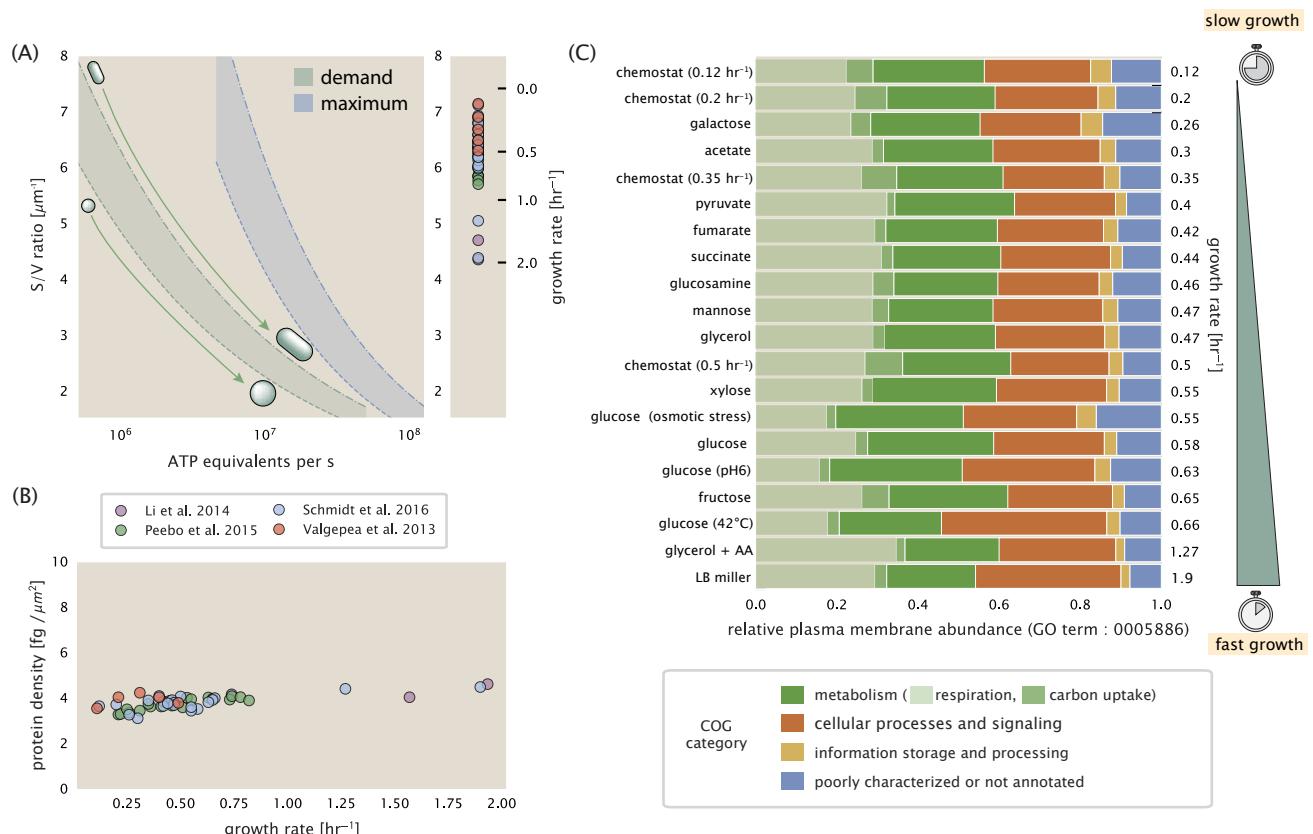
### 336 DNA Replication

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

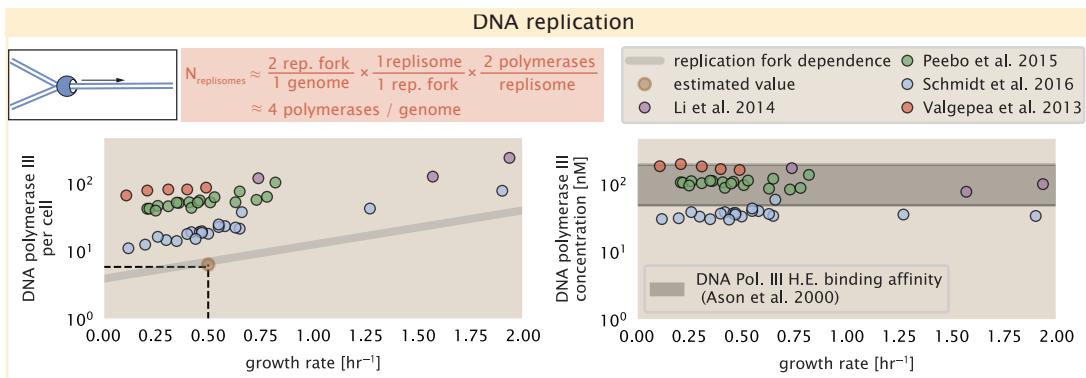
342 Replication is initiated at a single region of the chromosome termed the *oriC* locus at which a pair of replisomes,  
343 each consisting of two DNA polymerase III, begin their high-fidelity replication of the genome in opposite directions  
344 (*Fijalkowska et al., 2012*).

345 *In vitro* measurements have shown that DNA Polymerase III copies DNA at a rate of ≈ 600 nucleotides per  
346 second (BNID: 104120). Therefore, to replicate a single chromosome, two replisomes moving at their maximal rate  
347 would copy the entire genome in ≈ 4000 s. Thus, with a division time of 5000 s, there is sufficient time for a pair of  
348 replisomes complexes to replicate the entire genome.

349 In rapidly growing cultures, bacteria like *E. coli* can initiate as many as 10 - 12 replication forks at a given time  
350 (*Bremer and Dennis, 2008; Si et al., 2017*), we expect only a few DNA polymerases (≈ 10) are needed. However, as  
351 shown in **Figure 6** DNA polymerase III is nearly an order of magnitude more abundant. This discrepancy can be  
352 understood by considering its binding constant to DNA. *In vitro* characterization has quantified the  $K_D$  of DNA  
353 polymerase III holoenzyme to single-stranded and double-stranded DNA to be 50 and 200 nM, respectively (*Ason*  
354 *et al., 2000*). The right-hand plot in **Figure 6** shows that the concentration of DNA polymerase III across all data  
355 sets is within this range. Thus, its copy number appears to vary such that its concentration is approximately



**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  $\mu\text{m}^2$  calculated for proteins with inner membrane annotation (GO term: 0005886). (C) Relative protein abundances by mass based on COG annotation. Metabolic proteins are further separated into respiration ( $F_1$ - $F_0$  ATP synthase, NADH dehydrogenase I, succinate:quinone oxidoreductase, cytochrome bo<sub>3</sub> 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)).



**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 ( $[\text{DnaE}]_3[\text{DnaQ}]_3[\text{HolE}]_3[\text{DnaX}]_5[\text{HolB}][\text{HolA}][\text{DnaN}]_4[\text{HolC}]_4[\text{HolD}]_4$ ) per cell. Right-hand plot shows the effective concentration of DNA polymerase III holoenzyme (See Appendix Estimation of Cell Size and Surface Area for calculation of cell size). 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.

356 equal to the dissociation constant to the DNA. While the processes regulating the initiation of DNA replication are  
357 complex and involve more than just the holoenzyme, these data indicate that the kinetics of replication rather than  
358 the explicit copy number of the DNA polymerase III holoenzyme is the more relevant feature of DNA replication  
359 to consider. In light of this, the data in **Figure 6** suggests that for bacteria like *E. coli*, DNA replication does not  
360 represent a rate-limiting step in cell division. However, it is worth noting that for bacterium like *C. crescentus*  
361 whose chromosomal replication is initiated only once per cell cycle (Jensen et al., 2001), the time to double their  
362 chromosome indeed represents an upper limit to their growth rate.

### 363 RNA Synthesis

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

368 rRNA serves as the catalytic and structural framework along with myriad ribosomal proteins as part of a  
369 complete ribosomal complex. Each ribosome contains three rRNA molecules of lengths 120, 1542, and 2904  
370 nucleotides (BNID: 108093), meaning each ribosome contains  $\approx 4500$  nucleotides overall.

371 *In vivo* measurements of the kinetics of rRNA transcription have revealed that RNA polymerases are loaded  
372 onto the promoter of an rRNA gene at a rate of  $\approx 1$  per second (BNID: 111997, 102362). If RNA polymerases are  
373 constantly loaded at this rate, then we can assume that  $\approx 1$  functional rRNA unit is synthesized per second per rRNA  
374 operon. While *E. coli* possesses 7 of these operons per chromosome, the fact that chromosome replication can be  
375 parallelized means that the average dosage of rRNA genes can be substantially higher at fast growth rates. At a  
376 growth rate of  $\approx 0.5 \text{ hr}^{-1}$ , however, the average cell has  $\approx 1$  copy of its chromosome and therefore approximately  
377  $\approx 7$  copies of the rRNA operons, therefore producing  $\approx 7$  rRNA units per second. With a 5000 second division  
378 time, this means the cell is able to generate around  $3 \times 10^4$  functional rRNA units, comparable within an order of  
379 magnitude to the number of ribosomes per cell.

380 How many RNA polymerases are then needed to constantly transcribe the required rRNA? If one polymerase is  
381 loaded per second, and the transcription rate is  $\approx 40$  nucleotides per second (BNID: 101094), then the typical  
382 spacing between polymerases will be  $\approx 40$  nucleotides. However, we must note that the polymerase itself has a  
383 footprint of  $\approx 40$  nucleotides (BNID: 107873), meaning that one could expect to find one RNA polymerase per 80  
384 nucleotide stretch of an rRNA gene. With a total length of  $\approx 4500$  nucleotides per operon and 7 operons per cell,

385 the number of RNA polymerases transcribing rRNA at any given time is then  $\approx$  500 per cell.

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

392 Our estimates also neglect other mechanistic features of transcription and transcriptional initiation more  
393 broadly. For example, we acknowledge that some fraction of the RNAP pool is nonspecifically bound to DNA  
394 during its search for promoters from which to begin transcription. Furthermore, we ignore the obstacles that RNA  
395 polymerase and DNA polymerase present to each other as they move along the DNA (**Finkelstein and Greene,**  
396 **2013**). Finally, we neglect the fact that RNA polymerase also require  $\sigma$ -factors for promoter recognition and  
397 transcription initiation (**Browning and Busby, 2016**).

398 While they are the machinery for transcription, RNA polymerase is not sufficient to initiate transcription.  
399 Promoter recognition and initiation of transcription is dependent on the presence of  $\sigma$ -factors, protein cofactors  
400 which bind directly to the polymerase (**Browning and Busby, 2016**). In **Figure 7–Figure Supplement 1**, we show that  
401 the predicted RNA polymerase copy number indeed is more comparable with the abundance of  $\sigma$ -70 (RpoD), the  
402 primary sigma factor in *E. coli*. There therefore remains more to be investigated as to what sets the observed  
403 abundance of RNA polymerase in these proteomic data sets. However, we conclude that our the observed excess  
404 in abundance for RNA polymerase abundances are generally in excess of what appears to be needed for growth,  
405 suggesting that the abundance of RNA polymerase itself is not particularly limiting.

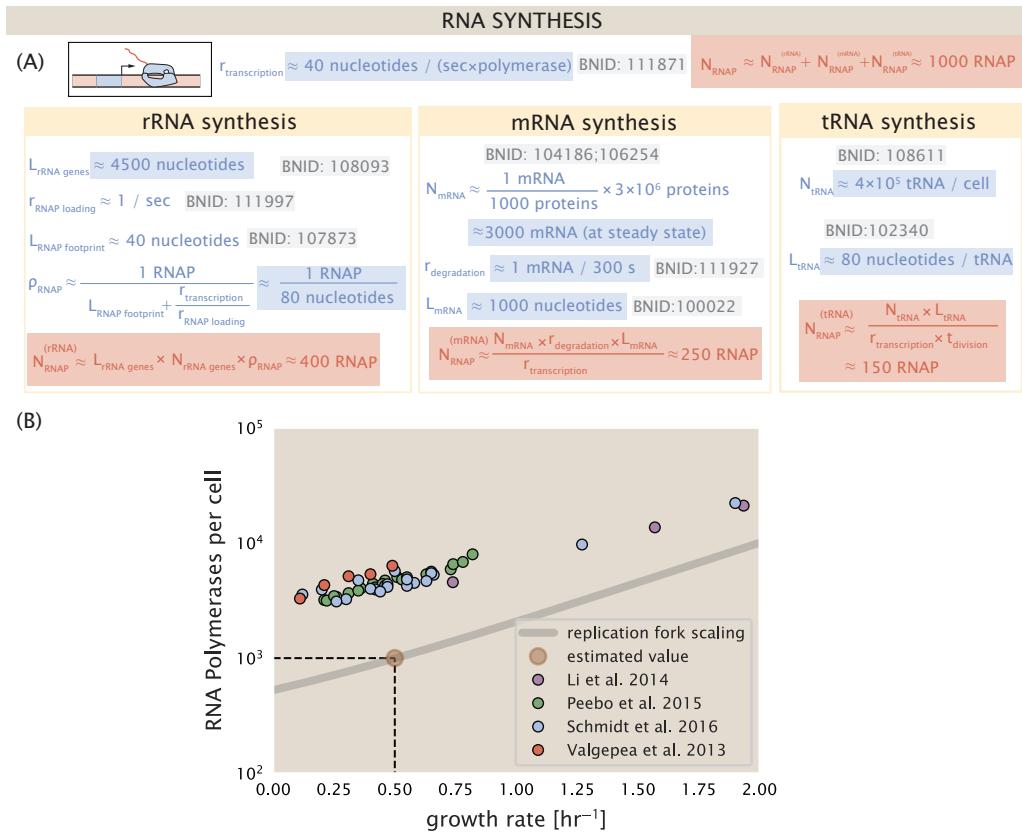
## 406 Protein Synthesis

407 The last process of the central dogma – the translation of RNA into protein – is the final subject in our dialogue  
408 between back-of-the-envelope estimates and comparison with proteomic data. Here we consider the number of  
409 ribosomes needed to replicate the cellular proteome. While the rate at which ribosomes translates is well known to  
410 have a growth rate dependence **Dai et al. (2018)**, here we make the approximation that translation occurs at a rate  
411 of  $\approx$  15 amino acids per second per ribosome (BNID: 100233). Under this approximation and assuming a division  
412 time of 5000 s, we can easily arrive at an estimate of  $\approx$   $10^4$  ribosomes needed to replicate the entire protein mass  
413 (**Figure 8**). This point estimate and the corresponding estimate across a continuum of growth rates proves to be  
414 notably comparable to the experimental observations, shown in **Figure 8(B)**. While the ribosome is responsible for  
415 the formation of peptide bonds, we do not diminish the importance of the charging of tRNAs with the appropriate  
416 amino acid, a process with occurs with remarkable accuracy. In the Appendix and in **Figure 8–Figure Supplement 1**,  
417 we consider the process of ligating tRNAs to their corresponding amino acid and again find notable comparability  
418 with the data.

419 Having now completed our circuit through key processes of cellular growth outlined in **Figure 1**, we can now  
420 take stock on our understanding of the observed growth rate dependence and abundances of various protein  
421 complexes. We note that in general, these simple estimates have been reasonably successful in quantitatively  
422 describing the observations in the proteomic data, suggesting that the cellular milieu is tuned in its size and  
423 composition to match their growth rate requirements, without any one process representing a particular bottleneck  
424 or rate limiting step in division. However, in our effort to identify key limitations on growth, there are two notable  
425 observations that we wish to emphasize.

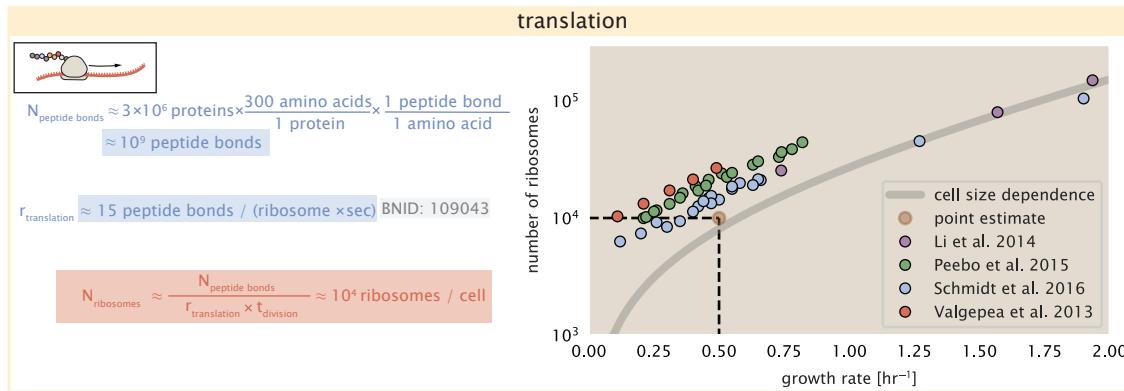
426 The first is a recurring theme throughout our estimates, which is that any inherent biochemical rate limitation  
427 can be overcome by expressing more proteins. We can view this as a parallelization of each biosynthesis task, and  
428 helps explain why bacteria tend to increase their protein content (and cell size) as they grow faster (**Ojikic et al.,**  
429 **2019**). The second, and ultimately the most significant in defining the cellular growth rate, is that the synthesis of  
430 ribosomal proteins presents a special case where parallelization is *not* possible and thereby imposes a limit on  
431 the fastest possible growth rate. This presents an optimization problem for the cell – how are the translational  
432 demands of the entire proteome met without investing resources in producing an excess of ribosomes?

433 This question, more frequently presented as a question of optimal resource allocation, has been the target of an  
434 extensive dialogue between experiment and phenomenological modeling over the past decade. In a now seminal



**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  $\sigma$ -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  $0.5 \text{ hr}^{-1}$ . 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.

work, *Scott et al. (2010)* presents an elegant treatment of resource allocation through partitioning of the proteome into sectors – one of which being ribosome-associated proteins – the relative sizes of which ultimately define the total cellular growth rate. In more recent years, this view has been more thoroughly dissected both experimentally and theoretically (*Klumpp and Hwa, 2014; Basan et al., 2015; Dai et al., 2018, 2016; Erickson et al., 2017*) and together represent a paradigm-shift in how we think of cellular physiology at the proteome-level. However, the quantitative description of the observations is often couched in terms of phenomenological constants and effective parameters with the key observable features of expression being computed in relative rather than absolute abundances. Furthermore, these approaches often exclude or integrate away effects of cell size and chromosome content, which we have illustrated in our estimates to have deep connections to the observed cellular growth rate.

In the closing sections of this work, we explore how ribosome content, cell size, and chromosome content are intertwined in their control over the cellular growth rate. To do so, we take a more detailed view of ribosome abundance exchanging order-of-magnitude estimates for a minimal mathematical model of growth rate control defined by parameters with tangible connections to the biological processes underlying cellular growth. Using this model, we interrogate how the size of the ribosome pool and its corresponding activity is regulated to balance the supply of amino acids via metabolism and catabolism with consumption of amino acids through peptide bond formation.

#### Maximum Growth Rate is Determined by the Ribosomal Mass Fraction

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

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

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

Alternatively, since we are most interested in the role of ribosomal synthesis on growth rate, we instead consider this in terms of the ribosomal mass fraction, denoted by  $\Phi_R$ .  $N_{\text{pep}}$  is related to the total protein mass through the molecular weight of each protein, we can also consider the growth rate in terms of the fraction of the total proteome mass dedicated to ribosomal proteins. By making the approximation that an average amino acid has a molecular weight of 110 Da (BNID: 104877), the total protein mass  $m_{\text{protein}}$  is related to  $N_{\text{pep}}$  by  $(m_{\text{protein}}/110 \text{ Da}) \times N_A$ ,

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

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

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

479 The growth rate defined by **Equation 2** reflects mass-balance under steady state growth and has long provided  
 480 a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function of growth rate (**Maaløe,  
 481 1979; Scott et al., 2010**). The maximum rate, when  $\Phi_R = 1$ , could only be achieved if a cell contained only  
 482 ribosomes (**Dill et al., 2011**). This corresponds to the synthesis time of all ribosomal subunits,  $L_R/r_t \approx 7$  minutes  
 483 and interestingly, is independent of the absolute number of ribosomes. To return to our earlier comments on  
 484 parallelization, it is this step that is rate-limiting, with each ribosome being required to produce a second ribosome.  
 485 Unless elongation rate increased, or cells could trim their total ribosomal protein mass, this dependency limits both  
 486 the maximum growth rate (when  $\Phi_R = 1$ ), and also the achievable growth rate under more moderate values of  $\Phi_R$ .

#### 487 rRNA Synthesis Presents a Potential Bottleneck during Rapid Growth

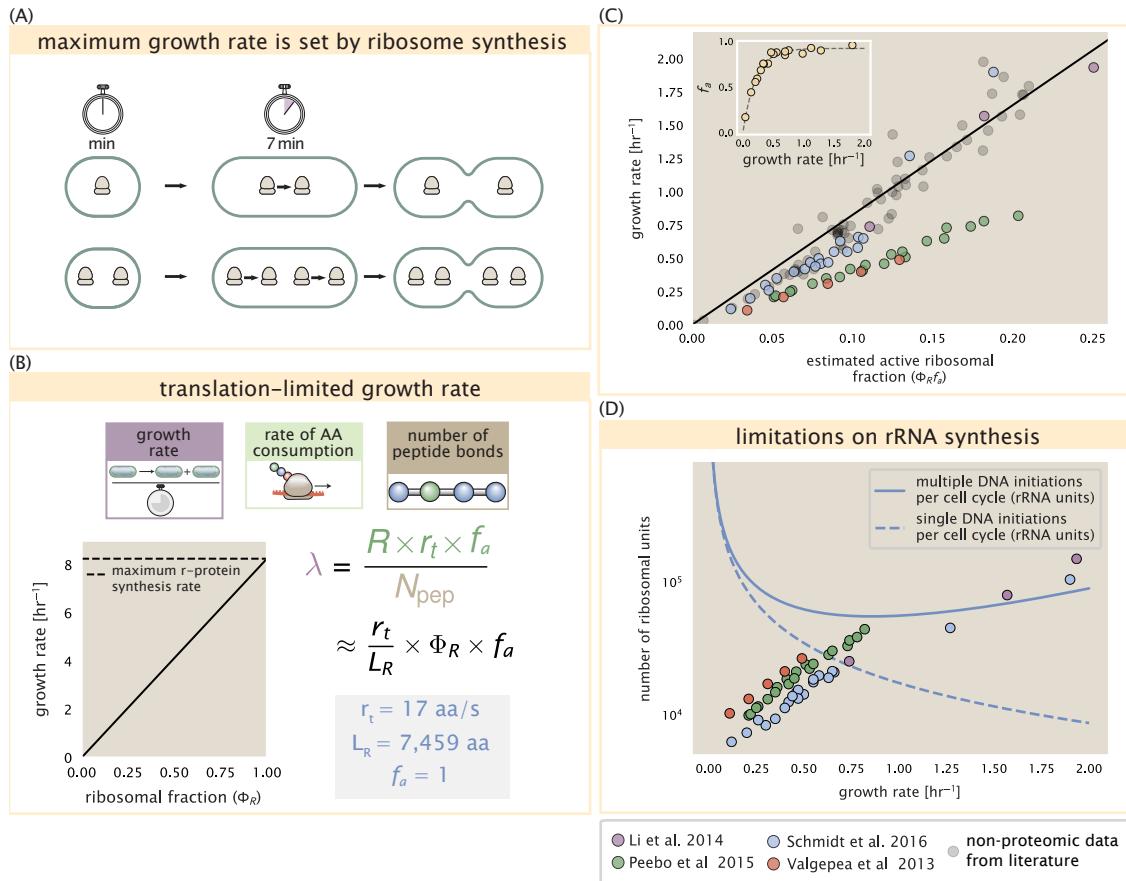
488 *E. coli* rarely exhibits growth rates above  $2 \text{ hr}^{-1}$  (**Bremer and Dennis, 2008**), which is still well-below the synthesis  
 489 rate of a single ribosome, and below the growth rates reported for several other bacteria (**Roller et al., 2016**). Here  
 490 we need to also consider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

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

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

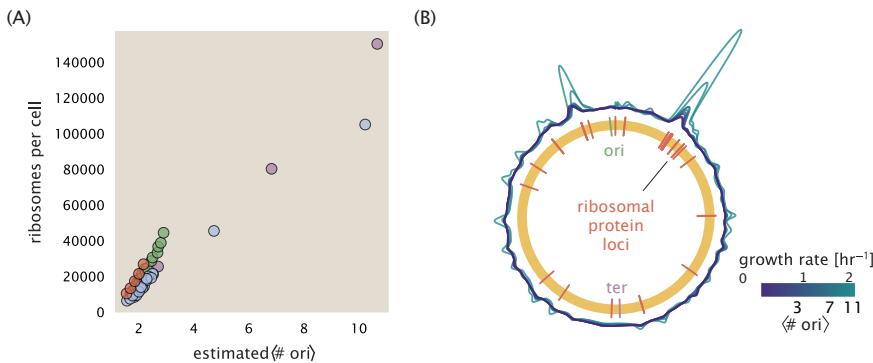
496 Here,  $t_{\text{cyc}}$  is the cell cycle time (referring to the time from replication initiation to cell division), and  $\tau$  is the cell  
 497 doubling time. We used the experimental measurements of  $t_{\text{cyc}}$  and  $\tau$  from **Si et al. (2017)** (**Figure 9–Figure  
 498 Supplement 1(B)**) to calculate  $\langle \# \text{ ori} \rangle$  with **Equation 3** as a function of growth rates. For growth rates above about  
 499 0.5  $\text{hr}^{-1}$ ,  $t_{\text{cyc}}$  is approximately constant at about 70 minutes, which means that  $\langle \# \text{ ori} \rangle$  will grow exponentially  
 500 with growth rate. Since the rRNA operons are predominantly located near to origin of replication (BNID: 100352,  
 501 **Dennis et al. (2004)**), we make the simplifying assumption that that the number of rRNA operons will be directly  
 502 proportional to  $\langle \# \text{ ori} \rangle$ .

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

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

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

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

### 528 A Minimal Model of Nutrient-Mediated Growth Rate Control

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

533 Here we propose that cells modulate their protein abundance in direct response to the availability of nutrients  
 534 in their environment. As noted earlier, bacteria can modulate ribosomal activity through the secondary-messenger  
 535 molecules like (p)ppGpp in poorer nutrient conditions (Figure 9(C) - inset; Dai et al. (2016)). Importantly, these  
 536 secondary-messengers also cause global changes in transcriptional and translational activity (Hauryliuk et al.,  
 537 2015; Zhu and Dai, 2019; Büke et al., 2020). In *E. coli*, amino acid starvation leads to the accumulation of de-acetylated  
 538 tRNAs at the ribosome's A-site and a strong increase in (p)ppGpp synthesis activity by the enzyme RelA (Hauryliuk  
 539 et al., 2015). Along with this, there is increasing evidence that (p)ppGpp also acts to inhibit the initiation of DNA  
 540 replication (Kraemer et al., 2019), providing a potential mechanism to lower (# ori) and maintain a smaller cell size  
 541 in poorer growth conditions (Fernández-Coll et al., 2020).

542 To consider this quantitatively, we assume that cells modulate their proteome ( $N_{\text{pep}}$ ,  $R$ ,  $\Phi_R$ ) 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 ???. 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.

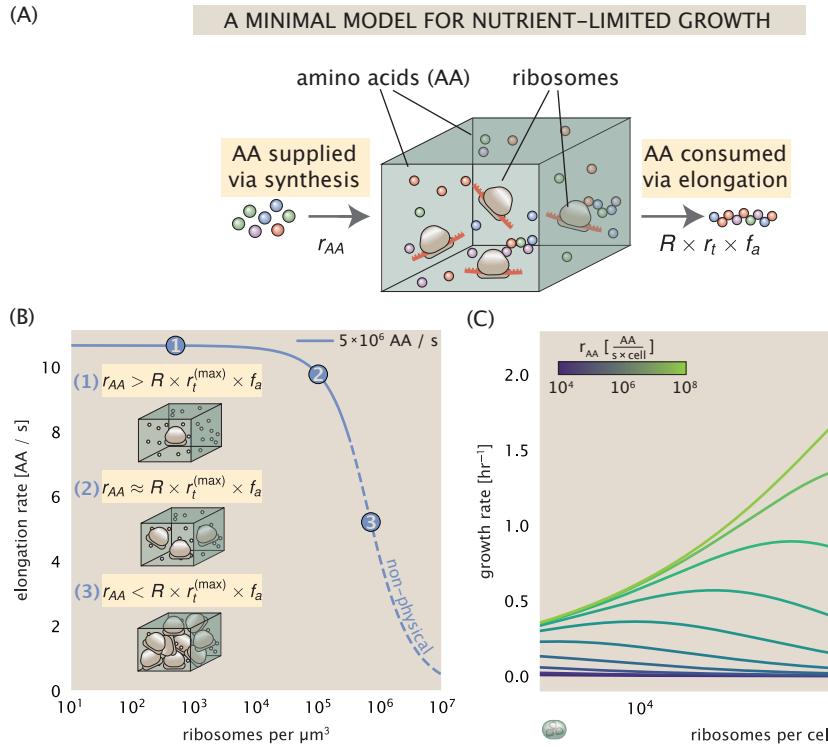
In **Figure 11(B)**, we illustrate how the elongation rate will depend on the ribosomal copy number. Here, we have considered an arbitrarily chosen  $r_{AA} = 5 \times 10^6 \text{ AA} \cdot \text{s}^{-1} \cdot \mu\text{m}^{-3}$  and  $f_a = 1$  for a unit cell volume  $V = 1\text{fL}$ . At low ribosome 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  $R$ ,  $N_{\text{pep}}$ , 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  $R$ ,  $N_{\text{pep}}$ , 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. 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  $r_{AA} = 5 \times 10^6 \text{ AA} / \text{s}$ ,  $r_t^{(\max)} = 17.1 \text{ AA} / \text{s}$ ,  $f_a = 1$ , and a unit cell volume of  $V = 1 \mu\text{l}$ . See Appendix ?? for additional model details. (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 size and total protein abundance  $N_{\text{pep}}$ . We direct the reader to the Supplemental Information for discussion on the inference of the relationship between cell size, number of peptide bonds, and ribosome copy number.

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

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

607 Of the processes illustrated in *Figure 1*, we arrive at a ribosome-centric view of cellular growth rate control.  
608 This is in some sense unsurprising given the long-held observation that *E. coli* and many other organisms vary  
609 their ribosomal abundance as a function of growth conditions and growth rate (*Scott et al., 2010; Metzl-Raz et al.,  
610 2017*). However, through our dialogue with the proteomic data, two additional key points emerge. The first relates  
611 to our question of what process sets the absolute speed limit of bacterial growth. While a cell can parallelize  
612 many of its processes simply by increasing the abundance of specific proteins or firing multiple rounds of DNA  
613 replication, this is not so for synthesis of ribosomes (*Figure 9(A)*). The translation time for each ribosome [ $\approx$  7 min,  
614 *Dill et al. (2011)*] places an inherent limit on the growth rate that can only be surpassed if the cell were to increase  
615 their polypeptide elongation rate, or if they could reduce the total protein and rRNA mass of the ribosome. The  
616 second point relates to the long-observed correlations between growth rate and cell size (*Schaechter et al., 1958; Si  
617 et al., 2017*), and between growth rate and ribosomal mass fraction. While both trends have sparked tremendous  
618 curiosity and driven substantial amounts of research in their own regards, these relationships are themselves  
619 intertwined. In particular, it is the need for cells to increase their absolute number of ribosomes under conditions  
620 of rapid growth that require cells to also grow in size. Further experiments are needed to test the validity of this  
621 hypothesis. In particular, we believe that the change in growth rate in response to translation-inhibitory drugs  
622 (such as chloramphenicol) could be quantitatively predicted, given one had precision measurement of the relevant  
623 parameters, including the fraction of actively translating ribosomes  $f_a$  and changes in the metabolic capacity of the  
624 cell (i.e. the parameter  $r_{AA}$  in our minimal model) for a particular growth condition.

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

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

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

## 653 Methods

### 654 Data Analysis and Availability

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

## 660 Acknowledgements

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

## 666 Competing Interests

667 The authors declare no competing interests.

# 668 Appendix for: Fundamental limits on the 669 rate of bacterial cell division

670 **Nathan M. Belliveau<sup>†, 1</sup>, Griffin Chure<sup>†, 2</sup>, Christina L. Hueschen<sup>3</sup>, Hernan G. Garcia<sup>4</sup>, Jane**  
671 **Kondev<sup>5</sup>, Daniel S. Fisher<sup>6</sup>, Julie A. Theriot<sup>1, 7, \*</sup>, Rob Phillips<sup>8, 9, \*</sup>**

672 <sup>1</sup>Department of Biology, University of Washington, Seattle, WA, USA; <sup>2</sup>Department of Applied Physics,  
673 California Institute of Technology, Pasadena, CA, USA; <sup>3</sup>Department of Chemical Engineering, Stanford  
674 University, Stanford, CA, USA; <sup>4</sup>Department of Molecular Cell Biology and Department of Physics,  
675 University of California Berkeley, Berkeley, CA, USA; <sup>5</sup>Department of Physics, Brandeis University,  
676 Waltham, MA, USA; <sup>6</sup>Department of Applied Physics, Stanford University, Stanford, CA, USA; <sup>7</sup>Allen  
677 Institute for Cell Science, Seattle, WA, USA; <sup>8</sup>Division of Biology and Biological Engineering, California  
678 Institute of Technology, Pasadena, CA, USA; <sup>9</sup>Department of Physics, California Institute of Technology,  
679 Pasadena, CA, USA; \*Co-corresponding authors. Address correspondence to phillips@pboc.caltech.edu  
680 and jtheriot@uw.edu; <sup>†</sup>These authors contributed equally to this work

<b>681</b>	<b>Contents</b>	
<b>682</b>	<b>Introduction</b>	<b>1</b>
<b>683</b>	<b>Nutrient Transport</b>	<b>2</b>
684	Limits on Transporter Expression . . . . .	6
<b>685</b>	<b>Cell Envelope Biogenesis</b>	<b>6</b>
686	Lipid Synthesis . . . . .	6
687	Peptidoglycan Synthesis . . . . .	7
688	Limits on Cell Wall Biogenesis . . . . .	7
<b>689</b>	<b>Energy Production</b>	<b>7</b>
690	ATP Synthesis . . . . .	9
691	Generating the Proton Electrochemical Gradient . . . . .	9
692	Limits on Biosynthesis in a Crowded Membrane . . . . .	9
<b>693</b>	<b>Processes of the Central Dogma</b>	<b>11</b>
694	DNA Replication . . . . .	11
695	RNA Synthesis . . . . .	13
696	Protein Synthesis . . . . .	14
697	Maximum Growth Rate is Determined by the Ribosomal Mass Fraction . . . . .	16
698	rRNA Synthesis Presents a Potential Bottleneck during Rapid Growth . . . . .	17
699	Rapid Growth Requires <i>E. coli</i> to Increase Both Cell Size and Ribosomal Mass Fraction . . . . .	19
<b>700</b>	<b>A Minimal Model of Nutrient-Mediated Growth Rate Control</b>	<b>19</b>
701	Optimal Ribosomal Content and Cell Size Depend on Nutrient Availability and Metabolic Capacity . . . . .	20
<b>702</b>	<b>Discussion</b>	<b>20</b>
<b>703</b>	<b>Methods</b>	<b>23</b>
704	Data Analysis and Availability . . . . .	23
<b>705</b>	<b>Acknowledgements</b>	<b>23</b>
<b>706</b>	<b>Competing Interests</b>	<b>23</b>
<b>707</b>	<b>Experimental Details Behind Proteomic Data</b>	<b>27</b>
708	Fluorescence based measurements . . . . .	27
709	Ribosomal profiling measurements . . . . .	27
710	Mass spectrometry measurements . . . . .	28
<b>711</b>	<b>Summary of Proteomic Data</b>	<b>28</b>
<b>712</b>	<b>Estimation of Cell Size and Surface Area</b>	<b>29</b>
<b>713</b>	<b>Estimation of Total Protein Content per Cell</b>	<b>31</b>
714	Estimating Volume and Number of Amino Acids from Ribosome Copy Number . . . . .	31
<b>715</b>	<b>Additional Considerations of Schmidt <i>et al.</i> Data Set</b>	<b>31</b>
716	Effect of cell volume on reported absolute protein abundances . . . . .	33
717	Relaxing assumption of constant protein concentration across growth conditions . . . . .	35
718	Comparison with total protein measurements from Basan <i>et al.</i> 2015. . . . .	35
<b>719</b>	<b>Calculation of Complex Abundance</b>	<b>36</b>

720	<b>Extending Estimates to a Continuum of Growth Rates</b>	<b>38</b>
721	Estimation of the total cell mass . . . . .	38
722	Complex Abundance Scaling With Cell Volume . . . . .	38
723	A Relation for Complex Abundance Scaling With Surface Area . . . . .	39
724	Number of Lipids . . . . .	39
725	Number of Murein Monomers . . . . .	39
726	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis . . . . .	40
727	<b>Calculation of active ribosomal fraction.</b>	<b>40</b>
728	<b>Estimation of <math>\langle \#ori \rangle / \langle \#ter \rangle</math> and <math>\langle \#ori \rangle</math>.</b>	<b>40</b>

**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 <sup>a</sup>
Soufi <i>et al.</i> (2015)	mass spectrometry	fg/copies per cell
Schmidt <i>et al.</i> (2016)	mass spectrometry	fg/copies per cell <sup>b</sup>

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).

## 729 Experimental Details Behind Proteomic Data

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

### 736 Fluorescence based measurements

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

### 744 Ribosomal profiling measurements

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

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

759 **Mass spectrometry measurements**

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

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

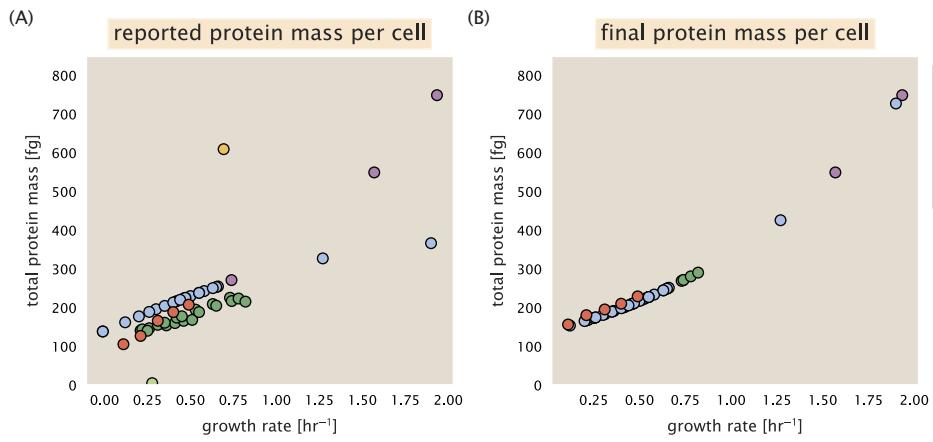
778 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide fragments of known  
779 abundance. These can serve as a direct way to calibrate mass spectrometry signal intensities to absolute mass.  
780 This is the approach taken by *Schmidt et al. (2016)*. In addition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

788 **Summary of Proteomic Data**

789 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)*;  
790 *Schmidt et al. (2016)*. As shown in *Figure 12(A)*, the reported total protein abundances in the work of *Taniguchi*  
791 *et al. (2010)* and *Soufi et al. (2015)* differed quite substantially from the other work. For the work of *Taniguchi et al.*  
792 (*2010*) this is in part due to a lower coverage in total proteomic mass quantified, though we also noticed that most  
793 proteins appear undercounted when compared to the other data.

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

807 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).

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

### 816 Estimation of Cell Size and Surface Area

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

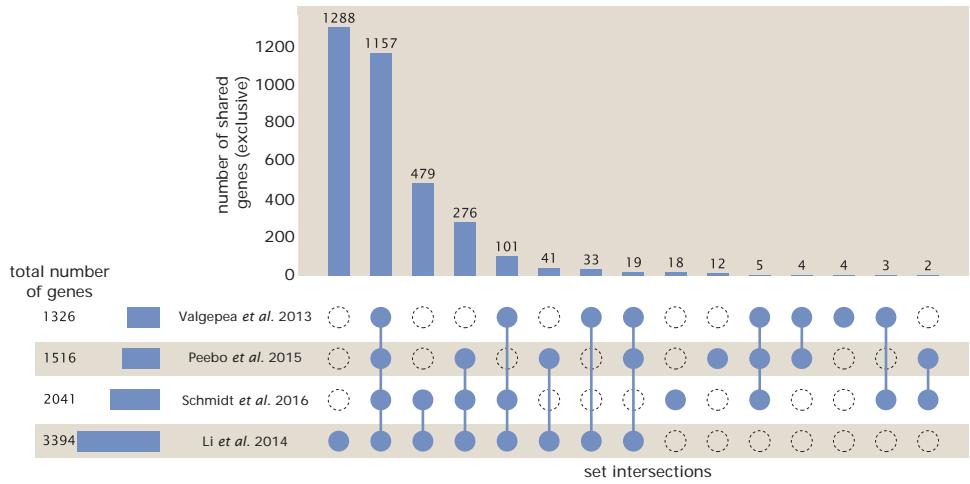
822 The average size measurements from each of their experiments are shown in Figure **Figure 14**, with cell length  
 823 and width shown in (A) and (B), respectively. The length data was well described by the exponential function  $0.5$   
 824  $e^{1.09 \cdot \lambda} + 1.76$   $\mu\text{m}$ , while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

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

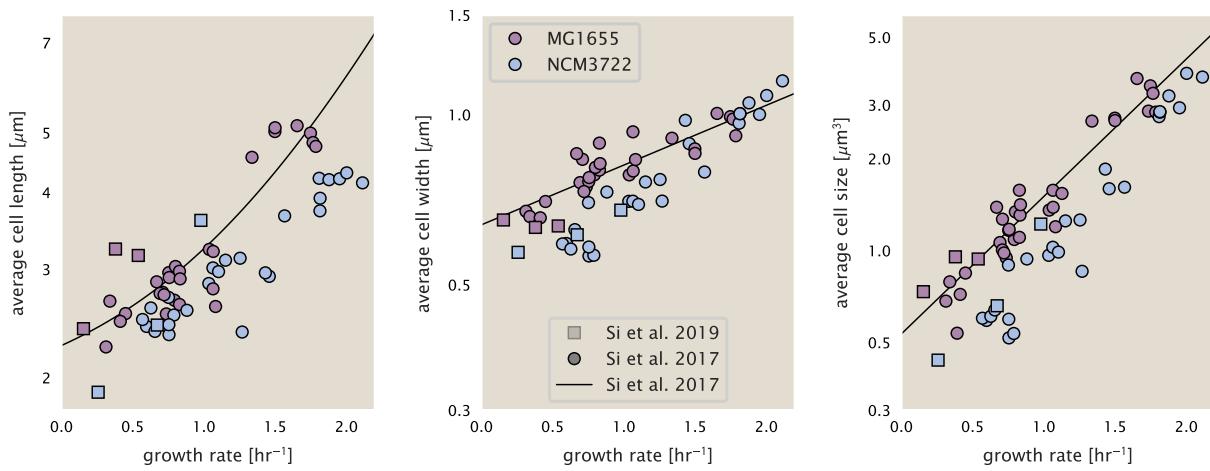
827 where  $r$  is half the cell width. A best fit to the data is described by  $0.533 e^{1.037 \cdot \lambda}$   $\mu\text{m}^3$ . Calculation of the cell surface  
 828 area is given by,

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

829 where  $\eta$  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  $0.533 e^{1.037 \cdot i} \mu\text{m}^3$ .

## 830 Estimation of Total Protein Content per Cell

831 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell size from the  
832 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  
833 mass of the cell. Here we begin by noting that for almost the entire range of growth rates considered here, protein,  
834 DNA, and RNA were reported to account for at least 90 % of the dry mass (*Basan et al. (2015)*). The authors also  
835 found that the total dry mass concentration was roughly constant across growth conditions. Under such a scenario,  
836 we can calculate the total dry mass concentration for protein, DNA, and RNA, which is given by  $1.1 \text{ g/ml} \times 30 \% \times 90$   
837 % or about  $[M_p] = 300 \text{ fg per fl}$ . Multiplying this by our prediction of cell size gives the total dry mass per cell.

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

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

$$847 M_p = \frac{[M_p + M_{RNA}]}{1 + (RP_{ratio})}. \quad (6)$$

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

## 851 Estimating Volume and Number of Amino Acids from Ribosome Copy Number

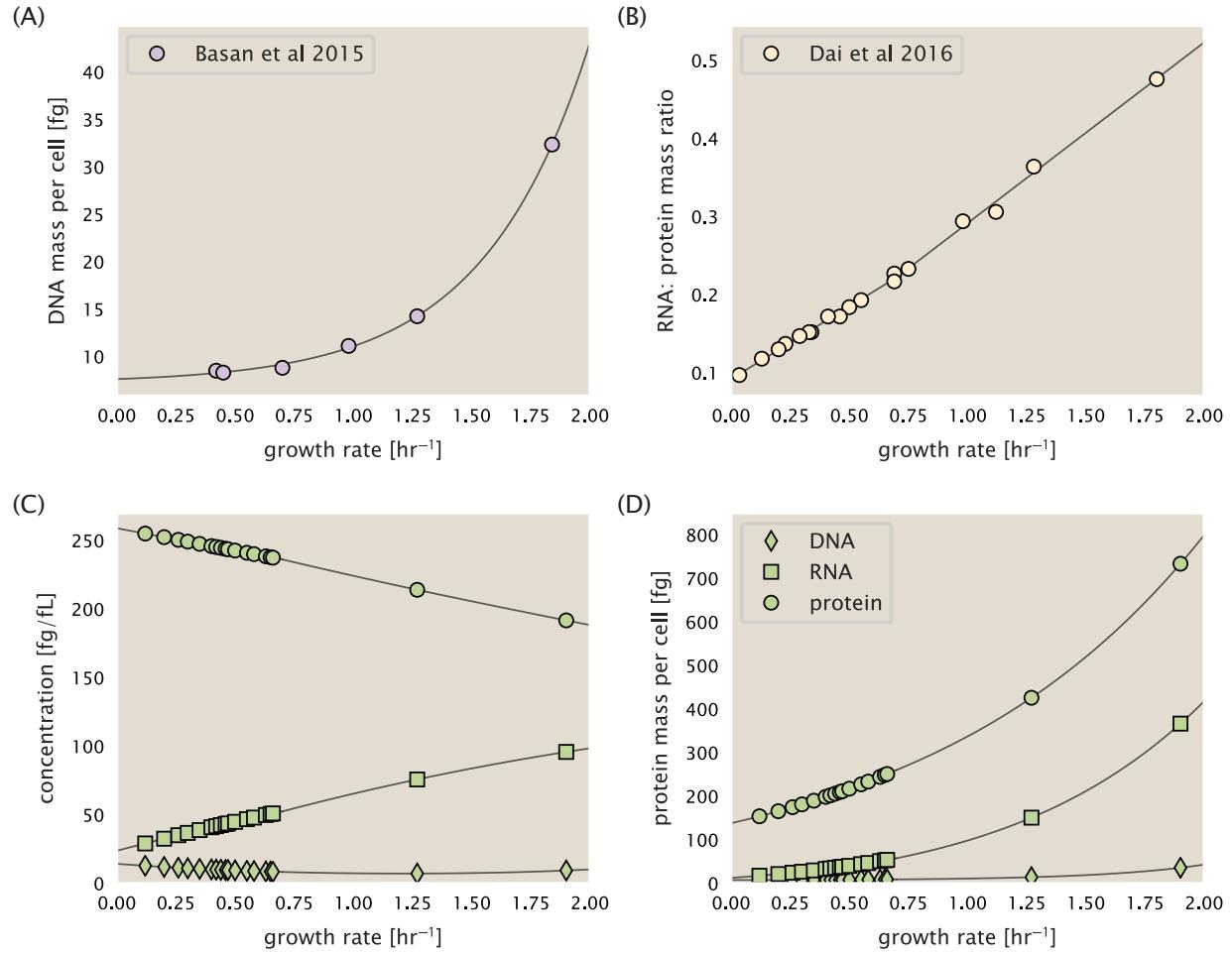
852 Towards the end of the main text, we examine a coarse-grained model of nutrient-limited growth. A key point  
853 in our analysis was to consider how elongation rate  $r_e$  and growth rate  $\lambda$  vary with respect to the experimentally  
854 observed changes in cell size, total number of peptide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

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

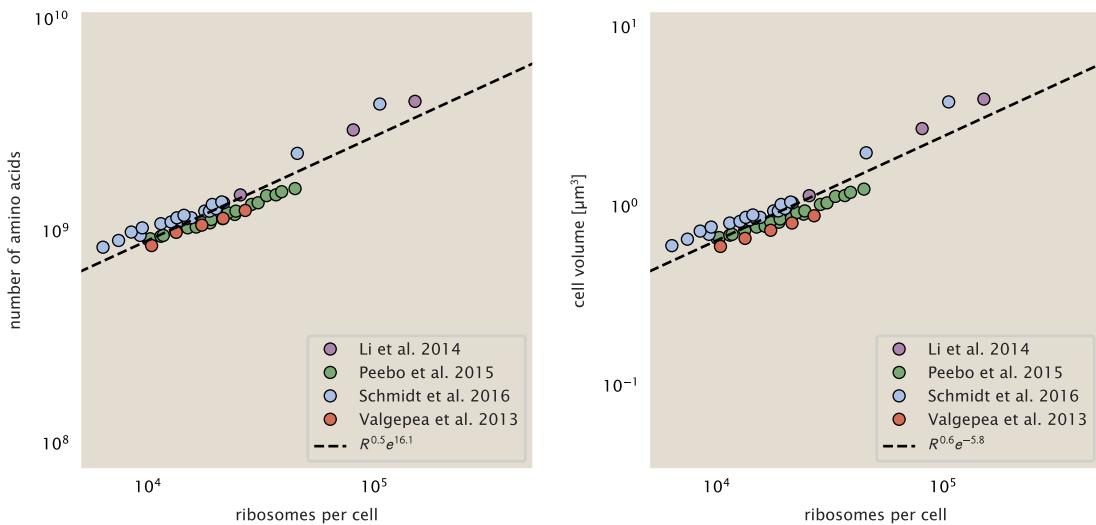
## 864 Additional Considerations of Schmidt *et al.* Data Set

865 While the data set from *Schmidt et al. (2016)* remains a heroic effort that our labs continue to return to as a resource,  
866 there were steps taken in their calculation of protein copy number that we felt needed further consideration. In  
867 particular, the authors made an assumption of constant cellular protein concentration across all growth conditions  
868 and used measurements of cell volume that appear inconsistent with an expected exponential scaling of cell size  
869 with growth rate that is well-documented in *E. coli* (*Schaechter et al. (1958); Taheri-Araghi et al. (2015); Si et al.*  
870 *(2017)*).

871 We begin by looking at their cell volume measurements, which are shown in blue in Figure *Figure 17*. As a  
872 comparison, we also plot cell sizes reported in three other recent papers: measurements from Taheri-Araghi *et*  
873 *al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from Basan *et al.* come from the lab of Terence  
874 Hwa. Each set of measurements used microscopy and cell segmentation to determine the length and width, and  
875 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  $\lambda$  is the growth rate in hr<sup>-1</sup>). (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<sup>-1</sup>, the RNA/protein ratio is  $0.18 \cdot \lambda + 0.093$ , while for growth rates faster than 0.7 hr<sup>-1</sup> 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).



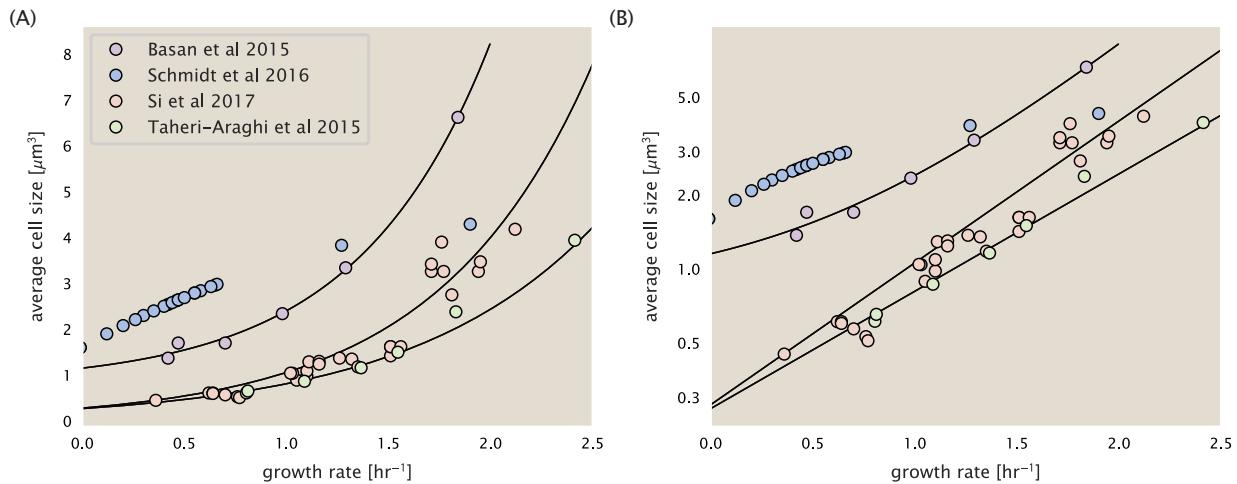
**Figure 16. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number.** Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line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

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

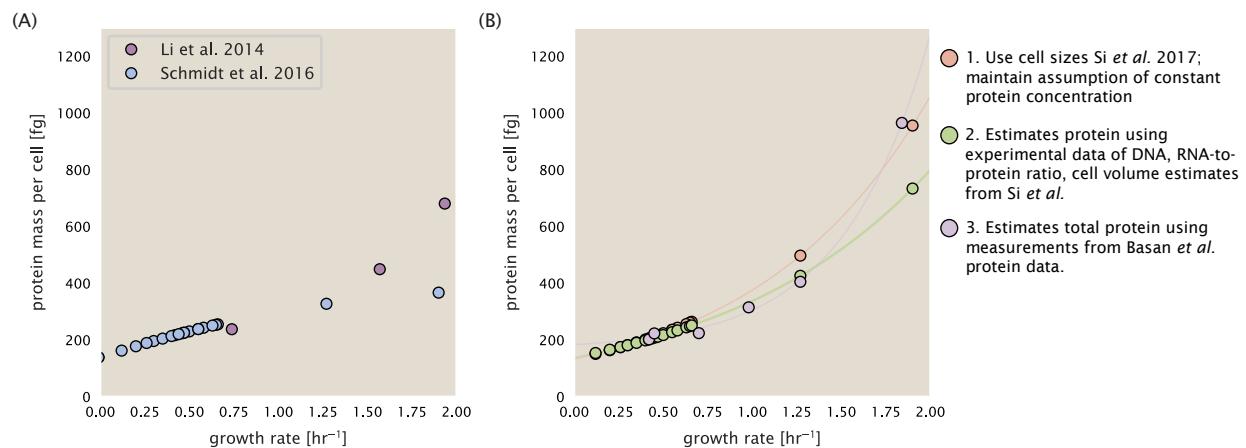
884 Since it is not obvious how measurements of cell size influenced their reported protein abundances, in the  
 885 following subsections we begin by considering this calculation. We then consider three different approaches to  
 886 estimate the growth-rate dependent total protein mass to compare with those values reported from **Schmidt *et al.***  
 887 (2016). The results of this are summarized in **Figure 17(B)**, with the original values from both **Schmidt *et al.* (2016)**  
 888 and **Li *et al.* (2014)** shown in **Figure 17(A)** for reference. For most growth conditions, we find that total protein per  
 889 cell is still in reasonable agreement. However, for the fastest growth conditions, with glycerol + supplemented  
 890 amino acids, and LB media, all estimates are substantially higher than those originally reported. This is the main  
 891 reason why we chose to readjusted protein abundance as shown in **Figure 12(B)** (with the calculation described in  
 892 section Estimation of Total Protein Content per Cell).

### 893 Effect of cell volume on reported absolute protein abundances

894 As noted in section Experimental Details Behind Proteomic Data, the authors calculated proteome-wide protein  
 895 abundances by first determining absolute abundances of 41 pre-selected proteins, which relied on adding syn-  
 896 synthetic heavy reference peptides into their protein samples at known abundance. This absolute quantitation was  
 897 performed in replicate for each growth condition. Separately, the authors also performed a more conventional  
 898 mass spectrometry measurement for samples from each growth condition, which attempted to maximize the  
 899 number of quantified proteins but only provided relative abundances based on peptide intensities. Finally, using  
 900 their 41 proteins with absolute abundances already determined, they then created calibration curves with which  
 901 to relate their relative intensity to absolute protein abundance for each growth condition. This allowed them to  
 902 estimate absolute protein abundance for all proteins detected in their proteome-wide data set. Combined with  
 903 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.

904 on a per cell basis.

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

912 For cells grown in M9 minimal media + glucose an average total mass of  $M_p = 240$  fg per cell was measured.  
913 Using their reported cell volume, reported as  $V_{orig} = 2.84$  fl, a cellular protein concentration of  $[M_p]_{orig} = M_p/V_{orig} =$   
914 85 fg/fl. Now, taking the assumption that cellular protein concentration is relatively independent of growth rate,  
915 they could then estimate the total protein mass for all other growth conditions from,

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

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

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

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

926 where  $M'_{P_i}$  is the new protein mass prediction, and  $V_{Si\_i}$  refers to the new volume prediction for each condition  $i$ ,  
927 These are shown as red data points in Figure [Figure 18\(B\)](#).

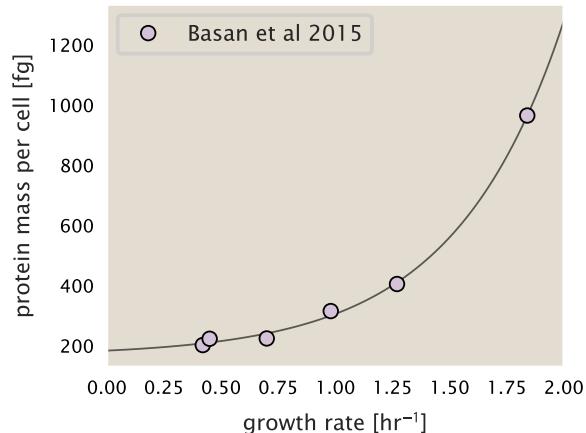
## 928 Relaxing assumption of constant protein concentration across growth conditions

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

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

935 One of the challenges in our estimates in the preceding sections is the need to estimate protein concentration  
936 and cell volumes. These are inherently difficult to accurately due to the small size of *E. coli*. Indeed, for all the  
937 additional measurements of cell volume included in Figure [Figure 17](#), no measurements were performed for cells  
938 growing at rates below  $0.5\text{ hr}^{-1}$ . It therefore remains to be determined whether our extrapolated cell volume  
939 estimates are appropriate, with the possibility that the logarithmic scaling of cell size might break down for slower  
940 growth.

941 In our last approach we therefore attempt to estimate total protein using experimental data that required no  
942 estimates of concentration or cell volume. Specifically, in the work of Basan *et al*, the authors measured total  
943 protein per cell for a broad range of growth rates (reproduced in Figure [Figure 19](#)). These were determined by  
944 first measuring bulk protein from cell lysate, measured by the colorimetric Biuret method ([You \*et al.\* \(2013\)](#)), and  
945 then abundance per cell was calculated from cell counts from either plating cells or a Coulter counter. While it  
946 is unclear why Schmidt *et al.* was unable to take a similar approach, the results from Basan *et al* appear more  
947 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  $\lambda$  is the growth rate in  $hr^{-1}$ .

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

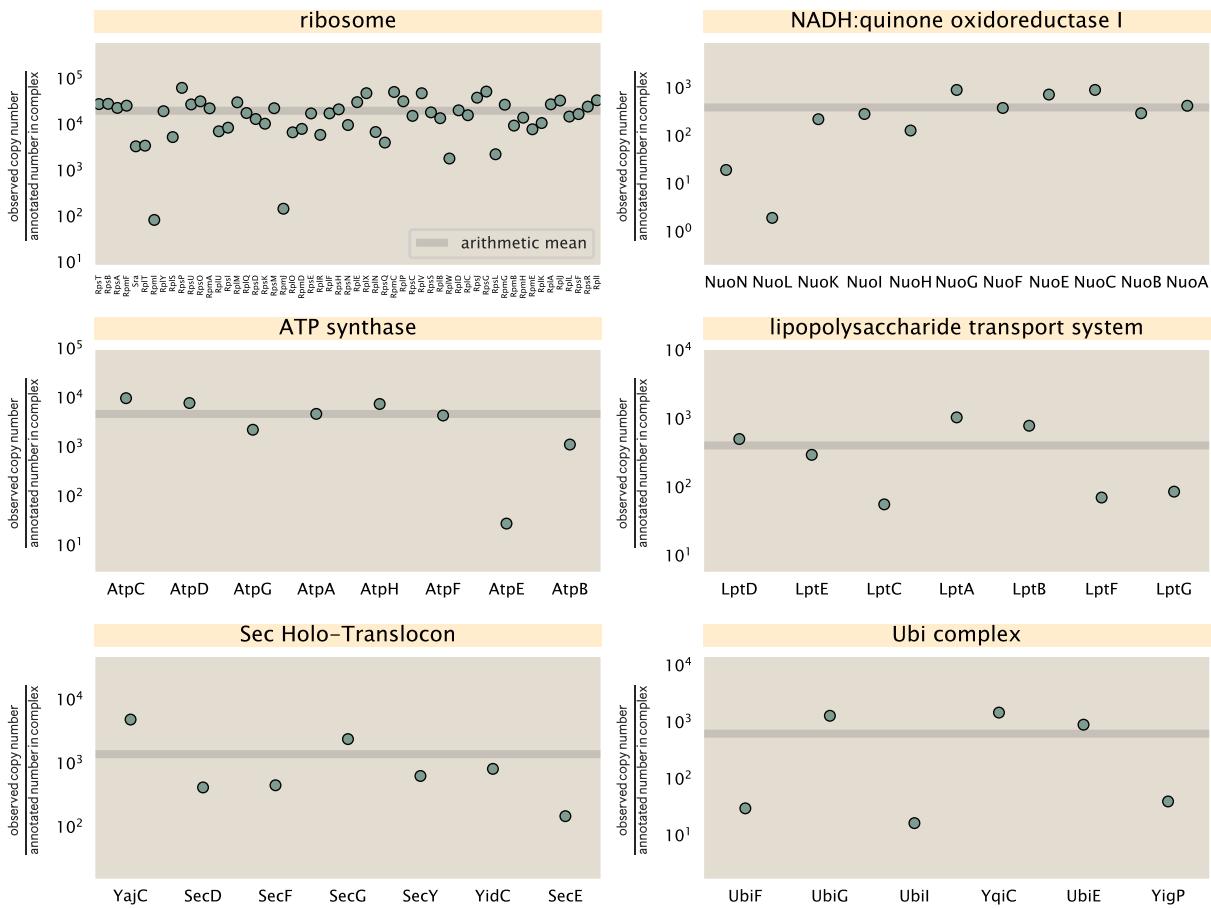
### 953 Calculation of Complex Abundance

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

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

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

975 In addition to the annotated complex, we collected information on the stoichiometry of each macromolecular  
 976 complex. For a complex with  $N_{\text{subunits}}$  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.

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

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

979 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  
980 present for that protein in the complex. For example, the 70S mature ribosome complex has 55 protein components,  
981 all of which are present in a single copy except L4 (RplL), which is present in 4 copies ( $m = 4$ ). For each ribosomal  
982 protein, we then calculate the maximum number of complexes that could be formed using Equation 9. This  
983 example, along with example from 5 other macromolecular complexes, can be seen in Figure 20.

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

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

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

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

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

995 **Estimation of the total cell mass**

996 For many of the processes estimated in the main text we relied on a cellular dry mass of  $\approx 300$  fg from which we  
997 computed elemental and protein fractions using knowledge of fractional composition of the dry mass. At modest  
998 growth rates, such as the 5000 s doubling time used in the main text, this is a reasonable number to use as the  
999 typical cell mass is  $\approx 1$  pg and *E. coli* cells can approximated as 70% water by volume. However, as we have shown  
1000 in the preceding sections, the cell size is highly dependent on the growth rate. This means that a dry mass of 300 fg  
1001 cannot be used reliably across all growth rates.

1002 Rather, using the phenomenological description of cell volume scaling exponentially with growth rate, and  
1003 using a rule-of-thumb of a cell buoyant density of  $\approx 1.1$  pg / fL (BNID: 103875), we can calculate the cell dry mass  
1004 across a range of physiological growth rates as

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

1005 where  $a$  and  $b$  are constants with units of  $\mu\text{m}^3$  and hr, respectively. The value of these constants can be estimated  
1006 from the careful volume measurements performed by *Si et al. (2017,?)*, as considered in Appendix Estimation of  
1007 Cell Size and Surface Area earlier.

1008 **Complex Abundance Scaling With Cell Volume**

1009 Several of the estimates performed in the main text are implicitly dependent on the cell volume. This includes  
1010 processes such as ATP utilization and, most prominently, the transport of nutrients, whose demand will be  
1011 proportional to the volume of the cell. Of the latter, we estimated the number of transporters that would be  
1012 needed to shuttle enough carbon, phosphorus, and sulfur across the membrane to build new cell mass. To do so,  
1013 we used elemental composition measurements combined with a 300 fg cell dry mass to make the point estimate.  
1014 As we now have a means to estimate the total cell mass as a function of volume, we can generalize these estimates  
1015 across growth rates.

1016 Rather than discussing the particular details of each transport system, we will derive this scaling expression in  
1017 very general terms. Consider that we wish to estimate the number of transporters for some substance  $X$ , which  
1018 has been measured to be made up some fraction of the dry mass,  $\theta_X$ . If we assume that, irrespective of growth  
1019 rate, the cell dry mass is relatively constant (*Basan et al., 2015*) and  $\approx 30\%$  of the total cell mass, we can state that  
1020 the total mass of substance  $X$  as a function of growth rate is

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

1021 where we have used  $\rho V(\lambda)$  as an estimate of the total cell mass, defined in **Equation 11**. To convert this to the  
1022 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  
1023 with **Equation 12**,

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

1024 To estimate the number of transporters needed, we make the approximation that loss of units of  $X$  via diffusion  
1025 through porins or due to the permeability of the membrane is negligible and that a single transporter complex  
1026 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  
1027 a time window over which the transport process can occur. This is related to the cell doubling time  $\tau$ , which can  
1028 be calculated from the the growth rate  $\lambda$  as  $\tau = \log(2)/\lambda$ . Putting everything together, we arrive at a generalized  
1029 transport scaling relation of

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

1030 This function is used to draw the continuum estimates for the number of transporters seen in Figures 2 and  
 1031 3 as transparent grey curves. Occasionally, this continuum scaling relationship will not precisely agree with the  
 1032 point estimate outlined in the main text. This is due to the choice of  $\approx 300$  fg total dry mass per cell for the point  
 1033 estimate, whereas we considered more precise values of cell mass in the continuum estimate. We note, however,  
 1034 that both this scaling relation and the point estimates are meant to describe the order-of-magnitude observed,  
 1035 and not the predict the exact values of the abundances.

1036 **Equation 14** is a very general relation for processes where the cell volume is the "natural variable" of the  
 1037 problem. This means that, as the cell increases in volume, the requirements for substance  $X$  also scale with  
 1038 volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the  
 1039 dry mass attributable to the substance, and the formula mass of the substance is known, **Equation 14** can be used  
 1040 to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell,  
 1041 **Equation 14** can be slightly modified to the form

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

1042 where we have included the term  $N_{\text{ATP}}$  to account for the number of ATP equivalents needed per amino acid for  
 1043 translation ( $\approx 4$ , BNID: 114971), and  $w_{AA}$  is the average mass of an amino acid. The grey curves in Figure 4 o the  
 1044 main text were made using this type of expression.

#### 1045 A Relation for Complex Abundance Scaling With Surface Area

1046 In our estimation for the number of complexes needed for lipid synthesis and peptidoglycan maturation, we used  
 1047 a particular estimate for the cell surface area ( $\approx 5 \mu\text{m}$ , BNID: 101792) and the fraction of dry mass attributable to  
 1048 peptidoglycan ( $\approx 3\%$ , BNID: 101936). Both of these values come from glucose-fed *E. coli* in balance growth. As we  
 1049 are interested in describing the scaling as a function of the growth rate, we must also consider how these values  
 1050 scale with cell surface area, which is the natural variable for these types of processes. In the coming paragraphs,  
 1051 we highlight how we incorporate a condition-dependent surface area into our calculation of the number of lipids  
 1052 and murein monomers that need to be synthesized and crosslinked, respectively.

#### 1053 Number of Lipids

1054 To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as  
 1055 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  
 1056 of lipids ( $\approx 40\%$ , BNID: 100078) are independent of the growth rate. Using these approximations combined with  
 1057 **Equation 5**, and recognizing that each membrane is composed of two leaflets, we can compute the number of  
 1058 lipids as a function of growth rate as

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

1059 where  $\eta$  is the length-to-width aspect ratio and  $V$  is the cell volume.

#### 1060 Number of Murein Monomers

1061 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we used an empirical  
 1062 measurement that  $\approx 3\%$  of the dry mass is attributable to peptidoglycan and that a single murein monomer is  
 1063  $m_{\text{murein}} \approx 1000 \text{ Da}$ . While the latter is independent of growth rate, the former is not. As the peptidoglycan exists as  
 1064 a thin shell with a width of  $w \approx 10 \text{ nm}$  encapsulating the cell, one would expect the number of murein monomers  
 1065 scales with the surface area of this shell. In a similar spirit to our calculation of the number of lipids, the total  
 1066 number of murein monomers as a function of growth rate can be calculated as

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

1067 where  $\rho_{\text{pg}}$  is the density of peptidoglycan.

## 1068 Complex Abundance Scaling With Number of Origins, and rRNA Synthesis

1069 While the majority of our estimates hinge on the total cell volume or surface area, processes related to the central  
1070 dogma, namely DNA replication and synthesis of rRNA, depend on the number of chromosomes present in the  
1071 cell. As discussed in the main text, the ability of *E. coli* to parallelize the replication of its chromosome by having  
1072 multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived in  
1073 *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  
1074 average number of origins of replication at a given growth rate can be calculated as

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

1075 where  $t_{cyc}$  is the total time of replication and division. We can make the approximation that  $t_{cyc} \approx 70$  min, which is  
1076 the time it takes two replisomes to copy an entire chromosome.

1077 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of replication. Thus,  
1078 at a given growth rate  $\lambda$ , the average dosage of rRNA operons per cell  $D_{rRNA}$  is

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

1079 This makes the approximation that *all* rRNA operons are localized around the origin. In reality, the operons are  
1080 some distance away from the origin, making **Equation 19** an approximation (*Dennis et al., 2004*).

1081 In the main text, we stated that at a growth rate of  $0.5 \text{ hr}^{-1}$ , there is  $\approx 1$  chromosome per cell. While a fair  
1082 approximation, **Equation 18** illustrates that is not precisely true, even at slow growth rates. In estimating the  
1083 number of RNA polymerases as a function of growth rate, we consider that regardless of the number of rRNA  
1084 operons, they are all sufficiently loaded with RNA polymerase such that each operon produces one rRNA per  
1085 second. Thus, the total number of RNA polymerase as a function of the growth rate can be calculated as

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

1086 where  $L_{\text{operon}}$  is the total length of an rRNA operon ( $\approx 4500$  bp) and  $\rho_{\text{RNA polymerase}}$  is packing density of RNA  
1087 polymerase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

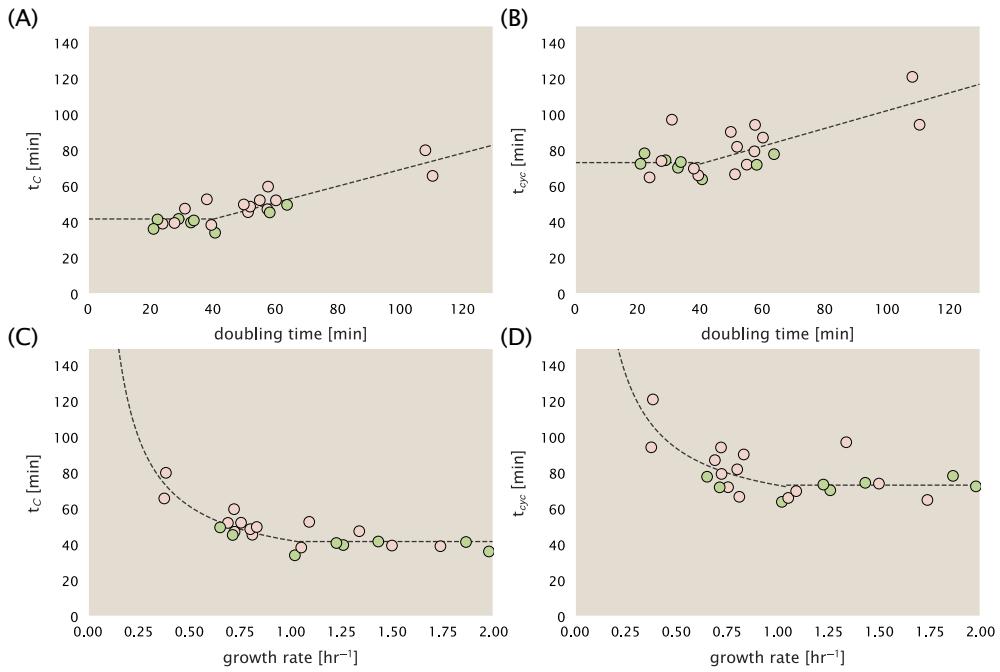
## 1088 Calculation of active ribosomal fraction.

1089 In the main text we used the active ribosomal fraction  $f_a$  that was reported in the work of *Dai et al. (2016)* to  
1090 estimate the active ribosomal mass fraction  $\Phi_R \times f_a$  across growth conditions. We lacked any specific model to  
1091 consider how  $f_a$  should vary with growth rate, and instead find that the data is well-approximated by fitting to an  
1092 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  $f_a$   
1093 for each of the data points shown in **Figure 9(C)**.

## 1094 Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$ .

1095 *E. coli* shows robust scaling of cell size with the average the number of origins  $\langle \#ori \rangle$  per cell (*Si et al., 2017*).  
1096 Since protein makes up a majority of the cell's dry mass, the change in cell size is also a reflection of the changes  
1097 in proteomic composition and total abundance across growth conditions. Given the potential constraints on  
1098 rRNA synthesis and changes in ribosomal copy number with  $\langle \#ori \rangle$ , it becomes important to also consider how  
1099 protein copy numbers vary with the state of chromosomal replication. This is particularly true when trying to make  
1100 sense of the changes in ribosomal fraction and growth-rate dependent changes in proteomic composition at a  
1101 mechanistic level. As considered in the main text, it is becoming increasingly apparent that regulation through the  
1102 secondary messengers (p)ppGpp may act to limit DNA replication and also reduce ribosomal activity in poorer  
1103 nutrient conditions. In this context, both  $\langle \#ori \rangle$ , as well as the  $\langle \#ori \rangle / \langle \#ter \rangle$  ratio become important parameters to  
1104 consider and keep tract of. An increase in  $\langle \#ori \rangle / \langle \#ter \rangle$  ratio in particular, causes a relatively higher gene dosage  
1105 in rRNA and r-protein genes due to skew in genes near the origin, where the majority of these are located

1106 In the main text we estimated the change in  $\langle \#ori \rangle$  with growth rate using the nutrient-limited wild-type cell  
1107 data from *Si et al. (2017)*. We consider their measurements of DNA replication time ( $t_C$ , 'C' period of cell division),  
1108 total cell cycle time ( $t_{cyc}$ , 'C' + 'D' period of cell division), and doubling time  $\tau$  from wild-type *E. coli* growing across  
1109 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  $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$  and  $\langle \# \text{ori} \rangle$  using data from Si et al. (2017).** (A) and (B) plot the reported  $t_C$  and  $t_{cyc}$  as a function of cell doubling time  $\tau$ , 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  $\langle \# \text{ori} \rangle$ . If the cell cycle time takes longer than the time of cell division, the cell will need to initiate DNA replication more often than its rate of division,  $2^{\lambda t} = 2^{\ln(2) \cdot t/\tau}$  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  $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$  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  $\tau$ . 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  $\lambda$  along with our piecewise linear fits, which match the plots in the main text.

1125 **References**

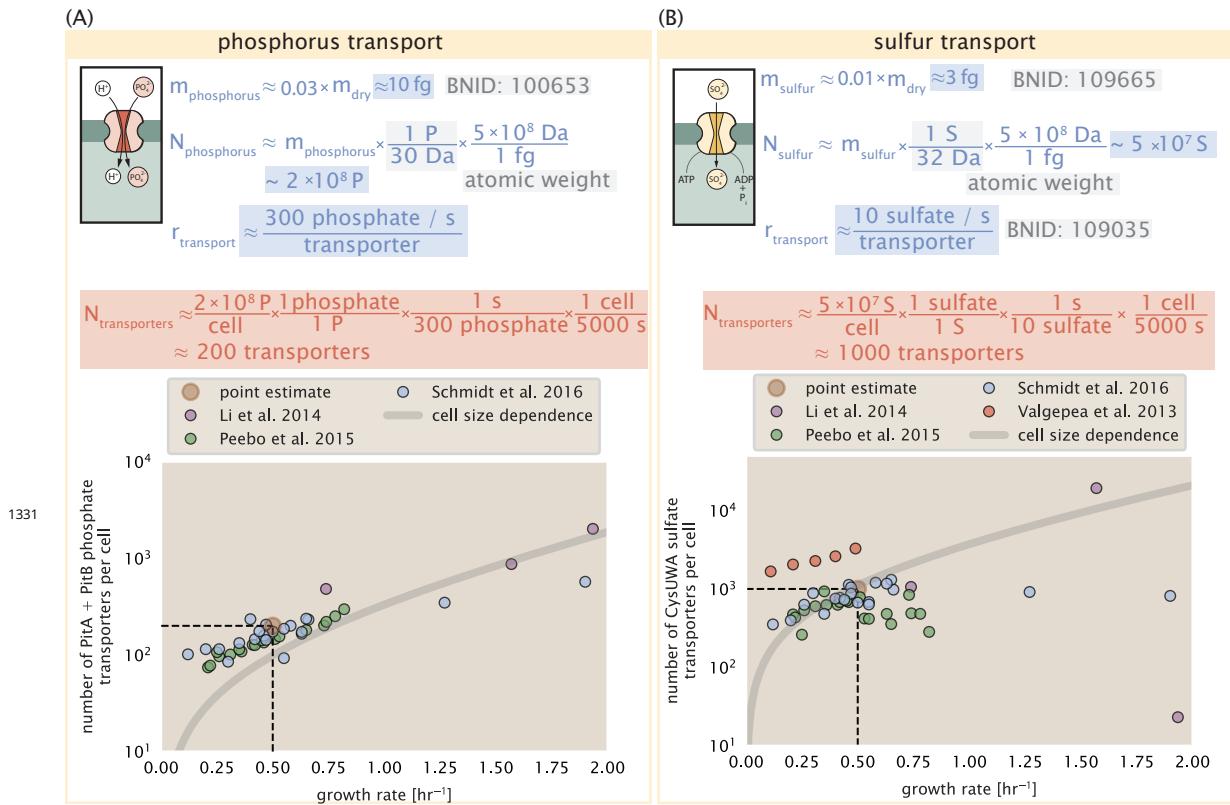
- 1126 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia*  
1127 *coli*. *BMC Systems Biology*, 8(1):133.
- 1128 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by  
1129 ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1130 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,  
1131 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.  
1132 (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1133 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model  
1134 for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding  
1135 Specificity of the  $\gamma$  Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- 1136 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH<sub>3</sub> Channel.  
1137 *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1138 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006).  
1139 Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*,  
1140 2(1):2460.
- 1141 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein  
1142 synthesis. *Molecular Systems Biology*, 11(10):836.
- 1143 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*,  
1144 18(1):81–94. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 1145 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.  
1146 (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proceedings of  
1147 the National Academy of Sciences*, 115(21):E4796–E4805.
- 1148 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential  
1149 Growth Rates. *EcoSal Plus*, 3(1).
- 1150 Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews  
1151 Microbiology*, 14(10):638–650.
- 1152 Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*,  
1153 266:2020.06.16.154187.
- 1154 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.,  
1155 Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan  
1156 Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- 1157 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in *Escherichia coli*.  
1158 *Biochemical Journal*, 117(3):551–562.
- 1159 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Translational  
1160 Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- 1161 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T.  
1162 (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature  
1163 Microbiology*, 2(2):16231.
- 1164 Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.  
1165 *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1166 Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.  
1167 *Microbiology and Molecular Biology Reviews*, 68(4):639–668.
- 1168 Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*,  
1169 108(44):17876–17882.
- 1170 Erickson, D. W., Schink, S. J., Patsalo, V., Williamson, J. R., Gerland, U., and Hwa, T. (2017). A global resource allocation strategy  
1171 governs growth transition kinetics of *Escherichia coli*. *Nature*, 551(7678):119–123.

- 1172 Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation. *Applied Microbiology and Biotechnology*, 94(6):1483–1494.
- 1173  
1174
- 1175 Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M., and Dunny, G. M. (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA Synthesis Independent of Growth Rates. *mBio*, 11(2):45.
- 1176  
1177
- 1178 Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA polymerase affair. *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1179
- 1180 Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–263.
- 1181 Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of the National Academy of Sciences*.
- 1182
- 1183 Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S., Alquicira-Hernández, K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-Lira, H., Bonavides-Martínez, C., Pérez-Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes, A., Hernández-Koutoucheva, A., Moral-Chávez, V. D., Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB version 9.0: High-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Research*, 44(D1):D133–D143.
- 1184  
1185  
1186  
1187
- 1188 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 Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390.
- 1189
- 1190 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-Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- 1191  
1192
- 1193 Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in microbiology*, 26(10):815–832.
- 1194
- 1195 Hauriyliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- 1196
- 1197 Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1198
- 1199 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). Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying Oxygen Conditions. *PLoS ONE*, 9(9):e107640.
- 1200
- 1201 Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2):e784–e784.
- 1202
- 1203 Ingledeew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*, 48(3):222–271.
- 1204 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B., and Phillips, R. (2020). Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a time. *bioRxiv*.
- 1205
- 1206 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*, 20(17):4952–4963.
- 1207
- 1208 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1209
- 1210 Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E., Ong, Q., Ong, W. K., Paley, S. M., and Subhraveti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics*, 20(4):1085–1093.
- 1211  
1212
- 1213 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. (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–401.
- 1214
- 1215 Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-Ramírez, D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge about *Escherichia coli*K-12. *Nucleic Acids Research*, 45(D1):D543–D550.
- 1216  
1217  
1218

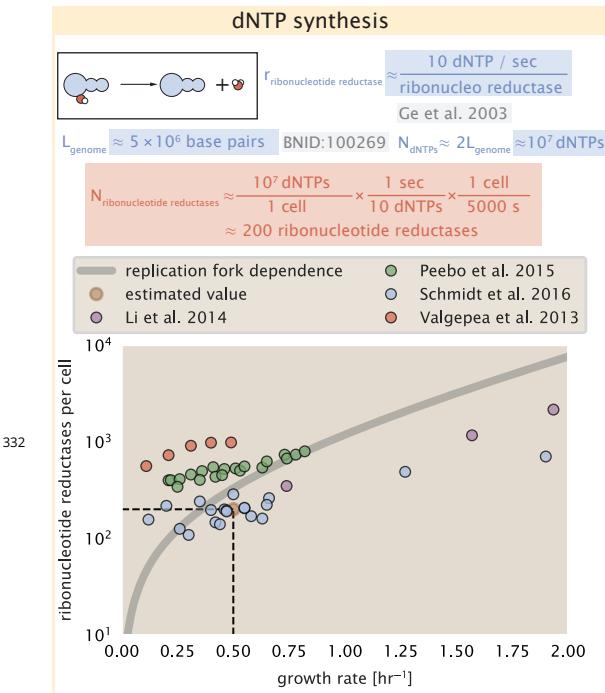
- 1219 Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that enables the use of  
1220 hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of Sciences*, 114(33):E6922–E6931.
- 1221 Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition.  
1222 *Current Opinion in Biotechnology*, 28:96–102.
- 1223 Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139(7):1366–  
1224 1375.
- 1225 Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*,  
1226 125(2):028103.
- 1227 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation in *E. coli* by  
1228 Modulating Supercoiling of oriC. *mBio*, 10(4):822.
- 1229 Lascu, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and*  
1230 *Biomembranes*, 32(3):237–246.
- 1231 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical  
1232 investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1233 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting sets. *IEEE Transactions*  
1234 *on Visualization and Computer Graphics*, 20(12):1983–1992.
- 1235 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles  
1236 underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1237 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in  
1238 cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1239 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon  
1240 Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- 1241 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facilitator  
1242 GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1243 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*,  
1244 112(51):15690–15695.
- 1245 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On*. Gene Expression. Springer.
- 1246 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by  
1247 condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1248 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  
1249 Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.
- 1250 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular  
1251 and cell biology. *Nucleic Acids Research*, 38(suppl\_1):D750–D753.
- 1252 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation.  
1253 *Growth Symposium*, 9:223–289.
- 1254 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1255 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall  
1256 synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–  
1257 12515.
- 1258 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 1259 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria.  
1260 *eLife*, 8:642.
- 1261 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- 1262 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with  
1263 increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.

- 1264 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1265 Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., and Maranas, C. D. (2012). An  
1266 integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering*,  
1267 14(6):687–704.
- 1268 Rogers, H., Perkins, H., and Ward, J. (1980). *Microbial Cell Walls and Membranes*. Chapman and Hall, London.
- 1269 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive  
1270 strategies. *Nature microbiology*, 1(11):1–7.
- 1271 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of*  
1272 *Bacteriology*, 131(2):505–511.
- 1273 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–  
1274 11734.
- 1275 Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical  
1276 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- 1277 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann,  
1278 M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 1279 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression:  
1280 origins and consequences. *Science*, 330(6007):1099–1102.
- 1281 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.  
1282 *Cell*, 172(6):1294–1305.
- 1283 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  
1284 Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 1285 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  
1286 Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- 1287 Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews*,  
1288 40(1):133–159.
- 1289 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth  
1290 and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1291 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*, 19.
- 1292 Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van*  
1293 *Leeuwenhoek*, 39(1):545–565.
- 1294 Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*.  
1295 *Archives of Microbiology*, 113(3):185–189.
- 1296 Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the  
1297 Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 1298 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  
1299 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 1300 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  
1301 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538.
- 1302 Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of  
1303 protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- 1304 Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic  
1305 models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*,  
1306 107(2):369–381. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1307 The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*,  
1308 47(D1):D330–D338.

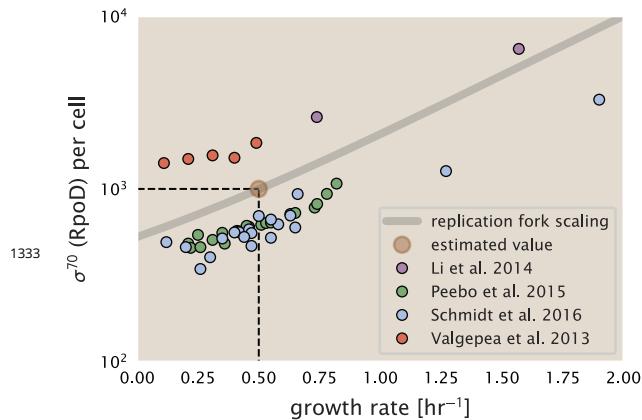
- 1309 Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and  
1310 translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 1311 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,  
1312 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,  
1313 E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero,  
1314 E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0:  
1315 Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1316 Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data  
1317 Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.
- 1318 Vollmer, W., Blanot, D., and De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*,  
1319 32(2):149–167.
- 1320 Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Letters*, 545(1):61–70.
- 1321 Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two  
1322 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- 1323 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  
1324 bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- 1325 Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from  
1326 *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 108(46):18643–18648.
- 1327 Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014). Regulation of crp gene expression by the catabolite repressor/activator, cra, in  
1328 *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1329 Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in  
1330 *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.



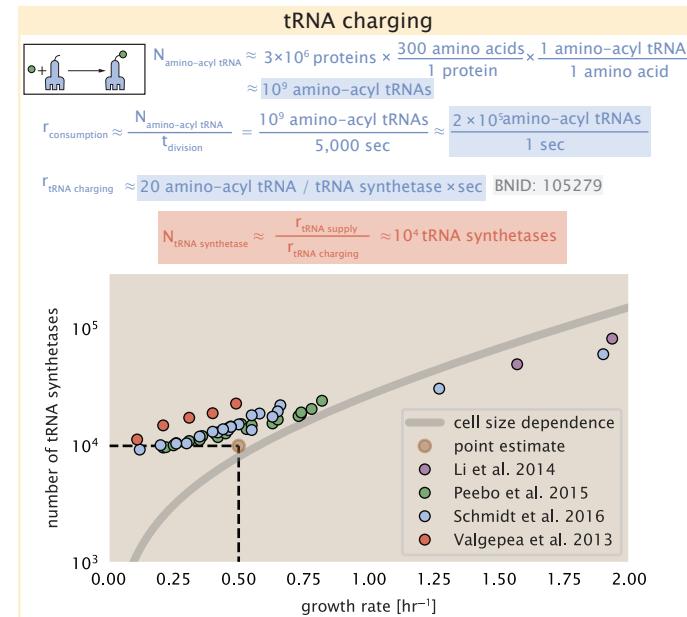
**Figure 2-Figure supplement 1.** (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (B) Estimate of the number of CysUWA complexes necessary to maintain a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



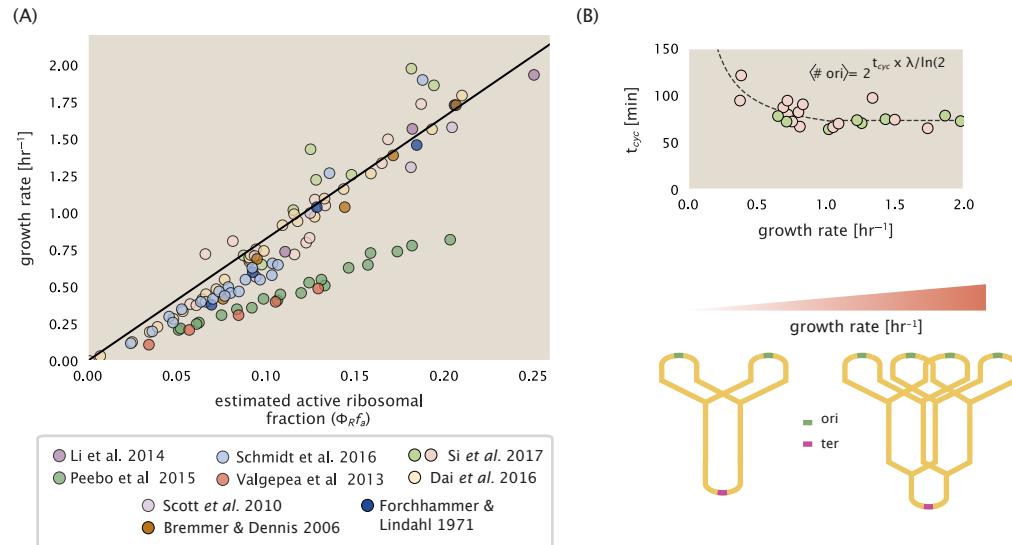
**Figure 6–Figure supplement 1.** Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of  $\approx 10^7$  dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ( $[\text{NrdA}]_2[\text{NrdB}]_2$ ) 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  $\sigma^{70}$  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]<sub>2</sub>, [AsnS]<sub>2</sub>, [AspS]<sub>2</sub>, [TyrS]<sub>2</sub>, [TrpS]<sub>2</sub>, [ThrS]<sub>2</sub>, [SerS]<sub>2</sub>, [ProS]<sub>2</sub>, [PheS]<sub>2</sub>[PheT]<sub>2</sub>, [MetG]<sub>2</sub>, [lysS]<sub>2</sub>, [HisS]<sub>2</sub>, [GlyS]<sub>2</sub>[GlyQ]<sub>2</sub>).



**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  $\tau$  and cell cycle time  $t_{\text{cyc}}$  from **Si et al. (2017)**. Dashed line shows fit to the data, which were used to estimate  $\langle \# \text{ori} \rangle$ .  $t_{\text{cyc}}$  was assumed to vary in proportion to  $\tau$  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.