

Fundamental limits on the rate of bacterial cell division

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

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

14

15 **Abstract** Recent years have seen a deluge of experiments dissecting the relationship between bacterial
16 growth rate, cell size, and protein content, quantifying the abundance of proteins across growth conditions with
17 unprecedented resolution. However, we still lack a rigorous understanding of what sets the scale of these
18 quantities and when protein abundances should (or should not) depend on growth rate. Here, we seek to
19 quantitatively understand this relationship across a collection of *Escherichia coli* proteomic data sets covering \approx
20 4000 proteins and 36 growth rates. We estimate the basic requirements for steady-state growth by considering
21 key processes in nutrient transport, energy generation, cell envelope biogenesis, and the central dogma. From
22 these estimates, ribosome biogenesis emerges as a primary determinant of growth rate. We expand on this
23 assessment by exploring a model of ribosomal regulation as a function of the nutrient supply, revealing a
24 mechanism that ties cell size and growth rate to ribosomal content.

25

26 Introduction

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

35 Jacques Monod once remarked that “the study of the growth of bacterial cultures does not constitute a special-
36 ized subject or branch of research, it is the basic method of Microbiology.” Those words ring as true today as they
37 did when they were written 70 years ago (?) with the quantitative power of this “method” recently undergoing
38 renaissance. Many of the key questions addressed by the pioneering efforts in the middle of the last century can
39 be revisited by examining them through the lens of the increasingly refined molecular census that is available for
40 bacteria such as the microbial workhorse *E. coli*.

41 Several of the evergreen questions about bacterial growth and physiology that were originally raised by mi-

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 begin by considering these two ques-
46 tions from two distinct angles. First, as a result of an array of high-quality proteome-wide measurements of *E. coli*
47 under diverse growth conditions, we have a census that allows us to explore how the number of key molecular
48 players change as a function of growth rate. Here, we have assembled a singular data set using measurements
49 collected over the past decade via mass spectrometry (???) or ribosomal profiling (?) of the composition of the
50 *E. coli* proteome across 36 unique growth rates (see Appendix Experimental Details Behind Proteomic Data for
51 further discussion of the data). Second, by compiling molecular turnover rate measurements for many of the fun-
52 damental processes associated with bacterial growth, we make quantitative estimates of key cellular processes
53 (schematized in **Figure 1**) to determine whether our current understanding of the dynamics of these processes
54 are sufficient to explain the magnitude of the observed protein copy numbers across conditions. The census, com-
55 bined with these estimates, provide a window into the question of whether the rates of central processes such
56 as energy generation or DNA synthesis are regulated systematically as a function of cell growth rate by altering
57 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 ≈ 5000 seconds, as the data sets examined here heavily sample this growth regime. While we formulate
60 point estimates for the protein abundances at this division time, we also consider how these values will vary at
61 other growth rates due to changes in cell size, surface area, and chromosome copy number (?). Broadly, we find
62 that the protein copy numbers appeared tuned for the task of cell doubling across a continuum of growth rates
63 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 plausi-
66 ble candidate that limits the rate of cell division. We reach this conclusion by considering that ribosome synthesis
67 is 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
69 the long-observed correlation between growth rate and cell size (??) can be simply attributed to the increased ab-
70 solute number of ribosomes per cell under conditions supporting extremely rapid growth. To better understand
71 how the observed alterations in absolute protein abundances, and in particular, changes in ribosome copy num-
72 ber, influence growth rate across different nutrient conditions we consider a minimal model of cellular growth.
73 Our conclusions from these analyses provide important insight into how *E. coli* regulates growth across conditions
74 of differing nutrient availability and identifies fundamental constraints in bacterial growth more broadly.

114 Nutrient Transport

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

123 The elemental composition of *E. coli* has received much quantitative attention over the past half century (????),
124 providing us with a starting point for estimating how many atoms of each element must be scavenged from the
125 environment. A synthesis of these studies presents an approximate dry mass composition of $\approx 50\%$ carbon (BNID:
126 100649, see **Box 1**), $\approx 15\%$ nitrogen (BNID: 106666), $\approx 3\%$ phosphorus (BNID: 100653), and 1% sulfur (BNID: 100655)
127 with remainder being attributable to oxygen, hydrogen, and various transition metals. We use this stoichiometric
128 breakdown to estimate the abundance and growth rate dependence of a variety of transporters responsible for
129 carbon uptake, and provide more extensive investigation of the other critical elements – phosphorus, sulfur, and
130 nitrogen – in the Appendix Additional Estimates of Fundamental Biological Processes.

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

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

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

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

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

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

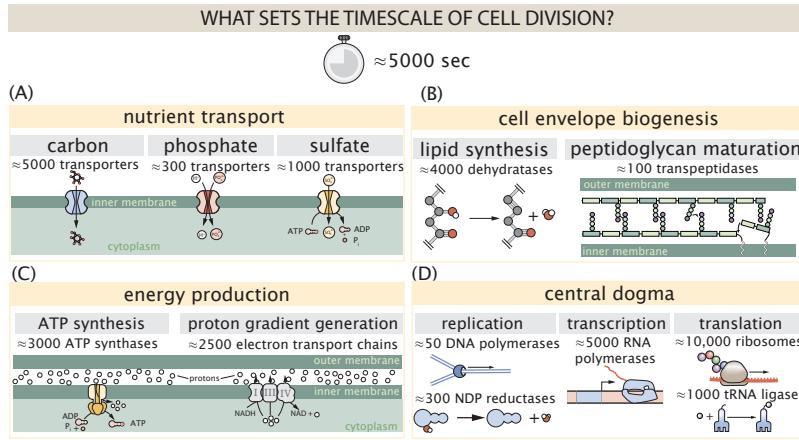


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) cell envelope biogenesis, (C)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

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

However, we find this estimate to be exceeded by several fold by experimental measurements (*Figure 2(A)*, bottom), implying that the cell is capable of transporting more carbon atoms than strictly needed for biosynthesis. While we estimate ≈ 1000 transporters are needed with a 5000 second division time, we can abstract this calculation to consider any particular growth rate given knowledge of the cell density and volume as a function of growth rate and direct the reader to the Appendix Extending Estimates to a Continuum of Growth Rates for more information. This abstraction, shown as a grey line in *Figure 2(A)*, reveals an excess of transporters even at faster growth rates. This contrasts with our observations for uptake of phosphorus and sulfur, which align well with our expectations across different growth conditions (*Figure 2-Figure Supplement 1* and discussed further in Appendix Additional Estimates of Fundamental Biological Processes).

It is important to note, however, that this estimate neglects any specifics of the regulation of the carbon transport system. Using the diverse array of growth conditions available in the data, we can explore how individual carbon transport systems depend on specific carbon availability. In *Figure 2(B)*, we show the total number of carbohydrate transporters specific to different carbon sources. A striking observation, shown in the top-left plot of *Figure 2(B)*, is the constancy in the expression of the glucose-specific transport systems, an observation that stands in contrast with other species of transporters. Additionally, we note that the total number of glucose-specific transporters is tightly distributed at $\approx 10^4$ per cell, the approximate number of transporters needed to sustain rapid growth of several divisions per hour. This illustrates that *E. coli* maintains a substantial number of complexes present for transporting glucose regardless of growth condition, which is known to be the preferential carbon source (???).

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 (?????). Points colored in red in *Figure 2(B)* (labeled by red text-boxes) correspond to growth conditions in which the specific carbon source (glyc-

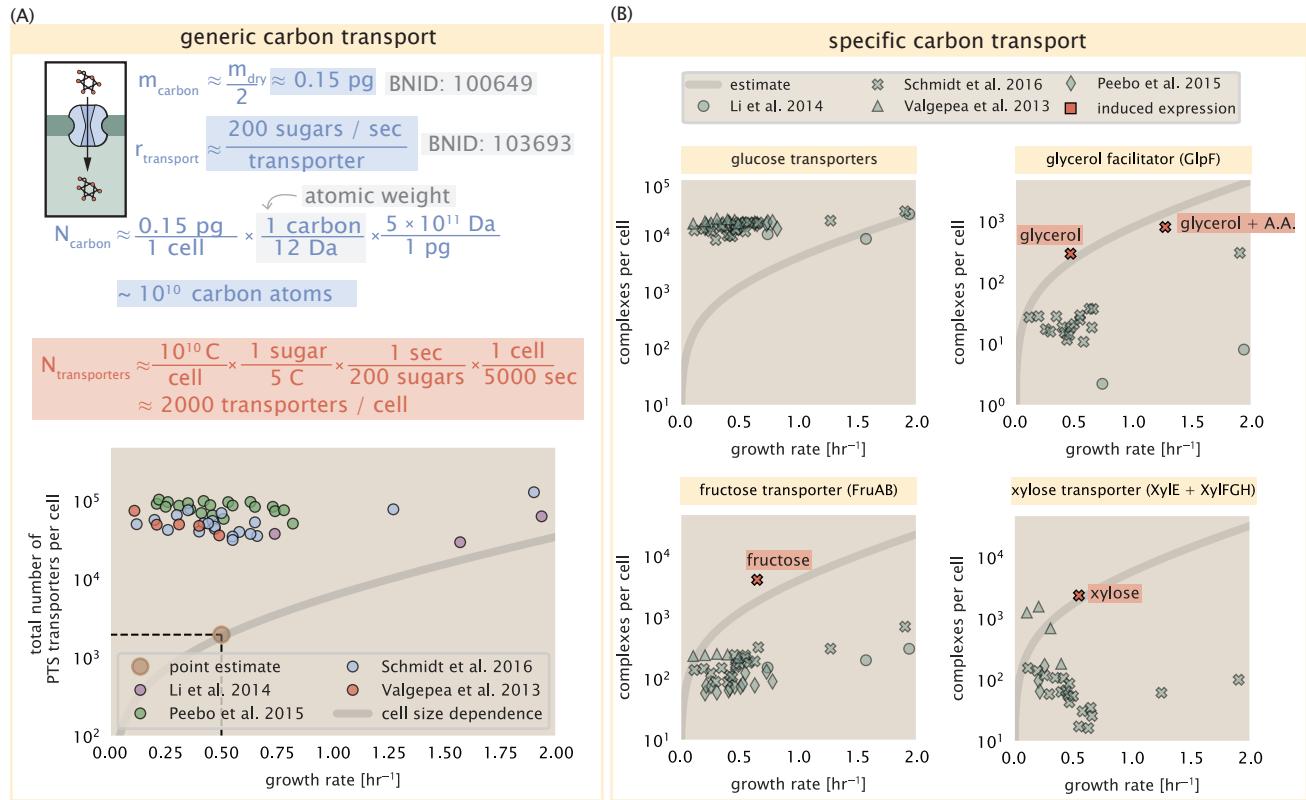


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules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

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

erol, xylose, or fructose) is present as the sole source of carbon. 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/ μm^2 . For a typical transporter occupying about 50 nm^2 , this amounts to about only $\approx 1\%$ of the total inner membrane area (?). Additionally, bacterial cell membranes typically have densities of 10^5 proteins/ μm^2 (?), 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. *E. coli* is a rod-shaped bacterium with a remarkably robust length-to-width aspect ratio of $\approx 4:1$ (?). At modest growth rates, the total cell surface area is $\approx 5 \mu\text{m}^2$ (BNID: 101792). 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 $\approx 20 \mu\text{m}^2$ (see Appendix Estimation of Cell Size and Surface Area for a description of the calculation of cell surface area as a function of cell size). In this section, 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 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 (?). Recently, a combination of stochastic kinetic modeling (?) and *in vitro* kinetic measurements (??) have revealed remarkably slow steps in the fatty acid synthesis pathways which may serve as the rate limiting reactions for making new membrane phospholipids. 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* (?), have been estimated to have kinetic turnover rates of ≈ 1 dehydration per second per enzyme (?). Thus, given this rate and the need to synthesize $\approx 2 \times 10^7$ lipids over 5000 seconds, one can estimate that a typical cell requires ≈ 4000 ACP dehydratases. This is in reasonable agreement with the experimentally observed copy numbers of FabZ and FabA (**Figure 3(A)**). Furthermore, we can extend this estimate to account for the change in membrane surface area as a function of the growth rate (grey line in **Figure 3(A)**), which captures the observed growth rate dependent expression of these two enzymes.

Peptidoglycan Synthesis

Bacterial cells demonstrate exquisite control over their cell shape. This is primarily due to the cell wall, a stiff, several nanometer thick meshwork of polymerized discaccharides. The formation of the peptidoglycan is an intricate process involving many macromolecular players (?), whose coordinated action maintains cell shape and integrity

even in the face of large-scale perturbations (??). The peptidoglycan alone comprises \approx 3% of the cellular dry mass (BNID: 1019360, making it the most massive molecule in *E. coli*. The polymerized unit of the peptidoglycan is a N-acetylglucosamine and N-acetylmuramic acid disaccharide, of which the former is functionalized with a short pentapeptide. With a mass of \approx 1000 Da, this unit, which we refer to as a murein monomer, it is polymerized to form long strands in the periplasm which are then attached to each other via their peptide linkers. Together, these quantities provide an estimate of $\approx 5 \times 10^6$ murein monomers per cell.

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

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

231 Limits on Cell Wall Biogenesis

While the processes we have considered represent only a small portion of proteins devoted to cell envelope biogenesis, we find it unlikely that they limit cellular growth in general. The relative amount of mass required for lipid and peptidoglycan components decrease at faster growth rates due to a decrease in their surface area to volume (S/V) ratio (?). Furthermore, despite the slow catalytic rate of FabZ and FabA in lipid synthesis, experimental data and recent computational modeling has shown that the rate of fatty-acid synthesis can be drastically increased by increasing the concentration of FabZ (??). With a proteome size of $\approx 3 \times 10^6$ proteins, a hypothetical 10-fold increase in expression from 4000 to 40,000 ACP dehydratases would result in a paltry $\approx 1\%$ increase in the size of the proteome. In the context of peptidoglycan synthesis, we note that our estimate considers only the transpeptidase enzymes that are involved lateral and longitudinal elongation of the peptidoglycan. This neglects the presence of other transpeptidases that are present in the periplasm and also involved in remodeling and maturation of the peptidoglycan. It is therefore possible that if this was setting the speed limit for cell division, the simple expression of more transpeptidases may be sufficient to maintain the structural integrity of the cell wall.

244 Energy Production

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

250 ATP Synthesis

Hydrolysis of the terminal phosphodiester bond of ATP into ADP (or alternatively GTP and GDP) and an inorganic phosphate provides the thermodynamic driving force in a wide array of biochemical reactions. One such reaction is the formation of peptide bonds during translation, which requires ≈ 2 ATPs for the charging of an amino acid to the tRNA and ≈ 2 GTP for the formation of each peptide bond. Assuming the ATP costs associated with error correction and post-translational modifications of proteins are negligible, we can make the approximation that each peptide bond has a net cost of ≈ 4 ATP (BNID: 101442, ?). Formation of GTP from ATP is achieved via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy investment (?) and

CELL ENVELOPE BIOSYNTHESIS

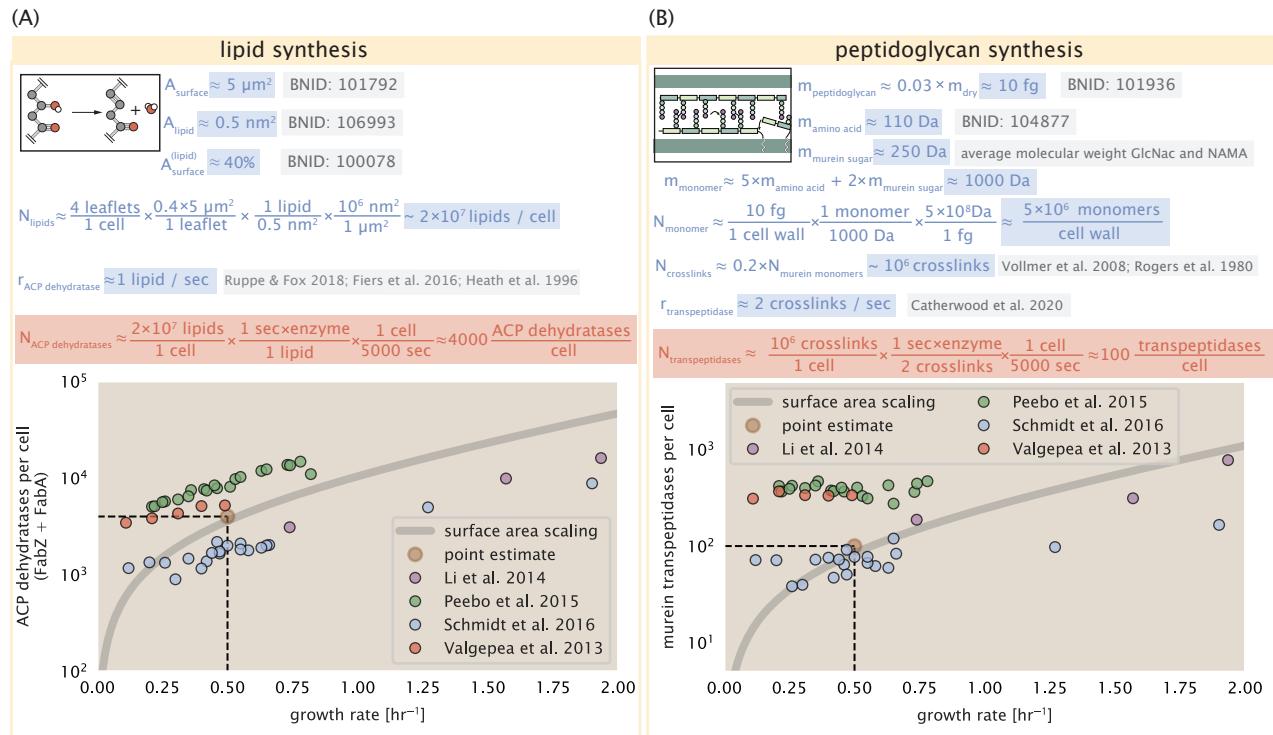


Figure 3. (A) Top panel shows an estimation for the number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis. The rate of ACP dehydratases was inferred from experimental measurements via a stochastic kinetic model described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

therefore consider all NTP requirements of the cell to be functionally equivalent to being exclusively ATP. In total, the energetic costs of peptide bond formation consume \approx 80% of the cells ATP budget (BNID: 107782; 106158; 101637; 111918, ??). The pool of ATP is produced by the F_1 - F_0 ATP synthase – a membrane-bound rotary motor which under ideal conditions can yield \approx 300 ATP per second (BNID: 114701; ?).

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

Generating the Proton Electrochemical Gradient

In order to produce ATP, the F_1 - F_0 ATP synthase itself must consume energy. Rather than burning through its own product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the electrochemical potential established across the inner membrane through cellular respiration. This electrochemical gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains as they reduce NADH. In *E. coli*, this potential difference is ≈ -200 mV (BNID: 102120). A simple estimate of the inner membrane as a capacitor with a working voltage of -200 mV reveals that $\approx 2 \times 10^4$ protons must be present in the intermembrane space. However, each rotation of an ATP synthase shuttles \approx 4 protons into the cytosol (BNID: 103390). With a few thousand ATP synthases producing ATP at their maximal rate, the potential difference would be rapidly abolished in a few milliseconds if it were not being actively maintained.

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

Limits on Biosynthesis in a Crowded Membrane

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

In our estimate of ATP production above we found that a cell demands about 5×10^9 ATP per cell cycle or 10^6 ATP/s. With a cell volume of roughly 1 fL (BNID: 100004), this corresponds to about 2×10^{10} ATP per fL of cell volume, in line with previous estimates (??). In *Figure 5* (A) we plot this ATP demand as a function of the S/V ratio in green, where we have considered a range of cell shapes from spherical to rod-shaped with an aspect ratio (length/width) equal to 4. In order to consider the maximum ATP that could be produced, we consider the amount of ATP that can be generated by a membrane filled with ATP synthase and electron transport complexes, which provides a maximal production of about 3 ATP / ($\text{nm}^2 \cdot \text{s}$) (?). This is shown in blue in *Figure 5*(A), which shows that at least for the growth rates observed (right column in plot), the energy demand is roughly an order of magnitude less. Interestingly, ? also found that ATP production by respiration is less efficient than by fermentation per membrane area occupied due to the additional proteins of the electron transport chain. This suggests that, even under

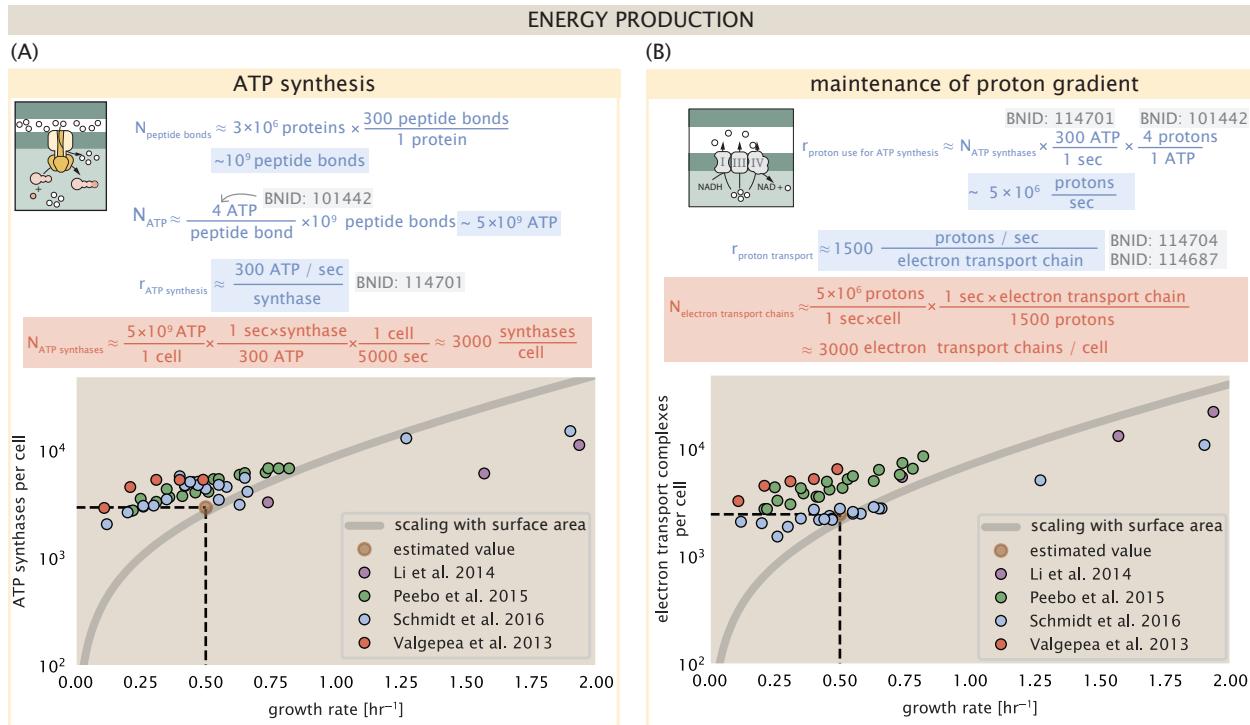


Figure 4. The abundance of F₁-F₀ ATP synthases and electron transport chain complexes as a function of growth rate.

(A) Estimate of the number of F₁-F₀ ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes. Points in plot correspond to the mean number of complete F₁-F₀ ATP synthase complexes that can be formed given proteomic measurements and the subunit stoichiometry [AtpE]₁₀[AtpF]₂[AtpB][AtpC][AtpH][AtpA]₃[AtpG][AtpD]₃. (B) Estimate of the number of electron transport chain complexes needed to maintain a membrane potential of -200 mV given estimate of number of F₁-F₀ ATP synthases from (A). Points in plot correspond to the average number of complexes identified as being involved in aerobic respiration by the Gene Ontology identifier GO:0019646 that could be formed given proteomic observations. These complexes include cytochromes *bd1* ([CydA][CydB][CydX][CydH]), *bdII* ([AppC][AppB]), *bo₃*,([CyoD][CyoA][CyoB][CyoC]) and NADH:quinone oxioreductase I ([NuoA][NuoH][NuoJ][NuoK][NuoL][NuoM][NuoN][NuoB][NuoC][NuoE][NuoF][NuoG][NuoI]) and II ([Ndh]). Grey lines in both (A) and (B) correspond to the estimate procedure described, but applied to a continuum of growth rates. We direct the reader to the Supporting Information for a more thorough description of this approach.

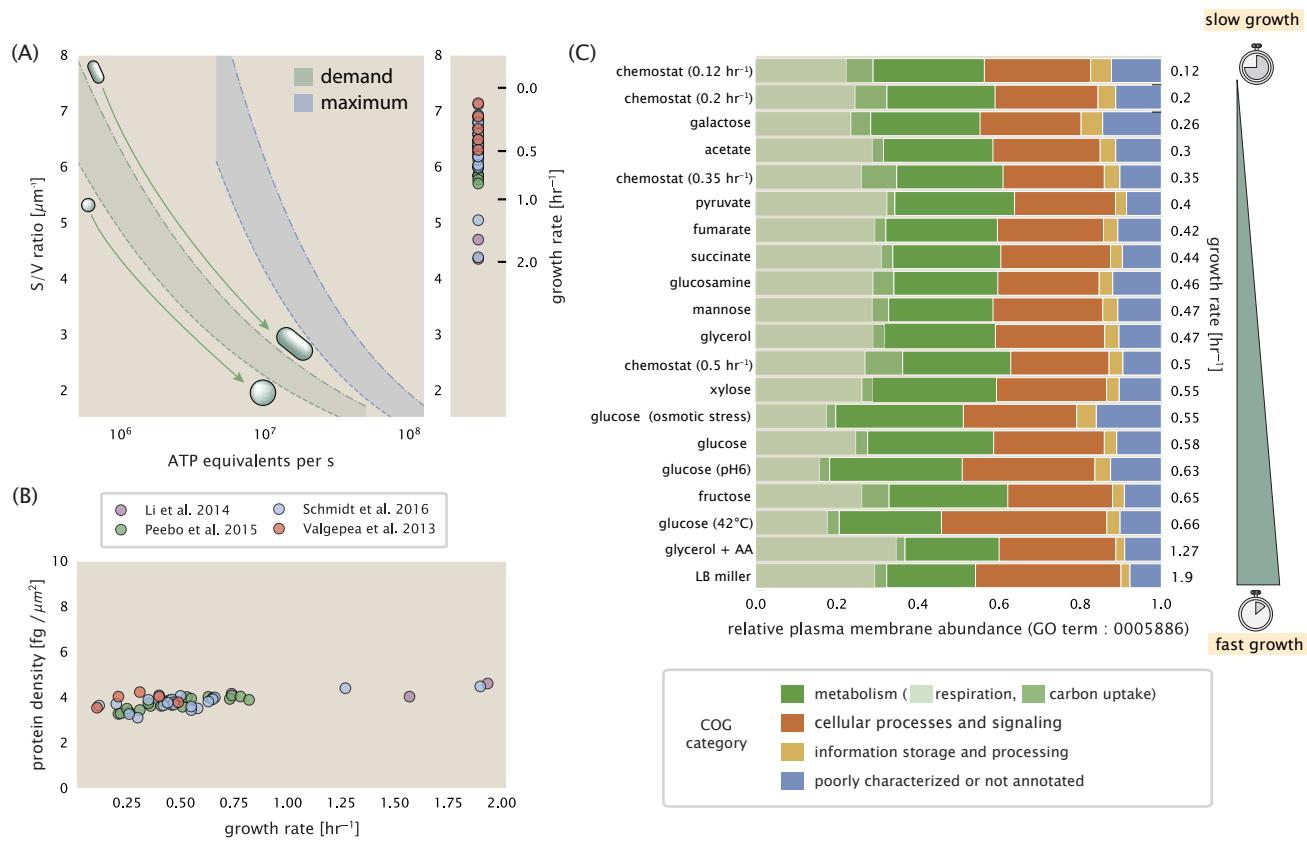


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (---) reflecting a rod-shaped bacterium like *E. coli* with a typical aspect ratio (length / width) of 4 (?). 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 ? (See Appendix Estimation of Cell Size and Surface Area on calculation). (B) Total protein mass per μm^2 calculated for proteins with inner membrane annotation (GO term: 0005886). (C) Relative protein abundances by mass based on COG annotation. Metabolic proteins are further separated into respiration (F_1 - F_0 ATP synthase, NADH dehydrogenase I, succinate:quinone oxidoreductase, cytochrome bo_3 , ubiquinol oxidase, cytochrome bd-I ubiquinol oxidase) and carbohydrate transport (GO term: GO:0008643). Note that the elongation factor EF-Tu can also associate with the inner membrane, but was excluded in this analysis due to its high relative abundance (roughly identical to the summed protein shown in part (B)).

307 anaerobic growth, there will be sufficient membrane space for ATP production.

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

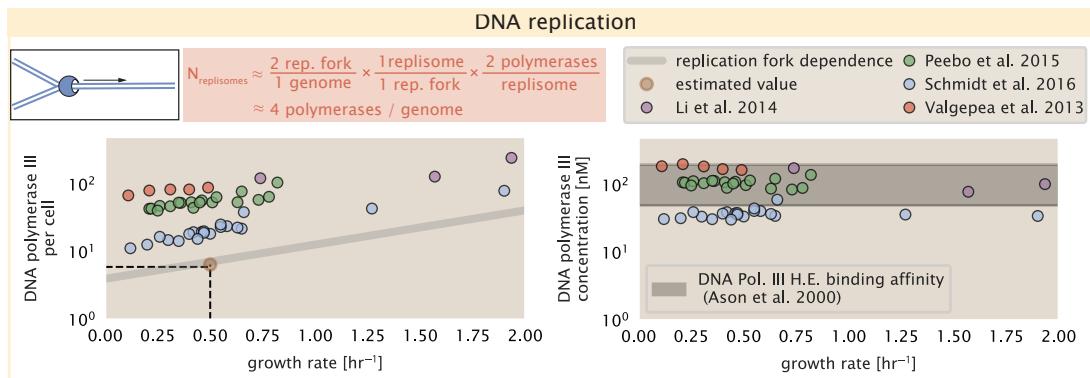


Figure 6. Complex abundance estimates for dNTP synthesis and DNA replication. An estimate for the minimum number of DNA polymerase holoenzyme complexes needed to facilitate replication of a single genome. Points in the left-hand plot correspond to the total number of DNA polymerase III holoenzyme complexes ($[DnaE]_3[DnaQ]_3[HolE]_3[DnaX]_5[HolB][HolA][DnaN]_4[HolC]_4[HolD]_4$) per cell. Right-hand plot shows the effective concentration of DNA polymerase III holoenzyme (See Appendix Estimation of Cell Size and Surface Area for 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.

Processes of the Central Dogma

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

DNA Replication

Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal plasmids up to ~ 100 kbp in length. While we consider the starting material dNTPs in **Figure 6–Figure Supplement 1** and discussed further in Appendix Additional Process of the Central Dogma, here we focus our quantitative thinking on the chromosome of *E. coli* which harbors ≈ 5000 genes and $\approx 5 \times 10^6$ base pairs.

To successfully divide and produce viable progeny, this chromosome must be faithfully replicated and segregated into each nascent cell. Replication is initiated at a single region of the chromosome termed the *oriC* locus at which a pair of replisomes, each consisting of two DNA polymerase III, begin their high-fidelity replication of the genome in opposite directions (?). *In vitro* measurements have shown that DNA Polymerase III copies DNA at a rate of ≈ 600 nucleotides per second (BNID: 104120). Therefore, to replicate a single chromosome, two replisomes moving at their maximal rate would copy the entire genome in ≈ 4000 s. Thus, with a division time of 5000 s, there is sufficient time for a pair of replisomes complexes to replicate the entire genome.

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

348 the time to double their chromosome indeed represents an upper limit to their growth rate.

349 RNA Synthesis

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

354 rRNA serves as the catalytic and structural component of the ribosome, comprising approximately 2/3 of the total
355 ribosomal mass, and is decorated with \approx 50 ribosomal proteins. Each ribosome contains three rRNA molecules
356 of lengths 120, 1542, and 2904 nucleotides (BNID: 108093), meaning each ribosome contains \approx 4500 nucleotides
357 overall. *In vivo* measurements of the kinetics of rRNA transcription have revealed that RNA polymerases are loaded
358 onto the promoter of an rRNA gene at a rate of \approx 1 per second (BNID: 111997, 102362). If RNA polymerases are
359 constantly loaded at this rate, then we can assume that \approx 1 functional rRNA unit is synthesized per second per
360 rRNA operon. While *E. coli* possesses 7 of these operons per chromosome, the fact that chromosome replication
361 can be parallelized means that the average dosage of rRNA genes can be substantially higher (up to \approx 70 copies)
362 at fast growth rates. At a growth rate of \approx 0.5 hr⁻¹, however, the average cell has \approx 1 copy of its chromosome
363 and therefore approximately \approx 7 copies of the rRNA operons, therefore producing \approx 7 rRNA units per second.
364 With a 5000 second division time, this means the cell is able to generate around 3×10^4 functional rRNA units,
365 comparable within an order of magnitude to the number of ribosomes per cell.

366 How many RNA polymerases are then needed to constantly transcribe the required rRNA? If one polymerase is
367 loaded per second, and the transcription rate is \approx 40 nucleotides per second (BNID: 101094), then the typical
368 spacing between polymerases will be \approx 40 nucleotides. However, we must note that the polymerase itself has a
369 footprint of \approx 40 nucleotides (BNID: 107873), meaning that one could expect to find one RNA polymerase per 80
370 nucleotide stretch of an rRNA gene. With a total length of \approx 4500 nucleotides per operon and 7 operons per cell,
371 the number of RNA polymerases transcribing rRNA at any given time is then \approx 500 per cell.

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

379 Our estimates also neglect other mechanistic features of transcription and transcriptional initiation more
380 broadly. For example, we acknowledge that some fraction of the RNAP pool is nonspecifically bound to DNA
381 during its search for promoters from which to begin transcription. Furthermore, we ignore the obstacles that RNA
382 polymerase and DNA polymerase present to each other as they move along the DNA (?). Finally, we neglect the
383 fact that RNA polymerase also require σ -factors for promoter recognition and transcription initiation (?).

384 While they are the machinery for transcription, RNA polymerase is not sufficient to initiate transcription. Pro-
385 moter recognition and initiation of transcription is dependent on the presence of σ -factors, protein cofactors which
386 bind directly to the polymerase (?). In *Figure 7-Figure Supplement 1*, we show that the predicted RNA polymerase
387 copy number indeed is more comparable with the abundance of σ -70 (RpoD), the primary sigma factor in *E. coli*.
388 There therefore remains more to be investigated as to what sets the observed abundance of RNA polymerase in
389 these proteomic data sets. However, we conclude that the observed excess in abundance for RNA polymerase
390 abundances are generally in excess of what appears to be needed for growth, suggesting that the abundance of
391 RNA polymerase itself is not particularly limiting.

392 Protein Synthesis

393 We conclude our dialogue between back-of-the-envelope estimates and comparison with the proteomic data by
394 examining the final process in the central dogma – translation. In doing so, we will begin with an estimate of the
395 number of ribosomes needed to replicate the cellular proteome. While the rate at which ribosomes translate is
396 well known to be dependent on the growth rate (? a phenomenon we consider later in this work) we will make

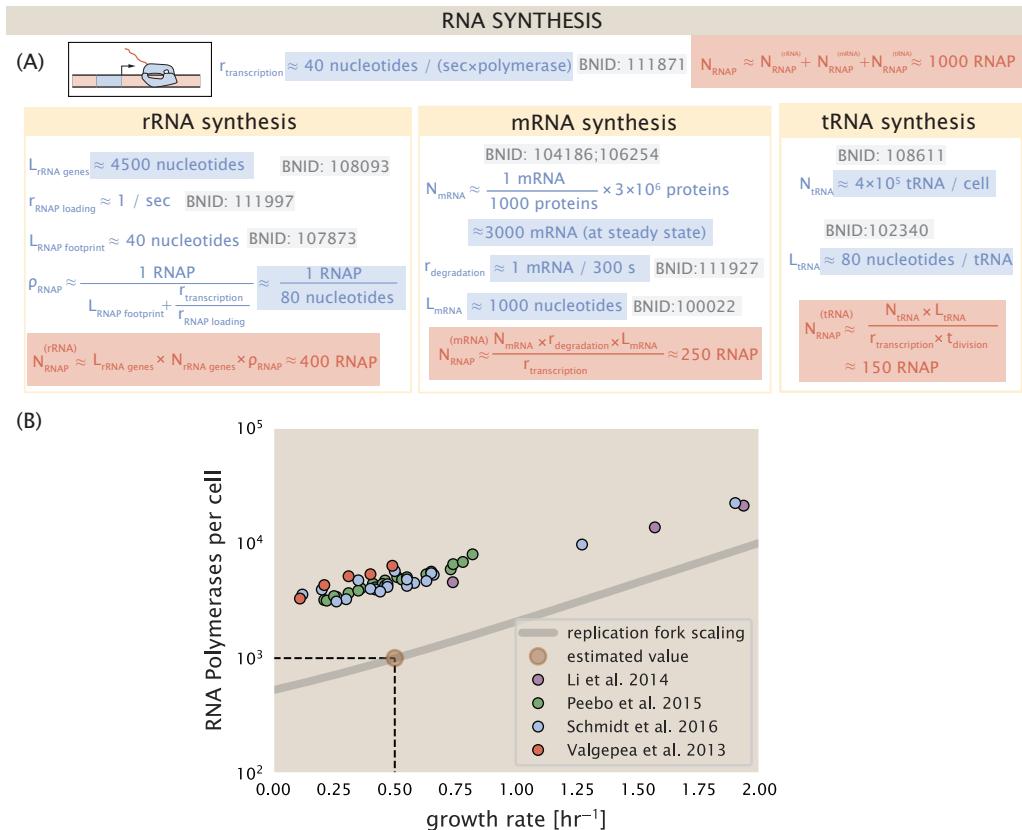


Figure 7. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. (B) The RNA polymerase core enzyme copy number as a function of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

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

397 the approximation that translation occurs at a modest rate of \approx 15 amino acids per second per ribosome (BNID:
398 100233) Under this approximation and our previous estimate of 10^9 peptide bonds per cell at a growth rate of 0.5
399 hr^{-1} , we can easily arrive at an estimate of $\approx 10^4$ ribosomes needed per cell to replicate the entire protein mass
400 (**Figure 8(A)**, top). This point estimate, as well as the corresponding estimate across a continuum of growth rates,
401 proves to be notably comparable to the experimental observations, shown in the bottom panel of **Figure 8(A)**.
402 While the ribosome is responsible for the formation of peptide bonds, we do not diminish the importance of
403 charging tRNAs with their appropriate amino acid, a process which occurs with remarkable fidelity. In the Appendix
404 and in **Figure 8–Figure Supplement 1**, we consider the process of ligating tRNAs to their corresponding amino acid
405 and again find notable accord between the data and our quantitative expectations.

406 Having completed our circuit through key processes of cellular growth outlined in **Figure 1**, we can now take
407 stock of our understanding of the observed growth rate dependence and abundances of various protein com-
408 plexes. We note that, broadly speaking, these simple estimates have been reasonably successful in quantitatively
409 describing the observations in the proteomic data, suggesting that the proteome is tuned in composition and
410 absolute abundance to match the growth rate requirements without any one process representing a singular bot-
411 tleneck or rate limiting step in division. However, in our effort to identify key limitations on growth, there are two
412 notable observations that we wish to emphasize.

413 The first is a recurring theme throughout our estimates. Of those investigated here, any inherent biochemical
414 rate limitation can be overcome by expressing more proteins. We can view this as a parallelization of each biosyn-
415 thesis task, which helps explain why bacteria tend to increase their protein content (and cell size) as growth rate
416 increases (?). The second, and ultimately the most significant in defining the cellular growth rate, is that the synthe-
417 sis of ribosomal proteins presents a special case where parallelization is *not* possible and thereby imposes a limit
418 on the fastest possible growth rate. Each ribosome has \approx 7500 amino acids across all of its protein components
419 which must be strung together as peptide bonds through the action of another ribosome. Once again using a
420 modest elongation rate of \approx 15 amino acids per second, we arrive an estimate of \approx 500 seconds or \approx 7 minutes to
421 replicate a single ribosome. This limit, as remarked upon by others (?), serves as a hard theoretical boundary for
422 how quickly *E. coli* could replicate. As each ribosome would therefore need to copy itself, this 7 minute speed limit
423 is independent of the number of ribosomes per cell (**Figure 8(B)**), yet assumes that the only proteins that need
424 to be replicated for division to occur are ribosomal proteins, an unrealistic regime not met in biological reality.
425 This poses an optimization problem for the cell – how are the translational demands of the entire proteome met
426 without investing resources in the production of an excess of ribosomes?

427 This question, more frequently presented as a question of optimal resource allocation, has been the target of
428 an extensive dialogue between experiment and theory over the past decade. In a now seminal work, ? present an
429 elegant treatment of resource allocation through partitioning of the proteome into sectors – one of which being
430 ribosome-associated proteins whose relative size ultimately defines the total cellular growth rate. In more recent
431 years, this view has been more thoroughly dissected experimentally (?????) and together have led to a paradigm-
432 shift in how we think of cellular physiology at the proteomic-level. However, the quantitative description of these
433 observations is often couched in terms of phenomenological constants and effective parameters with the key
434 observable features of expression often computed in relative, rather than absolute, abundances. Furthermore,
435 these approaches often exclude or integrate away effects of cell size and chromosome content, which we have
436 found through our estimates to have important connections to the observed cellular growth rate.

437 In the closing sections of this work, we explore how ribosomal content, total protein abundance, and chromo-
438 somal replication are intertwined in their control over the cellular growth rate. To do so, we take a more careful
439 view of ribosome abundance, increasing the sophistication of our analysis by exchanging our order-of-magnitude
440 estimates for a minimal mathematical model of growth rate control. This is defined by parameters with tangible
441 connections to the biological processes underlying cellular growth and protein synthesis. Using this model, we in-
442 terrogate how the size of the ribosome pool and its corresponding translational capacity enable cells to maintain
443 a balance between the of amino acids via metabolism and catabolism and their consumption through the peptide
444 bond formation required for growth.

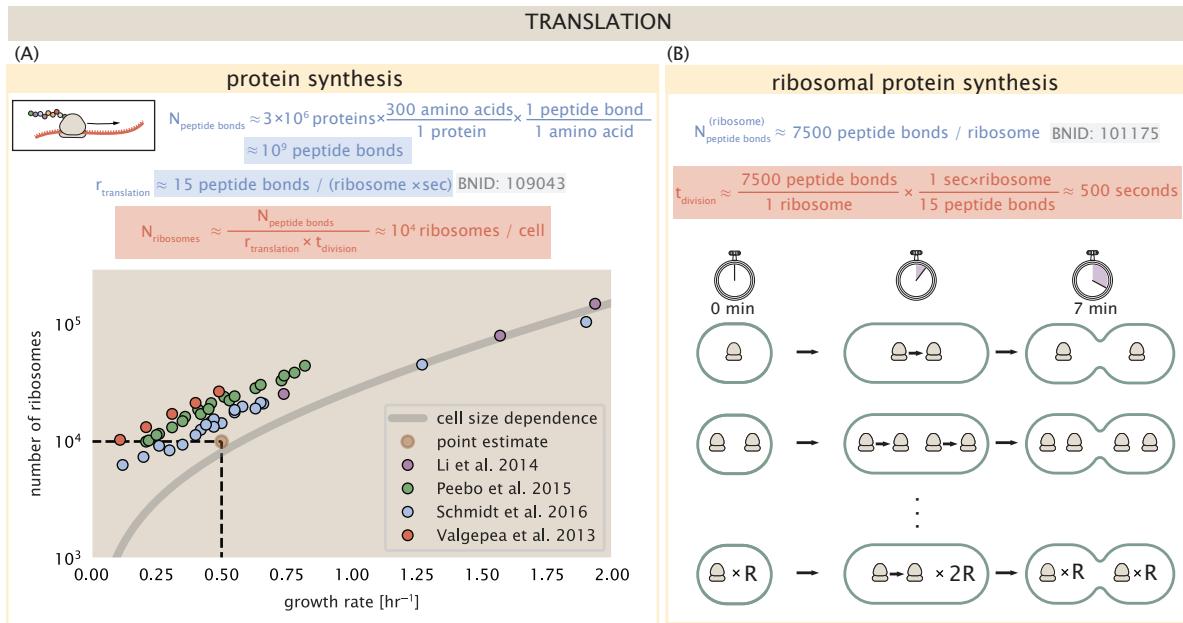


Figure 8. Estimation of the required number of ribosomes and the speed limit for bacterial replication. (A) Estimation of the number of ribosomes required to synthesize 10^9 peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr^{-1} . Grey lines correspond to the estimated complex abundance calculated at different growth rates. (B) Estimation for the time to replicate a ribosome. This rate is independent of the number of ribosomes R and instead is limited by the time required to double an individual ribosome.

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

445 Maximum Growth Rate is Determined by the Ribosomal Mass Fraction

446 The 7 minute speed limit shown in **Figure 8(B)** assumes all proteins in the cell are ribosomes. In order to connect
 447 this to the experimental data (and physiological reality more broadly), we first need to relax this assumption and
 448 determine a translation-limited growth rate. Here, we will assume that the cell is composed of N_{pep} peptide bonds
 449 and R ribosomes, whose precise values will depend on the growth rate λ . The protein subunits of each ribosomal
 450 protein sum to a total of ≈ 7500 amino acids as noted earlier, which we denote by L_R . With an average mass of
 451 an amino acid of $m_{\text{AA}} \approx 110 \text{ Da}$ (BNID: 104877), the total ribosomal mass fraction Φ_R is given by

$$\Phi_R = \frac{m_{\text{ribosomes}}}{m_{\text{proteome}}} \approx \frac{m_{\text{AA}} \times R \times L_R}{m_{\text{AA}} \times N_{\text{pep}}} = \frac{R \times L_R}{N_{\text{pep}}}. \quad (1)$$

452 For exponentially growing cells (?), the rate of cellular growth will be related to the rate of protein synthesis via

$$\lambda N_{\text{pep}} = r_t \times R \times f_a, \quad (2)$$

453 where r_t is the translation rate. Here, we've introduced a multiplicative factor f_a which represents the fraction of
 454 the ribosomes that are actively translating. This term allows us to account for immature or non-functional ribo-
 455 somes or active sequestration of ribosomes through the action of the secondary messenger alarmone (p)ppGpp
 456 in poorer nutrient conditions ?.

457 Combining **Equation 1** and **Equation 2** results in an expression for a translation-limited growth rate, which is
 458 given by

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

459 This result, derived in a similar manner in ?, reflects mass-balance under steady state growth and has long provided
 460 a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function growth rate (??). The
 461 left-hand panel of **Figure 9(A)** shows this growth rate plotted as a function of the ribosomal mass fraction. In
 462 the regime where all ribosomes are active ($f_a = 1$) and the entire proteome is composed of ribosomal proteins
 463 ($\Phi_R = 1$), indeed, we arrive at the maximum theoretical growth rate of r_t/L_R , and $\approx 7 \text{ min}$ for *E. coli*.

464 Connecting **Equation 3** to the proteomic data serving as the centerpiece of our work, however, requires knowl-
465 edge of f_a at each growth rate as proteomic measurements only provide a measure of Φ_R . Recently, ? determined
466 f_a as a function of the growth rate (**Figure 9(A)**, right-hand panel, inset), revealing that $f_a \approx 1$ at growth rates above
467 0.75 hr^{-1} and $f_a < 1$ as the growth rate slows. Using these data, we inferred the approximate active fraction (see Ap-
468 pendix Calculation of active ribosomal fraction) at each growth rate and used this to compute $\Phi_R \times f_a$ (**Figure 9(A)**,
469 colored points in right-hand panel). In general, these data skirt the translation-limited growth rate determined
470 using **Equation 3** with r_t taken to be the maximal elongation rate of 17 amino acids per second measured by ?. There is a notable discrepancy between the data collected in ?? and that collected from ???. When compared
471 to other measurements (non-proteomic with significantly lower resolution) of the active ribosome mass fraction
472 (**Figure 9(B)**, grey points in right-hand panel), the data from ? and ? are notably aberrant, suggesting a systematic
473 error in these data. These additional measurements come from a number of recent studies and are determined
474 from measurements of total RNA to total protein mass ratios (**Figure 9-Figure Supplement 1**).
475

476 Together, these results illustrate that the growth rates observed across the amalgamated data sets are close
477 to the translation-limited growth rate determined through their ribosomal activity, at least for the data reported
478 in ? and ?. While this is a useful framework to consider how the relative abundance of ribosomes (compared to
479 all other proteins) defines the growth rate, it is worth noting that as growth rate increases, so does the cell size
480 and therefore so will the total proteomic mass (?). With a handle on how elongation rate and the total number
481 of peptide bonds per proteome is related to the growth rate, we now expand this description to account for the
482 increasing cell size and ribosome copy number at faster growth rates, enabling us to identify a potential bottleneck
483 in the synthesis of rRNA.

484 **rRNA Synthesis Presents a Potential Bottleneck During Rapid Growth**

485 Even under idealized experimental conditions, *E. coli* rarely exhibits growth rates above 2 hr^{-1} (?), which is still
486 well-below the synthesis rate of a single ribosome, and below the maximum growth rates reported for several
487 other bacteria (?). While we have considered potential limits imposed by translation of ribosomal *proteins*, here
488 we consider potential limiting regimes for the production of rRNA.

489 Due to multiple initiations of chromosomal replication per cell doubling, the effective number of rRNA operons
490 increases with growth rate and will do so in proportion to the average number of origins per cell, $\langle \# \text{ ori} \rangle$. This later
491 parameter is set by how often replication must be initiated in order to keep up with cell doubling times τ whose
492 time may be shorter than the cell cycle time t_{cyc} (referring to the time from replication initiation to cell division) ?.
493 This is quantified by

$$\langle \# \text{ ori} \rangle = 2^{t_{\text{cyc}}/\tau} = 2^{t_{\text{cyc}}/4/\log(2)}. \quad (4)$$

494 We used the experimental measurements of t_{cyc} (the timescale of chromosome replication and cell division) and
495 τ (the timescale of a cell doubling) from ? (**Figure 9-Figure Supplement 1(B)**) to calculate $\langle \# \text{ ori} \rangle$ with **Equation 4**
496 as a function of growth rates. For growth rates above about 0.5 hr^{-1} , t_{cyc} is approximately constant at about 70
497 minutes, implying that $\langle \# \text{ ori} \rangle$ will grow exponentially with growth rates beyond 0.5 hr^{-1} . As the rRNA operons are
498 predominantly located close to origin of replication (BNID: 100352), we make the simplifying assumption that that
499 the number of rRNA operons will be directly proportional to $\langle \# \text{ ori} \rangle$.

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

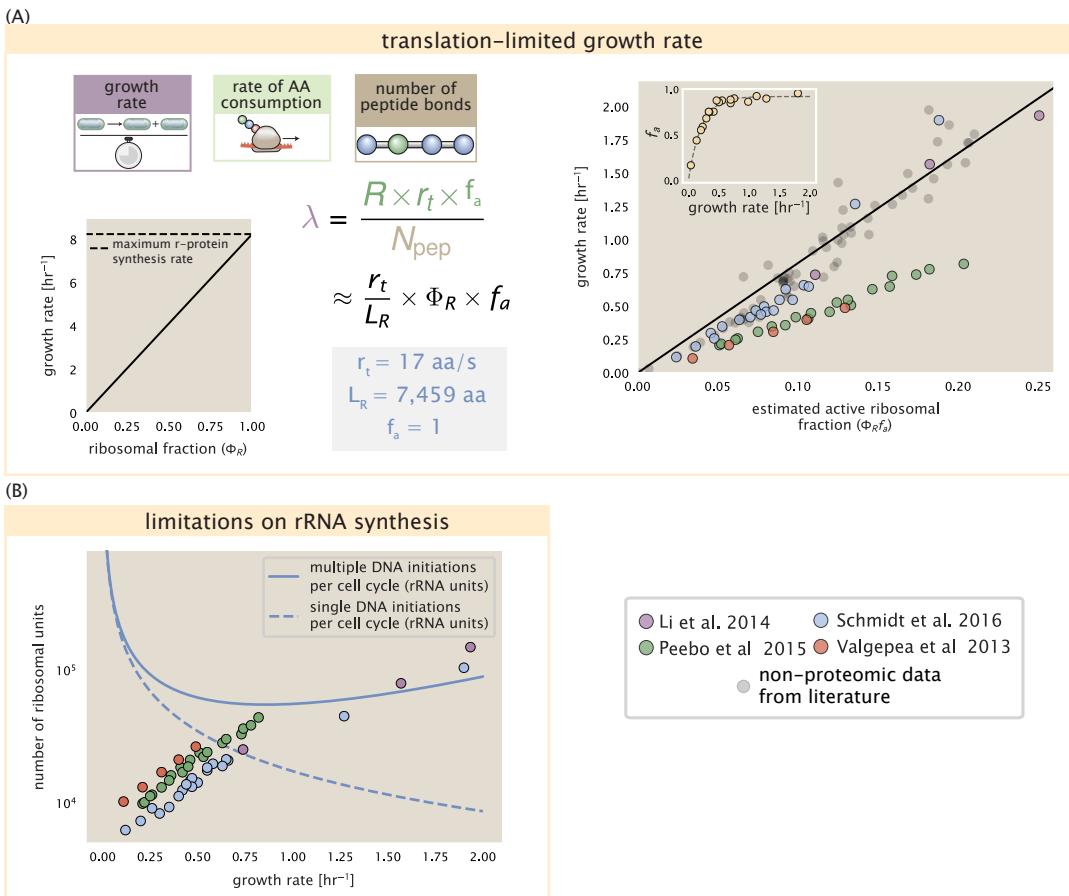


Figure 9. Translation-limited growth rate. (A) *left*: Translation-limited growth as a function of the ribosomal fraction. The solid line is calculated for an elongation rate of 17 aa per second. The dashed line corresponds to the maximum rate of ribosomal protein synthesis (≈ 7 min). *right*: Translation-limited growth rate as a function of the actively translating ribosomal fraction. The actively translating ribosomal fraction is calculated using the estimated values of f_a from ? (shown in inset; see Appendix Calculation of active ribosomal fraction for additional detail). Gray data points show additional measurements from literature and consider further in the supplemental figure part (A). (B) 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 ?. The dashed line shows the maximal number of functional rRNA units produced from a single chromosomal initiation per cell cycle.

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

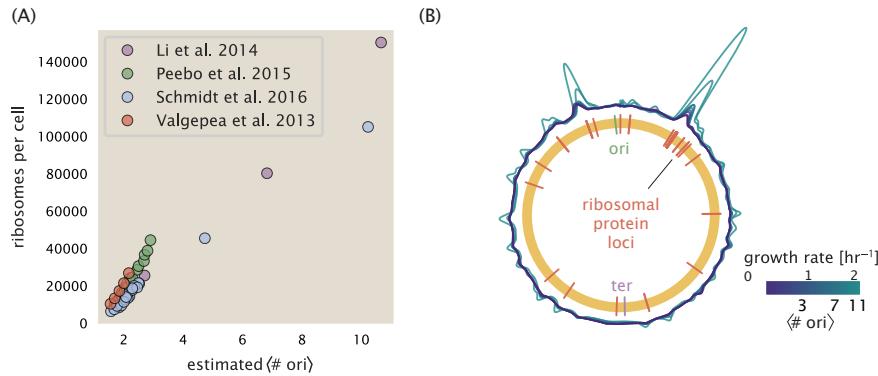


Figure 10. Cells increase both absolute ribosome abundance and Φ_R with $\langle \# \text{ori} \rangle$. (A) Plot of the ribosome copy number estimated from the proteomic data against the estimated $\langle \# \text{ori} \rangle$ (see Appendix Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ 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 (?) 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. $\langle \# \text{ori} \rangle$ are estimated using the data in (A) and Equation 4.

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

511 In *Figure 9(B*, right-hand side) we find that above about 0.75 hr^{-1} , the growth rate is determined by the ribosomal
 512 mass fraction Φ_R , since f_a is close to 1, and r_t is near its maximal rate (?). While Φ_R will need to increase in order for
 513 cells to grow faster, the fractional dependence in *Equation 3* gives little insight into how this is actually achieved
 514 in the cell.

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

526 A Minimal Model of Nutrient-Mediated Growth Rate Control

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

531 Here we propose that cells modulate their protein abundance in direct response to the availability of nutrients
 532 in their environment. As noted earlier, bacteria can modulate ribosomal activity through the secondary-messenger
 533 molecules like (p)ppGpp in poorer nutrient conditions (*Figure 9(C*) - inset; ?). Importantly, these secondary-messengers
 534 also cause global changes in transcriptional and translational activity (???). In *E. coli*, amino acid starvation leads
 535 to the accumulation of de-acylated tRNAs at the ribosome's A-site and a strong increase in (p)ppGpp synthesis
 536 activity by the enzyme RelA (?). Along with this, there is increasing evidence that (p)ppGpp also acts to inhibit the
 537 initiation of DNA replication (?), providing a potential mechanism to lower $\langle \# \text{ori} \rangle$ and maintain a smaller cell size
 538 in poorer growth conditions (?).

539 To consider this quantitatively, we assume that cells modulate their proteome (N_{pep} , R , Φ_R) to better maxi-
 540 mize their rate of peptide elongation r_t . The elongation rate r_t will depend on how quickly the ribosomes can

541 match codons with their correct amino-acyl tRNA, along with the subsequent steps of peptide bond formation
542 and translocation. This ultimately depends on the cellular concentration of amino acids, which we treat as a single
543 effective species, $[AA]_{\text{eff}}$. In our model, we determine the the rate of peptide elongation r_t and achievable growth
544 rate as simply depending on the supply of amino acids (and, therefore, also amino-acyl tRNAs), through a param-
545 eter 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
546 shown schematically in **Figure 11(A)** and derived in Appendix Derivation of Minimal Model for Nutrient-Mediated
547 Growth Rate Control. Given our observation that protein synthesis and energy production are not limiting, we
548 assume that other molecular players required by ribosomes such as elongation factors and GTP are available in
549 sufficient abundance.

550 In **Figure 11(B)**, we illustrate how the elongation rate will depend on the ribosomal copy number. Here, we
551 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 $V = 1\text{fL}$.
552 At low ribosome copy numbers, the observed elongation rate is dependent primarily on $[AA]_{\text{eff}}$ through r_{AA} [as
553 $r_t^{\max} \times R \times f_a << r_{AA}$, point (1) in **Figure 11(B)**]. As the ribosome copy number is increased such that the amino acid
554 supply rate and consumption rate are nearly equal [point (2) in **Figure 11(B)**], the observed elongation rate begins
555 to decrease sharply. When the ribosome copy number is increased even further, consumption at the maximum
556 elongation rate exceeds the supply rate, yielding a significantly reduced elongation rate [point (3) in **Figure 11B**].
557 While the elongation rate will always be dominated by the amino acid supply rate at sufficiently low ribosome
558 copy numbers, the elongation rate at larger ribosome abundances can be increased by tuning f_a such that not all
559 ribosomes are elongating, reducing the total consumption rate.

560 Optimal Ribosomal Content and Cell Size Depend on Nutrient Availability and Metabolic Capacity
561 To relate elongation rate to growth rate, we constrain the set of parameters based on our available proteomic mea-
562 surements; namely, we restrict the values of R , N_{pep} , and cell size to those associated with the amalgamated pro-
563 teomic data (described in Appendix Estimation of Total Protein Content per Cell). We then consider how changes
564 in the nutrient conditions, through the parameter r_{AA} , influence the maximum growth rate as determined by
565 **Equation 3**. **Figure 11(C)** shows how the observed growth rate depends on the rate of amino acid supply r_{AA} as
566 a function of the cellular ribosome copy number. A feature immediately apparent is the presence of a maximal
567 growth rate whose dependence on R (and consequently, the cell size) increases with increasing r_{AA} . Importantly,
568 however, there is an optimum set of R , N_{pep} , and V that are strictly dependent on the value of r_{AA} . Increasing
569 the ribosomal concentration beyond the cell's metabolic capacity has the adverse consequence of depleting the
570 supply of amino acids and a concomitant decrease in the elongation rate r_t , [**Figure 11(B)**].

571 Also of note is the growth rate profiles shown for low amino acid supply rates [purple and blue lines in **Fig-**
572 **ure 11(C)**], representing growth in nutrient-poor media. In these conditions, there no longer exists a peak in
573 growth, at least in the range of physiologically-relevant ribosome copy numbers. Instead, cells limit their pool of
574 actively translating ribosomes by decreasing f_a (?), which would help maintain the pool of available amino acids
575 $[AA]_{\text{eff}}$ and increase the achievable elongation rate. This observation is in agreement with the central premise of
576 the cellular resource allocation principle proposed by ??? and ?.

577 Discussion

578 Continued experimental and technological improvements have led to a treasure trove of quantitative biological
579 data (??????), and an ever advancing molecular view and mechanistic understanding of the constituents that sup-
580 port bacterial growth (?????). In this work we have compiled what we believe to be the state-of-the-art knowledge
581 on proteomic copy number across a broad range of growth conditions in *E. coli*. We have made this data accessible
582 through a [GitHub repository](#), and an [interactive figure](#) that allows exploration of specific protein and protein com-
583 plex copy numbers. Through a series of order-of-magnitude estimates that traverse key steps in the bacterial cell
584 cycle, this proteomic data has been a resource to guide our understanding of two key questions: what biological
585 processes limit the absolute speed limit of bacterial growth, and how do cells alter their molecular constituents
586 as a function of changes in growth rate or nutrient availability? While not exhaustive, our series of estimates pro-
587 vide insight on the scales of macromolecular complex abundance across four classes of cellular processes – the
588 transport of nutrients, the production of energy, the synthesis of the membrane and cell wall, and the numerous
589 steps of the central dogma.

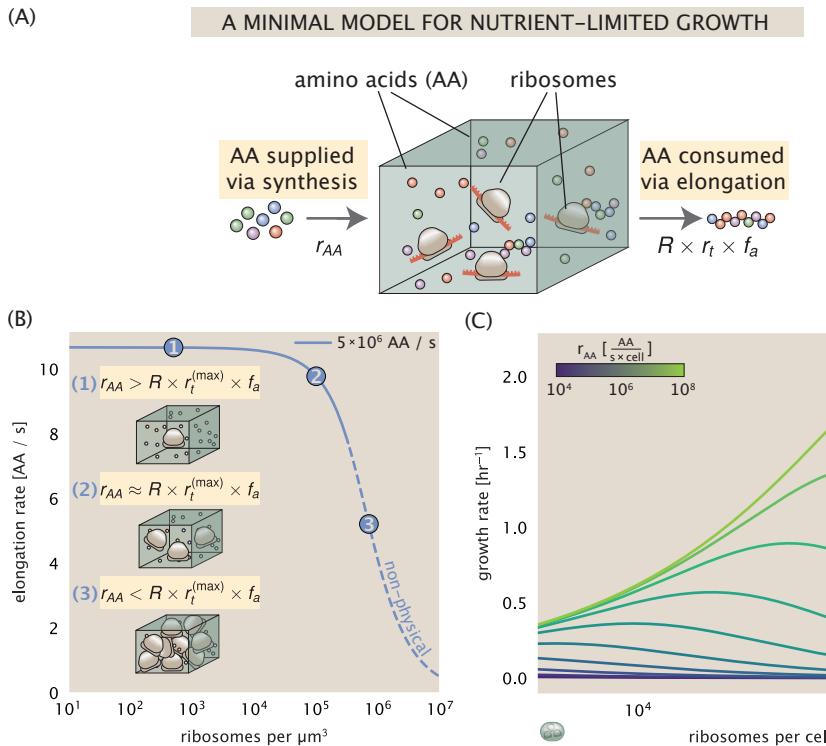


Figure 11. A minimal model of growth rate control under nutrient limitation. (A) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate r_{AA} . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate $r_t \times R \times f_a$, 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 an amino acid supply rate of 5×10^6 AA / s, a maximal elongation rate of 17.1 AA / s, $f_a = 1$, and a unit cell volume of $V = 1$ fL. See Appendix Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control 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_{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.

Figure 11–Figure supplement 1. An interactive figure for exploration of the model parameter space.

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

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

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

631 This work is by no means meant to be an exhaustive dissection of bacterial growth rate control, and there are
632 many aspects of the bacterial proteome and growth that we neglected to consider. For example, other recent
633 work (???) has explored how the proteome is structured and how that structure depends on growth rate. In the
634 work of ?, the authors coarse-grained the proteome into six discrete categories being related to either translation,
635 catabolism, anabolism, and others related to signaling and core metabolism. The relative mass fraction of the
636 proteome occupied by each sector could be modulated by external application of drugs or simply by changing
637 the nutritional content of the medium. While we have explored how the quantities of individual complexes are
638 related to cell growth, we acknowledge that higher-order interactions between groups of complexes or metabolic
639 networks at a systems-level may reveal additional insights into how these growth-rate dependences are mecha-
640 nistically achieved. Furthermore, while we anticipate the conclusions summarized here are applicable to a wide

641 collection of bacteria with similar lifestyles as *E. coli*, other bacteria and archaea may have evolved other strate-
642 gies that were not considered. Further experiments with the level of rigor now possible in *E. coli* will need to be
643 performed in a variety of microbial organisms to learn more about how regulation of proteomic composition and
644 growth rate control has evolved over the past 3.5 billion years.

645 Methods

646 Data Analysis and Availability

647 All proteomic measurements come from the experimental work of ??? (mass spectrometry) and ? (ribosomal
648 profiling). Data curation and analysis was done programmatically in Python, and compiled data and analysis
649 files are accessible through a [GitHub repository] (DOI:XXX) associated with this paper as well as on the associated
650 [paper website](#). An interactive figure that allows exploration of specific protein and protein complex copy numbers
651 is available at [link].

652 Acknowledgements

653 We thank Matthias Heinemann, Alexander Schmidt, and Gene-Wei Li for additional input regarding their data. We
654 also thank members of the Phillips, Theriot, Kondev, and Garcia labs for useful discussions. We thank Suzannah
655 M. Beeler, Avi Flamholz, Soichi Hirokawa, and Manuel Razo-Mejia for reading and providing comments on drafts
656 of this manuscript. R.P. is supported by La Fondation Pierre-Gilles de Gennes, the Rosen Center at Caltech, and
657 the NIH 1R35 GM118043 (MRA). J.A.T. is supported by the Howard Hughes Medical Institute, and NIH Grant R37-
658 AI036929. N.M.B is a HHMI Fellow of The Jane Coffin Childs Memorial Fund.

659 Competing Interests

660 The authors declare no competing interests.

⁶⁶¹ Appendix for: Fundamental limits on the ⁶⁶² rate of bacterial cell division

⁶⁶³ **Nathan M. Belliveau^{†, 1}, Griffin Chure^{†, 2}, Christina L. Hueschen³, Hernan G. Garcia⁴, Jane**
⁶⁶⁴ **Kondev⁵, Daniel S. Fisher⁶, Julie A. Theriot^{1, 7, *}, Rob Phillips^{8, 9, *}**

⁶⁶⁵ ¹Department of Biology, University of Washington, Seattle, WA, USA; ²Department of Applied Physics,
⁶⁶⁶ California Institute of Technology, Pasadena, CA, USA; ³Department of Chemical Engineering,
⁶⁶⁷ Stanford University, Stanford, CA, USA; ⁴Department of Molecular Cell Biology and Department of
⁶⁶⁸ Physics, University of California Berkeley, Berkeley, CA, USA; ⁵Department of Physics, Brandeis
⁶⁶⁹ University, Waltham, MA, USA; ⁶Department of Applied Physics, Stanford University, Stanford, CA,
⁶⁷⁰ USA; ⁷Allen Institute for Cell Science, Seattle, WA, USA; ⁸Division of Biology and Biological Engineering,
⁶⁷¹ California Institute of Technology, Pasadena, CA, USA; ⁹Department of Physics, California Institute of
⁶⁷² Technology, Pasadena, CA, USA; *Co-corresponding authors. Address correspondence to
⁶⁷³ phillips@pboc.caltech.edu and jtheriot@uw.edu; [†]These authors contributed equally to this work

674	Contents	
675	Introduction	1
676	Nutrient Transport	4
677	Limits on Transporter Expression	6
678	Cell Envelope Biogenesis	6
679	Lipid Synthesis	6
680	Peptidoglycan Synthesis	6
681	Limits on Cell Wall Biogenesis	7
682	Energy Production	7
683	ATP Synthesis	7
684	Generating the Proton Electrochemical Gradient	9
685	Limits on Biosynthesis in a Crowded Membrane	9
686	Processes of the Central Dogma	11
687	DNA Replication	12
688	RNA Synthesis	13
689	Protein Synthesis	13
690	Maximum Growth Rate is Determined by the Ribosomal Mass Fraction	15
691	rRNA Synthesis Presents a Potential Bottleneck During Rapid Growth	18
692	Rapid Growth Requires <i>E. coli</i> to Increase Both Cell Size and Ribosomal Mass Fraction	18
693	A Minimal Model of Nutrient-Mediated Growth Rate Control	18
694	Optimal Ribosomal Content and Cell Size Depend on Nutrient Availability and Metabolic Capacity	19
695	Discussion	21
696	Methods	22
697	Data Analysis and Availability	22
698	Acknowledgements	22
699	Competing Interests	22
700	Additional Estimates of Fundamental Biological Processes	25
701	Nutrient Transport	25
702	Nitrogen	25
703	Phosphorus	25
704	Sulfur	25
705	Additional Process of the Central Dogma	26
706	dNTP synthesis	26
707	mRNA and tRNA Synthesis	26
708	tRNA Charging	27
709	Experimental Details Behind Proteomic Data	28
710	Fluorescence based measurements	28
711	Ribosomal profiling measurements	28
712	Mass spectrometry measurements	29
713	Summary of Proteomic Data	29
714	Estimation of Cell Size and Surface Area	30

715	Estimation of Total Protein Content per Cell	31
716	Estimating Volume and Number of Amino Acids from Ribosome Copy Number	32
717	Additional Considerations of Schmidt <i>et al.</i> Data Set	32
718	Effect of cell volume on reported absolute protein abundances	34
719	Relaxing assumption of constant protein concentration across growth conditions	36
720	Comparison with total protein measurements from Basan <i>et al.</i> 2015.	36
721	Calculation of Complex Abundance	36
722	Extending Estimates to a Continuum of Growth Rates	38
723	Estimation of the total cell mass	39
724	Complex Abundance Scaling With Cell Volume	39
725	A Relation for Complex Abundance Scaling With Surface Area	40
726	Number of Lipids	40
727	Number of Murein Monomers	40
728	Complex Abundance Scaling With Number of Origins, and rRNA Synthesis	40
729	Calculation of active ribosomal fraction.	41
730	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	41
731	Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control	42

732 Additional Estimates of Fundamental Biological Processes

733 In the main text of this work, we present estimates for a significant number of fundamental biological processes
734 that are necessary for cell division. While we believe the estimates provided in the main text provide a succinct
735 summary of the corresponding process, we left out additional estimates of related processes for brevity. In this
736 section of the appendix, we present these additional estimates in full.

737 Nutrient Transport

738 In the main text, we make passing mention that while transport carbon often comes in the form of carbohydrates
739 and sugar alcohols while other critical elements – such as nitrogen, sulfur, and phosphorus – are transported as
740 inorganic ions. Below, we present estimates for the transport requirements of these materials.

741 Nitrogen

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

753 Phosphorus

754 Phosphorus is critical to the cellular energy economy in the form of high-energy phosphodiester bonds making
755 up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-translational modification of
756 proteins and defining the polar-heads of lipids. In total, phosphorus makes up $\approx 3\%$ of the cellular dry mass which
757 in typical experimental conditions is in the form of inorganic phosphate. The cell membrane has remarkably low
758 permeability to this highly-charged and critical molecule, therefore requiring the expression of active transport
759 systems. In *E. coli*, the proton electrochemical gradient across the inner membrane is leveraged to transport
760 inorganic phosphate into the cell (?). Proton-solute symporters are widespread in *E. coli* (??) and can have rapid
761 transport rates of 50 to 100 molecules per second for sugars and other solutes (BNID: 103159; 111777). As a
762 more extreme example, the proton transporters in the F₁-F₀ ATP synthase, which use the proton electrochemical
763 gradient for rotational motion, can shuttle protons across the membrane at a rate of ≈ 1000 per second (BNID:
764 104890; 103390). In *E. coli* the PitA phosphate transport system has been shown to be very tightly coupled with the
765 proton electrochemical gradient with a 1:1 proton:phosphate stoichiometric ratio (??). Taking the geometric mean
766 of the aforementioned estimates gives a plausible rate of phosphate transport on the order of 300 per second.
767 Illustrated in *Figure 2–Figure Supplement 1(A)*, we can estimate that ≈ 200 phosphate transporters are necessary
768 to maintain an $\approx 3\%$ dry mass with a 5000 s division time. This estimate is consistent with observation when we
769 examine the observed copy numbers of PitA in proteomic data sets (plot in *Figure 2–Figure Supplement 1(A)*). While
770 our estimate is very much in line with the observed numbers, we emphasize that this is likely a slight overestimate
771 of the number of transporters needed as there are other phosphorous scavenging systems, such as the ATP-
772 dependent phosphate transporter Pst system which we have neglected.

773 Sulfur

774 Similar to phosphate, sulfate is highly-charged and not particularly membrane permeable, requiring active trans-
775 port. While there exists a H⁺/sulfate symporter in *E. coli*, it is in relatively low abundance and is not well charac-
776 terized (?). Sulfate is predominantly acquired via the ATP-dependent ABC transporter CysUWA system which also
777 plays an important role in selenium transport (??). While specific kinetic details of this transport system are not
778 readily available, generic ATP transport systems in prokaryotes transport on the order of 1 to 10 molecules per

779 second (BNID: 109035). Combining this generic transport rate, measurement of sulfur comprising 1% of dry mass,
780 and a 5000 second division time yields an estimate of \approx 1000 CysUWA complexes per cell (**Figure 2–Figure Supple-**
781 **ment 1(B)**). Once again, this estimate is in notable agreement with proteomic data sets, suggesting that there are
782 sufficient transporters present to acquire the necessary sulfur. In a similar spirit of our estimate of phosphorus
783 transport, we emphasize that this is likely an overestimate of the number of necessary transporters as we have
784 neglected other sulfur scavenging systems that are in lower abundance.

785 Additional Process of the Central Dogma

786 In the main text, we consider the processes underlying the backbone of the central dogma, namely DNA replication,
787 RNA transcription, and protein translation. In this section we turn our attention to additional processes related
788 to the central dogma, primarily dNTP synthesis for DNA replication and amino-acyl tRNA synthesis for translation.
789 Additionally, we explore in more detail the estimates shown in **Figure 7(A)** for the RNA polymerase requirements
790 of mRNA and tRNA synthesis.

791 dNTP synthesis

792 The four major dNTPs (dATP, dTTP, dCTP, and dGTP) serve as the fundamental units of the genetic code. Thus, to
793 faithfully replicate the chromosome, the cell must be able to synthesize enough of these bases in the first place.
794 All dNTPs are synthesized *de novo* in separate pathways, requiring different building blocks. However, a critical
795 step present in all dNTP synthesis pathways is the conversion from ribonucleotide to deoxyribonucleotide via the
796 removal of the 3' hydroxyl group of the ribose ring (?). This reaction is mediated by a class of enzymes termed
797 ribonucleotide reductases, of which *E. coli* possesses two aerobically active complexes (termed I and II) and a
798 single anaerobically active enzyme. Due to their peculiar formation of a radical intermediate, these enzymes have
799 received much biochemical, kinetic, and structural characterization. One such work (?) performed a detailed *in*
800 *vitro* measurement of the steady-state kinetic rates of these complexes, revealing a turnover rate of \approx 10 dNTP
801 per second.

802 Since this reaction is central to the synthesis of all dNTPs, it is reasonable to consider the abundance of these
803 complexes as a measure of the total dNTP production in *E. coli*. Illustrated schematically in **Figure 6 (A)**, we consider
804 the fact that to replicate the cell's genome, on the order of \approx 10⁷ dNTPs must be synthesized. Assuming a pro-
805 duction rate of 10 per second per ribonucleotide reductase complex and a cell division time of 5000 seconds, we
806 arrive at an estimate of \approx 200 complexes needed per cell. As shown in the bottom panel of **Figure Supplement 1**
807 (A), this estimate agrees with the experimental measurements of these complexes abundances within \approx 1/2 an
808 order of magnitude. Extension of this estimate across a continuum of growth rate, including the fact that multiple
809 chromosomes can be replicated at a given time, is shown as a grey transparent line in **Figure Supplement 1**. Sim-
810 ilarly to our point estimate, this refinement agrees well with the data, accurately describing both the magnitude
811 of the complex abundance and the dependence on growth rate.

812 Recent work has revealed that during replication, the ribonucleotide reductase complexes coalesce to form
813 discrete foci colocalized with the DNA replisome complex (?). This is particularly pronounced in conditions where
814 growth is slow, indicating that spatial organization and regulation of the activity of the complexes plays an impor-
815 tant role.

816 mRNA and tRNA Synthesis

817 In **Figure 7** of the main text, we present with limited explanation estimates for the number of RNA polymerases
818 needed to synthesize enough mRNA and tRNA molecules. Here, we present a rationalization for these estimates.

819 To form a functional protein, all protein coding genes must first be transcribed from DNA to form an mRNA
820 molecule. While each protein requires an mRNA blueprint, many copies of the protein can be synthesized from
821 a single mRNA. Factors such as strength of the ribosomal binding site, mRNA stability, and rare codon usage
822 frequency dictate the number of proteins that can be made from a single mRNA, with yields ranging from 10¹
823 to 10⁴ (BNID: 104186; 100196; 106254). Computing the geometric mean of this range yields \approx 1000 proteins
824 synthesized per mRNA, a value that agrees with experimental measurements of the number of proteins per cell
825 (\approx 3 \times 10⁶, BNID: 100088) and total number of mRNA per cell (\approx 3 \times 10³, BNID: 100064).

826 This estimation captures the *steady-state* mRNA copy number, meaning that at any given time, there will exist
827 approximately 3000 unique mRNA molecules. To determine the *total* number of mRNA that need to be synthesized
828 over the cell's lifetime, we must consider degradation of the mRNA. In most bacteria, mRNAs are rather unstable
829 with life times on the order of several minutes (BNID: 104324; 106253; 111927; 111998). For convenience, we
830 assume that the typical mRNA in our cell of interest has a typical lifetime of \approx 300 seconds. Using this value, we
831 can determine the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell.
832 While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional dynamics, we
833 state here that the total mRNA production rate must be on the order of \approx 15 mRNA per second. In *E. coli*, the
834 average protein is \approx 300 amino acids in length (BNID: 108986), meaning that the corresponding mRNA is \approx 900
835 nucleotides which we will further approximate as \approx 1000 nucleotides to account for the non-protein coding regions
836 on the 5' and 3' ends. This means that the cell must have enough RNA polymerase molecules about to sustain a
837 transcription rate of \approx 1.5×10^4 nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA
838 at a clip of 40 nucleotides per second, we arrive at a comfortable estimate of \approx 250 RNA polymerase complexes
839 needed to satisfy the mRNA demands of the cell. It is worth noting that this number is approximately half of that
840 required to synthesize enough rRNA, as we saw in the previous section. We find this to be a striking result as these
841 250 RNA polymerase molecules are responsible for the transcription of the \approx 4000 protein coding genes that are
842 not ribosome associated.

843 We now turn our attention to the synthesis of tRNA. Unlike mRNA or rRNA, each individual tRNA is remarkably
844 short, ranging from 70 to 95 nucleotides each (BNID: 109645; 102340). What they lack in length, they make up
845 for in abundance, with reported values ranging from $\approx 5 \times 10^4$ (BNID: 105280) to $\approx 5 \times 10^5$ (BNID: 108611). To test
846 tRNA synthesis as a possible growth-rate limiting stage, we will err towards a higher abundance of $\approx 5 \times 10^5$ per
847 cell. Combining the abundance and tRNA length measurements, we make the estimate that $\approx 5 \times 10^7$ nucleotides
848 are sequestered in tRNA per cell. Unlike mRNA, tRNA is remarkably stable with typical lifetimes *in vivo* on the
849 order of \approx 48 hours (??) – well beyond the timescale of division. Once again using our rule-of-thumb for the
850 rate of transcription to be 40 nucleotides per second and assuming a division time of \approx 5000 seconds, we arrive
851 at an estimate of \approx 200 RNA polymerases to synthesize enough tRNA. This requirement pales in comparison to
852 the number of polymerases needed to generate the rRNA and mRNA pools and can be neglected as a significant
853 transcriptional burden.

854 tRNA Charging

855 In the previous subsection, we focused solely on estimating the number of RNA polymerases needed for the
856 generation of the tRNA molecule itself. We now explore the protein complex requirements for ligation of the
857 appropriate amino acid to each tRNA. We begin by again using an estimate of $\approx 3 \times 10^6$ proteins per cell at a 5000
858 s division time (BNID: 115702) and a typical protein length of \approx 300 amino acids (BNID: 100017), we can estimate
859 that a total of $\approx 10^9$ amino acids are stitched together by peptide bonds.

860 How many tRNAs are needed to facilitate this remarkable number of amino acid delivery events to the trans-
861 lating ribosomes? It is important to note that tRNAs are recycled after they've passed through the ribosome and
862 can be recharged with a new amino acid, ready for another round of peptide bond formation. While some *in vitro*
863 data exists on the turnover of tRNA in *E. coli* for different amino acids, we can make a reasonable estimate by
864 comparing the number of amino acids to be polymerized to cell division time. Using our stopwatch of 5000 s and
865 10^9 amino acids, we arrive at a requirement of $\approx 2 \times 10^5$ tRNA molecules to be consumed by the ribosome per
866 second.

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

Table 1. Overview of proteomic data sets.

Author	Method	Reported Quantity
Taniguchi <i>et al.</i> (2010)	YFP-fusion, cell fluorescence	fg/copies per cell
Valgepea <i>et al.</i> (2012)	mass spectrometry	fg/copies per cell
Peebo <i>et al.</i> (2014)	mass spectrometry	fg/copies per fl
Li <i>et al.</i> (2014)	ribosomal profiling	fg/copies per cell ^a
Soufi <i>et al.</i> (2015)	mass spectrometry	fg/copies per cell
Schmidt <i>et al.</i> (2016)	mass spectrometry	fg/copies per cell ^b

a. The reported values assume that the proteins are long-lived compared to the generation time but are unable to account for post-translational modifications that may alter absolute protein abundances.

b. This mass spectrometry approach differs substantially from the others since in addition to the relative proteome-wide abundance measurements, the authors performed absolute quantification of 41 proteins across all growth conditions (see section on Additional Considerations of Schmidt *et al.* Data Set for more details on this).

We can make an assumption that amino-acyl tRNAs are in steady-state where they are produced at the same rate they are consumed, meaning that 2×10^5 tRNAs must be charged per second. Combining these estimates together, as shown schematically in *Figure 8–Figure Supplement 1*, yields an estimate of $\sim 10^4$ tRNA synthetases per cell with a division time of 5000 s. This point estimate is in very close agreement with the observed number of synthetases (the sum of all 20 tRNA synthetases in *E. coli*). This estimation strategy seems to adequately describe the observed growth rate dependence of the tRNA synthetase copy number (shown as the grey line in *Figure 8–Figure Supplement 1*, suggesting that the copy number scales with the cell volume.

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

Experimental Details Behind Proteomic Data

Here we provide a brief summary of the experiments behind each proteomic data set considered. The purpose of this section is to identify how the authors arrived at absolute protein abundances. In the following section (see section on Summary of Proteomic Data) we will then provide a summary of the protein abundance measurements. Table 1 provides an overview of the publications we considered. These are predominately mass spectrometry-based, with the exception of the work from ? which used ribosomal profiling, and the fluorescence-based counting done in ?. After having compiled and comparing these measurements, we noted substantial deviations in the measurements from ? and ? (shown in the following section), and decided to only use the data from ??? in the main text. For completeness, we include these additional datasets in our discussion of the experimental data.

Fluorescence based measurements

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

Ribosomal profiling measurements

The work of ? takes a sequencing based approach to estimate protein abundance. Ribosomal profiling, which refers to the deep sequencing of ribosome-protected mRNA fragments, can provide a quantitative measurement

908 of the protein synthesis rate. As long as the protein life-time is long relative to the cell doubling time, it is possible to
909 estimate absolute protein copy numbers. The absolute protein synthesis rate has units of proteins per generation,
910 and for stable proteins will also correspond to the protein copy number per cell.

911 In the experiments, ribosome-protected mRNA is extracted from cell lysate and selected on a denaturing poly-
912 acrylamide gel for deep sequencing (15–45 nt long fragments collected and sequenced by using an Illumina HiSeq
913 2000 in ?). Counts of ribosome footprints from the sequencing data were then corrected empirically for position-
914 dependent biases in ribosomal density across each gene, as well as dependencies on specific sequences including
915 the Shine-Dalgarno sequence. These data-corrected ribosome densities represent relative protein synthesis rates.
916 Absolute protein synthesis rates are obtained by multiplying the relative rates by the total cellular protein per cell.
917 The total protein per unit volume was determined with the Lowry method to quantify total protein, calibrated
918 against bovine serum albumin (BSA). By counting colony-forming units following serial dilution of their cell cul-
919 tures, they then calculated the total protein per cell.

920 Mass spectrometry measurements

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

928 Most quantitative experiments rely on labeling protein with stable isotopes, which allow multiple samples to
929 be measured together by the mass spectrometer. By measuring samples of known total protein abundance sim-
930 taneously (i.e. one sample of interest, and one reference), it is possible to determine relative protein abundances.
931 Absolute protein abundances can be estimated following the same approach used above for ribosomal profiling,
932 which is to multiply each relative abundance measurement by the total cellular protein per cell. This is the ap-
933 proach taken by ?? and ?, with relative protein abundances determined based on the relative peptide intensities
934 (label free quantification 'LFQ' intensities). For the data of ?, total protein per cell was determined by measuring
935 total protein by the Lowry method, and counting colony-forming units following serial dilution. For the data from
936 ?, the authors did not determine cell quantities and instead report the cellular protein abundances in protein per
937 unit volume by assuming a mass density of 1.1 g/ml, with a 30% dry mass fraction.

938 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide fragments of known
939 abundance. These can serve as a direct way to calibrate mass spectrometry signal intensities to absolute mass.
940 This is the approach taken by ?. In addition to a set of shotgun proteomic measurements to determine proteome-
941 wide relative abundances, the authors also performed absolute quantification of 41 proteins covering over four
942 orders of magnitude in cellular abundance. Here, a synthetic peptide was generated for each of the proteins,
943 doped into each protein sample, and used these to determine absolute protein abundances of the 41 proteins.
944 These absolute measurements, determined for every growth condition, were then used as a calibration curve to
945 convert proteomic-wide relative abundances into absolute protein abundance per cell. A more extensive discus-
946 sion of the ? data set can be found in Section Additional Considerations of Schmidt *et al.* Data Set.

947 Summary of Proteomic Data

948 In the work of the main text we only used the data from ?????. As shown in **Figure 12(A)**, the reported total protein
949 abundances in the work of ? and ? differed quite substantially from the other work. For the work of ? this is in part
950 due to a lower coverage in total proteomic mass quantified, though we also noticed that most proteins appear
951 undercounted when compared to the other data.

952 **Figure 12(B)** summarizes the total protein mass for each data set used in our final compiled data set. Our
953 inclination initially was to leave reported copy numbers untouched, but a notable descrepency between the scaling
954 of the total protein per cell between ? and the other data sets forced us to dig deeper into those measurements
955 (compare ? and ? data in **Figure 12(A)**). The particular trend in ? appears to be due to assumptions made about cell
956 size and we provide a more extensive discussion and analysis of their data in Additional Considerations of Schmidt

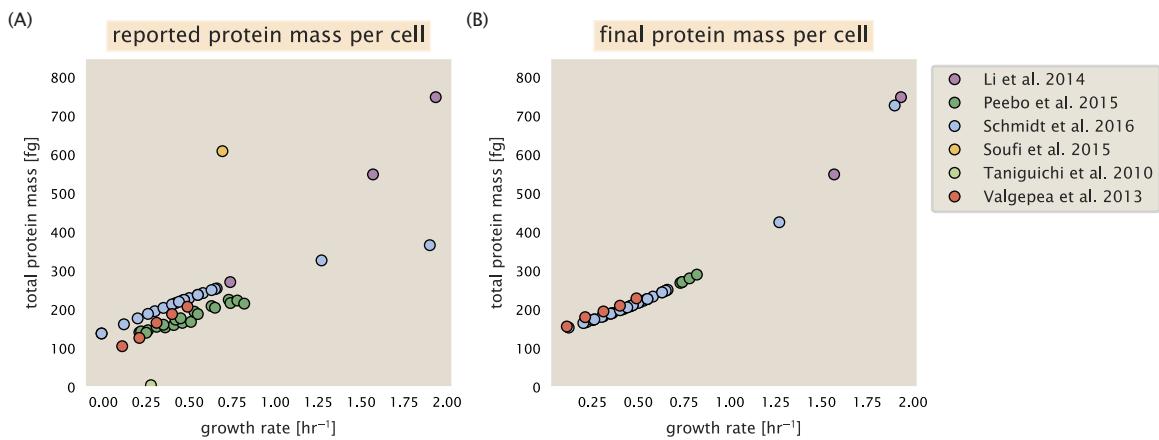


Figure 12. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as originally reported in the data sets of ??????. Note that the data from ? only reported protein abundances per unit volume and total protein per cell was found by multiplying these by the growth-rate dependent cell size as determined by ?. (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 ? (see section on Estimation of Cell Size and Surface Area for further details).

957 et al. Data Set. As a compromise, and in an effort to treat all data equally, we instead applied an correction factor
 958 to all protein abundance values based on a data-driven estimate of total protein per cell. Here we used cell size
 959 measurements from ??, and an estimate of total protein content through expected dry mass. Total protein per cell
 960 was then determined using available data on total DNA, RNA, and protein from ??, which account for the majority
 961 of dry mass in the cell. We describe these details further in sections on Estimation of Cell Size and Surface Area
 962 and Estimation of Total Protein Content per Cell that follows.

963 Lastly, in *Figure 13* we show the total proteomic coverage and overlap of proteins quantified across each data
 964 set. Here we have used an UpSet diagram (?) to compare the data. Overall, the overlap in quantified proteins is
 965 quite high, with 1157 proteins quantified across all data sets. The sequencing based approach of ? has substan-
 966 tially higher coverage compared to the mass spectrometry data sets (3394 genes versus the 2041 genes quantified
 967 in the work of ?). However, in terms of total protein mass, the data from ??? each quantify roughly equivalent total
 968 protein mass. An exception to this is in the data from ?, where we find that the total protein quantified in ? is 90-95
 969 % of the total protein mass (when using the data from ? as a reference).

970 Estimation of Cell Size and Surface Area

971 Since most of the proteomic data sets lack cell size measurements, we chose instead to use a common estimate
 972 of size for any analysis requiring cell size or surface area. Since each of the data sets used either K-12 MG1655 or
 973 its derivative, BW25113 (from the lab of Barry L. Wanner; the parent strain of the Keio collection (??)), below we
 974 fit the MG1655 cell size data from the supplemental material of ?? using the optimize.curve_fit function from the
 975 Scipy python package (?). A quick comment on nomenclature: throughout the text, we usually refer to cell size, in
 976 units of μm^3 ; however, on occasion we will mention size as a volume in units of fL.

977 The average size measurements from each of their experiments are shown in *Figure 14*, with cell length and
 978 width shown in (A) and (B), respectively. The length data was well described by the exponential function $0.5 e^{1.09 \cdot \lambda}$
 979 + 1.76 μm , while the width data was well described by $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. In order to estimate cell size we take the cell
 980 as a cylinder with two hemispherical ends (??). Specifically, cell size is estimated from,

$$V = \pi \cdot r^2 \cdot (l - 2r/3), \quad (5)$$

981 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

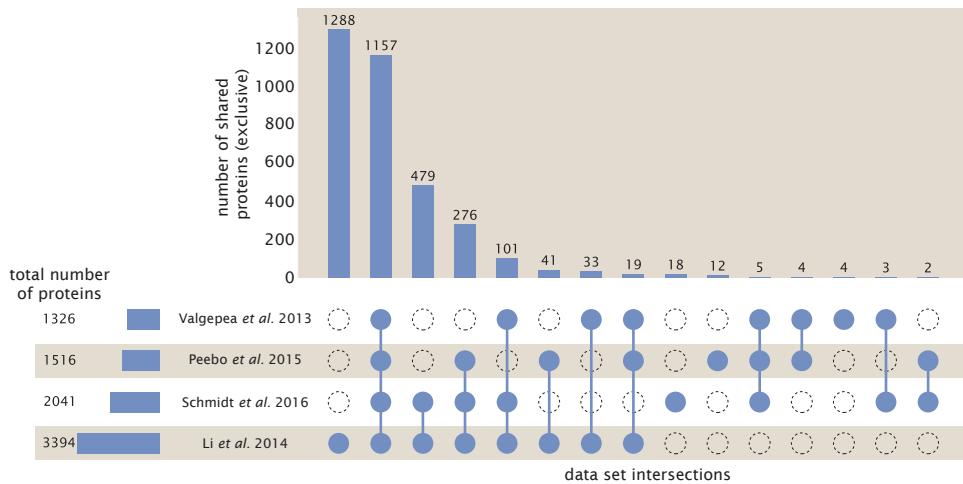


Figure 13. Comparison of proteomic coverage across different data sets. An UpSet diagram (?) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of ??. Bar plot on bottom left indicates the total number of genes reported in each individual data set. The main bar plot summarizes the number of unique proteins identified across overlapping subsets of the data. For example, in the first column only the data from ? is considered (indicated by solid blue circle) and 1288 proteins are identified as exclusive to the data set. In the second column, the intersection of all four data sets is considered, with 1157 proteins quantified across them. This follows for each additional column in the plot, with the subset under consideration denoted by the solid blue circles.

982 area is given by,

$$S = \eta \cdot \pi \left(\frac{\eta \cdot \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V^{2/3}, \quad (6)$$

983 where η is the aspect ratio ($\eta = l/w$) (?).

984 Estimation of Total Protein Content per Cell

985 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell size from the
 986 fit shown in **Figure 14(C)** (cell size = $0.533 e^{1.037 \cdot \lambda}$ μm^3 , as noted in the previous section). We then estimate the total
 987 protein content from the total dry mass of the cell. Here we begin by noting that for almost the entire range of
 988 growth rates considered here, protein, DNA, and RNA were reported to account for at least 90 % of the dry mass
 989 (?). The authors also found that the total dry mass concentration was roughly constant across growth conditions.
 990 Under such a scenario, we can calculate the total dry mass concentration for protein, DNA, and RNA, which is
 991 given by $1.1 \text{ g/ml} \times 30 \% \times 90 \% \text{ or about } [M_p] = 300 \text{ fg per fL}$. Multiplying this by our prediction of cell size gives
 992 the total dry mass per cell.

993 However, even if dry mass concentration is relatively constant across growth conditions, it is not obvious how
 994 protein concentration might vary due to the substantial increase in rRNA at faster growth rates (?). The increase in
 995 rRNA increases from the linear increase in ribosomal content with faster growth rate (?), since it makes up about
 996 about 2/3 of the ribosomal mass. To proceed we therefore relied on experimental measurements of total DNA
 997 content per cell from ?, and RNA to protein ratios that were measured in Dai *et al.* (and cover the entire range of
 998 growth conditions considered here). These are reproduced in **Figure 15(A)** and (B), respectively.

999 Assuming that the protein, DNA, and RNA account for 90 % of the total dry mass, the protein mass can then de-
 1000 termined by first subtracting the experimentally measured DNA mass, and then using the experimental estimate
 1001 of the RNA to protein ratio. The total protein per cell is will be related to the summed RNA and protein mass by,

$$M_p = \frac{[M_p + M_{RNA}]}{1 + (RP_{ratio})}. \quad (7)$$

1002 (RP_{ratio} refers to the RNA to protein ratio as measured by Dai *et al.*). In **Figure 15(C)** we plot the estimated cellular
 1003 concentrations for protein, DNA, and RNA from these calculations, and in **Figure 15(D)** we plot their total expected

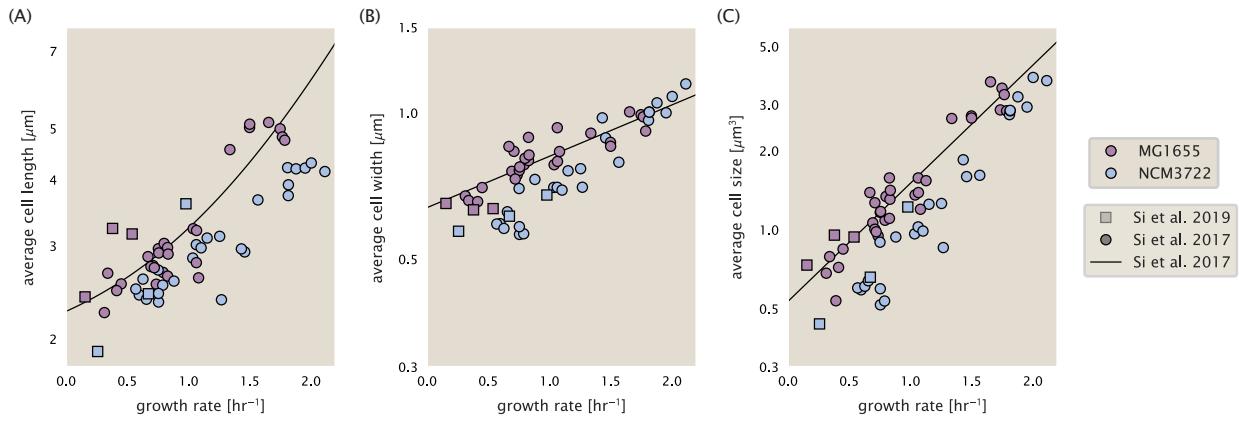


Figure 14. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 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: $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, width: $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. (C) Cell size was calculated as cylinders with two hemispherical ends (Equation 5). The MG1655 strain data gave a best fit of $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$.

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)*).

1006 Estimating Volume and Number of Amino Acids from Ribosome Copy Number

1007 Towards the end of the main text, we examine a coarse-grained model of nutrient-limited growth. A key point
 1008 in our analysis was to consider how elongation rate r_i and growth rate λ vary with respect to the experimentally
 1009 observed changes in cell size, total number of peptide bonds per cell N_{pep} , and ribosomal content. In order to
 1010 restrict parameters to those observed experimentally, but otherwise allow us to explore the model, we performed
 1011 a phenomenological fit of N_{pep} and V as a function of the measured ribosomal copy number R . As has been
 1012 described in the preceding sections of this supplement, we estimate cell volume for each growth condition using
 1013 the size measurements from ??, and N_{pep} is approximated by taking the total protein mass and dividing this
 1014 number by the average mass of an amino acid, 110 Da (BNID: 104877).

1015 Given the exponential scaling of V and N_{pep} with growth rate, we performed a linear regression of the log trans-
 1016 form of these parameters as a function of the log transform of the ribosome copy number. Using optimization
 1017 by minimization, we estimated the best-fit values of the intercept and slope for each regression. *Figure 16* shows
 1018 the result of each regression as a dashed line.

1019 Additional Considerations of Schmidt et al. Data Set

1020 While the data set from ? remains a heroic effort that our labs continue to return to as a resource, there were
 1021 steps taken in their calculation of protein copy number that we felt needed further consideration. In particular,
 1022 the authors made an assumption of constant cellular protein concentration across all growth conditions and
 1023 used measurements of cell volume that appear inconsistent with an expected exponential scaling of cell size
 1024 with growth rate that is well-documented in *E. coli* (??).

1025 We begin by looking at their cell volume measurements, which are shown in blue in *Figure 17*. As a
 1026 comparison, we also plot cell sizes reported in three other recent papers: measurements from Taheri-Araghi *et*
 1027 *al.* and Si *et al.* come from the lab of Suckjoon Jun, while those from Basan *et al.* come from the lab of Terence
 1028 Hwa. Each set of measurements used microscopy and cell segmentation to determine the length and width,
 1029 and then calculated cell size by treating the cell as a cylinder with two hemispherical ends, as we considered
 1030 in the previous section. While there is notable discrepancy between the two research groups, which are both
 1031 using strain NCM3722, Basan *et al.* found that this came specifically from uncertainty in determining the cell
 1032 width. This is prone to inaccuracy given the small cell size and optical resolution limits (further described in their

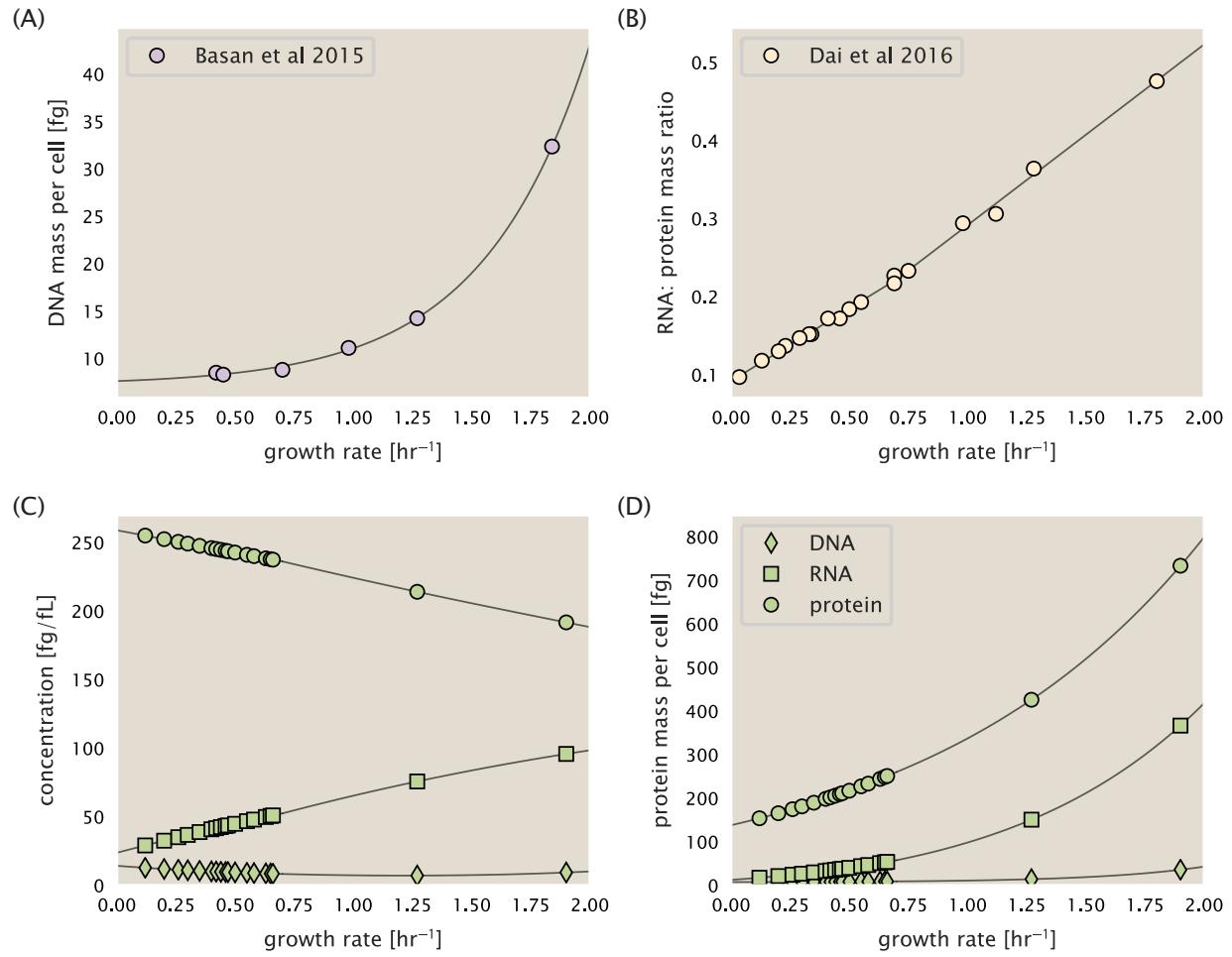


Figure 15. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23 \cdot \lambda} + 7.2$ fg per cell, where λ is the growth rate in hr^{-1}). (B) RNA to protein measurements as a function of growth rate. The data was fit to two lines (shown in black) due to the change in slope at slower growth rates ?? . For growth rates below 0.7 hr^{-1} , the RNA/protein ratio is $0.18 \cdot \lambda + 0.093$, while for growth rates faster than 0.7 hr^{-1} the RNA/protein ratio is given by $0.25 \cdot \lambda + 0.035$. For (A) and (B) cells are grown under varying levels of nutrient limitation, with cells grown in minimal media with different carbon sources for the slowest growth conditions, and rich-defined media for fast growth rates. (C) Estimation of cellular protein, DNA, and RNA concentration. (D) Total cellular mass estimated for protein, DNA, and RNA using the cell size calculated in Estimation of Cell Size and Surface Area. Symbols (diamond: DNA, square: RNA, circle: protein) show estimated values of mass concentration and mass per cell for the specific growth rates in ? .

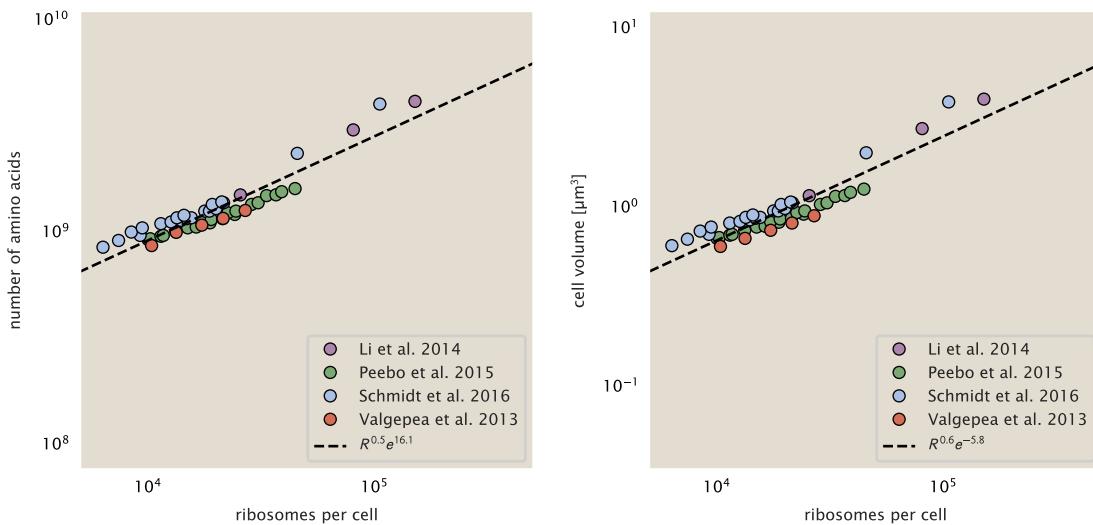


Figure 16. Phenomenological regression of cell volume and number of amino acids per cell as a function of the ribosome copy number. Colored points correspond to the measured value (or calculated value in the case of the cell volume) with colors denoting different data sets. The dashed black line shows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

supplemental text). Perhaps the more concerning point is that while each of these alternative measurements show an exponential increase in cell size at faster growth rates, the measurements used by Schmidt *et al.* appear to plateau. This resulted in an analogous trend in their final reported total cellular protein per cell as shown in **Figure 18** (purple data points), and is in disagreement with other measurements of total protein at these growth rates (?).

Since it is not obvious how measurements of cell size influenced their reported protein abundances, in the following subsections we begin by considering how the authors determined total protein mass per cell. We then consider three different approaches to estimate the growth-rate dependent total protein mass and compare these estimates with those reported by ?. Those results are summarized in **Figure 17(B)**, with the original values from both ? and ? shown in **Figure 17(A)** for reference. For most growth conditions, we find reasonable agreement between our estimates and the reported total protein per cell. However, for the fastest growth conditions, with glycerol + supplemented amino acids, and LB media, all estimates are substantially higher than those originally reported. This is the main reason why we chose to readjust protein abundance as shown in **Figure 12(B)** (with the calculation described in section Estimation of Total Protein Content per Cell).

Effect of cell volume on reported absolute protein abundances

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

While this approach provided absolute abundances, another necessary step to arrive at total cellular protein was to account for any protein loss during their various protein extraction steps. Here the authors attempted

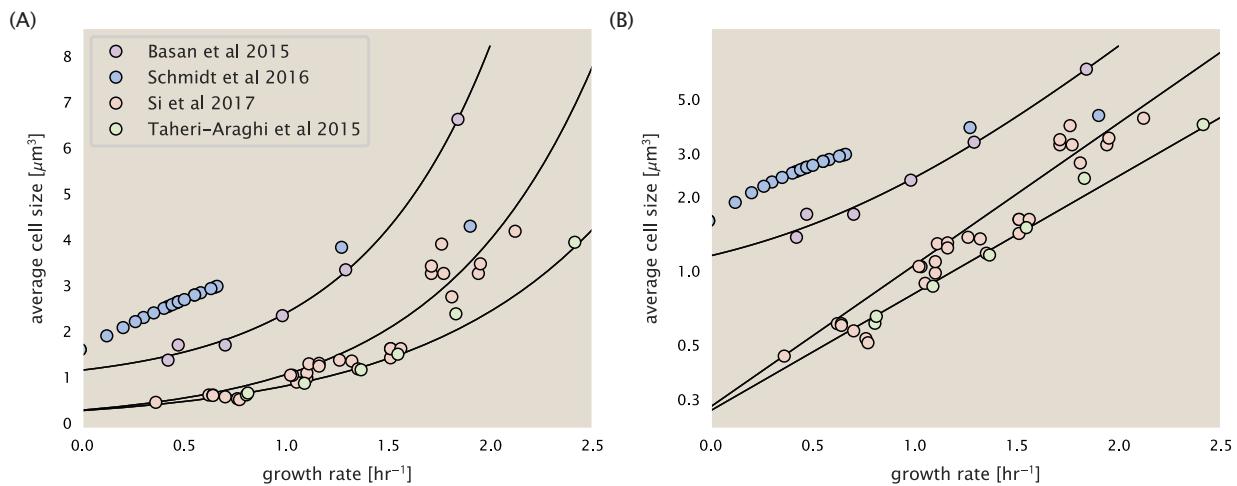


Figure 17. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (?) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are shown in purple (?), while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

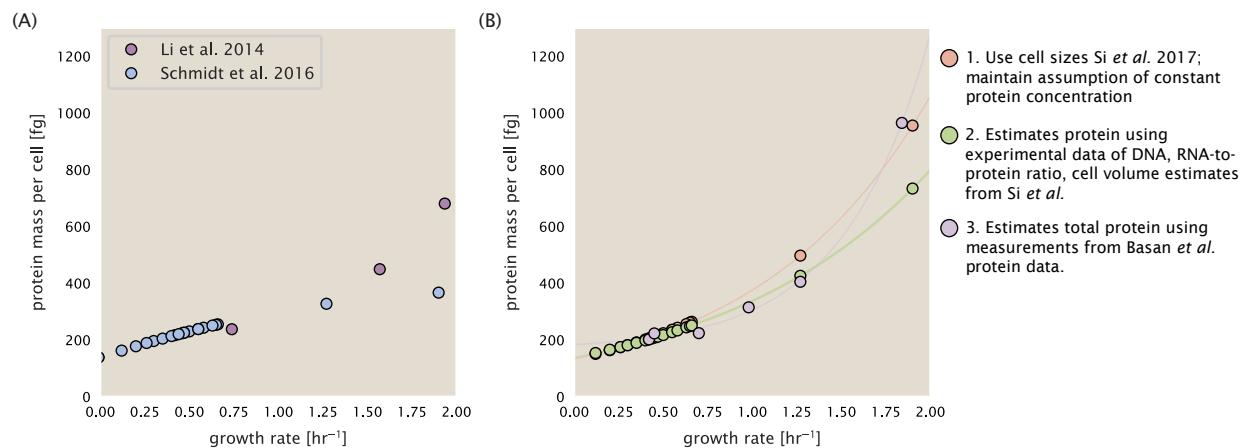


Figure 18. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt *et al.* (A) The original protein mass from Schmidt *et al.* and Li *et al.* are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. *light red*: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

1061 to determine total protein separately using a BCA protein assay. In personal communications, it was noted that
1062 determining reasonable total protein abundances by BCA across their array of growth conditions was particularly
1063 troublesome. Instead, they noted confidence in their total protein measurements for cells grown in M9 minimal
1064 media + glucose and used this as a reference point with which to estimate the total protein for all other growth
1065 conditions.

1066 For cells grown in M9 minimal media + glucose an average total mass of $M_P = 240$ fg per cell was measured.
1067 Using their reported cell volume, reported as $V_{orig} = 2.84$ fl, a cellular protein concentration of $[M_P]_{orig} = M_P/V_{orig} =$
1068 85 fg/fl. Now, taking the assumption that cellular protein concentration is relatively independent of growth rate,
1069 they could then estimate the total protein mass for all other growth conditions from,

$$M_{P_i} = [M_P]_{orig} \cdot V_i \quad (8)$$

1070 where M_{P_i} represents the total protein mass per cell and V_i is the cell volume for each growth condition i as mea-
1071 sured in Volkmer and Heinemann, 2011. Here the thinking is that the values of M_{P_i} reflects the total cellular protein
1072 for growth condition i , where any discrepancy from their absolute protein abundance is assumed to be due to
1073 protein loss during sample preparation. The protein abundances from their absolute abundance measurements
1074 noted above were therefore scaled to their estimates and are shown in Figure **Figure 18** (purple data points).

1075 If we instead consider the cell volumes predicted in the work of Si *et al.*, we again need to take growth in M9
1076 minimal media + glucose as a reference with known total mass, but we can follow a similar approach to estimate
1077 total protein mass for all other growth conditions. Letting $V_{Si_glu} = 0.6$ fl be the predicted cell volume, the cellular
1078 protein concentration becomes $[M_P]_{Si} = M_P/V_{Si_glu} = 400$ fg/fl. The new total protein mass per cell can then be
1079 calculated from,

$$M'_{P_i} = [M_P]_{Si} \cdot V_{Si_i} \quad (9)$$

1080 where M'_{P_i} is the new protein mass prediction, and V_{Si_i} refers to the new volume prediction for each condition i ,
1081 These are shown as red data points in Figure **Figure 18(B)**.

1082 Relaxing assumption of constant protein concentration across growth conditions

1083 We next relax the assumption that cellular protein concentration is constant and instead, attempt to estimate
1084 it using experimental data. Here we use the estimation of total protein mass per cell detailed in Estimation of
1085 Total Protein Content per Cell for all data points in the ? data set. The green data points in **Figure 18(B)** show this
1086 prediction, and this represents the approach used to estimate total protein per cell for all data sets.

1087 Comparison with total protein measurements from Basan *et al.* 2015.

1088 One of the challenges in our estimates in the preceding sections is the need to estimate protein concentration
1089 and cell volumes. These are inherently difficult to measure accurately due to the small size of *E. coli*. Indeed, for all
1090 the additional measurements of cell volume included in Figure **Figure 17**, no measurements were performed for
1091 cells growing at rates below 0.5 hr^{-1} . It therefore remains to be determined whether our extrapolated cell volume
1092 estimates are appropriate, with the possibility that the logarithmic scaling of cell size might break down for slower
1093 growth.

1094 In our last approach we therefore attempt to estimate total protein using experimental data that required no
1095 estimates of concentration or cell volume. Specifically, in the work of Basan *et al*, the authors measured total
1096 protein per cell for a broad range of growth rates (reproduced in Figure **Figure 19**). These were determined by
1097 first measuring bulk protein from cell lysate, measured by the colorimetric Biuret method (?), and then abundance
1098 per cell was calculated from cell counts from either plating cells or a Coulter counter. While it is unclear why
1099 Schmidt *et al.* was unable to take a similar approach, the results from Basan *et al* appear more consistent with
1100 our expectation that cell mass will increase exponentially with faster growth rates. In addition, although they do
1101 not consider growth rates below about 0.5 hr^{-1} , it is interesting to note that the protein mass per cell appears to
1102 plateau to a minimum value at slow growth. In contrast, our estimates using cell volume so far have predicted
1103 that total protein mass should continue to decrease slightly for slower growing cells. By fitting this data to an
1104 exponential function dependent on growth rate, we could then estimate the total protein per cell for each growth
1105 condition considered by ?. These are plotted as red data points in **Figure 18(B)**.

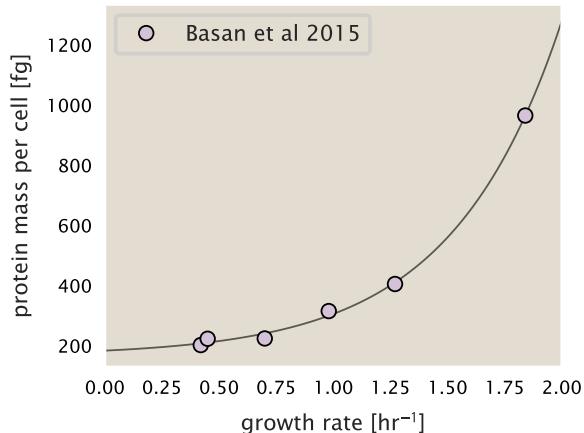


Figure 19. Total cellular protein reported in Basan et al. 2015. Measured protein mass as a function of growth rate as reproduced from Basan *et al.* 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given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

1106 Calculation of Complex Abundance

1107 All protein data quantified the abundance of individual proteins per cell. However, this work requires estimates
 1108 on the abundance of individual protein *complexes*, rather than the copy number of individual proteins. In our
 1109 analysis of the protein copy number data, it became clear that the reported copy numbers do not always align
 1110 with those based on reported stoichiometry. As one example of this, the F-O subunit of ATP synthase consists
 1111 of three protein subunits with a stoichiometry of [AtpB][AtpF]₂[AtpE]₁₀ (also referred to as subunits a, b, and c,
 1112 respectively). In the experimental data of ?, the values deviate from this quite substantially, with approximately
 1113 1000 AtpB, 9000 AtpF, and 300 AtpE reported per cell (minimal media + glucose growth condition). This highlights
 1114 the technical challenges that still remain in our ability to quantify cellular composition, particularly for membrane-
 1115 bound proteins like the ATP synthase complex considered here. In this section, we outline the approach we used
 1116 to annotate proteins as part of each macromolecular complex and how we used averaging across the individual
 1117 protein measurements to estimate an absolute complex abundances per cell.

1118 Protein complexes, and proteins individually, often have a variety of names, both longform and shorthand. As
 1119 individual proteins can have a variety of different synonyms, we sought to ensure that each protein annotated in
 1120 the data sets used the same synonym. To do use, we relied heavily on the EcoCyc Genomic Database (?). Each
 1121 protein in available data sets included an annotation of one of the gene name synonyms as well as an accession
 1122 ID – either a UniProt or Blattner “b-number”. We programmatically matched up individual accession IDs between
 1123 the proteins in different data sets. In cases where accession IDs matched but the gene names were different, we
 1124 manually verified that the gene product was the same between the datasets and chose a single synonym. All code
 1125 used in the data cleaning and unification procedures can be found on the associated [GitHub repository](#) (DOI:XXX)
 1126 associated with this paper as well as on the associated [paper website](#).

1127 With each protein conforming to a single identification scheme, we then needed to identify the molecular
 1128 complexes each protein was a member of. Additionally, we needed to identify how many copies of each protein
 1129 were present in each complex (i.e. the subunit copy number) and compute the estimated abundance complex that
 1130 accounted for fluctuations in subunit stoichiometry. To map proteins to complexes, we accessed the EcoCyc *E. coli*
 1131 database ? using PathwayTools version 23.0 ?. With a license for PathWay Tools, we mapped each unique protein
 1132 to its annotated complexes via the BioCyc Python package. As we mapped each protein with *all* of its complex
 1133 annotations, there was redundancy in the dataset. For example, ribosomal protein L20 (RplT) is annotated to be a
 1134 component of the 50S ribosome (EcoCyc complex CPLX-03962) as well as a component of the mature 70S ribosome
 1135 (EcoCyc complex CPLX-03964).

1136 In addition to the annotated complex, we collected information on the stoichiometry of each macromolecular
 1137 complex. For a complex with N_{subunits} protein species, for each protein subunit i we first calculate the number of

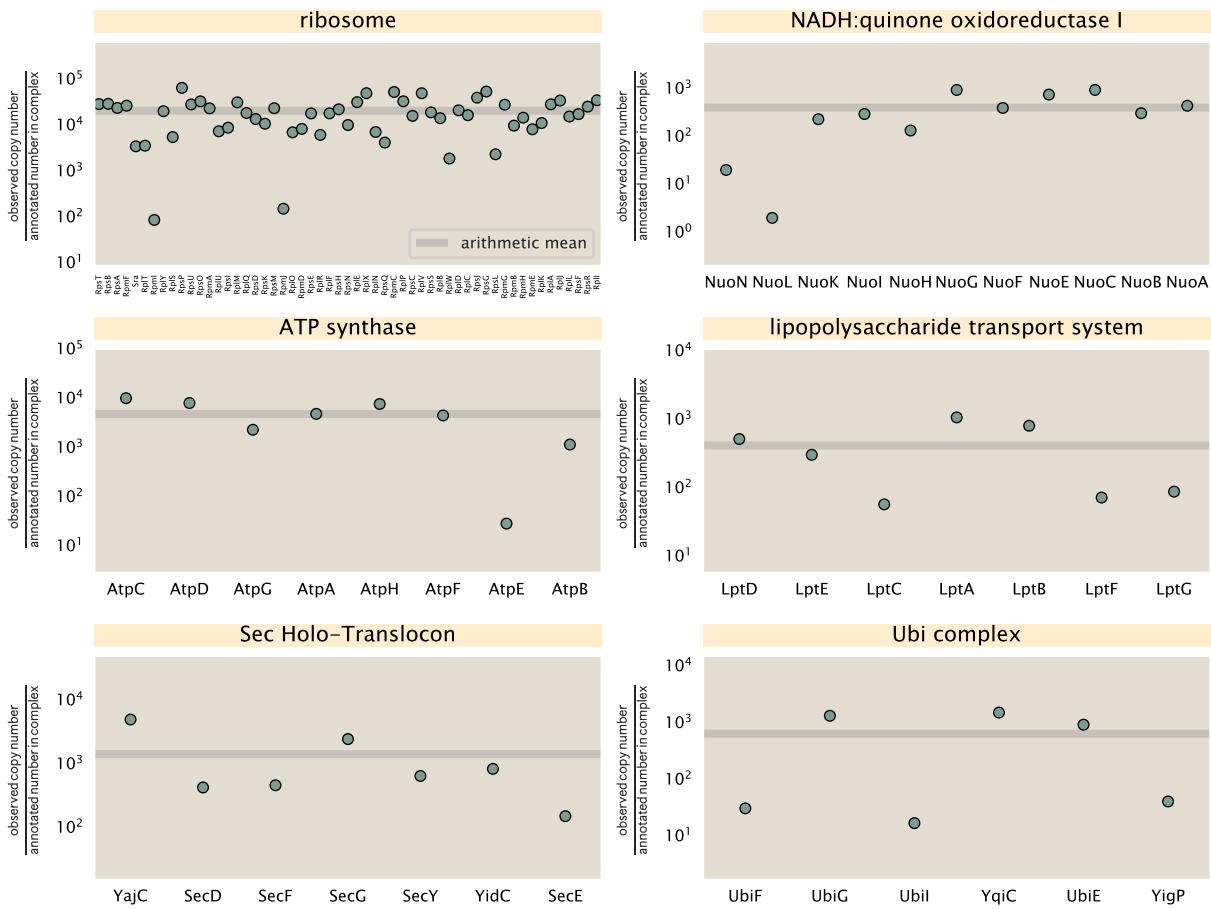


Figure 20. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds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

1138 complexes that could be formed given the measured protein copy numbers per cell,

$$N_{\text{complex}}(\text{subunit } i) = \frac{P_{\text{subunit } i}^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (10)$$

1139 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
1140 present for that protein in the complex. For example, the 70S mature ribosome complex has 55 protein components,
1141 all of which are present in a single copy except L4 (RplL), which is present in 4 copies ($m = 4$). For each
1142 ribosomal protein, we then calculate the maximum number of complexes that could be formed using Equation 10.
1143 This example, along with example from 5 other macromolecular complexes, can be seen in Figure 20.

1144 It is important to note that measurement noise, efficiency of protein extraction, and differences in protein
1145 stability will mean that the precise value of each calculation will be different for each component of a given complex.
1146 Thus, to report the total complex abundance, we use the arithmetic mean of across all subunits in the complex,

$$\langle N_{\text{complex}} \rangle = \frac{1}{N_{\text{subunits}}} \sum_i^{N_{\text{subunits}}} \frac{P_i^{(\text{measured})}}{m_{\text{subunit } i}}. \quad (11)$$

1147 in Figure 20, we show this mean value as a grey line for a variety of different complexes. Additionally, we have
1148 built an interactive figure accessible on the [paper website](#) where the validity of this approach can be examined
1149 for any complex with more than two subunits (thus, excluding monomers and dimers).

1150 **Extending Estimates to a Continuum of Growth Rates**

1151 In the main text, we considered a standard stopwatch of 5000 s to estimate the abundance of the various protein
1152 complexes considered. In addition to point estimates, we also showed the estimate as a function of growth rate
1153 as transparent grey curves. In this section, we elaborate on this continuum estimate, giving examples of estimates
1154 that scale with either cell volume, cell surface area, or number of origins of replication.

1155 **Estimation of the total cell mass**

1156 For many of the processes estimated in the main text we relied on a cellular dry mass of ≈ 300 fg from which we
1157 computed elemental and protein fractions using knowledge of fractional composition of the dry mass. At modest
1158 growth rates, such as the 5000 s doubling time used in the main text, this is a reasonable number to use as the
1159 typical cell mass is ≈ 1 pg and *E. coli* cells can approximated as 70% water by volume. However, as we have shown
1160 in the preceding sections, the cell size is highly dependent on the growth rate. This means that a dry mass of 300
1161 fg cannot be used reliably across all growth rates.

1162 Rather, using the phenomenological description of cell volume scaling exponentially with growth rate, and
1163 using a rule-of-thumb of a cell buoyant density of ≈ 1.1 pg / fL (BNID: 103875), we can calculate the cell dry mass
1164 across a range of physiological growth rates as

$$m_{\text{cell}} \approx \rho V(\lambda) \approx \rho a e^{\lambda * b} \quad (12)$$

1165 where a and b are constants with units of μm^3 and hr, respectively. The value of these constants can be estimated
1166 from the careful volume measurements performed by ??, as considered in Appendix Estimation of Cell Size and
1167 Surface Area earlier.

1168 **Complex Abundance Scaling With Cell Volume**

1169 Several of the estimates performed in the main text are implicitly dependent on the cell volume. This includes
1170 processes such as ATP utilization and, most prominently, the transport of nutrients, whose demand will be pro-
1171 portional to the volume of the cell. Of the latter, we estimated the number of transporters that would be needed
1172 to shuttle enough carbon, phosphorus, and sulfur across the membrane to build new cell mass. To do so, we
1173 used elemental composition measurements combined with a 300 fg cell dry mass to make the point estimate. As
1174 we now have a means to estimate the total cell mass as a function of volume, we can generalize these estimates
1175 across growth rates.

1176 Rather than discussing the particular details of each transport system, we will derive this scaling expression in
1177 very general terms. Consider that we wish to estimate the number of transporters for some substance X , which
1178 has been measured to be made up some fraction of the dry mass, θ_X . If we assume that, irrespective of growth
1179 rate, the cell dry mass is relatively constant (?) and $\approx 30\%$ of the total cell mass, we can state that the total mass
1180 of substance X as a function of growth rate is

$$m_X \approx 0.3 \times \rho V(\lambda) \theta_X, \quad (13)$$

1181 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in **Equation 12**. To convert this to the
1182 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
1183 with **Equation 13**,

$$N_X \approx \frac{m_X}{w_X}. \quad (14)$$

1184 To estimate the number of transporters needed, we make the approximation that loss of units of X via diffusion
1185 through porins or due to the permeability of the membrane is negligible and that a single transporter complex
1186 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
1187 a time window over which the transport process can occur. This is related to the cell doubling time τ , which can
1188 be calculated from the the growth rate λ as $\tau = \log(2)/\lambda$. Putting everything together, we arrive at a generalized
1189 transport scaling relation of

$$N_{\text{transporters}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_X}{w_X r_X \tau}. \quad (15)$$

1190 This function is used to draw the continuum estimates for the number of transporters seen in Figures 2 and
 1191 3 as transparent grey curves. Occasionally, this continuum scaling relationship will not precisely agree with the
 1192 point estimate outlined in the main text. This is due to the choice of ≈ 300 fg total dry mass per cell for the point
 1193 estimate, whereas we considered more precise values of cell mass in the continuum estimate. We note, however,
 1194 that both this scaling relation and the point estimates are meant to describe the order-of-magnitude observed,
 1195 and not the predict the exact values of the abundances.

1196 **Equation 15** is a very general relation for processes where the cell volume is the "natural variable" of the
 1197 problem. This means that, as the cell increases in volume, the requirements for substance X also scale with
 1198 volume rather than scaling with surface area, for example. So long as the rate of the process, the fraction of the
 1199 dry mass attributable to the substance, and the formula mass of the substance is known, **Equation 15** can be used
 1200 to compute the number of complexes needed. For example, to compute the number of ATP synthases per cell,
 1201 **Equation 15** can be slightly modified to the form

$$N_{\text{ATP synthases}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_{\text{protein}} N_{\text{ATP}}}{w_{AA} r_{\text{ATP}} \tau}, \quad (16)$$

1202 where we have included the term N_{ATP} to account for the number of ATP equivalents needed per amino acid for
 1203 translation (≈ 4 , BNID: 114971), and w_{AA} is the average mass of an amino acid. The grey curves in Figure 4 of the
 1204 main text were made using this type of expression.

1205 A Relation for Complex Abundance Scaling With Surface Area

1206 In our estimation for the number of complexes needed for lipid synthesis and peptidoglycan maturation, we used
 1207 a particular estimate for the cell surface area ($\approx 5 \mu\text{m}$, BNID: 101792) and the fraction of dry mass attributable to
 1208 peptidoglycan ($\approx 3\%$, BNID: 101936). Both of these values come from glucose-fed *E. coli* in balanced growth. As we
 1209 are interested in describing the scaling as a function of the growth rate, we must also consider how these values
 1210 scale with cell surface area, which is the natural variable for these types of processes. In the coming paragraphs,
 1211 we highlight how we incorporate a condition-dependent surface area into our calculation of the number of lipids
 1212 and murein monomers that need to be synthesized and crosslinked, respectively.

1213 Number of Lipids

1214 To compute the number of lipids as a function of growth rate, we make the assumption that some features, such as
 1215 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
 1216 of lipids ($\approx 40\%$, BNID: 100078) are independent of the growth rate. Using these approximations combined with
 1217 **Equation 6**, and recognizing that each membrane is composed of two leaflets, we can compute the number of
 1218 lipids as a function of growth rate as

$$N_{\text{lipids}}(\lambda) \approx \frac{4 \text{ leaflets} \times 0.4 \times \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{A_{\text{lipid}}} \quad (17)$$

1219 where η is the length-to-width aspect ratio and V is the cell volume.

1220 Number of Murein Monomers

1221 In calculation of the number of transpeptidases needed for maturation of the peptidoglycan, we used an empirical
 1222 measurement that $\approx 3\%$ of the dry mass is attributable to peptidoglycan and that a single murien monomer is
 1223 $m_{\text{murein}} \approx 1000 \text{ Da}$. While the latter is independent of growth rate, the former is not. As the peptidoglycan exists as
 1224 a thin shell with a width of $w \approx 10 \text{ nm}$ encapsulating the cell, one would expect the number of murein monomers
 1225 scales with the surface area of this shell. In a similar spirit to our calculation of the number of lipids, the total
 1226 number of murein monomers as a function of growth rate can be calculated as

$$N_{\text{murein monomers}}(\lambda) \approx \frac{\rho_{\text{pg}} w \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{m_{\text{murein}}}, \quad (18)$$

1227 where ρ_{pg} is the density of peptidoglycan.

1228 **Complex Abundance Scaling With Number of Origins, and rRNA Synthesis**

1229 While the majority of our estimates hinge on the total cell volume or surface area, processes related to the central
1230 dogma, namely DNA replication and synthesis of rRNA, depend on the number of chromosomes present in the
1231 cell. As discussed in the main text, the ability of *E. coli* to parallelize the replication of its chromosome by having
1232 multiple active origins of replication is critical to synthesize enough rRNA, especially at fast growth rates. Derived
1233 in ? and reproduced in the main text and Appendix Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$ below, the average
1234 number of origins of replication at a given growth rate can be calculated as

$$\langle \#ori \rangle \approx 2^{t_{\text{cyc}} \lambda / \ln 2} \quad (19)$$

1235 where t_{cyc} is the total time of replication and division. We can make the approximation that $t_{\text{cyc}} \approx 70$ min, which is
1236 the time from the initiation of chromosomal replication until division. This time corresponds to the sum of the so-
1237 called C and D periods of the cell cycle, which correspond to the time it takes to replicate the entire chromosome
1238 (C period) and the time from completion to eventual division (D period) ?.

1239 In the case of rRNA synthesis, the majority of the rRNA operons are surrounding the origin of replication. Thus,
1240 at a given growth rate λ , the average dosage of rRNA operons per cell D_{rRNA} is

$$D_{\text{rRNA}}(\lambda) \approx N_{\text{rRNA operons}} \times 2^{t_{\text{cyc}} \lambda / \ln 2}. \quad (20)$$

1241 This makes the approximation that *all* rRNA operons are localized around the origin. In reality, the operons
1242 are some distance away from the origin, making **Equation 20** an approximation (?).

1243 In the main text, we stated that at a growth rate of 0.5 hr^{-1} , there is ≈ 1 chromosome per cell. While a fair
1244 approximation, **Equation 19** illustrates that is not precisely true, even at slow growth rates. In estimating the
1245 number of RNA polymerases as a function of growth rate, we consider that regardless of the number of rRNA
1246 operons, they are all sufficiently loaded with RNA polymerase such that each operon produces one rRNA per
1247 second. Thus, the total number of RNA polymerase as a function of the growth rate can be calculated as

$$N_{\text{RNA polymerase}}(\lambda) \approx L_{\text{operon}} D_{\text{rRNA}} \rho_{\text{RNA polymerase}}, \quad (21)$$

1248 where L_{operon} is the total length of an rRNA operon (≈ 4500 bp) and $\rho_{\text{RNA polymerase}}$ is packing density of RNA poly-
1249 merase on a given operon, taken to be 1 RNA polymerase per 80 nucleotides.

1250 **Calculation of active ribosomal fraction.**

1251 In the main text we used the active ribosomal fraction f_a that was reported in the work of ? to estimate the active
1252 ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We lacked any specific model to consider how f_a should
1253 vary with growth rate, and instead find that the data is well-approximated by fitting to an exponential curve (f_a
1254 = $-0.889 e^{4.6 \cdot \lambda} + 0.922$; dashed line in inset of **Figure 9(C)**). We use this function to estimate f_a for each of the data
1255 points shown in **Figure 9(C)**.

1256 **Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.**

1257 *E. coli* shows robust scaling of cell size with the average number of origins per cell, $\langle \#ori \rangle$ (?). Since protein makes
1258 up a majority of the cell's dry mass, the change in cell size is also a reflection of the changes in proteomic com-
1259 position and total abundance across growth conditions. Given the potential constraints on rRNA synthesis and
1260 changes in ribosomal copy number with $\langle \#ori \rangle$, it becomes important to also consider how protein copy numbers
1261 vary with the state of chromosomal replication. This is particularly true when trying to make sense of the changes
1262 in ribosomal fraction and growth-rate dependent changes in proteomic composition at a mechanistic level. As
1263 considered in the main text, it is becoming increasingly apparent that regulation through the secondary messen-
1264 gers (p)ppGpp may act to limit DNA replication and also reduce ribosomal activity in poorer nutrient conditions.
1265 In this context, both $\langle \#ori \rangle$, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio become important parameters to consider and keep
1266 track of. An increase in $\langle \#ori \rangle / \langle \#ter \rangle$ ratio in particular, causes a relatively higher gene dosage in rRNA and
1267 r-protein genes due to skew in genes near the origin, where the majority of these are located

1268 In the main text we estimated the change in $\langle \#ori \rangle$ with growth rate using the nutrient-limited wild-type cell
1269 data from ?. We consider their measurements of DNA replication time (t_C , 'C' period of cell division), total cell cycle

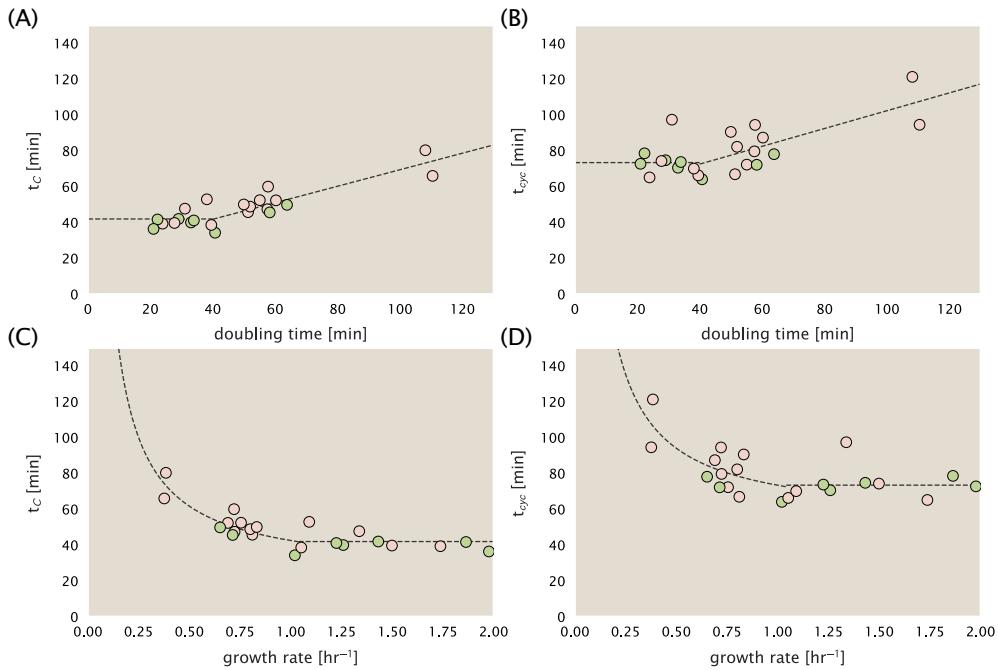


Figure 21. Estimation of $\langle \#ori \rangle / \langle \# ter \rangle$ and $\langle \#ori \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling times (rich media), t_C and t_{cyc} are assumed constant ($t_C = 42$ minutes, $t_{cyc} = 73$ minutes). At the transition, taken to occur at 40 minutes, the dashed line corresponds to an assumed proportional increase in each parameter as a function of the doubling time ($t_C = 0.46\tau + 23.3$ minutes, $t_{cyc} = 0.50\tau + 52.7$ minutes). (C) and (D) plot the same data as in (A) and (B), but as a function of growth rate, given by $\lambda = \ln(2)/\tau$.

time (t_{cyc} , 'C' + 'D' period of cell division), and doubling time τ from wild-type *E. coli* growing across a range of growth conditions. Here we show how we estimate this parameter, as well as the $\langle \#ori \rangle / \langle \# ter \rangle$ ratio from their data. We begin by considering $\langle \#ori \rangle$. If the cell cycle time takes longer than the time of cell division, the cell will need to initiate DNA replication more often than its rate of division, $2^{\lambda\tau} = 2^{\ln(2)\cdot\tau/\tau}$ to maintain steady state growth. Cells will need to do this in proportion to the ratio $\lambda_{cyc}/\lambda = t_{cyc}/\tau$, and the number of origins per cell (on average) is then given by $2^{t_{cyc}/\tau}$. The average number of termini will in contrast depend on the lag time between DNA replication and cell division, t_D , with $\langle \#ori \rangle / \langle \# ter \rangle$ ratio = $2^{t_{cyc}/\tau-t_D/\tau} = 2^{t_C/\tau}$.

In Figure 21(A) and (B) we plot the measured t_C and t_{cyc} values versus the doubling time from ?. The authors estimated t_C by marker frequency analysis using qPCR, while $t_{cyc} = t_C + t_D$ were inferred from t_C and τ . In the plots we see that both t_C and t_{cyc} reach a minimum at around 40 and 75 minutes, respectively. For a C period of 40 minutes, this would correspond to a maximum rate of elongation of about 1,000 bp/sec. Since we lacked a specific model to describe how each of these parameters vary with growth condition, we assumed that they were linearly dependent on the doubling time. For each parameter, t_C and t_{cyc} , we split them up into two domains corresponding to poorer nutrient conditions and rich nutrient conditions (cut off at $\tau \approx 40$ minutes where chromosomal replication becomes nearly constant). The fit lines are shown as solid black lines. In Figure 21(C) and (D) we also show t_C and t_{cyc} as a function of growth rate λ along with our piecewise linear fits, which match the plots in the main text.

Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control

Here we provide a derivation of the minimal model for growth rate control under nutrient-limited growth. By growth rate control, we are specifically referring to the ability of bacteria to modulate their proteome (N_{pep} , R , Φ_R) and cell size as nutrient conditions change, with slower growing cells generally being smaller in size (?). This capability provides bacteria with a particular benefit when nutrients are more scarce since it will mean there is a smaller net demand on carbon, phosphorus, sulfur, and nitrogen. The specific goal of developing this model is

1293 to help us better explore the overall constraints on growth that follow from 1) our observation that many of the
 1294 cellular processes we've considered require increased protein abundance at faster growth rates, and 2) a strict
 1295 limit on growth rate that is governed by the ribosomal synthesis rate and ribosomal mass fraction Φ_R .

1296 In **Figure 11(A)** of the main text we provide a schematic of the model, where we consider growth as simply
 1297 governed by the rate of protein synthesis ($r_t \times R \times f_a$). In order to grow rapidly, at least to the extent possible,
 1298 these three parameters need to be maximized (with $r_t \leq 17$ amino acids per second, and $f_a \leq 1$ reported in the
 1299 work of ?). The elongation rate r_t will depend on how quickly ribosomes can match codons with their correct
 1300 amino-acyl tRNA, along with the subsequent steps of peptide bond formation and translocation. This ultimately
 1301 depends on the cellular concentration amino acids, which we treat as a single effective species, $[AA]_{\text{eff}}$.

1302 In our model, we need to determine the rate of peptide elongation r_t , which we consider as simply depending
 1303 on the supply of amino acids (and, therefore, also amino-acyl tRNAs) through a parameter r_{AA} in units of AA per
 1304 second, and the rate of amino acid consumption by protein synthesis ($r_t \times R \times f_a$). The balance between these
 1305 two rates will determine the effective amino acid concentration in the cell $[AA]_{\text{eff}}$. An important premise for this
 1306 formulation is growing evidence that cells are able to modulate their biosynthesis activity according to nutrient
 1307 availability (i.e. extent of chromosomal replication, transcriptional, and translation activity) through secondary-
 1308 messenger molecules like (p)ppGpp (?????). Given our observation that protein synthesis and energy production
 1309 are not limiting, we assume that other molecular players required by ribosomes like elongation factors and GTP
 1310 are available in sufficient abundance. In addition, experimentally, the relative number of tRNA and elongation
 1311 factor EF-Tu per ribosome have been found to increase in poorer nutrient conditions ???).

1312 We begin by considering a coarse-grained description of peptide elongation, which includes 1) the time re-
 1313 quired to find and bind each correct amino-acyl tRNA, and 2) the remaining steps in peptide elongation that will
 1314 not depend on the amino acid availability. These time scales will be related to the inverse of the elongation rate

1315 r_t ,

$$\frac{1}{r_t} = \frac{1}{k_{on}\alpha[AA]_{\text{eff}}} + \frac{1}{r_t^{\max}}. \quad (22)$$

1316 where we have assumed that the rate of binding by amino-acyl tRNA k_{on} is proportional to $[AA]_{\text{eff}}$ by a constant
 1317 α . r_t^{\max} refers to the maximum elongation rate. This leads to a Michaelis-Menten dependence of the elongation
 1318 rate r_t on the effective amino acid concentration $[AA]_{\text{eff}}$ (??). We can re-write this more succinctly in terms of an
 1319 effective dissociation constant,

$$K_D = \frac{r_t^{\max}}{\alpha k_{on}}, \quad (23)$$

1320 where the elongation rate r_t is now given by

$$r_t = \frac{r_t^{\max}}{1 + K_D/[AA]_{\text{eff}}}. \quad (24)$$

1321 The rate of amino acid supply r_{AA} will vary with changing nutrient conditions and the cell can maintain $[AA]_{\text{eff}}$
 1322 by tuning the rate of amino acid consumption, $r_t \times R \times f_a$. Thus, $[AA]_{\text{eff}}$ is determined by the difference in the
 1323 rate of amino acid synthesis (or import, for rich media) and/or tRNA charging, r_{AA} , and the rate of consumption,
 1324 $r_t \times R \times f_a$. Over an arbitrary length of time t of cellular growth, the cell will grow in volume, requiring us to consider
 1325 these rates in terms of concentration rather than absolute numbers, with $[AA]_{\text{eff}}$ given by,

$$\int_0^t \frac{d[AA]_{\text{eff}}}{dt} dt = \int_0^t ([r_{AA}] - [r_t \times R \times f_a]) dt. \quad (25)$$

1326 This considers the net change in amino acid concentration over a time from 0 to t , with the square brackets
 1327 indicating concentrations per unit time. Integrating **Equation 25** yields.

$$[AA]_{\text{eff}} = t([r_{AA}] - [r_t \times R \times f_a]). \quad (26)$$

1328 Alternatively, to connect to the experimental data in terms of absolute ribosome copy number R we can con-
 1329 sider a unit volume V ,

$$[AA]_{\text{eff}} = \frac{t(r_{AA} - r_t \times R \times f_a)}{V \times N_A}, \quad (27)$$

1330 where r_{AA} is in units of AA per unit time and r_t is in units of AA per unit time per ribosome. N_A refers to Avogadro's
 1331 number and is needed to convert between concentration and absolute numbers per cell. With an expression for
 1332 $[AA]_{\text{eff}}$ in hand, we can now solve **Equation 24** for r_t , which is a quadratic function with a physically-meaningful
 1333 root of

$$r_t = \frac{t(r_{AA} + r_t^{(\max)} R f_a) + K_D V N_A - \sqrt{(r_{AA} t + r_t^{(\max)} R f_a t + K_D V N_A)^2 - 4(R f_a t)(r_t^{(\max)} r_{AA} t)}}{2R f_a t}. \quad (28)$$

1334 This is the key equation that allows us to calculate growth rate for any combination of N_{pep} , R , f_a , and cell size V
 1335 as a function of amino acid supply r_{AA} (**Equation 3** of the main text). We refer the reader to A Minimal Model of
 1336 Nutrient-Mediated Growth Rate Control of the main text for our exploration of this model in the context of the
 1337 proteomic data.

1338 We end this section by noting several distinctions of this formulation with previous work. The first, as noted
 1339 in the main text, relates to the now seminal work of ?, which provides a treatment of resource allocation that
 1340 partitions of the proteome into sectors – including one for ribosome-associated proteins and one for metabolic
 1341 proteins. As cells grow faster, there is a notable change in the mass fraction of these sectors, with an increase
 1342 in ribosomal content that is predominantly achieved at the expense of a decrease in the metabolic sector. By
 1343 including an additional constraint through the phenomenological parameter v , which characterizes the quality of
 1344 the growth medium ???, the authors derive a model of growth rate, dependent on optimal resource allocation.
 1345 Here we have developed a model that considers the effect of changes in absolute protein abundance and ribo-
 1346 somal content, and consider how these influence the achievable growth rate. In addition, by accounting for the
 1347 metabolic supply of amino acids directly through their availability in the cell (i.e. $[AA]_{\text{eff}}$), we are able to consider
 1348 how the balance between translation-specific metabolic capacity and translational capacity influences both the
 1349 elongation rate r_t and growth rate λ .

1350 The second and last point we note is that the recent works from ? and ? also employ a similar coarse-graining of
 1351 translation elongation as we've considered above. Here, however, a notable distinction is that the authors consider
 1352 the entire ternary complex (i.e. the complex of amino-acyl tRNA, EF-Tu, and GTP) as rate limiting. Further, through
 1353 an assumed proportionality between ternary complex and ribosome abundance, they arrive at a formulation of
 1354 elongation rate r_t that exhibits a Michaelis-Menten dependence on the ribosomal fraction Φ_R . They demonstrate
 1355 that all their measurements of elongation rate, even upon addition of sublethal doses of chloramphenicol (which
 1356 cause an increase in both r_t and Φ_R), can be collapsed onto a single curve described by this Michaelis-Menten
 1357 dependence. There is always a benefit to increase their ribosomal fraction Φ_R on growth rate when nutrient
 1358 conditions allow (see Maximum Growth Rate is Determined by the Ribosomal Mass Fraction), and this trend in
 1359 the data in part follows from the tendency for cells to increase Φ_R and better maximize r_t as nutrient conditions
 1360 improve. In addition, it does not account for the decrease in the fraction of actively translating ribosome f_a that
 1361 was strikingly apparent at slow growth rates or in sublethal doses of chloramphenicol in the work of ?. Through
 1362 **Equation 28** we also account for changes in the fraction of actively translating ribosomes. Ultimately, we find that
 1363 cells are able to maximize both Φ_R , r_t , and their growth rate only to the extent allowed by the nutrient conditions
 1364 (i.e. via r_{AA}) and through the maintenance of the cellular pool of amino acids $[AA]_{\text{eff}}$, amino-acyl tRNA, GTP, as well
 1365 as the synthesis of other key molecular constituents like EF-Tu.

1366 **References**

- 1367 Abelson, H., Johnson, L., Penman, S., and Green, H. (1974). Changes in RNA in relation to growth of the fibroblast: II. The lifetime
1368 of mRNA, rRNA, and tRNA in resting and growing cells. *Cell*, 1(4):161–165.
- 1369 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia*
1370 *coli*. *BMC Systems Biology*, 8(1):133.
- 1371 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by ammo-
1372 nium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 1373 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,
1374 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,
1375 G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- 1376 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model
1377 for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding
1378 Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- 1379 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH_3 Channel.
1380 *Journal of Biological Chemistry*, 291(36):19184–19195.
- 1381 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006).
1382 Construction of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*,
1383 2(1):2460.
- 1384 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein
1385 synthesis. *Molecular Systems Biology*, 11(10):836.
- 1386 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*,
1387 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 1388 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.
1389 (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proceedings
1390 of the National Academy of Sciences*, 115(21):E4796–E4805.
- 1391 Booth, I. R., Mitchell, W. J., and Hamilton, W. A. (1979). Quantitative analysis of proton-linked transport systems. The lactose
1392 permease of *Escherichia coli*. *Biochemical Journal*, 182(3):687–696.
- 1393 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Expon-
1394 ential Growth Rates. *EcoSal Plus*, 3(1).
- 1395 Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews
1396 Microbiology*, 14(10):638–650.
- 1397 Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*,
1398 266:2020.06.16.154187.
- 1399 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.,
1400 Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan
1401 Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- 1402 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in *Escherichia coli*.
1403 *Biochemical Journal*, 117(3):551–562.
- 1404 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Transla-
1405 tional Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- 1406 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T.
1407 (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature
1408 Microbiology*, 2(2):16231.
- 1409 Datsenko, K. A. and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.
1410 *Proceedings of the National Academy of Sciences*, 97(12):6640–6645.
- 1411 Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach.
1412 *Microbiology and Molecular Biology Reviews*, 68(4):639–668.

- 1413 Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*, 108(44):17876–17882.
- 1415 Dong, H., Nilsson, L., and Kurland, C. G. (1996). Co-variation of tRNA Abundance and Codon Usage in *Escherichia coli* at Different Growth Rates. *Journal of Molecular Biology*, 260(5):649–663.
- 1417 Erickson, D. W., Schink, S. J., Patsalo, V., Williamson, J. R., Gerland, U., and Hwa, T. (2017). A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature*, 551(7678):119–123.
- 1419 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.
- 1422 Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., and Palsson, B. Ø. (2007). A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology*, 3(1):121.
- 1425 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.
- 1428 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.
- 1430 Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–263.
- 1431 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*.
- 1433 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., Porron-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.
- 1438 Ge, J., Yu, G., Ator, M. A., and Stubbe, J. (2003). Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class I Ribonucleotide Reductase. *Biochemistry*, 42(34):10071–10083.
- 1440 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.
- 1442 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.
- 1445 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.
- 1447 Harris, R. M., Webb, D. C., Howitt, S. M., and Cox, G. B. (2001). Characterization of PitA and PitB from *Escherichia coli*. *Journal of Bacteriology*, 183(17):5008–5014.
- 1449 Hauryliuk, 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.
- 1451 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.
- 1453 Helmstetter, C. E. and Cooper, S. (1968). DNA synthesis during the division cycle of rapidly growing *Escherichia coli* Br. *Journal of Molecular Biology*, 31(3):507–518.
- 1455 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.
- 1458 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.

- 1461 Ingledew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*, 48(3):222–271.
- 1462 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B., and Phillips, R. (2020).
1463 Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a time. *bioRxiv*.
- 1464 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO journal*,
1465 20(17):4952–4963.
- 1466 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the
1467 future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1468 Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., Keseler, I. M., Krummenacker, M., Midford, P. E.,
1469 Ong, Q., Ong, W. K., Paley, S. M., and Subhraveti, P. (2019). The BioCyc collection of microbial genomes and metabolic pathways.
1470 *Briefings in Bioinformatics*, 20(4):1085–1093.
- 1471 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.
1472 (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–401.
- 1473 Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S.,
1474 Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-
1475 Ramírez, D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge
1476 about *Escherichia coli*K-12. *Nucleic Acids Research*, 45(D1):D543–D550.
- 1477 Khademi, S., O'Connell, J., Remis, J., Robles-Colmenares, Y., Miercke, L. J. W., and Stroud, R. M. (2004). Mechanism of Ammonia
1478 Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*, 305(5690):1587–1594.
- 1479 Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that enables the use of
1480 hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy of Sciences*, 114(33):E6922–E6931.
- 1481 Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition.
1482 *Current Opinion in Biotechnology*, 28:96–102.
- 1483 Klumpp, S., Scott, M., Pedersen, S., and Hwa, T. (2013). Molecular crowding limits translation and cell growth. *Proceedings of the
1484 National Academy of Sciences*, 110(42):16754–16759.
- 1485 Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*,
1486 139(7):1366–1375.
- 1487 Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*,
1488 125(2):028103.
- 1489 Kraemer, J. A., Sanderlin, A. G., and Laub, M. T. (2019). The Stringent Response Inhibits DNA Replication Initiation in *E. coli*by
1490 Modulating Supercoiling of oriC. *mBio*, 10(4):822.
- 1491 Lascur, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and Biomem-
1492 branes*, 32(3):237–246.
- 1493 Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical
1494 investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- 1495 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: visualization of intersecting sets. *IEEE Transactions
1496 on Visualization and Computer Graphics*, 20(12):1983–1992.
- 1497 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles
1498 underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1499 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in
1500 cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1501 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon
1502 Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- 1503 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol Facili-
1504 tator GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1505 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*,
1506 112(51):15690–15695.

- 1507 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On*. Gene Expression.
1508 Springer.
- 1509 Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed
1510 by condition-dependent proteome profiling. *eLife*, 6:e03528.
- 1511 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).
1512 A Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.
- 1513 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular
1514 and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1515 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation.
1516 *Growth Symposium*, 9:223–289.
- 1517 Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- 1518 Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall
1519 synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–
1520 12515.
- 1521 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 1522 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria.
1523 *eLife*, 8:642.
- 1524 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- 1525 Pedersen, S. (1978). Patterns of protein synthesis in *E. coli*: a catalog of the amount of 140 individual proteins at different growth
1526 rates. *Cell*, 14(1):179–190.
- 1527 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with
1528 increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 1529 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham, Cham.
- 1530 Ramos, S. and Kaback, H. R. (1977). The relation between the electrochemical proton gradient and active transport in *Escherichia coli*
1531 membrane vesicles. *Biochemistry*, 16(5):854–859.
- 1532 Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., and Maranas, C. D. (2012). An
1533 integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering*,
1534 14(6):687–704.
- 1535 Rogers, H., Perkins, H., and Ward, J. (1980). *Microbial Cell Walls and Membranes*. Chapman and Hall, London.
- 1536 Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive
1537 strategies. *Nature microbiology*, 1(11):1–7.
- 1538 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of
1539 Bacteriology*, 131(2):505–511.
- 1540 Rudd, S. G., Valerie, N. C. K., and Helleday, T. (2016). Pathways controlling dNTP pools to maintain genome stability. *DNA Repair*,
1541 44:193–204.
- 1542 Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–
1543 11734.
- 1544 Sánchez-Romero, M. A., Molina, F., and Jiménez-Sánchez, A. (2011). Organization of ribonucleoside diphosphate reductase
1545 during multifork chromosome replication in *Escherichia coli*. *Microbiology*, 157(8):2220–2225.
- 1546 Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical
1547 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- 1548 Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann,
1549 M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 1550 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression:
1551 origins and consequences. *Science*, 330(6007):1099–1102.

- 1552 Sekowska, A., Kung, H.-F., and Danchin, A. (2000). Sulfur Metabolism in *Escherichia coli* and Related Bacteria: Facts and Fiction.
1553 *Journal of Molecular Microbiology and Biotechnology*, 2(2):34.
- 1554 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.
1555 *Cell*, 172(6):1294–1305.
- 1556 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
1557 Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 1558 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
1559 Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- 1560 Sirko, A., Zatyka, M., Sadowy, E., and Hulanicka, D. (1995). Sulfate and thiosulfate transport in *Escherichia coli* K-12: Evidence for
1561 a functional overlapping of sulfate- and thiosulfate-binding proteins. *Journal of Bacteriology*, 177(14):4134–4136.
- 1562 Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology
1563 Reviews*, 40(1):133–159.
- 1564 Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth
1565 and ethanol stress. *Frontiers in Microbiology*, 6:198.
- 1566 Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*,
1567 19.
- 1568 Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van
1569 Leeuwenhoek*, 39(1):545–565.
- 1570 Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*.
1571 *Archives of Microbiology*, 113(3):185–189.
- 1572 Svenningsen, S. L., Kongstad, M., Stenum, T. S. n., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017). Transfer RNA is highly
1573 unstable during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Research*, 45(2):793–804.
- 1574 Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1575 Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 1576 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
1577 and homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 1578 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
1579 transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538.
- 1580 Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of
1581 protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- 1582 Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic
1583 models for *Escherichia coli* mg1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*,
1584 107(2):369–381. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1585 The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*,
1586 47(D1):D330–D338.
- 1587 Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and
1588 translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 1589 van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen Assimilation in *Escherichia coli*: Putting Molecular
1590 Data into a Systems Perspective. *Microbiology and Molecular Biology Reviews*, 77(4):628–695.
- 1591 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W.,
1592 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,
1593 E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero,
1594 E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2020). SciPy 1.0:
1595 Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- 1596 Volkmer, B. and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate
1597 Data Conversion for Systems Biology Modeling. *PLOS ONE*, 6(7):e23126.

- ¹⁵⁹⁸ Vollmer, W., Blanot, D., and De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2):149–167.
- ¹⁶⁰⁰ Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Letters*, 545(1):61–70.
- ¹⁶⁰¹ Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- ¹⁶⁰³ You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C. W., Wang, Y.-P., Lenz, P., Yan, D., and Hwa, T. (2013). Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306.
- ¹⁶⁰⁵ Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 108(46):18643–18648.
- ¹⁶⁰⁷ Zhang, L., Jiang, W., Nan, J., Almqvist, J., and Huang, Y. (2014a). The *Escherichia coli* CysZ is a pH dependent sulfate transporter that can be inhibited by sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(7):1809–1816.
- ¹⁶⁰⁹ Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014b). Regulation of crp gene expression by the catabolite repressor/activator, cra, in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- ¹⁶¹¹ Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.

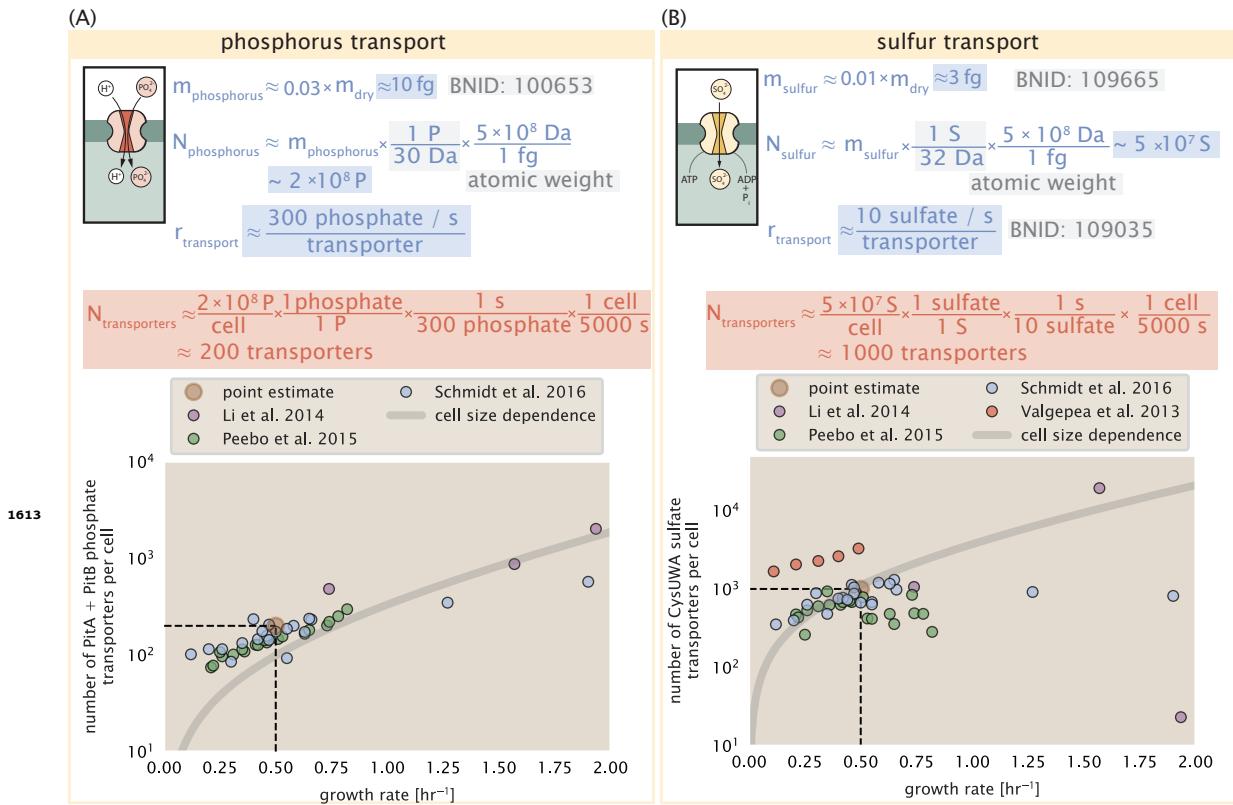


Figure 2-Figure supplement 1. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell. (B) Estimate of the number of CysUWA complexes necessary to maintain a 1% sulfur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

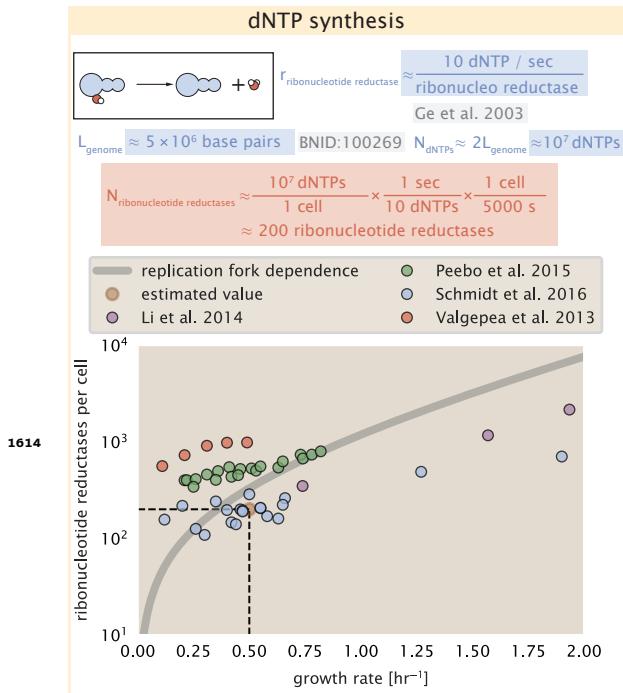


Figure 6–Figure supplement 1. Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ($[\text{NrdA}]_2[\text{NrdB}]_2$)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

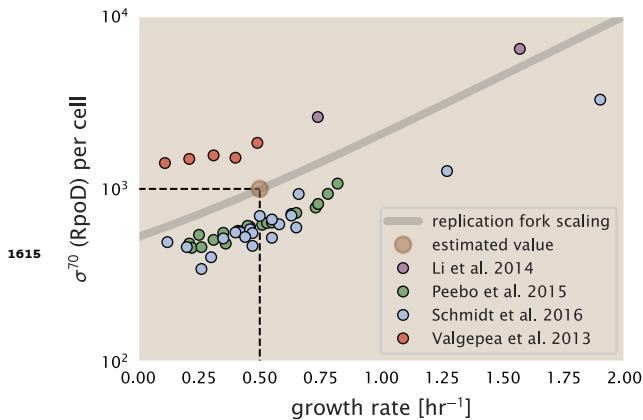


Figure 7–Figure supplement 1. The abundance of σ^{70} as a function of growth rate. Estimated value for the number of RNAP is shown as a translucent brown point and grey line.

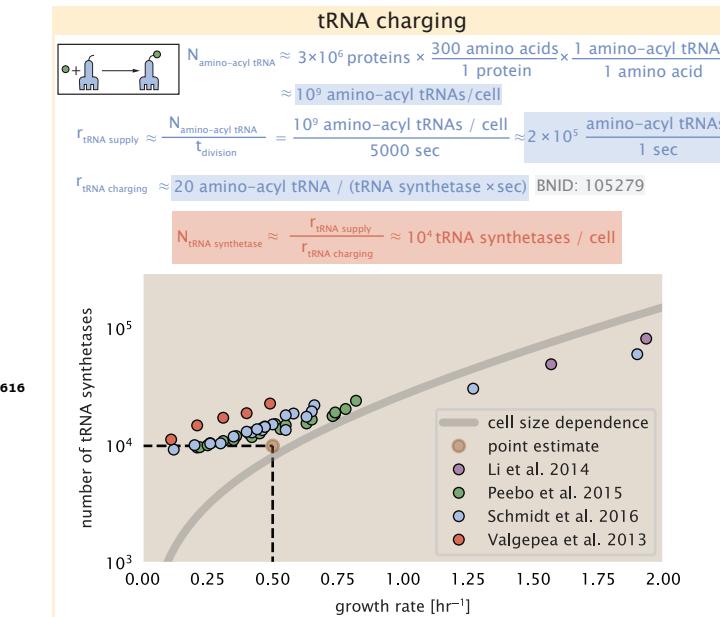


Figure 8-Figure supplement 1. Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [GltX], [IleS], [LeuS], [ValS], [AlaS]₂, [AsnS]₂, [AspS]₂, [TyrS]₂,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

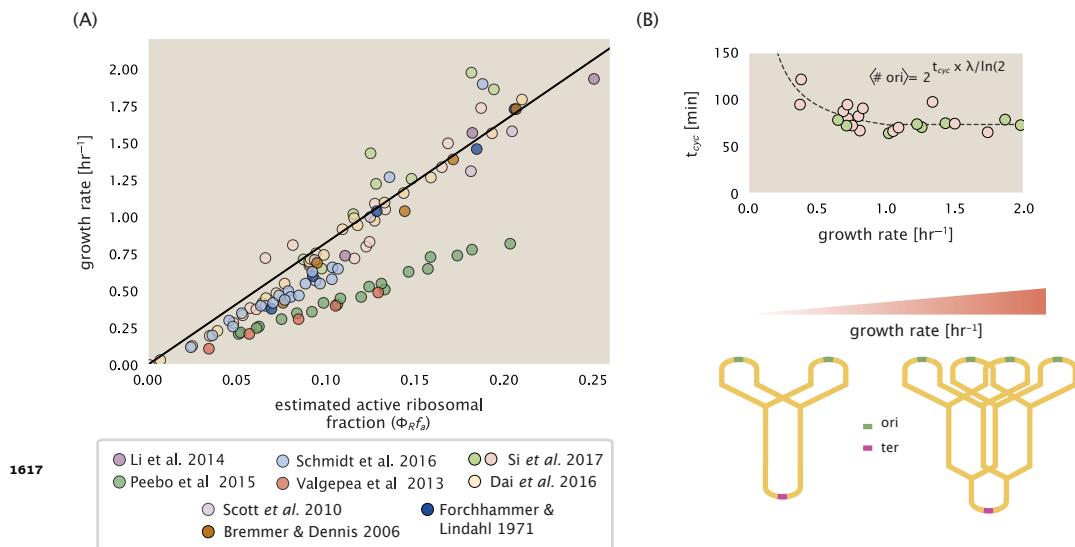


Figure 9-Figure supplement 1. (A) Actively translating ribosomal fraction versus growth rate. The actively translating ribosomal fraction is calculated using the estimated values of f_a from ? (shown in inset; see Appendix Calculation of active ribosomal fraction for additional detail). Additional measurements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

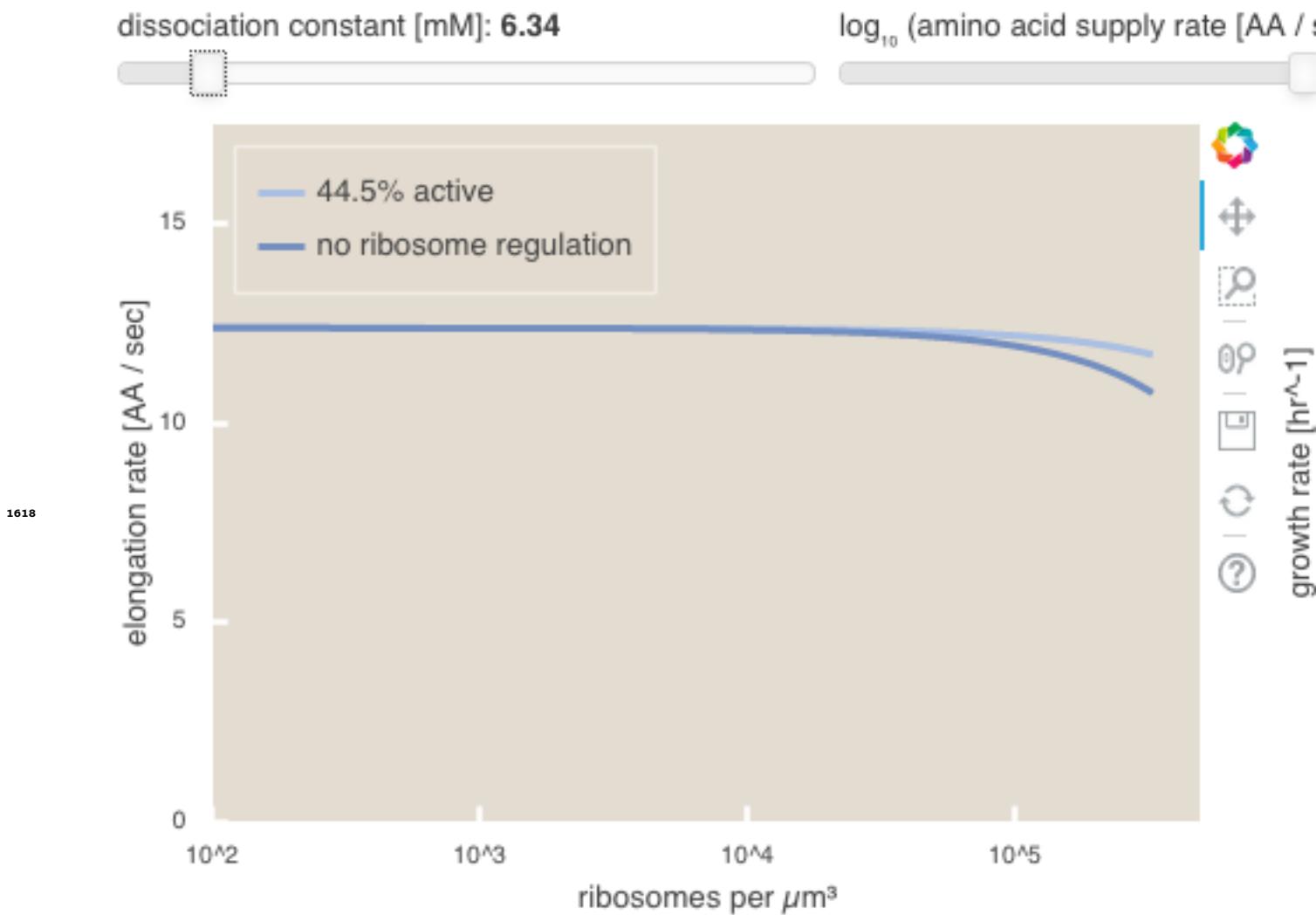


Figure 11–Figure supplement 1. An interactive version of parts (B) and (C) of **Figure 11** which permit the user to modulate the rate of amino acid supply, the dissociation constant of amino acids to the ribosome, and the fraction of the ribosome pool that is actively translating. This interactive figure, and the code used to generate it, is available on the [paper website](#).