

Fundamental limits on the rate of bacterial cell division

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

*For correspondence:

[†]These authors contributed equally to this work

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

15

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

27

Introduction

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

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

42 pioneering efforts in the middle of the last century can be revisited by examining them through the
 43 lens of the increasingly refined molecular census that is available for bacteria such as the microbial
 44 workhorse *Escherichia coli*. In this work, we explore an amalgamation of recent proteomic data sets
 45 to explore fundamental limits of bacterial growth.

46 Several of the evergreen questions about bacterial growth that were originally raised by micro-
 47 biologists in the middle of the 20th century can now be reframed in light of this newly available
 48 data. For example, what biological processes set the absolute speed limit for how fast bacterial
 49 cells can grow and reproduce? How do cells alter the absolute numbers and relative ratios of their
 50 molecular constituents as a function of changes in growth rate or nutrient availability? In this paper,
 51 we address these two questions from two distinct angles. First, as a result of an array of high-quality
 52 proteome-wide measurements of the *E. coli* proteome under myriad growth conditions, we have a
 53 census that allows us to explore how the number of key molecular players change as a function
 54 of growth rate. This census provides a window into the question of whether the rates of central
 55 processes such as energy generation or DNA synthesis are regulated systematically as a function of
 56 cell growth rate by altering protein copy number in individual cells. Second, by compiling molecular
 57 turnover rate measurements for many of the fundamental processes associated with bacterial
 58 growth, we can make quantitative estimates to determine whether the observed protein copy
 59 numbers under varying conditions appear to be in excess of what would be minimally required to
 60 support cell growth at the observed rates.

61 In this paper, we make a series of order-of-magnitude estimates for the copy numbers and
 62 growth rate dependent expression of a variety of different processes, schematized in **Figure 1**,
 63 informed by the collection of proteomic data sets. We use these estimates to explore which, if any,
 64 of the hypothesis illustrated in **Figure 1** may act as molecular bottlenecks that limit bacterial growth.
 65 Specifically, we leverage a combination of *E. coli* proteomic data sets collected over the past decade
 66 using either mass spectrometry (*Schmidt et al., 2016; Peebo et al., 2015; Valgepea et al., 2013*) or
 67 ribosomal profiling (*Li et al., 2014*) across 31 unique growth conditions. Throughout, our estimates
 68 we consider a modest growth rate of $\approx 0.5 \text{ hr}^{-1}$ corresponding to a doubling time of ≈ 5000 seconds,
 69 as the the data sets heavily sample this regime. While we formulate point estimates for the complex
 70 abundances at this division time, we consider how these values will vary at other growth rates due
 71 to changes in cell size, surface area, and chromosome copy number (*Taheri-Araghi et al., 2015*).

72 Broadly, we find that for the majority of these estimates the protein copy numbers appear
 73 well-tuned for the task of cell doubling at a given growth rate. From our analysis, it emerges that
 74 translation, particularly of ribosomal proteins, is the most plausible candidate for a molecular
 75 bottleneck. We reach this conclusion by considering that translation is 1) a rate limiting step for
 76 the *fastest* bacterial division, and 2) a major determinant of bacterial growth across the nutrient
 77 conditions we have considered under steady state, exponential growth. This enables us to suggest
 78 that the long-observed correlation between growth rate and cell size (*Schaechter et al., 1958; Si*
 79 *et al., 2017*) can be simply attributed to the increased absolute number of ribosomes per cell under
 80 conditions supporting extremely rapid growth, a hypothesis which we formally mathematize and
 81 explore.

127 Uptake of Nutrients

128 We begin our series of estimates by considering the critical transport processes diagrammed in
 129 **Figure 1(A)**. In order to build new cellular mass, the molecular and elemental building blocks must
 130 be scavenged from the environment in different forms. Carbon, for example, is acquired via the
 131 transport of carbohydrates and sugar alcohols with some carbon sources receiving preferential
 132 treatment in their consumption (*Monod, 1947*). Phosphorus, sulfur, and nitrogen, on the other hand,
 133 are harvested primarily in the forms of inorganic salts, namely phosphate, sulfate, and ammonia
 134 (*Jun et al., 2018; Assentoft et al., 2016; Stasi et al., 2019; Antonenko et al., 1997; Rosenberg et al.,*
 135 *1977; Willsky et al., 1973*). All of these compounds have different permeabilities across the cell
 136 membrane *Phillips (2018)* and most require some energetic investment either via ATP hydrolysis

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

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

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

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

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

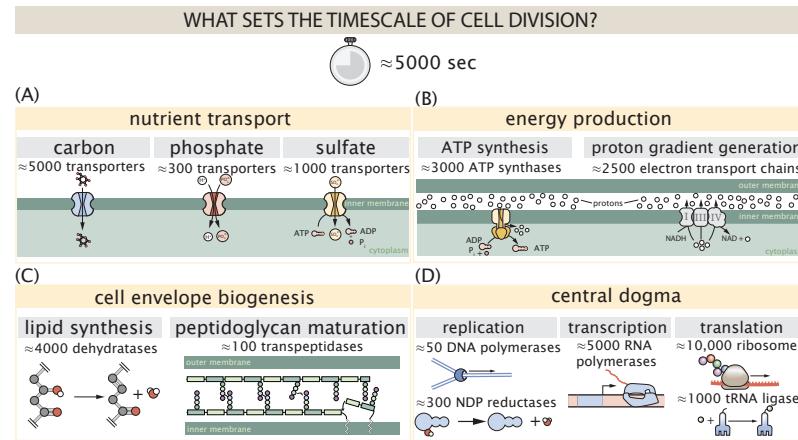


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP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

137 or through the proton electrochemical gradient to bring the material across the hydrophobic cell
 138 membrane. Given the diversity of biological transport mechanisms and the vast number of inputs
 139 needed to build a cell, we begin by considering transport of some of the most important cellular
 140 ingredients: carbon, nitrogen, oxygen, hydrogen, phosphorus, and sulfur.

141 The elemental composition of *E. coli* has received much quantitative attention over the past
 142 half century (*Neidhardt et al., 1991; Taymaz-Nikerel et al., 2010; Heldal et al., 1985; Bauer and Ziv,*
 143 *1976*), providing us with a starting point for estimating the copy numbers of various transporters.
 144 While there is some variability in the exact elemental percentages (with different uncertainties),
 145 we can estimate that the dry mass of a typical *E. coli* cell is $\approx 45\%$ carbon (BioNumber ID: 100649,
 146 see **Box 1**), $\approx 15\%$ nitrogen (BNID: 106666), $\approx 3\%$ phosphorus (BNID: 100653), and 1% sulfur (BNID:
 147 100655). In the coming paragraphs, we will engage in a dialogue between back-of-the-envelope
 148 estimates for the numbers of transporters needed to facilitate these chemical stoichiometries and
 149 the experimental proteomic measurements of the biological reality. Such an approach provides the
 150 opportunity to test if our biological knowledge is sufficient to understand the scale at which these
 151 complexes are produced. At the end of this section, we discuss physical limits as to the number
 152 of transporters that can be present, and comment on the plausibility of this process acting as a
 153 molecular bottleneck.

154 **Nitrogen Transport**

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

167 work, permitting us to neglect nitrogen transport as a potential rate limiting process in cell division
 168 in typical experimental conditions.

169 Carbon Transport

170 We begin with the most abundant element in *E. coli* by mass, carbon. Using ≈ 0.3 pg as the typical *E.*
 171 *coli* dry mass (BNID: 103904), we estimate that $\sim 10^{10}$ carbon atoms must be brought into the cell in
 172 order to double all of the carbon-containing molecules (**Figure 2(A, top)**). Typical laboratory growth
 173 conditions, such as those explored in the aforementioned proteomic data sets, provide carbon
 174 as a single class of sugar such as glucose, galactose, or xylose to name a few. *E. coli* has evolved
 175 myriad mechanisms by which these sugars can be transported across the cell membrane. One
 176 such mechanism of transport is via the PTS system which is a highly modular system capable of
 177 transporting a diverse range of sugars (*Escalante et al., 2012*). The glucose-specific component of
 178 this system transports ≈ 200 glucose molecules per second per transporter (BNID: 114686). Making
 179 the assumption that this is a typical sugar transport rate, coupled with the need to transport $\sim 10^{10}$
 180 carbon atoms, we arrive at the conclusion that on the order of 1,000 transporters must be expressed
 181 in order to bring in enough carbon atoms to divide in 5000 s, diagrammed in the top panel of
 182 **Figure 2(A)**. This estimate, along with the observed average number of the PTS system carbohydrate
 183 transporters present in the proteomic data sets (*Schmidt et al., 2016; Peebo et al., 2015; Valgepea*
 184 *et al., 2013; Li et al., 2014*), is shown in **Figure 2(A)**. While we estimate 1500 transporters are needed
 185 with a 5000 s division time, we can abstract this calculation to consider any particular growth rate
 186 given knowledge of the cell density and volume as a function of growth rate and direct the reader
 187 to the Supplemental Information for more information. As revealed in **Figure 2(A)**, experimental
 188 measurements exceed the estimate by several fold, illustrating that transport of carbon into the
 189 cell is not rate limiting for cell division. Abstracting this point estimate at 5000 s to a continuum
 190 of growth rates (grey line in **Figure 2(A)**) reveals an excess of transporters at other growth rates,
 191 though in rapid growth regimes, the abundance is below our simple estimate.

192 The estimate presented in **Figure 2(A)** neglects any specifics of the regulation of the carbon
 193 transport system and presents a view of how many carbohydrate transporters are present on
 194 average. Using the diverse array of growth conditions explored in the proteomic data sets, we
 195 can explore how individual carbon transport systems depend on the population growth rate. In
 196 **Figure 2(B)**, we show the total number of carbohydrate transporters specific to different carbon
 197 sources. A striking observation, shown in the top-left plot of **Figure 2(B)**, is the constancy in the
 198 expression of the glucose-specific transport systems. Additionally, we note that the total number
 199 of glucose-specific transporters is tightly distributed at $\approx 10^4$ per cell, the approximate number of
 200 transporters needed to sustain rapid growth of several divisions per hour. This illustrates that *E.*
 201 *coli* maintains a substantial number of complexes present for transporting glucose regardless of
 202 growth rate, which is known to be the preferential carbon source (*Monod, 1947; Liu et al., 2005;*
 203 *Adelberg et al., 2014*).

204 It is now understood that a large number of metabolic operons are regulated with dual-input
 205 logic gates that are only expressed when glucose concentrations are low (mediated by cyclic-AMP
 206 receptor protein CRP) and the concentration of other carbon sources are elevated (*Gama-Castro*
 207 *et al., 2016; Zhang et al., 2014b*). A famed example of such dual-input regulatory logic is in the
 208 regulation of the *lac* operon which is only natively activated in the absence of glucose and the
 209 presence of allolactose, an intermediate in lactose metabolism (*Jacob and Monod, 1961*), though we
 210 now know of many other such examples (*Ireland et al., 2020; Gama-Castro et al., 2016; Belliveau*
 211 *et al., 2018*). This illustrates that once glucose is depleted from the environment, cells have a means
 212 to dramatically increase the abundance of the specific transporter needed to digest the next sugar
 213 that is present. Several examples of induced expression of specific carbon-source transporters
 214 are shown in **Figure 2(B)**. Points colored in red (labeled by red text-boxes) correspond to growth
 215 conditions in which the specific carbon source (glycerol, xylose, or fructose) is present. These
 216 plots show that, in the absence of the particular carbon source, expression of the transporters

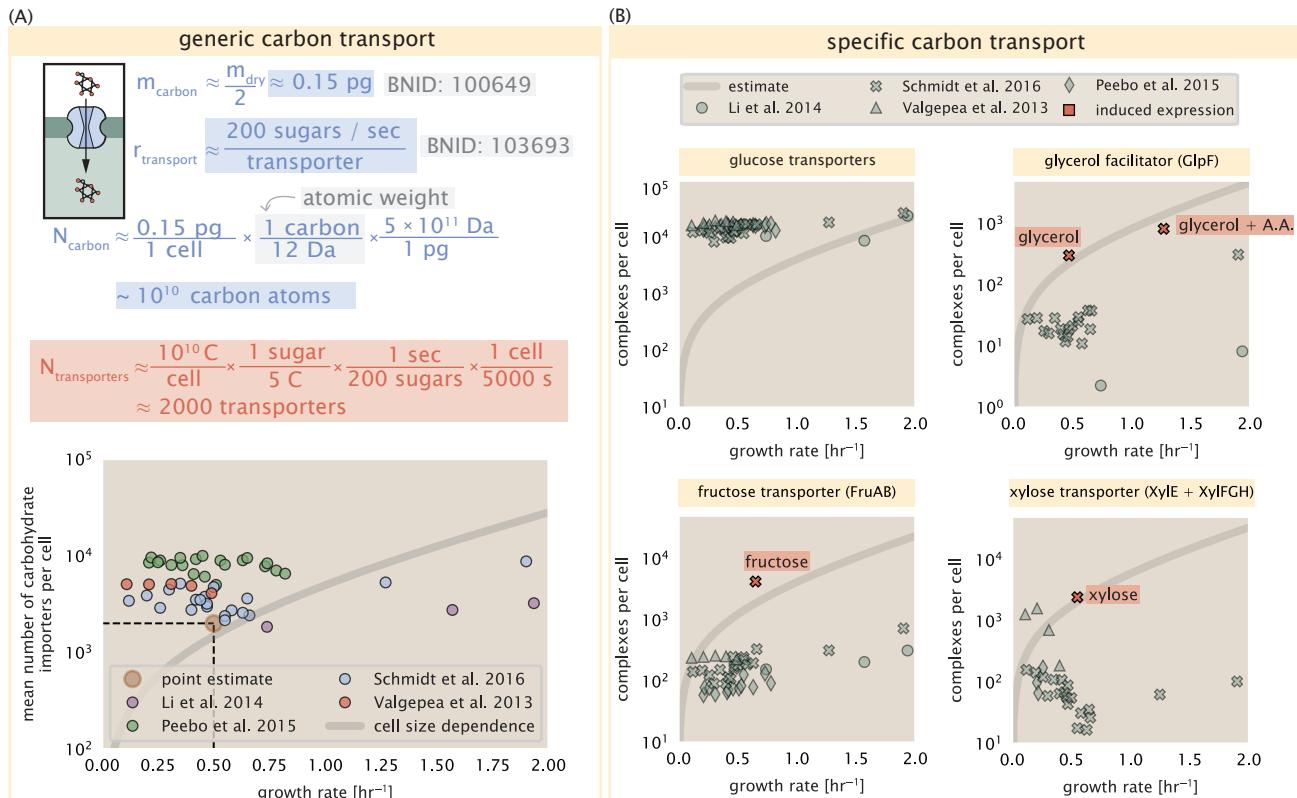


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules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

217 is maintained on the order of $\sim 10^2$ per cell. However, when the transport substrate is present,
 218 expression is induced and the transporters become highly-expressed. The grey lines in **Figure 2(B)**
 219 show the estimated number of transporters needed at each growth rate to satisfy the cellular
 220 carbon requirement. It is notable that in all cases, the magnitude of induced expression (shown in
 221 red) falls close to the estimate, illustrating the ability of the cell to tune expression in response to
 222 changing environments. Together, this generic estimation and the specific examples of induced
 223 expression suggest that transport of carbon across the cell membrane, while critical for growth, is
 224 not the rate-limiting step of cell division.

225 **Phosphorus and Sulfur Transport**

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

250 Satisfied that there are a sufficient number of phosphate transporters present in the cell, we now
 251 turn to sulfur transport as another potentially rate limiting process. Similar to phosphate, sulfate is
 252 highly-charged and not particularly membrane permeable, requiring active transport. While there
 253 exists a H⁺/sulfate symporter in *E. coli*, it is in relatively low abundance and is not well characterized
 254 (Zhang et al., 2014a). Sulfate is predominantly acquired via the ATP-dependent ABC transporter
 255 CysUWA system which also plays an important role in selenium transport (Sekowska et al., 2000;
 256 Sirko et al., 1995). While specific kinetic details of this transport system are not readily available,
 257 generic ATP transport systems in prokaryotes transport on the order of 1 to 10 molecules per
 258 second (BNID: 109035). Combining this generic transport rate, measurement of sulfur comprising
 259 1% of dry mass, and a 5000 second division time yields an estimate of ≈ 1000 CysUWA complexes
 260 per cell (**Figure 3(B)**). Once again, this estimate is in notable agreement with proteomic data sets,
 261 suggesting that there are sufficient transporters present to acquire the necessary sulfur. In a similar
 262 spirit of our estimate of phosphorus transport, we emphasize that this is likely an overestimate of
 263 the number of necessary transporters as we have neglected other sulfur scavenging systems that
 264 are in lower abundance.

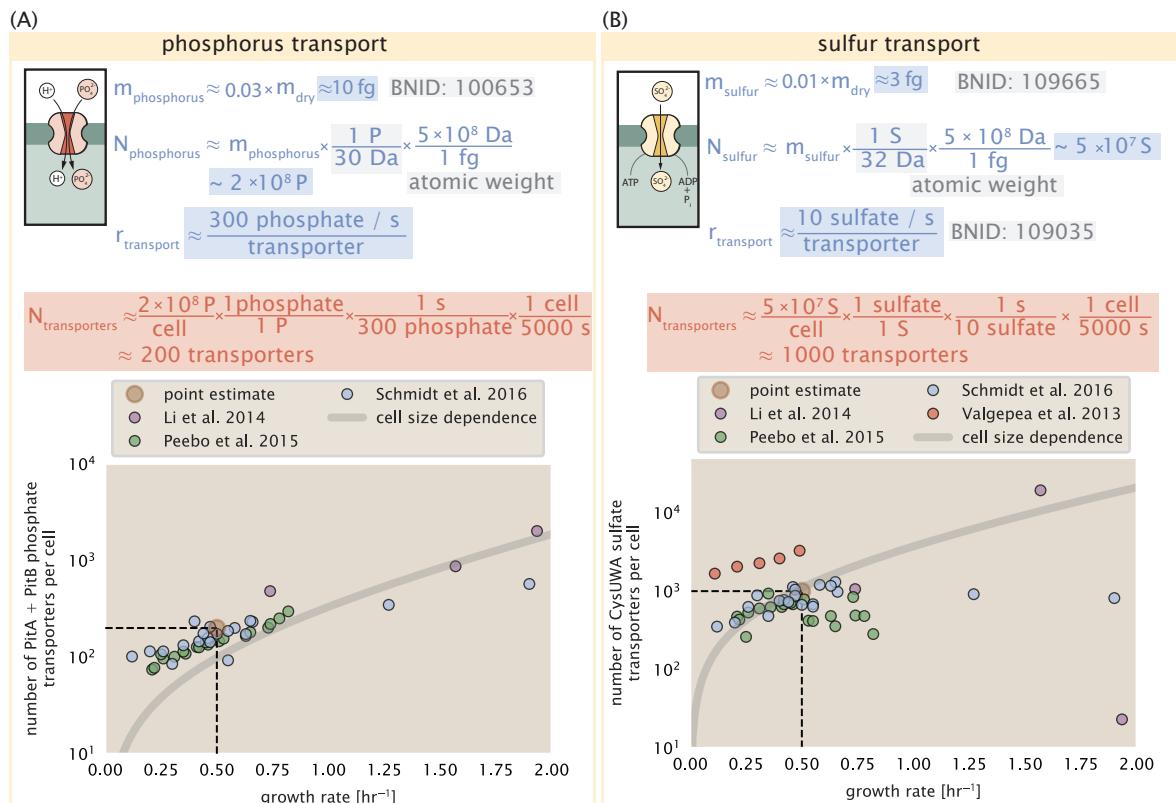


Figure 3. Estimates and measurements of phosphate and sulfate transport systems as a function of growth rate. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell.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

265 **Limits on Transporter Expression**

266 So which, if any, of these processes may be rate limiting for growth? As suggested by *Figure 2*
 267 (B), induced expression can lead to an order-of-magnitude (or more) increase in the amount of
 268 transporters needed to facilitate transport. Thus, if acquisition of nutrients was the limiting state
 269 in cell division, could expression simply be increased to accommodate faster growth? A way to
 270 approach this question is to compute the amount of space in the bacterial membrane that could be
 271 occupied by nutrient transporters. Considering a rule-of-thumb for the surface area of *E. coli* of
 272 about $5 \mu\text{m}^2$ (BNID: 101792), we expect an areal density for 1000 transporters to be approximately
 273 200 transporters/ μm^2 . For a typical transporter occupying about $50 \text{ nm}^2/\text{dimer}$, this amounts to
 274 about only 1 percent of the total inner membrane (*Szenk et al., 2017*). In addition, bacterial cell
 275 membranes typically have densities of 10^5 proteins/ μm^2 (*Phillips, 2018*), implying that the cell could
 276 accommodate more transporters of a variety of species if it were rate limiting. As we will see in the
 277 next section, however, occupancy of the membrane can impose other limits on the rate of energy
 278 production.

279 **Translation and Ribosomal Synthesis**

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

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

291 **tRNA Synthetases**

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

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

303 There are many processes which go into synthesizing a tRNA and ligating it with the appropriate
 304 amino acids. As we discussed previously, there appear to be more than enough RNA polymerases
 305 per cell to synthesize the needed pool of tRNAs. Without considering the many ways in which
 306 amino acids can be scavenged or synthesized *de novo*, we can explore ligation the as a potential
 307 rate limiting step. The enzymes which link the correct amino acid to the tRNA, known as tRNA
 308 synthetases or tRNA ligases, are incredible in their proofreading of substrates with the incorrect
 309 amino acid being ligated once out of every 10^4 to 10^5 events (BNID: 103469). This is due in part
 310 to the consumption of energy as well as a multi-step pathway to ligation. While the rate at which
 311 tRNA is ligated is highly dependent on the identity of the amino acid, it is reasonable to state that
 312 the typical tRNA synthetase has charging rate of ≈ 20 AA per tRNA synthetase per second (BNID:
 313 105279).

314 We can make an assumption that amino-acyl tRNAs are in steady-state where they are produced
 315 at the same rate they are consumed, meaning that 2×10^5 tRNAs must be charged per second.
 316 Combining these estimates together, as shown schematically in **Figure 4(A)**, yields an estimate of
 317 $\sim 10^4$ tRNA synthetases per cell with a division time of 5000 s. This point estimate is in very close
 318 agreement with the observed number of synthetases (the sum of all 20 tRNA synthetases in *E. coli*).
 319 This estimation strategy seems to adequately describe the observed growth rate dependence of
 320 the tRNA synthetase copy number (shown as the grey line in **Figure 4(B)**), suggesting that the copy
 321 number scales with the cell volume.

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

328 Protein Synthesis

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

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

342 Discussion

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

359 In general, the copy numbers of the complexes involved in these processes were reasonable
 360 agreement with our order-of-magnitude estimates. Since many of these estimates represent soft
 361 lower-bound quantities, this suggests that cells do not express proteins grossly in excess of what
 362 is needed for a particular growth rate. Several exceptions, however, also highlight the dichotomy

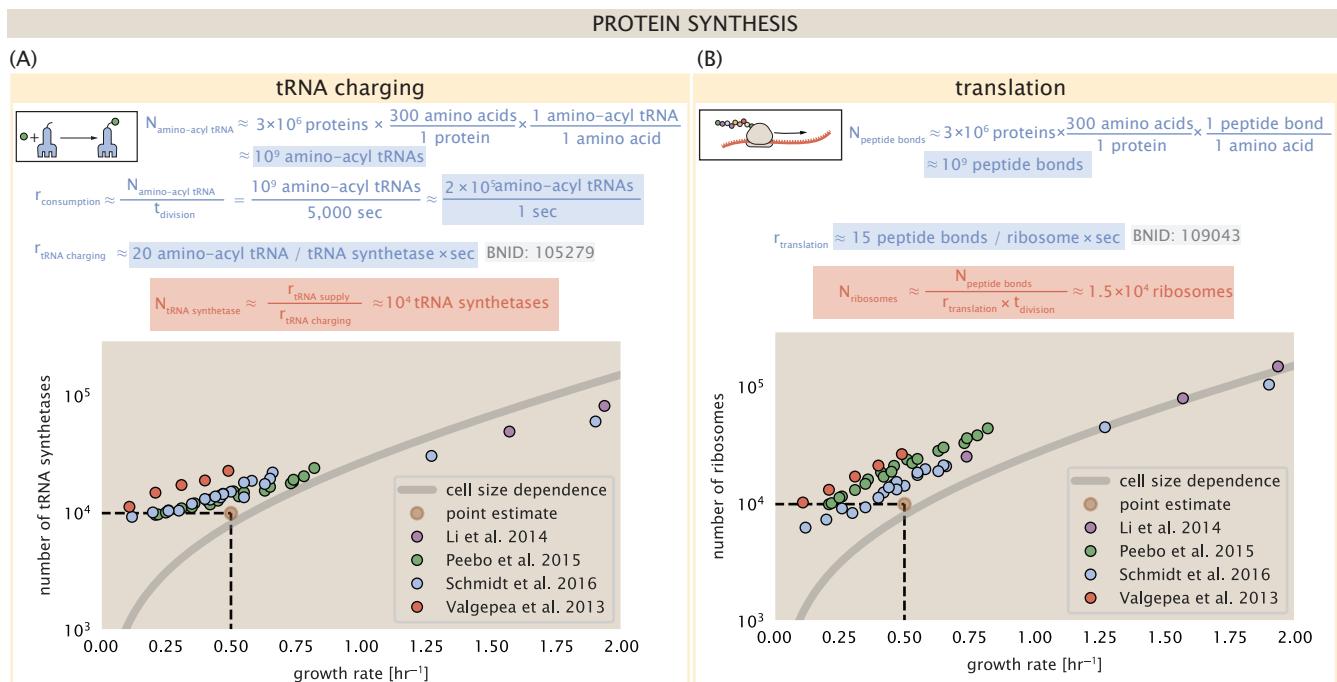


Figure 4. Estimation of the required tRNA synthetases and ribosomes. (A) Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [Glx], [IleS], [LeuS], [ValS], [AlaS]₂, [AsnS]₂, [AspS]₂, [TyrS]₂, [TrpS]₂, [ThrS]₂, [SerS]₂, [ProS]₂, [PheS]₂[PheT]₂, [MetG]₂, [IysS]₂, [HisS]₂, [GlyS]₂[GlyQ]₂). (B) Estimation of the number of ribosomes required to synthesize 10⁹ peptide bonds with an elongation rate of 15 peptide bonds per second. The average abundance of ribosomes is plotted as a function of growth rate. Our estimated values are shown for a growth rate of 0.5 hr⁻¹. Grey lines correspond to the estimated complex abundance calculated at different growth rates. See Supplemental Information XX for a more detail description of this calculation.

363 between a proteome that appears to "optimize" expression according to growth rate and one that
 364 must be able to quickly adapt to environments of different nutritional quality. Take, for example, the
 365 expression of carbon transporters. Shown in **Figure 2(B)**, we find that cells always express a similar
 366 number of glucose transporters irrespective of growth condition. At the same time, it is interesting
 367 to note that many of the alternative carbon transporters are still expressed in low but non-zero
 368 numbers (\approx 10-100 copies per cell) across growth conditions. This may relate to the regulatory
 369 configuration for many of these operons, which require the presence of a metabolite signal in
 370 order for alternative carbon utilization operons to be induced (Monod, 1949; Laxhuber et al., 2020).
 371 Furthermore, upon induction, these transporters are expressed and present in abundances in close
 372 agreement with a simple estimate.

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

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

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

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

425 Supplemental material for: 426 **Fundamental limits on the rate of 427 bacterial cell division**

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

*For correspondence:

[†]These authors contributed equally
to this work

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

440 **Contents**

441	Introduction	1
442	Uptake of Nutrients	2
443	Nitrogen Transport	4
444	Carbon Transport	5
445	Phosphorus and Sulfur Transport	7
446	Limits on Transporter Expression	9
447	Translation and Ribosomal Synthesis	9
448	tRNA Synthetases	9
449	Protein Synthesis	10
450	Discussion	10
451	Summary of Proteome Data: Experimental Details	16
452	Fluorescence based measurements	16
453	Ribosomal profiling measurements	16
454	Mass spectrometry measurements	17
455	Summary of Proteomic Data	17
456	Estimation of Cell Size, Surface Area	18
457	Estimation of Total Protein Content per Cell	19
458	Additional Considerations of Schmidt <i>et al.</i> Data Set	20
459	Effect of cell volume on reported absolute protein abundances	22
460	Relaxing assumption of constant protein concentration across growth conditions	23
461	Experimental measurements of total protein from Basan <i>et al.</i> 2015.	24
462	Calculation of Complex Abundance	24
463	Extending Estimates to a Continuum of Growth Rates	25
464	Estimation of the total cell mass	27
465	Complex Abundance Scaling With Cell Volume	27
466	A Relation for Complex Abundance Scaling With Surface Area	28
467	Number of Lipids	28
468	Number of Murein Monomers	28
469	Complex Abundance Scaling With Number of Origins	29
470	Estimation of $\langle \#ori \rangle / \langle \#ter \rangle$ and $\langle \#ori \rangle$.	29
471	Calculation of active ribosomal fraction.	30

Table 1. Overview of proteomic data sets.

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

- a. The reported values assume that the proteins are long-lived compared to the generation time but are unable to account for post-translational modifications that may alter absolute protein abundances.
- b. This mass spectrometry approach differs substantially from the others since in addition to the relative proteome-wide abundance measurements, the authors performed absolute quantification of 41 proteins across all growth conditions (see Section Additional Considerations of Schmidt *et al.* Data Set for more details on this).

472 Summary of Proteome Data: Experimental Details

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

480 Fluorescence based measurements

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

489 Ribosomal profiling measurements

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

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

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

506 Mass spectrometry measurements

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

515 Most quantitative experiments rely on labeling protein with stable isotopes, which allow multiple
 516 samples to be measured together by the mass spectrometer. By measuring samples of known total
 517 protein abundance simultaneously (i.e. one sample of interest, and one reference), it is possible to
 518 determine relative protein abundances. Absolute protein abundances can be estimated following
 519 the same approach used above for ribosomal profiling, which is to multiply each relative abundance
 520 measurement by the total cellular protein per cell. This is the approach taken by *Valgepea et al.*
 521 (*2013*) and *Peebo et al.* (*2015*), with relative protein abundances determined based on the relative
 522 peptide intensities (label free quantification 'LFQ' intensities). For the data of *Valgepea et al.* (*2013*),
 523 total protein per cell was determined by measuring total protein by the Lowry method, and counting
 524 colony-forming units following serial dilution. For the data from *Peebo et al.* (*2015*), the authors did
 525 not determine cell quantities and instead report the cellular protein abundances in protein per unit
 526 volume by assuming a mass density of 1.1 g/ml, with a 30% dry mass fraction.

527 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide
 528 fragments of known abundance. These can serve as a direct way to calibrate mass spectrometry
 529 signal intensities to absolute mass. This is the approach taken by *Schmidt et al.* (*2016*). In addition
 530 to a set of shotgun proteomic measurements to determine proteome-wide relative abundances,
 531 the authors also performed absolute quantification of 41 proteins covering over four orders of
 532 magnitude in cellular abundance. Here, a synthetic peptide was generated for each of the proteins,
 533 doped into each protein sample, and used these to determine absolute protein abundances of the
 534 41 proteins. These absolute measurements, determined for every growth condition, were then
 535 used as a calibration curve to convert proteomic-wide relative abundances into absolute protein
 536 abundance per cell. A more extensive discussion of the *Schmidt et al.* (*2016*) data set can be found
 537 in Section Additional Considerations of Schmidt *et al.* Data Set.

538 Summary of Proteomic Data

539 In the work of the main text we only used the data from *Valgepea et al.* (*2013*); *Li et al.* (*2014*); *Peebo*
 540 *et al.* (*2015*); *Schmidt et al.* (*2016*). As shown in *Figure 5(A)*, the reported total protein abundances
 541 in the work of *Taniguchi et al.* (*2010*) and *Soufi et al.* (*2015*) differed quite substantially from the
 542 other work. For the work of *Taniguchi et al.* (*2010*) this is in part due to a lower coverage in total
 543 proteomic mass quantified, though we also noticed that most proteins appear undercounted when
 544 compared to the other data.

545 *Figure 5(B)* summarizes the total protein mass for each data point in our final compiled data
 546 set. We note that protein abundances were all scaled so they followed a common growth rate-
 547 dependent change in total protein mass. While our inclination initially was to leave reported copy
 548 numbers untouched, a notable discrepancy in the scaling total protein per cell between *Schmidt*
 549 *et al.* (*2016*) and the other data sets forced us to dig deeper into those measurements (compare
 550 *Schmidt et al.* (*2016*) and *Li et al.* (*2014*) data in *Figure 5(A)*). The particular trend in *Schmidt et al.*
 551 (*2016*) appears to be due to assumptions of cell size and we provide a more extensive discussion

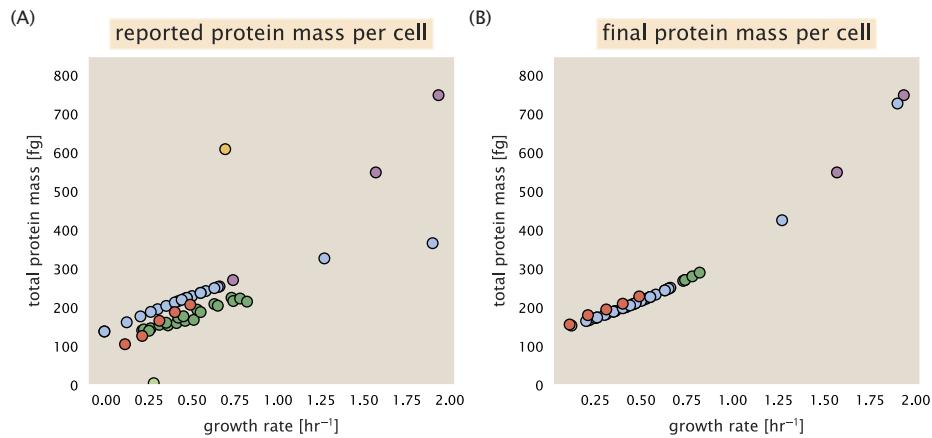


Figure 5. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as original reported in the data sets of *Taniguchi et al. (2010)*; *Valgepea et al. (2013)*; *Li et al. (2014)*; *Soufi et al. (2015)*; *Peebo et al. (2015)*; *Schmidt et al. (2016)*. Note that the data from *Peebo et al. (2015)* only reported protein abundances per unit volume and total protein per cell was found by multiplying these by the growth-rate dependent cell size as determined by *Si et al. (2017)*. (B) Adjusted total protein abundances across the proteomic data sets are summarized. Protein abundances were adjusted so that all data shared a common set of growth-rate dependent total protein per cell and cellular protein concentration following the cell size expectations of *Si et al. (2017)* (see section on Estimation of Cell Size, Surface Area for further details).

and analysis of that data set in section Additional Considerations of Schmidt *et al.* Data Set. As a compromise, and in an effort to treat all data equally, we instead scaled all protein abundance values to a data-driven estimate of total protein per cell. Here we used cell size measurements from *Si et al. (2017, 2019)*, and an estimate of total protein content through expected dry mass. Total protein per cell was estimated using available data on total DNA, RNA, and protein from *Basan et al. (2015)*; *Dai et al. (2016)*, which account for the majority of dry mass in the cell. We consider these details in sections Estimation of Cell Size, Surface Area and Estimation of Total Protein Content per Cell that follows.

Lastly, in *Figure 6* we show the total proteomic coverage and overlap of proteins quantified across each data set. Here we have used an UpSet diagram (*Lex et al., 2014*) to compare the data. Overall, the overlap in quantified proteins is quite high, with 1157 proteins quantified across all data sets. The sequencing based approach of *Li et al. (2014)* has substantially higher coverage compared to the mass spectrometry data sets (3394 genes versus the 2041 genes quantified in the work of *Schmidt et al. (2016)*). However, in terms of total protein mass, the data from *Li et al. (2014)*; *Schmidt et al. (2016)*; *Peebo et al. (2015)* each quantify roughly equivalent total protein mass. An exception to this is in the data from *Valgepea et al. (2013)*, where we find that the total protein quantified in *Valgepea et al. (2013)* is 90-95 % of the total protein mass (when using the data from *Schmidt et al. (2016)* as a reference).

570 Estimation of Cell Size, Surface Area

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

577 The average size measurements from each of their experiments are shown in Figure *Figure 7*,
 578 with cell length and width shown in (A) and (B), respectively. The length data was well described by
 579 the exponential function $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, while the width data was well described by $0.64 e^{0.24 \cdot \lambda}$.

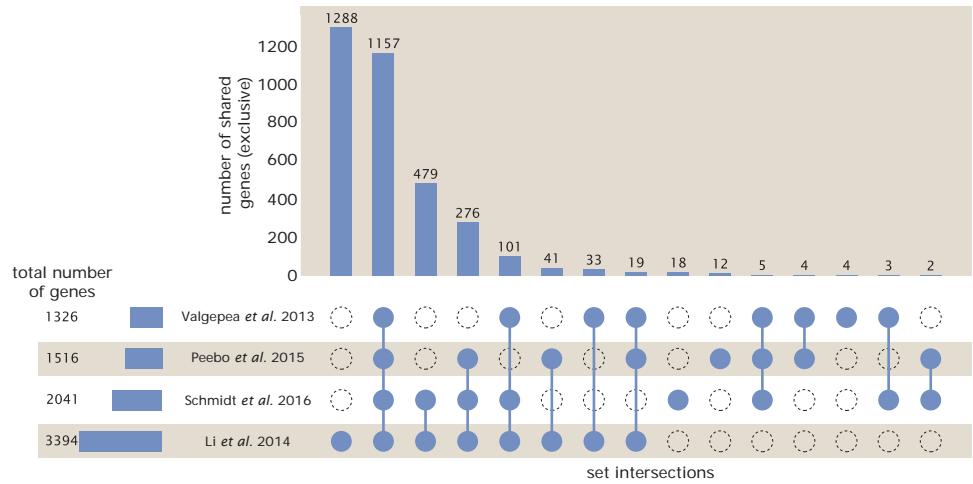


Figure 6. Comparison of proteomic coverage across different data sets. An UpSet diagram (Lex et al., 2014) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of Valgepea et al. (2013); Li et al. (2014); Schmidt et al. (2016); Peebo et al. (2015). The total number of genes reported in each individual data set are noted on the bottom left, while their overlap is summarized in the bar chart. Each column here refers to the intersection of a set of data sets identified by blue circles.

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

$$V = \pi \cdot r^2 \cdot (l - 2r/3), \quad (1)$$

582 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
583 the cell surface area is given by,

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

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

585 Estimation of Total Protein Content per Cell

586 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell
587 size from the fit shown in Figure 7(C) ($0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$). We then estimate the total protein
588 content from the total dry mass of the cell. Here we begin by noting that for almost the entire range
589 of growth rates considered here, protein, DNA, and RNA were reported to account for at least 90 %
590 of the dry mass (Basan et al. (2015)). The authors also found that the total dry mass concentration
591 was roughly constant across growth conditions. Under such a scenario, we can calculate the total
592 dry mass concentration for protein, DNA, and RNA, which is given by $1.1 \text{ g/ml} \times 30 \% \times 90 \% \text{ or about}$
593 $[M_p] = 300 \text{ fg per fl}$. Multiplying this by our prediction of cell size gives the total dry mass per cell.

594 However, even if dry mass concentration is relatively constant across growth conditions, it
595 is not obvious how protein concentration might vary due to the substantial increase in rRNA at
596 faster growth rates (Dai et al. (2016)). This is a well-documented result that arises from an increase
597 in ribosomal abundance at faster growth rates (Scott et al. (2010)). To proceed therefore rely on
598 experimental measurements of total DNA content per cell that also come from Basan et al., and RNA
599 to protein ratios that were measured in Dai et al. (and cover the entire range of growth conditions
600 considered here). These are reproduced in Figure 8(A) and (B), respectively.

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

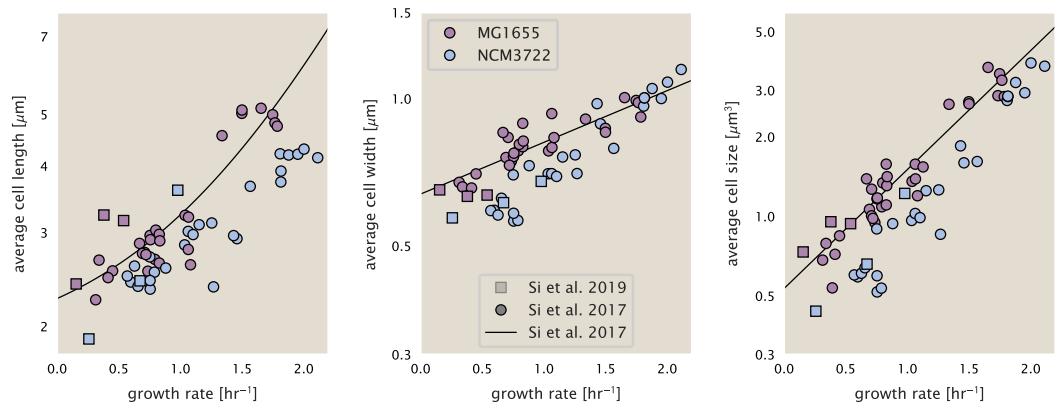


Figure 7. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 images and 5,000-30,000 for each set of samples; which likely means about 1,000-5,000 measurements per mean value reported here since they considered about 6 conditions at a time). Fits were made to the MG1655 strain data; length: $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, width: $0.64 e^{0.24 \cdot \lambda} \mu\text{m}$. (C) Cell size, V , was calculated as cylinders with two hemispherical ends (Equation 1). The MG1655 strain data gave a best fit of $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$.

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

$$M_P = \frac{[M_P + M_{RNA}]}{1 + (RP_{ratio})}. \quad (3)$$

605 (RP_{ratio} refers to the RNA to protein ratio as measured by Dai et al.. In Figure **Figure 8(C)** we plot
 606 the estimated cellular concentrations for protein, DNA, and RNA from these calculations, and in
 607 Figure **Figure 8(D)** we plot their total expected mass per cell. This later quantity is the growth
 608 rate-dependent total protein mass that was used to extimate total protein abundance across all
 609 data sets (and summaried in **Figure 5(B)**).

610 Additional Considerations of Schmidt et al. Data Set

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

617 We begin by looking at their cell volume measurements, which are shown in blue in Figure **Figure 9**. As a comparison, we also plot cell sizes reported in three other recent papers: measurements
 618 from Taheri-Araghi et al. and Si et al. come from the lab of Suckjoon Jun, while those from Basan
 619 et al. come from the lab of Terence Hwa. Each set of measurements used microscopy and cell
 620 segmentation to determine the length and width, and then calculated cell size by treating the cell
 621 is a cylinder with two hemispherical ends, as we considered in the previous section. While there
 622 is notable discrepancy between the two research groups, which are both using strain NCM3722,
 623 Basan et al. found that this came specifically from uncertainty in determining the cell width. This is
 624 prone to inaccuracy given the small cell size and optical resolution limits (further described in their
 625 supplemental text). Perhaps the more concerning point is that while each of these alternative mea-
 626 surements show an exponential increase in cell size at faster growth rates, the measurements used
 627 by Schmidt et al. appear to plateau. This resulted in an analogous trend in their final reported total

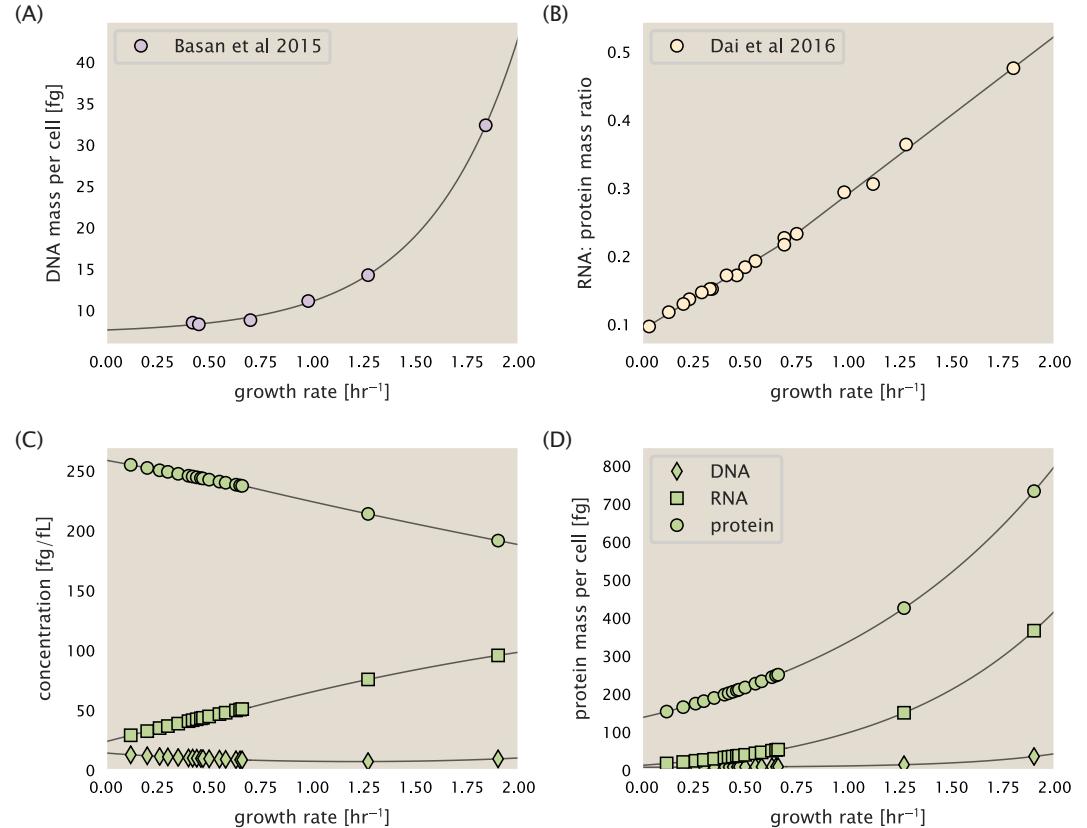


Figure 8. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23\lambda} + 7.2$ fg per cell, where λ is the growth rate in hr⁻¹). (B) RNA to protein measurements as a function of growth rate. The data was fit to two lines: for growth rates below 0.7 hr⁻¹, the RNA/protein ratio is $0.18\lambda + 0.093$, while for growth rates faster than 0.7 hr⁻¹ the RNA/protein ratio is given by $0.25\lambda + 0.035$. For (A) and (B) cells are grown under varying levels of nutrient limitation, with cells grown in minimal media with different carbon sources for the slowest growth conditions, and rich-defined media for fast growth rates. (C) Predictions of cellular protein, DNA, and RNA concentration. (D) Total cellular mass predicted for protein, DNA, and RNA using the cell size predictions from Si *et al.*. Symbols (diamond: DNA, square: RNA, circle: protein) show estimated values of mass concentration and mass per cell for the specific growth rates in Schmidt *et al.* (2016).

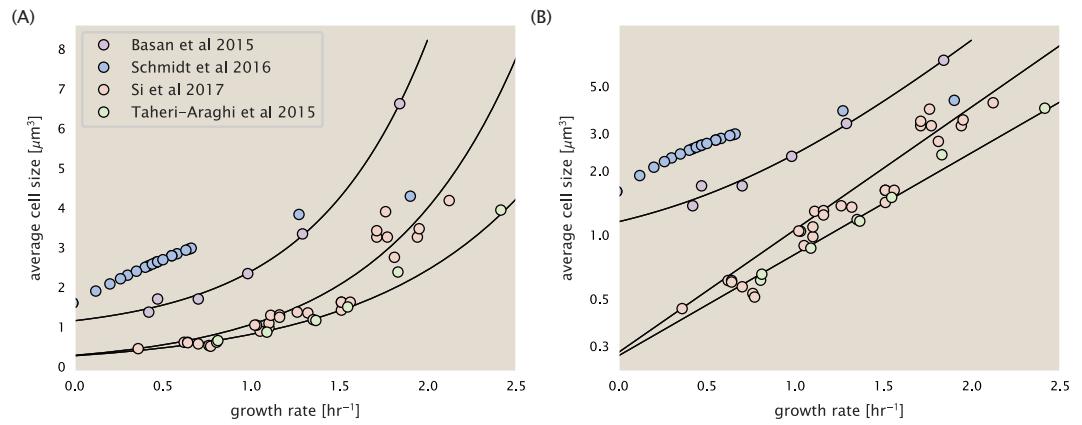


Figure 9. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are shown in purple (*Basan et al. (2015)*), while the two data sets shown in green and red come from the lab of Suckjoon Jun (*Taheri-Araghi et al. (2015); Si et al. (2017)*). (B) Same as in (A) but with the data plotted on a logarithmic y-axis to highlight the exponential scaling that is expected for *E. coli*.

629 cellular protein per cell as shown in Figure *Figure 10* (purple data points), and is in disagreement
630 with other measurements of total protein at these growth rates (*Basan et al., 2015*).

631 Since it is not obvious how measurements of cell size influenced their reported protein abundances,
632 in the following subsections we begin by considering this calculation. We then consider
633 three different approaches to estimate the growth-rate dependent total protein mass to compare
634 with those values reported from *Schmidt et al. (2016)*. The results of this are summarized in
635 *Figure 9(B)*, with the original values from both *Schmidt et al. (2016)* and *Li et al. (2014)* shown in
636 *Figure 9(A)* for reference. For most growth conditions, we find that total protein per cell is still in
637 reasonable agreement. However, for the fastest growth conditions, with glycerol + supplemented
638 amino acids, and LB media, all estimates are substantially higher than those originally reported.
639 This is the main reason why we chose to readjusted protein abundance as shown in *Figure 5(B)*
640 (with the calculation described in section Estimation of Total Protein Content per Cell).

641 Effect of cell volume on reported absolute protein abundances

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

654 While this approach provided absolute abundances, another necessary step to arrive at total
655 cellular protein was to account for any protein loss during their various protein extraction steps.
656 Here the authors attempted to determine total protein separately using a BCA protein assay. In
657 personal communications, it was noted that determining reasonable total protein abundances
658 by BCA across their array of growth conditions was particularly troublesome. Instead, they noted

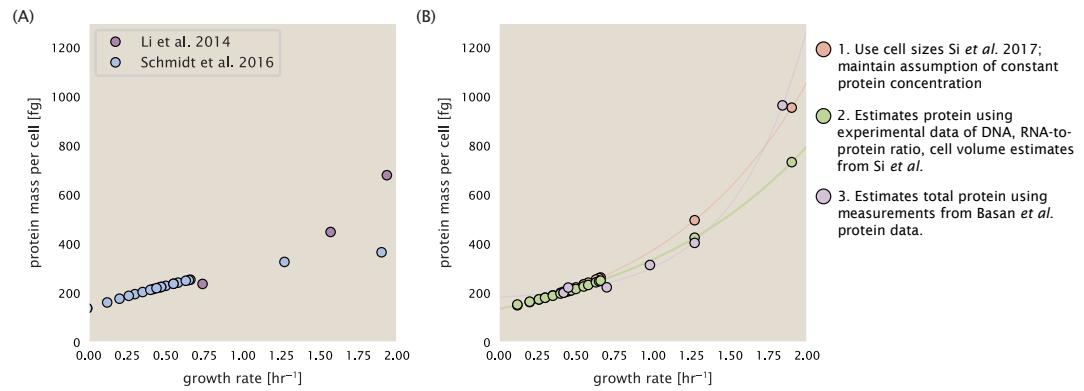


Figure 10. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt et al. (A) The original protein mass from Schmidt et al. and Li et al. are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. light red: Rescaling of total protein mass assuming a growth rate independent protein concentration and cell volumes estimated from Si et al. 2017. 2. light green: Rescaling of total protein mass using estimates of growth rate-dependent protein concentrations and cell volumes estimated from Si et al. 2017. Total protein per cell is calculated by assuming a 1.1 g/ml cellular mass density, 30% dry mass, with 90% of the dry mass corresponding to DNA, RNA, and protein (Basan et al., 2015). See Estimation of Total Protein Content per Cell for details on calculation. 3. light purple: Rescaling of total protein mass using the experimental measurements from Basan et al. 2015.

confidence in their total protein measurements for cells grown in M9 minimal media + glucose and used this as a reference point with which to estimate the total protein for all other growth conditions.

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

$$M_{p,i} = [M_p]_{orig} \cdot V_i \quad (4)$$

where $M_{p,i}$ represents the total protein mass per cell and V_i is the cell volume for each growth condition i as measured in Volkmer and Heinemann, 2011. Here the thinking is that the values of $M_{p,i}$ reflects the total cellular protein for growth condition i , where any discrepancy from their absolute protein abundance is assumed to be due to protein loss during sample preparation. The protein abundances from their absolute abundance measurements noted above were therefore scaled to their estimates and are shown in Figure **Figure 10** (purple data points).

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

$$M'_{p,i} = [M_p]_{Si} \cdot V_{Si,i} \quad (5)$$

where $M'_{p,i}$ is the new protein mass prediction, and $V_{Si,i}$ refers to the new volume prediction for each condition i . These are shown as red data points in Figure **Figure 10(B)**.

680 Relaxing assumption of constant protein concentration across growth conditions

681 We next relax the assumption that cellular protein concentration is constant and instead, attempt
682 to estimate it using experimental data. Here we use the estimation of total protein mass per cell

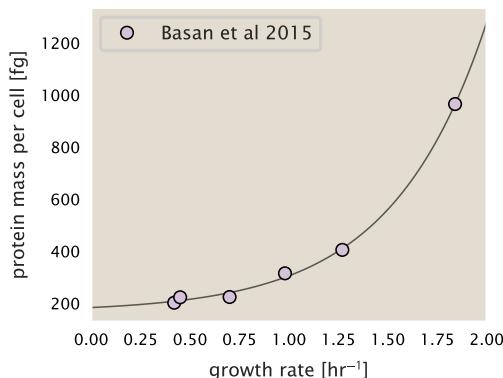


Figure 11. Total cellular protein reported in Basan *et al.* 2015. Measured protein mass as a function of growth rate as reproduced from Basan *et al.* 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given by $14.65 e^{2.180 \cdot \lambda} + 172$ fg per cell, where λ is the growth rate in hr^{-1} .

683 detailed in section Estimation of Total Protein Content per Cell for all data points in the *Schmidt*
 684 *et al.* (2016) data set. The green data points in *Figure 10(B)* show this prediction, and this represents
 685 the approach used to estimate total protein per cell for all data sets.

686 **Experimental measurements of total protein from Basan *et al.* 2015.**

687 One of the challenges in our estimates in the preceding sections is the need to estimate protein
 688 concentration and cell volumes. These are inherently difficult to accurately due to the small size
 689 of *E. coli*. Indeed, for all the additional measurements of cell volume included in Figure *Figure 9*, no
 690 measurements were performed for cells growing at rates below $0.5\ hr^{-1}$. It therefore remains to be
 691 determined whether our extrapolated cell volume estimates are appropriate, with the possibility
 692 that the logarithmic scaling of cell size might break down for slower growth.

693 In our last approach we therefore attempt to estimate total protein using experimental data
 694 that required no estimates of concentration or cell volume. Specifically, in the work of Basan *et al.*,
 695 the authors measured total protein per cell for a broad range of growth rates (reproduced in Figure
 696 *Figure 11*). These were determined by first measuring bulk protein from cell lysate, measured by
 697 the colorimetric Biuret method (*You et al. (2013)*), and then abundance per cell was calculated from
 698 cell counts from either plating cells or a Coulter counter. While it is unclear why Schmidt *et al.* was
 699 unable to take a similar approach, the results from Basan *et al* appear more consistent with our
 700 expectation that cell mass will increase exponentially with faster growth rates. In addition, although
 701 they do not consider growth rates below about $0.5\ hr^{-1}$, it is interesting to note that the protein
 702 mass per cell appears to plateau to a minimum value at slow growth. In contrast, our estimates
 703 using cell volume so far have predicted that total protein mass should continue to decrease slightly
 704 for slower growing cells. By fitting this data to an exponential function dependent on growth rate,
 705 we could then estimate the total protein per cell for each growth condition considered by *Schmidt*
 706 *et al.* (2016). These are plotted as red data points in *Figure 10(B)*.

707 **Calculation of Complex Abundance**

708 All data collected quantified the abundance of individual proteins with high resolution. After
 709 correcting for errors introduced from overestimated volumes and imposed boundaries on the
 710 protein concentration, we are left with a large data set, largely comparable between one another.
 711 However, this work is focused on estimating the abundance of individual protein *complexes*, rather
 712 than copies of individual proteins. In this section, we outline the procedure we used to annotate
 713 proteins as being part of a macromolecular complex as well as how we computed their absolute

714 abundance.

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

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

737 In addition to the annotated complex, we collected information of how many copies of each
 738 individual protein is present in each macromolecular complex. With this number in hand, we
 739 calculated the maximum number of complexes that *could* be formed given the observed abundance
 740 of each protein subunit as

$$N_{\text{complex}}^{(\max)}(\text{subunit}) = \frac{N_{\text{subunit}}^{(\text{observed})}}{N_{\text{subunit}}^{(\text{annotated})}}. \quad (6)$$

741 For example, the 70S mature ribosome complex has 55 protein components, all of which are
 742 present in a single copy except L4 (RplL), which is present in 4 copies. For each ribosomal protein,
 743 we then calculate the maximum number of complexes that could be formed using **Equation 6**. This
 744 example, along with example from 5 other macromolecular complexes, can be seen in **Figure 12**.

745 It is important to note that measurement noise, efficiency of protein extraction, stochastic errors
 746 will mean that the precise value of each calculation will be different for each component of a given
 747 complex. Thus, to report the total complex abundance, we computed the arithmetic mean of
 748 $N_{\text{complex}}^{(\max)}$ for all subunits as

$$\langle N_{\text{complex}} \rangle = \frac{1}{N_{\text{subunits}}} \sum_i^{N_{\text{subunits}}} \frac{N_i^{(\text{observed})}}{N_i^{(\text{annotated})}}. \quad (7)$$

749 in **Figure 12**, we show this mean value as a grey line for a variety of different complexes. Addi-
 750 tionally, we have built an interactive figure accessible on the [paper website](#) where the validity of
 751 this approach can be examined for any complex with more than two subunits (thus, excluding
 752 monomers and dimers).

753 Extending Estimates to a Continuum of Growth Rates

754 In the main text, we considered a standard stopwatch of 5000 s to estimate the abundance of
 755 the various protein complexes considered. In addition to point estimates, we also showed the
 756 estimate as a function of growth rate as transparent grey curves. In this section, we elaborate

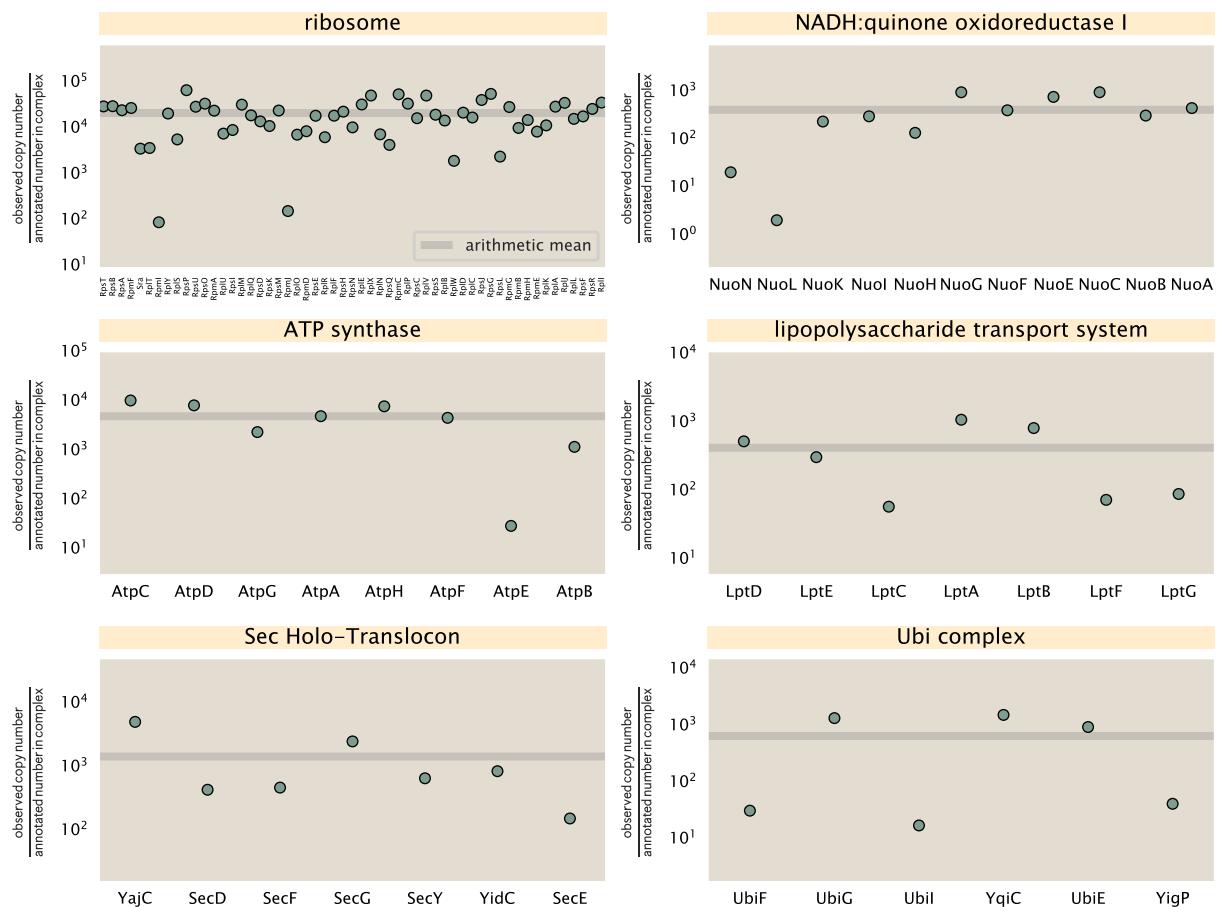


Figure 12. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds to the arithmetic mean across all subunits. These data correspond to measurements from Schmidt et al. (2016) in a glucose-supplemented minimal growth medium.

757 on this continuum estimate, giving examples of estimates that scale with either cell volume, cell
 758 surface area, or number of origins of replication.

759 **Estimation of the total cell mass**

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

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

$$m_{\text{cell}} \approx \rho V(\lambda) \approx \rho a e^{a*\lambda} \quad (8)$$

770 where a and b are constants with units of μm^3 and hr, respectively. The value of these constants can
 771 be estimated from the careful volume measurements performed by *Si et al. (2017)*, as is described
 772 in the previous section.

773 **Complex Abundance Scaling With Cell Volume**

774 Several of the estimates performed in the main text are implicitly dependent on the cell volume.
 775 This includes processes such as ATP synthesis and, most prominently, the transport of nutrients.
 776 Of the latter, we estimated the number of transporters that would be needed to shuttle enough
 777 carbon, phosphorus, and sulfur across the membrane to build new cell mass. To do so, we used
 778 elemental composition measurements combined with a 300 fg cell dry mass to make the point
 779 estimate. As we now have a means to estimate the total cell mass as a function of volume, we can
 780 generalize these estimates across growth rates.

781 Rather than discussing the particular details of each transport system, we will derive this scaling
 782 expression in very general terms. Consider we wish to estimate the number of transporters for
 783 some substance X , which has been measured to be make up some fraction of the dry mass θ_X . If
 784 we assume that, irrespective of growth rate, the cell dry mass is $\approx 30\%$ of the total cell mass, we can
 785 state that the total mass of substance X as a function of growth rate is

$$m_X \approx 0.3 \times \rho V(\lambda) \theta_X, \quad (9)$$

786 where we have used $\rho V(\lambda)$ as an estimate of the total cell mass, defined in *Equation 8*. To convert
 787 this to the number of units N_X of substance X in the cell, we can use the formula weight w_X of a
 788 single unit of X in conjunction with *Equation 9*,

$$N_X \approx \frac{m_X}{w_X}. \quad (10)$$

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

$$N_{\text{transporters}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_X}{w_X r_X \tau}. \quad (11)$$

795 This function is used to draw the continuum estimates for the number of transporters seen in
 796 Figures 2 and 3 as transparent grey curves. Occasionally, this continuum scaling relationship will

797 not precisely agree with the point estimate outlined in the main text. This is due to the fact that we
 798 make an initial approximation made of a dry cell mass of ≈ 300 fg for the point estimate while we
 799 consider more precise values in the continuum estimate. We note, however, that both this scaling
 800 relation and the point estimates are meant to describe the order-of-magnitude observed, and not
 801 to predict the exact values of the abundances.

802 **Equation 11** is a very general relation for processes where the cell volume is the "natural
 803 variable" of the problem. This means that, as the cell increases in volume, the requirements for
 804 substance X also scale with volume rather than scaling with surface area, for example. So long as
 805 the rate of the process, the fraction of the dry mass attributable to the substance, and the formula
 806 mass of the substance is known, **Equation 11** can be used to compute the number of complexes
 807 needed. For example, to compute the number of ATP synthases per cell, **Equation 11** can be slightly
 808 modified to the form

$$N_{\text{ATP synthases}}(\lambda) = \frac{0.3 \times \rho V(\lambda) \theta_{\text{protein}} N_{\text{ATP}}}{w_{AA} r_{\text{ATP}} \tau}, \quad (12)$$

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

812 A Relation for Complex Abundance Scaling With Surface Area

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

821 Number of Lipids

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

$$N_{\text{lipids}}(\lambda) \approx \frac{4 \text{ leaflets} \times 0.4 \times \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{A_{\text{lipid}}} \quad (13)$$

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

828 Number of Murein Monomers

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

$$N_{\text{murein monomers}}(\lambda) \approx \frac{\rho_{\text{pg}} w \eta \pi \left(\frac{\eta \pi}{4} - \frac{\pi}{12} \right)^{-2/3} V(\lambda)^{2/3}}{m_{\text{murein}}}, \quad (14)$$

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

837 **Complex Abundance Scaling With Number of Origins**

838 While the majority of our estimates hinge on the total cell volume or surface area, processes related
 839 to the central dogma, namely DNA replication and synthesis of rRNA, depend on the number of
 840 chromosomes present in the cell. As discussed in the main text, the ability of *E. coli* to parallelize
 841 the replication of its chromosome by having multiple active origins of replication at a given is
 842 critical to synthesize enough rRNA, especially at fast growth rates. Derived in *Si et al.* (2017) and
 843 reproduced in the main text, the average number of origins of replication at a given growth rate can
 844 be calculated as

$$\langle \#ori \rangle \approx 2^{t_{\text{cyc}}/\ln 2} \quad (15)$$

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

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

$$D_{\text{rRNA}}(\lambda) \approx N_{\text{rRNA operons}} \times 2^{t_{\text{cyc}}\lambda/\ln 2}. \quad (16)$$

849 This makes the approximation that *all* rRNA operons are localized around the origin. In reality,
 850 the operons are some distance away from the origin, making **Equation 16** an approximation.

851 In the main text, we stated that at the growth rate in question, there is ≈ 1 chromosome per cell.
 852 While a fair approximation, **Equation 15** illustrates that is not precisely true, even at slow growth
 853 rates. In estimating the number of RNA polymerases as a function of growth rate, we consider that
 854 regardless of the number of rRNA operons, they are all sufficiently loaded with RNA polymerase
 855 such that each operon produces one rRNA per second. Thus, the total number of RNA polymerase
 856 as a function of the growth rate can be calculated as

$$N_{\text{RNA polymerase}}(\lambda) \approx L_{\text{operon}} D_{\text{rRNA}} \rho_{\text{RNA polymerase}}, \quad (17)$$

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

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

860 *E. coli* shows robust scaling of cell size with average the number of origins $\langle \#ori \rangle$ per cell (*Si et al.*,
 861 2017). Since protein makes up a majority of the cell's dry mass, the change in cell size is also a
 862 reflection of the changes in proteomic composition and total abundance across growth conditions.
 863 Given the potential constraints on rRNA synthesis with $\langle \#ori \rangle$ (limited by the effective number of
 864 rRNA gene copies), it is important to consider how protein copy numbers vary with the state of
 865 chromosomal replication. This is particularly true when trying to connect the dependence of growth
 866 rate on ribosomal fraction with particular mechanism the cell employs to vary ribosomal synthesis.
 867 As considered in the main text, it is becoming increasingly apparent that regulation through the
 868 secondary messengers (p)ppGpp acts to throttle back DNA replication and ribosomal activity in
 869 poorer nutrient conditions. In this context, both $\langle \#ori \rangle$, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio become
 870 important parameters to consider and keep tract of. An increase in $\langle \#ori \rangle / \langle \#ter \rangle$ ratio in particular,
 871 causes a relatively higher gene dosage in rRNA and r-protein genes due to skew in genes near the
 872 origin, where the majority of these are located

873 In the main text we estimated the change in $\langle \#ori \rangle$ with growth rate using the nutrient-limited
 874 wild-type cell data from *Si et al.* (2017). We consider their measurements of DNA replication time
 875 (t_C , 'C' period of cell division), total cell cycle time (t_{cyc} , 'C' + 'D' period of cell division), and doubling
 876 time τ from wild-type *E. coli* growing across a range of growth conditions. Here we show how we
 877 estimate this parameter, as well as the $\langle \#ori \rangle / \langle \#ter \rangle$ ratio from their data. We begin by considering
 878 $\langle \#ori \rangle$. If the cell cycle time takes longer than the time of cell division, the cell will need to initiate
 879 DNA replication more often than its rate of division, $2^{\lambda t} = 2^{\ln(2) \cdot t / \tau}$ to maintain steady-state growth.
 880 Cells will need to do this in proportion to the ratio $\lambda_{\text{cyc}} / \lambda = t_{\text{cyc}} / \tau$, and the number of origins per cell

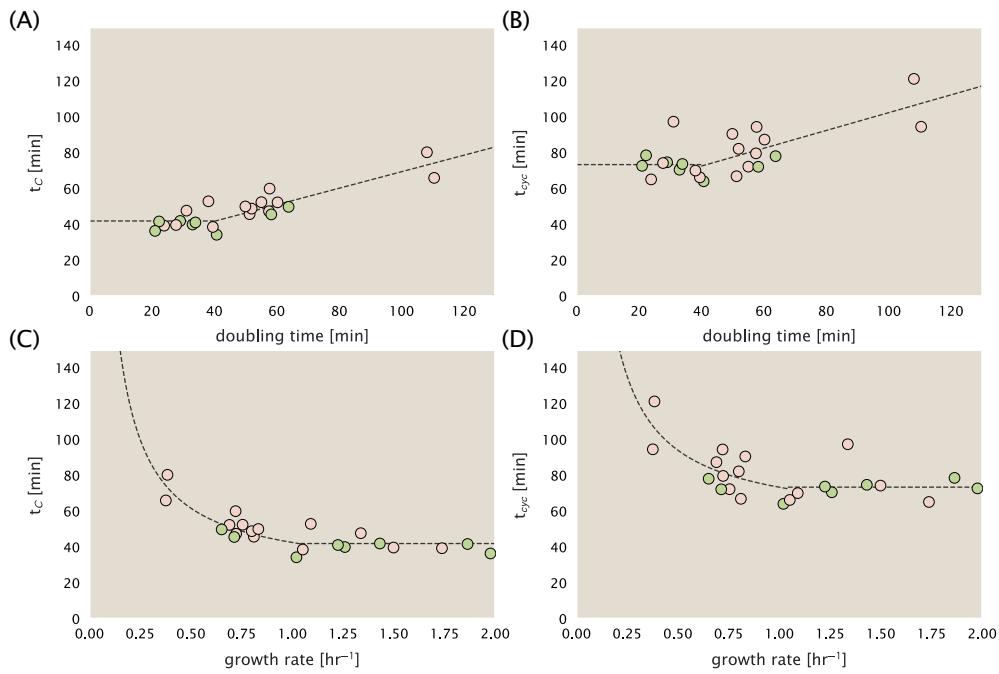


Figure 13. Estimation of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$ using data from Si et al. (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ , respectively. The dashed lines show a piecewise fit to the data. For short doubling times (rich media), t_C and t_{cyc} are assumed constant. 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. (C) and (D) plot the same data as in (A) and (B), but as a function of growth rate, given by $\lambda = \ln(2)/\tau$.

881 (on average) is then given by $2^{t_{cyc}/\tau}$. The average number of termini will in contrast depend on the
 882 lag time between DNA replication and cell division, t_D , with $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ ratio = $2^{t_{cyc}/\tau - t_D/\tau} = 2^{t_C/\tau}$.
 883 In Figure ??(A) and (B) we plot the measured t_C and t_{cyc} values versus the doubling time from
 884 **Si et al. (2017)**. The authors estimated t_C by marker frequency analysis using qPCR, while $t_{cyc} =$
 885 $t_C + t_D$ were inferred from t_C and τ . In the plots we see that both t_C and t_{cyc} reach a minimum at
 886 around 40 and 75 minutes, respectively. For a C period of 40 minutes, this would correspond to a
 887 maximum rate of elongation of about 1,000 bp/sec. Since we lacked a specific model to describe
 888 how each of these parameters vary with growth condition, we assumed that they were linearly
 889 dependent on the doubling time. For each parameter, t_C and t_{cyc} , we split them up into two domains
 890 corresponding to poorer nutrient conditions and rich nutrient conditions (cut off at $\tau \approx 40$ minutes
 891 where chromosomal replication becomes nearly constant). The fit lines are shown as solid black
 892 lines. In Figure ??(C) and (D) we also show t_C and t_{cyc} as a function of growth rate λ along with our
 893 piecewise linear fits, which match the plots in the main text.

894 Calculation of active ribosomal fraction.

895 In the main text we used the active ribosomal fraction f_a that was reported in the work of **Dai et al.**
 896 (2016) to estimate the active ribosomal mass fraction $\Phi_R \times f_a$ across growth conditions. We lacked
 897 any specific model to consider how f_a should vary with growth rate, and instead find that the data
 898 is well-approximated by fitting to an exponential curve ($f_a = -0.889 e^{4.6 \cdot \lambda} + 0.922$; dashed line in inset
 899 of ??(C)). We use this function to estimate f_a for each of the data points shown in ??(C).

900 **References**

- 901 **Aidellberg G**, Towbin BD, Rothschild D, Dekel E, Bren A, Alon U. Hierarchy of Non-Glucose Sugars in *Escherichia*
902 *coli*. BMC Systems Biology. 2014 Dec; 8(1):133. doi: 10.1186/s12918-014-0133-z.
- 903 **Antonenko YN**, Pohl P, Denisov GA. Permeation of Ammonia across Bilayer Lipid Membranes Studied by
904 Ammonium Ion Selective Microelectrodes. Biophysical Journal. 1997 May; 72(5):2187–2195.
- 905 **Assentoft M**, Kaptan S, Schneider HP, Deitmer JW, de Groot BL, MacAulay N. Aquaporin 4 as a NH₃ Channel.
906 Journal of Biological Chemistry. 2016 Sep; 291(36):19184–19195. doi: 10.1074/jbc.M116.740217.
- 907 **Baba T**, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction
908 of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular Systems
909 Biology. 2006 Jun; 2(1):2460.
- 910 **Basan M**, Zhu M, Dai X, Warren M, Sévin D, Wang YP, Hwa T. Inflating bacterial cells by increased protein
911 synthesis. Molecular Systems Biology. 2015 Oct; 11(10):836.
- 912 **Bauer S**, Ziv E. Dense Growth of Aerobic Bacteria in a Bench-Scale Fermentor. Biotechnology
913 and Bioengineering. 1976; 18(1):81–94. doi: 10.1002/bit.260180107, _eprint:
914 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 915 **Belliveau NM**, Barnes SL, Ireland WT, Jones DL, Sweredoski MJ, Moradian A, Hess S, Kinney JB, Phillips R. Systematic Approach for Dissecting the Molecular Mechanisms of Transcriptional Regulation in Bacteria. Proceedings
916 of the National Academy of Sciences. 2018 May; 115(21):E4796–E4805. doi: 10.1073/pnas.1722055115.
- 917 **Booth IR**, Mitchell WJ, Hamilton WA. Quantitative Analysis of Proton-Linked Transport Systems. The Lactose
918 Permease of *Escherichia coli*. Biochemical Journal. 1979 Sep; 182(3):687–696.
- 919 **Dai X**, Zhu M, Warren M, Balakrishnan R, Okano H, Williamson JR, Fredrick K, Hwa T. Slowdown of Translational
920 Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. mBio. 2018 Feb; 9(1):281.
- 921 **Dai X**, Zhu M, Warren M, Balakrishnan R, Patsalo V, Okano H, Williamson JR, Fredrick K, Wang YP, Hwa T. Reduction
922 of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. Nature
923 Microbiology. 2016 Dec; 2(2):16231.
- 924 **Datsenko KA**, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR
925 products. Proceedings of the National Academy of Sciences. 2000 Jun; 97(12):6640–6645.
- 926 **Dill KA**, Ghosh K, Schmit JD. Physical Limits of Cells and Proteomes. Proceedings of the National Academy of
927 Sciences. 2011 Nov; 108(44):17876–17882. doi: 10.1073/pnas.1114477108.
- 928 **Escalante A**, Salinas Cervantes A, Gosset G, Bolívar F. Current Knowledge of the *Escherichia coli* Phosphoenolpyruvate–Carbohydrate Phosphotransferase System: Peculiarities of Regulation and Impact on Growth and Product
929 Formation. Applied Microbiology and Biotechnology. 2012 Jun; 94(6):1483–1494. doi: 10.1007/s00253-
930 012-4101-5.
- 931 **Feist AM**, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BØ. A
932 Genome-Scale Metabolic Reconstruction for *Escherichia coli* K-12 MG1655 That Accounts for 1260 ORFs and
933 Thermodynamic Information. Molecular Systems Biology. 2007 Jan; 3(1):121. doi: 10.1038/msb4100155.
- 934 **Gallagher LA**, Bailey J, Manoil C. Ranking Essential Bacterial Processes by Speed of Mutant Death. Proceedings
935 of the National Academy of Sciences. 2020 Jul; doi: 10.1073/pnas.2001507117.
- 936 **Gama-Castro S**, Salgado H, Santos-Zavaleta A, Ledezma-Tejeida D, Muñiz-Rascado L, García-Sotelo JS, Alquicira-
937 Hernández K, Martínez-Flores I, Pannier L, Castro-Mondragón JA, Medina-Rivera A, Solano-Lira H, Bonavides-
938 Martínez C, Pérez-Rueda E, Alquicira-Hernández S, Porrón-Sotelo L, López-Fuentes A, Hernández-Koutoucheva
939 A, Moral-Chávez VD, Rinaldi F, et al. RegulonDB Version 9.0: High-Level Integration of Gene Regulation,
940 Coexpression, Motif Clustering and Beyond. Nucleic Acids Research. 2016 Jan; 44(D1):D133–D143. doi:
941 10.1093/nar/gkv1156.
- 942 **Guo Y**, Li D, Zhang S, Yang Y, Liu JJ, Wang X, Liu C, Milkie DE, Moore RP, Tulu US, Kiehart DP, Hu J, Lippincott-
943 Schwartz J, Betzig E, Li D. Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale
944 Resolution on Millisecond Timescales. Cell. 2018 Nov; 175(5):1430–1442.e17.
- 945 **Harris LK**, Theriot JA. Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. Trends in
946 microbiology. 2018 Oct; 26(10):815–832.

- 949 **Harris RM**, Webb DC, Howitt SM, Cox GB. Characterization of PitA and PitB from *Escherichia coli*. *Journal of*
 950 *Bacteriology*. 2001 Sep; 183(17):5008–5014. doi: [10.1128/JB.183.17.5008-5014.2001](https://doi.org/10.1128/JB.183.17.5008-5014.2001).
- 951 **van Heeswijk WC**, Westerhoff HV, Boogerd FC. Nitrogen Assimilation in *Escherichia coli*: Putting Molecular
 952 Data into a Systems Perspective. *Microbiology and Molecular Biology Reviews*. 2013 Dec; 77(4):628–695. doi:
 953 [10.1128/MMBR.00025-13](https://doi.org/10.1128/MMBR.00025-13).
- 954 **Heldal M**, Norland S, Tumyr O. X-Ray Microanalytic Method for Measurement of Dry Matter and Elemental
 955 Content of Individual Bacteria. *Applied and Environmental Microbiology*. 1985 Nov; 50(5):1251–1257.
- 956 **Hui S**, Silverman JM, Chen SS, Erickson DW, Basan M, Wang J, Hwa T, Williamson JR. Quantitative proteomic
 957 analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*. 2015
 958 Feb; 11(2):e784–e784.
- 959 **Ireland WT**, Beeler SM, Flores-Bautista E, Belliveau NM, Sweredoski MJ, Moradian A, Kinney JB, Phillips R.
 960 Deciphering the Regulatory Genome of *Escherichia coli*, One Hundred Promoters at a Time. *bioRxiv*. 2020 Jan;
 961 doi: [10.1101/2020.01.18.910323](https://doi.org/10.1101/2020.01.18.910323).
- 962 **Jacob F**, Monod J. Genetic Regulatory Mechanisms in the Synthesis of Proteins. *Journal of Molecular Biology*.
 963 1961 Jun; 3(3):318–356. doi: [10.1016/S0022-2836\(61\)80072-7](https://doi.org/10.1016/S0022-2836(61)80072-7).
- 964 **Jun S**, Si F, Pugatch R, Scott M. Fundamental Principles in Bacterial Physiology - History, Recent Progress, and
 965 the Future with Focus on Cell Size Control: A Review. *Reports on Progress in Physics*. 2018 May; 81(5):056601.
 966 doi: [10.1088/1361-6633/aaa628](https://doi.org/10.1088/1361-6633/aaa628).
- 967 **Karp PD**, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, Keseler IM, Krummenacker M, Midford PE,
 968 Ong Q, Ong WK, Paley SM, Subhraveti P. The BioCyc collection of microbial genomes and metabolic pathways.
 969 *Briefings in Bioinformatics*. 2019 Aug; 20(4):1085–1093.
- 970 **Karr JR**, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW. A
 971 Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*. 2012 Jul; 150(2):389–401. doi:
 972 [10.1016/j.cell.2012.05.044](https://doi.org/10.1016/j.cell.2012.05.044).
- 973 **Keseler IM**, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, Fulcher C, Gama-Castro
 974 S, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti
 975 P, Velázquez-Ramírez DA, Weaver D, Collado-Vides J, Paulsen I, et al. The EcoCyc database: reflecting new
 976 knowledge about *Escherichia coli*K-12. *Nucleic Acids Research*. 2017 Jan; 45(D1):D543–D550.
- 977 **Khademi S**, O'Connell J, Remis J, Robles-Colmenares Y, Miercke LJW, Stroud RM. Mechanism of Ammonia
 978 Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*. 2004 Sep; 305(5690):1587–1594. doi:
 979 [10.1126/science.1101952](https://doi.org/10.1126/science.1101952).
- 980 **Kostinski S**, Reuveni S. Ribosome Composition Maximizes Cellular Growth Rates in *E. Coli*. *Physical Review*
 981 *Letters*. 2020 Jul; 125(2):028103. doi: [10.1103/PhysRevLett.125.028103](https://doi.org/10.1103/PhysRevLett.125.028103).
- 982 **Laxhuber KS**, Morrison MJ, Chure G, Belliveau NM, Strandkvist C, Naughton KL, Phillips R. Theoretical investigation
 983 of a genetic switch for metabolic adaptation. *PLOS ONE*. 2020 May; 15(5):e0226453.
- 984 **Lex A**, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: visualization of intersecting sets. *IEEE Transactions*
 985 *on Visualization and Computer Graphics*. 2014; 20(12):1983–1992.
- 986 **Li GW**, Burkhardt D, Gross C, Weissman JS. Quantifying Absolute Protein Synthesis Rates Reveals Principles
 987 Underlying Allocation of Cellular Resources. *Cell*. 2014 Apr; 157(3):624–635. doi: [10.1016/j.cell.2014.02.033](https://doi.org/10.1016/j.cell.2014.02.033).
- 988 **Liebermeister W**, Noor E, Flamholz A, Davidi D, Bernhardt J, Milo R. Visual account of protein investment in
 989 cellular functions. *Proceedings of the National Academy of Sciences*. 2014 Jun; 111(23):8488–8493.
- 990 **Liu M**, Durfee T, Cabrera JE, Zhao K, Jin DJ, Blattner FR. Global Transcriptional Programs Reveal a Carbon Source
 991 Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*. 2005 Apr; 280(16):15921–15927. doi:
 992 [10.1074/jbc.M414050200](https://doi.org/10.1074/jbc.M414050200).
- 993 **Lu D**, Grayson P, Schulten K. Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol
 994 Facilitator GlpF. *Biophysical Journal*. 2003 Nov; 85(5):2977–2987. doi: [10.1016/S0006-3495\(03\)74718-3](https://doi.org/10.1016/S0006-3495(03)74718-3).
- 995 **Metzl-Raz E**, Kafri M, Yaakov G, Soifer I, Gurvich Y, Barkai N. Principles of cellular resource allocation revealed
 996 by condition-dependent proteome profiling. *eLife*. 2017; 6:e03528.

- 997 **Mikucki JA**, Pearson A, Johnston DT, Turchyn AV, Farquhar J, Schrag DP, Anbar AD, Priscu JC, Lee PA. A Contem-
998 porary Microbially Maintained Subglacial Ferrous "Ocean". *Science*. 2009 Apr; 324(5925):397–400.
- 999 **Milo R**, Jorgensen P, Moran U, Weber G, Springer M. BioNumbers—the Database of Key Numbers in Molecular
1000 and Cell Biology. *Nucleic Acids Research*. 2010 Jan; 38(suppl_1):D750–D753. doi: 10.1093/nar/gkp889.
- 1001 **Monod J**. The Phenomenon of Enzymatic Adaptation and Its Bearings on Problems of Genetics and Cellular
1002 Differentiation. *Growth Symposium*. 1947; 9:223–289.
- 1003 **Monod J**. The Growth of Bacterial Cultures. *Annual Review of Microbiology*. 1949 Oct; 3(1):371–394.
- 1004 **Morgenstein RM**, Bratton BP, Nguyen JP, Ouzounov N, Shaevitz JW, Gitai Z. RodZ Links MreB to Cell Wall
1005 Synthesis to Mediate MreB Rotation and Robust Morphogenesis. *Proceedings of the National Academy of
1006 Sciences*. 2015 Oct; 112(40):12510–12515. doi: 10.1073/pnas.1509610112.
- 1007 **Neidhardt FC**, Ingraham JL, Schaechter M. *Physiology of the Bacterial Cell - A Molecular Approach*, vol. 1.
1008 Elsevier; 1991.
- 1009 **Ojkic N**, Serbanescu D, Banerjee S. Surface-to-volume scaling and aspect ratio preservation in rod-shaped
1010 bacteria. *eLife*. 2019 Aug; 8:642.
- 1011 **Peebo K**, Valgepea K, Maser A, Nahku R, Adamberg K, Vilu R. Proteome Reallocation in *Escherichia coli* with
1012 Increasing Specific Growth Rate. *Molecular BioSystems*. 2015; 11(4):1184–1193. doi: 10.1039/C4MB00721B.
- 1013 **Phillips R**. Membranes by the Numbers. In: *Physics of Biological Membranes* Cham: Springer, Cham; 2018.p.
1014 73–105.
- 1015 **Ramos S**, Kaback HR. The Relation between the Electrochemical Proton Gradient and Active Transport in
1016 *Escherichia coli* Membrane Vesicles. *Biochemistry*. 1977 Mar; 16(5):854–859. doi: 10.1021/bi00624a007.
- 1017 **Rosenberg H**, Gerdes RG, Chegwidden K. Two Systems for the Uptake of Phosphate in *Escherichia coli*. *Journal
1018 of Bacteriology*. 1977 Aug; 131(2):505–511.
- 1019 **Schaechter M**, Maaløe O, Kjeldgaard NO. Dependency on Medium and Temperature of Cell Size and Chemical
1020 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*. 1958 Dec; 19(3):592–606.
- 1021 **Schmidt A**, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heine-
1022 mann M. The Quantitative and Condition-Dependent *Escherichia coli* Proteome. *Nature Biotechnology*. 2016
1023 Jan; 34(1):104–110. doi: 10.1038/nbt.3418.
- 1024 **Scott M**, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. Interdependence of cell growth and gene expression:
1025 origins and consequences. *Science*. 2010 Nov; 330(6007):1099–1102.
- 1026 **Sekowska A**, Kung HF, Danchin A. Sulfur Metabolism in *Escherichia coli* and Related Bacteria: Facts and Fiction.
1027 *Journal of Molecular Microbiology and Biotechnology*. 2000; 2(2):34.
- 1028 **Si F**, Le Treut G, Sauls JT, Vadia S, Levin PA, Jun S. Mechanistic Origin of Cell-Size Control and Homeostasis in
1029 Bacteria. *Current Biology*. 2019 Jun; 29(11):1760–1770.e7. doi: 10.1016/j.cub.2019.04.062.
- 1030 **Si F**, Li D, Cox SE, Sauls JT, Azizi O, Sou C, Schwartz AB, Erickstad MJ, Jun Y, Li X, Jun S. Invariance of Initiation Mass
1031 and Predictability of Cell Size in *Escherichia coli*. *Current Biology*. 2017 May; 27(9):1278–1287.
- 1032 **Sirko A**, Zatyka M, Sadowy E, Hulanicka D. Sulfate and Thiosulfate Transport in *Escherichia coli* K-12: Evidence
1033 for a Functional Overlapping of Sulfate- and Thiosulfate-Binding Proteins. *Journal of Bacteriology*. 1995 Jul;
1034 177(14):4134–4136. doi: 10.1128/jb.177.14.4134-4136.1995.
- 1035 **Soufi B**, Krug K, Harst A, Macek B. Characterization of the *E. coli* proteome and its modifications during growth
1036 and ethanol stress. *Frontiers in Microbiology*. 2015 Feb; 6:198.
- 1037 **Stasi R**, Neves HI, Spira B. Phosphate Uptake by the Phosphonate Transport System PhnCDE. *BMC Microbiology*.
1038 2019 Apr; 19. doi: 10.1186/s12866-019-1445-3.
- 1039 **Szenk M**, Dill KA, de Graff AMR. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
1040 Membrane Real Estate Hypothesis. *Cell Systems*. 2017 Aug; 5(2):95–104. doi: 10.1016/j.cels.2017.06.005.
- 1041 **Taheri-Araghi S**, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, Vergassola M, Jun S. Cell-size control and
1042 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*. 2015 Feb; 25(3):385–391.

- 1043 **Taniguchi Y**, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS. Quantifying *E. coli* proteome and
1044 transcriptome with single-molecule sensitivity in single cells. *Science (New York, NY)*. 2010 Jul; 329(5991):533–
1045 538.
- 1046 **Taymaz-Nikerel H**, Borujeni AE, Verheijen PJT, Heijnen JJ, van Gulik WM. Genome-Derived Minimal
1047 Metabolic Models for *Escherichia coli* MG1655 with Estimated in Vivo Respiratory ATP Stoichiometry.
1048 *Biotechnology and Bioengineering*. 2010; 107(2):369–381. doi: [10.1002/bit.22802](https://doi.org/10.1002/bit.22802), _eprint:
1049 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 1050 **Valgepea K**, Adamberg K, Seiman A, Vilu R. *Escherichia coli* Achieves Faster Growth by Increasing Catalytic and
1051 Translation Rates of Proteins. *Molecular BioSystems*. 2013; 9(9):2344. doi: [10.1039/c3mb70119k](https://doi.org/10.1039/c3mb70119k).
- 1052 **Virtanen P**, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser
1053 W, Bright J, van der Walt SJ, Brett M, Wilson J, Jarrod Millman K, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson
1054 E, Carey C, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*.
1055 2020; 17:261–272. doi: <https://doi.org/10.1038/s41592-019-0686-2>.
- 1056 **Volkmer B**, Heinemann M. Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate
1057 Data Conversion for Systems Biology Modeling. *PLOS ONE*. 2011 Jul; 6(7):e23126.
- 1058 **Willsky GR**, Bennett RL, Malamy MH. Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
1059 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*. 1973 Feb; 113(2):529–
1060 539.
- 1061 **You C**, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, Wang YP, Lenz P, Yan D, Hwa T. Coordination of bacterial
1062 proteome with metabolism by cyclic AMP signalling. *Nature*. 2013 Aug; 500(7462):301–306.
- 1063 **Zhang L**, Jiang W, Nan J, Almqvist J, Huang Y. The *Escherichia coli* CysZ Is a pH Dependent Sulfate Transporter That
1064 Can Be Inhibited by Sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2014 Jul; 1838(7):1809–1816.
1065 doi: [10.1016/j.bbamem.2014.03.003](https://doi.org/10.1016/j.bbamem.2014.03.003).
- 1066 **Zhang Z**, Aboulwafa M, Saier MH. Regulation of Crp Gene Expression by the Catabolite Repressor/Activator,
1067 Cra, in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*. 2014; 24(3):135–141. doi:
1068 [10.1159/000362722](https://doi.org/10.1159/000362722).