

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) 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.

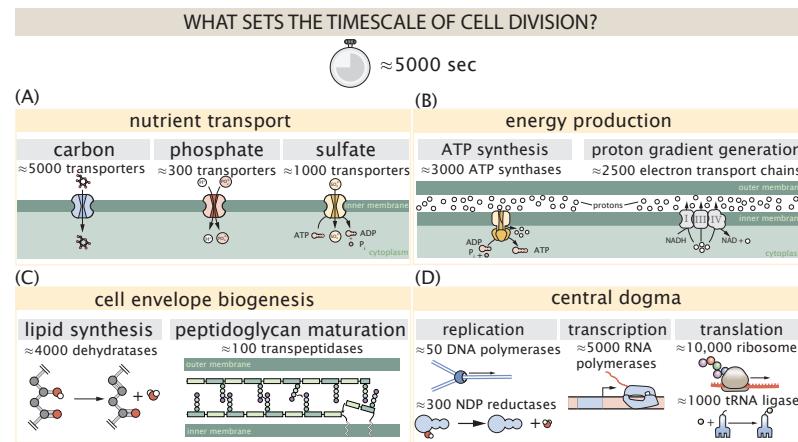


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are (A) nutrient transport across the cell membrane, (B) energy production (namely, ATP 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

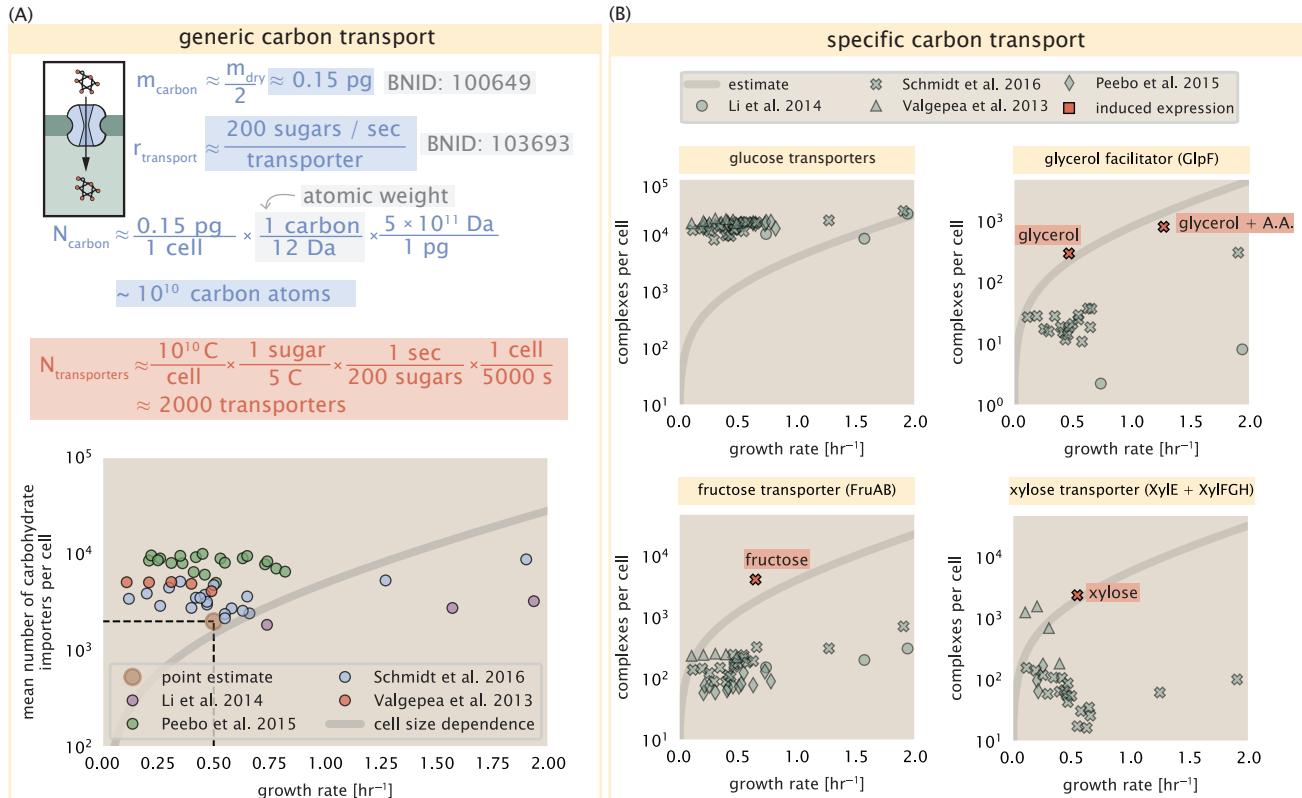


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\sim 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 5 C, and each transporter conducts sugar molecules 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.

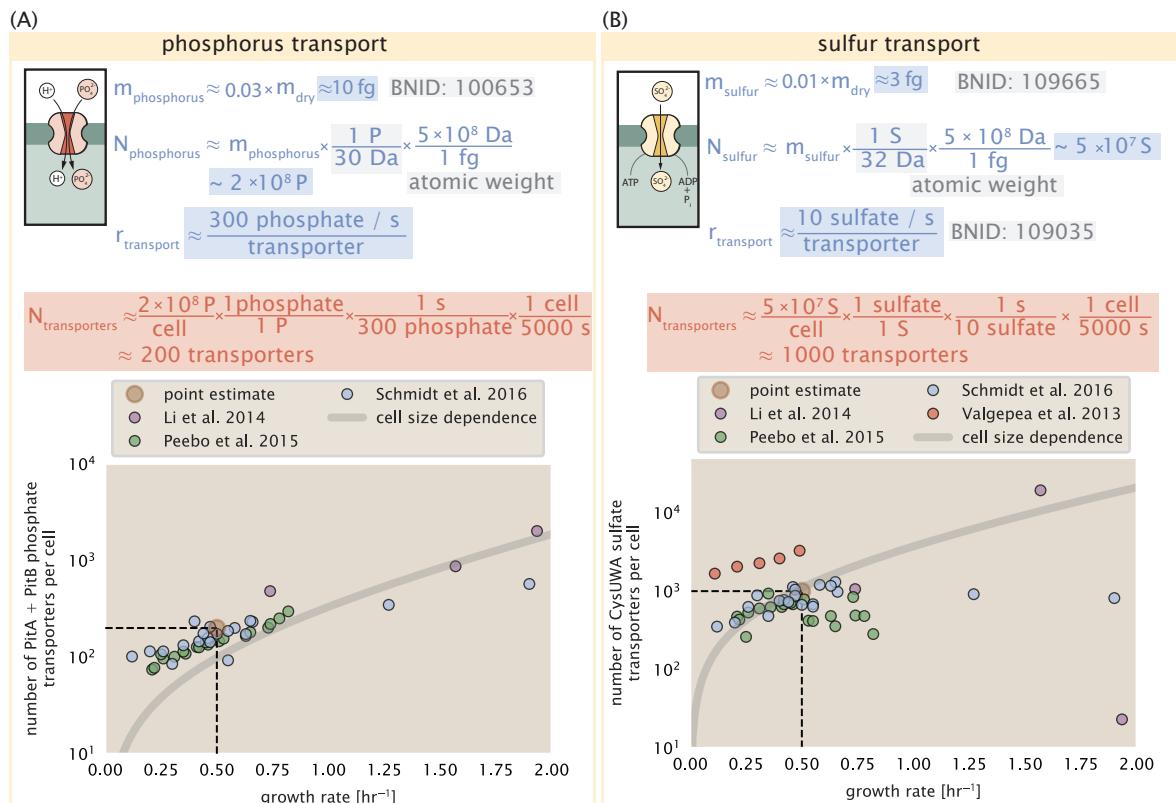


Figure 3. Estimates and measurements of phosphate and sulfate transport systems as a function of growth rate. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell.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

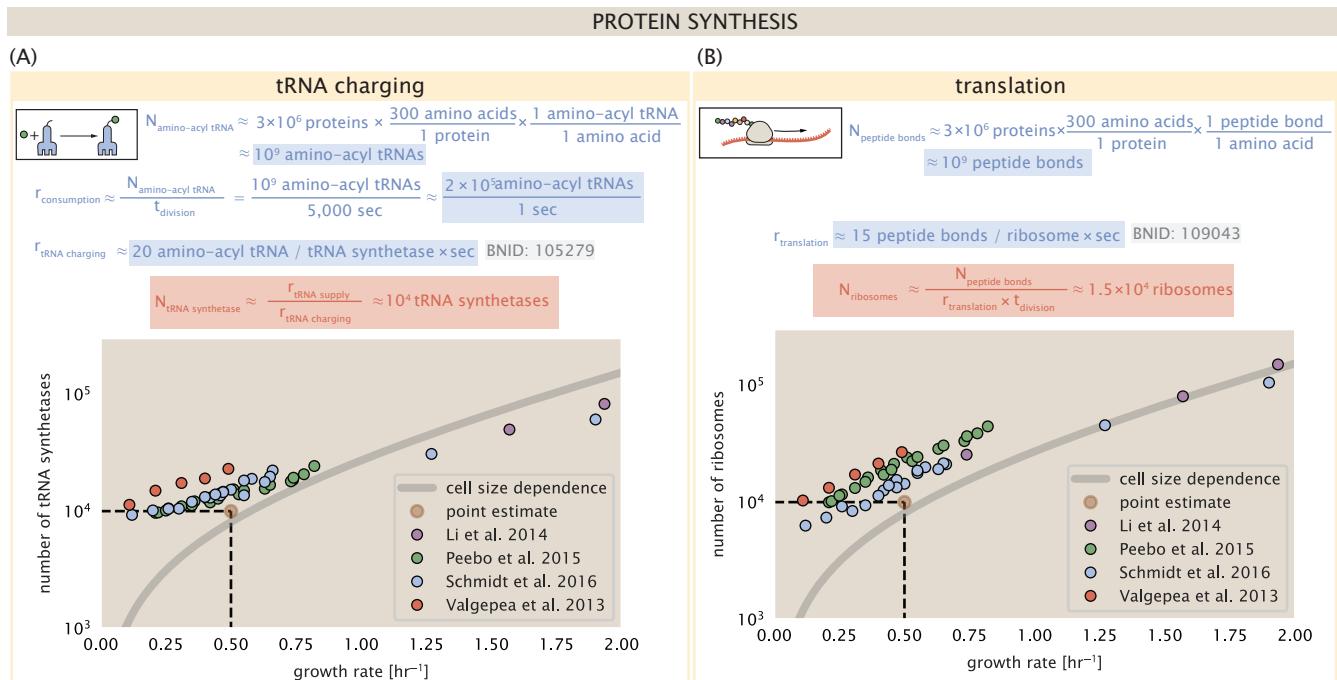


Figure 4. Estimation of the required tRNA synthetases and ribosomes. (A) Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [Glx], [IleS], [LeuS], [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.

342 Supplemental material for: 343 Fundamental limits on the rate of 344 bacterial cell division

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

*For correspondence:

[†]These authors contributed equally
to this work

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

357 **Contents**

358	Introduction	1
359	Uptake of Nutrients	2
360	Nitrogen Transport	4
361	Carbon Transport	5
362	Phosphorus and Sulfur Transport	7
363	Limits on Transporter Expression	9
364	Translation and Ribosomal Synthesis	9
365	tRNA Synthetases	9
366	Protein Synthesis	10
367	Summary of Proteome Data: Experimental Details	14
368	Fluorescence based measurements	14
369	Ribosomal profiling measurements	14
370	Mass spectrometry measurements	15
371	Summary of Proteomic Data	15
372	Estimation of Cell Size, Surface Area	16
373	Estimation of Total Protein Content per Cell	17
374	Additional Considerations of Schmidt <i>et al.</i> Data Set	18
375	Effect of cell volume on reported absolute protein abundances	20
376	Relaxing assumption of constant protein concentration across growth conditions	21
377	Experimental measurements of total protein from Basan <i>et al.</i> 2015.	22
378	Calculation of Complex Abundance	22
379	Extending Estimates to a Continuum of Growth Rates	23
380	Estimation of the total cell mass	25
381	Complex Abundance Scaling With Cell Volume	25
382	A Relation for Complex Abundance Scaling With Surface Area	26
383	Number of Lipids	26
384	Number of Murein Monomers	26
385	Complex Abundance Scaling With Number of Origins	27

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).

386 **Summary of Proteome Data: Experimental Details**

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

394 **Fluorescence based measurements**

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

403 **Ribosomal profiling measurements**

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

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

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

420 Mass spectrometry measurements

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

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

441 An alternative way to arrive at absolute protein abundances is to dope in synthetic peptide
 442 fragments of known abundance. These can serve as a direct way to calibrate mass spectrometry
 443 signal intensities to absolute mass. This is the approach taken by *Schmidt et al. (2016)*. In addition
 444 to a set of shotgun proteomic measurements to determine proteome-wide relative abundances,
 445 the authors also performed absolute quantification of 41 proteins covering over four orders of
 446 magnitude in cellular abundance. Here, a synthetic peptide was generated for each of the proteins,
 447 doped into each protein sample, and used these to determine absolute protein abundances of the
 448 41 proteins. These absolute measurements, determined for every growth condition, were then
 449 used as a calibration curve to convert proteomic-wide relative abundances into absolute protein
 450 abundance per cell. A more extensive discussion of the *Schmidt et al. (2016)* data set can be found
 451 in Section Additional Considerations of Schmidt et al. Data Set.

452 Summary of Proteomic Data

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

459 *Figure 5(B)* summarizes the total protein mass for each data point in our final compiled data
 460 set. We note that protein abundances were all scaled so they followed a common growth rate-
 461 dependent change in total protein mass. While our inclination initially was to leave reported copy
 462 numbers untouched, a notable discrepancy in the scaling total protein per cell between *Schmidt*
 463 *et al. (2016)* and the other data sets forced us to dig deeper into those measurements (compare
 464 *Schmidt et al. (2016)* and *Li et al. (2014)* data in *Figure 5(A)*). The particular trend in *Schmidt et al.*
 465 *(2016)* appears to be due to assumptions of cell size and we provide a more extensive discussion

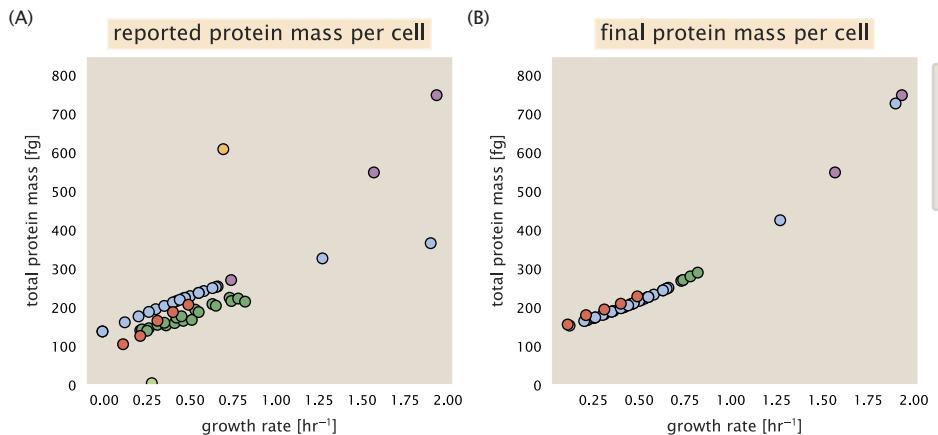


Figure 5. Summary of the growth-rate dependent total protein abundance for each data set. (A) Total protein abundance per cell as originally reported in the data sets of [Taniguchi et al. \(2010\)](#); [Valgepea et al. \(2013\)](#); [Li et al. \(2014\)](#); [Soufi et al. \(2015\)](#); [Peebo et al. \(2015\)](#); [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).

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

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

484 Estimation of Cell Size, Surface Area

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

491 The average size measurements from each of their experiments are shown in Figure [Figure 7](#),
 492 with cell length and width shown in (A) and (B), respectively. The length data was well described by

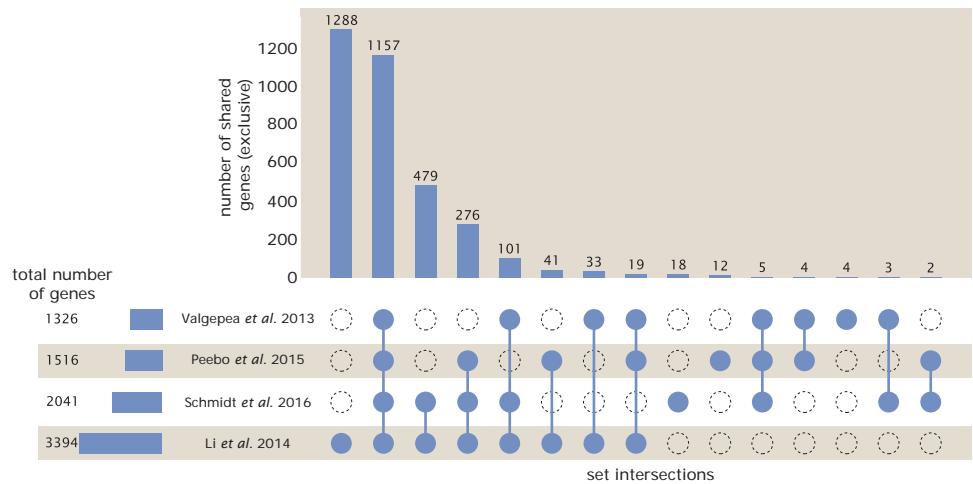


Figure 6. Comparison of proteomic coverage across different data sets. An UpSet diagram (Lex et al., 2014) summarizes the total number of protein coding genes whose protein abundance was reported in the data sets of Valgepea et al. (2013); Li et al. (2014); Schmidt et al. (2016); Peebo et al. (2015). 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.

493 the exponential function $0.5 e^{1.09 \cdot \lambda} + 1.76 \mu\text{m}$, while the width data was well described by $0.64 e^{0.24 \cdot \lambda}$
 494 μm . In order to estimate cell size we take the cell as a cylinders with two hemispherical ends (Si
 495 et al., 2017; Basan et al., 2015). Specifically, cell size is estimated from,

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

496 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
 497 the cell surface area is given by,

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

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

499 Estimation of Total Protein Content per Cell

500 In order to estimate total protein per cell for a particular growth rate, we begin by estimating the cell
 501 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
 502 content from the total dry mass of the cell. Here we begin by noting that for almost the entire range
 503 of growth rates considered here, protein, DNA, and RNA were reported to account for at least 90 %
 504 of the dry mass (Basan et al. (2015)). The authors also found that the total dry mass concentration
 505 was roughly constant across growth conditions. Under such a scenario, we can calculate the total
 506 dry mass concentration for protein, DNA, and RNA, which is given by $1.1 \text{ g/ml} \times 30 \% \times 90 \% \text{ or about}$
 507 $[M_p] = 300 \text{ fg per fl}$. Multiplying this by our prediction of cell size gives the total dry mass per cell.

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

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

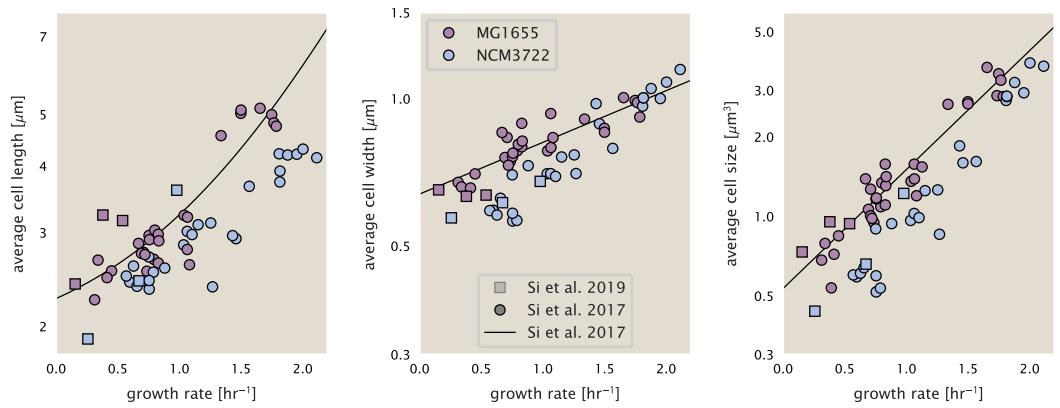


Figure 7. Summary of size measurements from Si et al. 2017, 2019. Cell lengths and widths were measured from cell contours obtained from phase contrast images, and refer to the long and short axis respectively. (A) Cell lengths and (B) cell widths show the mean measurements reported (they report 140-300 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 $0.533 e^{1.037 \cdot \lambda} \mu\text{m}^3$.

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

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

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

524 Additional Considerations of Schmidt et al. Data Set

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

531 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
 532 from Taheri-Araghi et al. and Si et al. come from the lab of Suckjoon Jun, while those from Basan
 533 et al. come from the lab of Terence Hwa. Each set of measurements used microscopy and cell
 534 segmentation to determine the length and width, and then calculated cell size by treating the cell
 535 is a cylinder with two hemispherical ends, as we considered in the previous section. While there
 536 is notable discrepancy between the two research groups, which are both using strain NCM3722,
 537 Basan et al. found that this came specifically from uncertainty in determining the cell width. This is
 538 prone to inaccuracy given the small cell size and optical resolution limits (further described in their
 539 supplemental text). Perhaps the more concerning point is that while each of these alternative mea-
 540 surements show an exponential increase in cell size at faster growth rates, the measurements used
 541 by Schmidt et al. appear to plateau. This resulted in an analogous trend in their final reported total

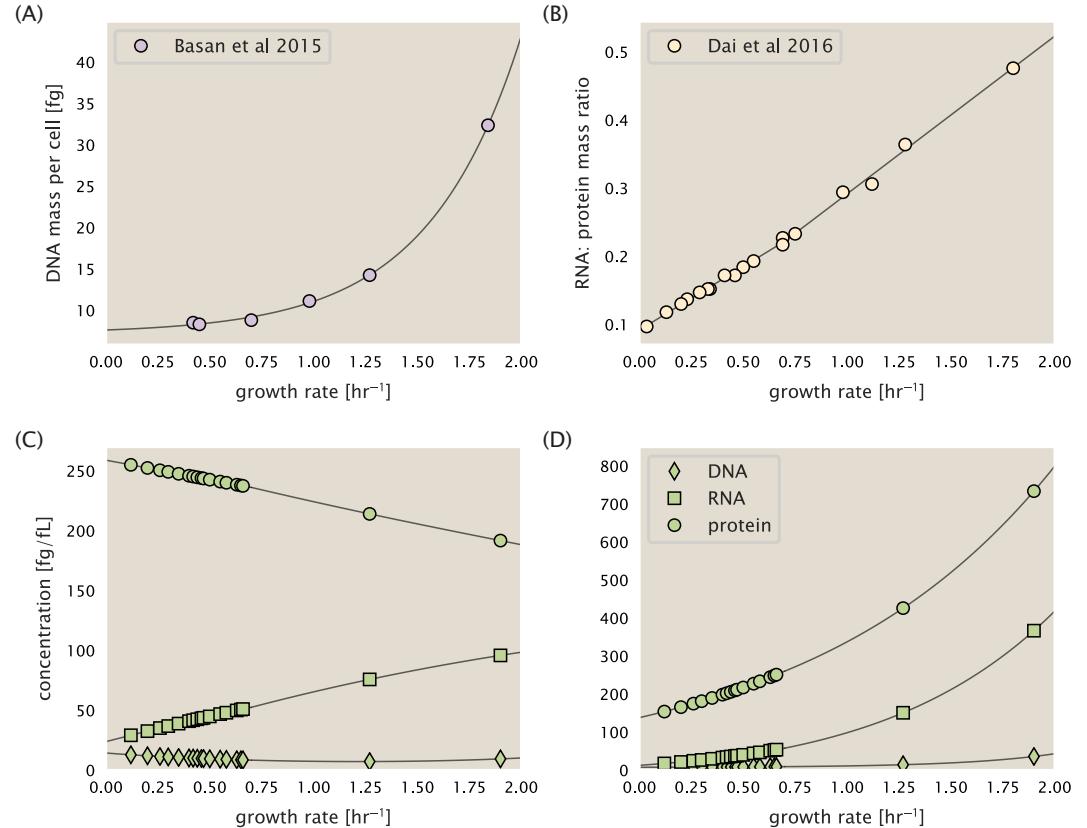


Figure 8. Empirical estimate of cellular protein, DNA, and RNA as a function of growth rate. (A) Measured DNA mass per cell as a function of growth rate, reproduced from Basan *et al.* 2015. The data was fit to an exponential curve (DNA mass in fg per cell is given by $0.42 e^{2.23\lambda} + 7.2$ fg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

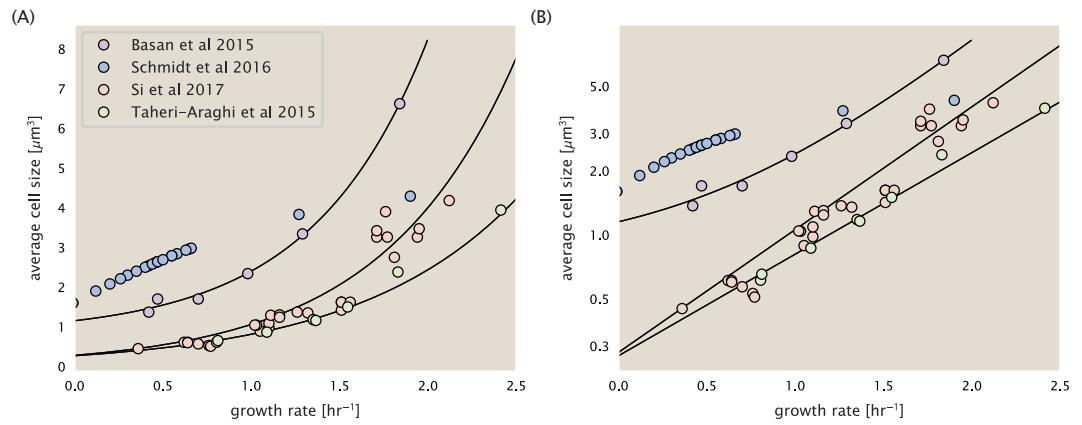


Figure 9. Measurements of cell size as a function of growth rate. (A) Plot of the reported cell sizes from several recent papers. The data in blue come from Volkmer and Heinemann, 2011 (*Volkmer and Heinemann (2011)*) and were used in the work of Schmidt *et al.*. Data from the lab of Terence Hwa are 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*.

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

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

555 Effect of cell volume on reported absolute protein abundances

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

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

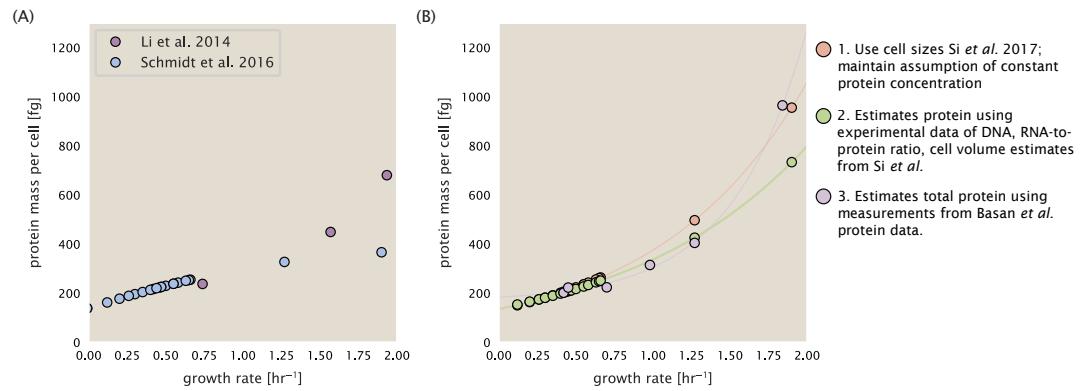


Figure 10. Alternative estimates of total cellular protein for the growth conditions considered in Schmidt et al. (A) The original protein mass from Schmidt et al. and Li et al. are shown in purple and blue, respectively. (B) Three alternative estimates of total protein per cell. 1. *light red*: Rescaling 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)**.

594 Relaxing assumption of constant protein concentration across growth conditions

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

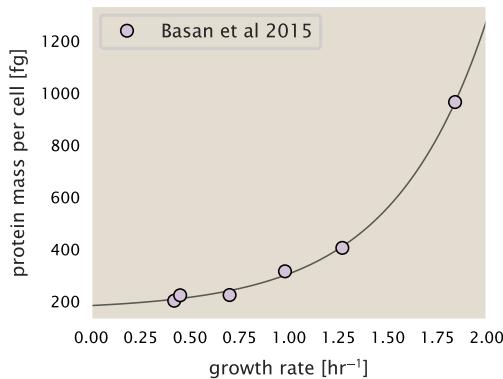


Figure 11. Total cellular protein reported in Basan *et al.* 2015. Measured protein mass as a function of growth rate as reproduced from Basan *et al.* 2015, with cells grown under different levels of nutrient limitation. The data was fit to an exponential curve where protein mass in fg per cell is given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

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

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

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

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

621 **Calculation of Complex Abundance**

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

628 abundance.

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

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

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

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

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

659 It is important to note that measurement noise, efficiency of protein extraction, stochastic errors
 660 will mean that the precise value of each calculation will be different for each component of a given
 661 complex. Thus, to report the total complex abundance, we computed the arithmetic mean of
 662 $N_{\text{complex}}^{(\max)}$ for all subunits as

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

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

667 Extending Estimates to a Continuum of Growth Rates

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

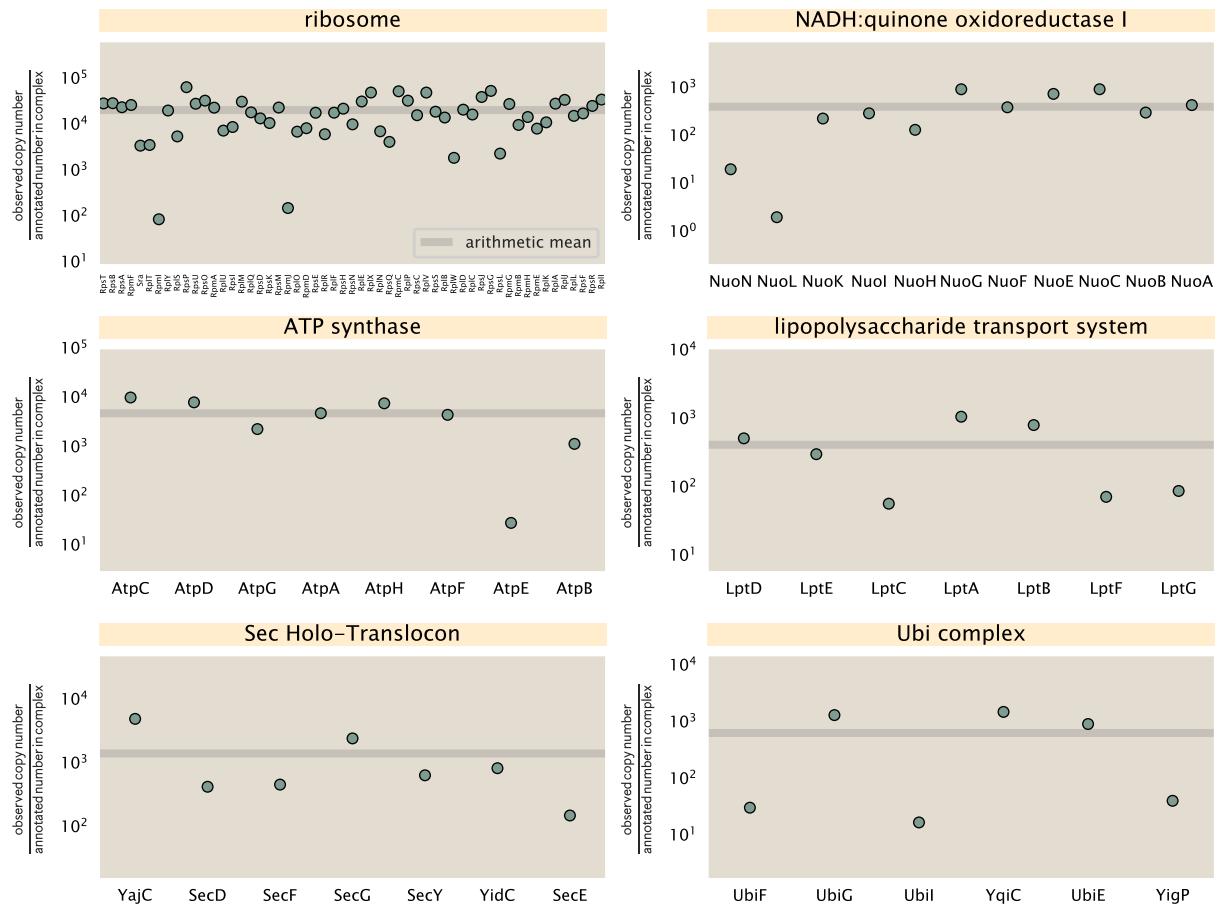


Figure 12. Calculation of the mean complex abundance from measurements of single subunits. Six of the largest complexes (by number of subunits) in *E. coli*. Points correspond to the maximum number of complexes that can be formed given measurement of that individual protein. Solid grey line corresponds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

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

673 **Estimation of the total cell mass**

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

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

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

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

687 **Complex Abundance Scaling With Cell Volume**

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

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

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

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

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

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

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

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

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

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

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

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

726 A Relation for Complex Abundance Scaling With Surface Area

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

735 Number of Lipids

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

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

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

742 Number of Murein Monomers

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

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

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

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

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

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

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

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

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

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

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

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

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

773 References

- 774 **Aidellberg G**, Towbin BD, Rothschild D, Dekel E, Bren A, Alon U. Hierarchy of Non-Glucose Sugars in *Escherichia*
775 *coli*. BMC Systems Biology. 2014 Dec; 8(1):133. doi: 10.1186/s12918-014-0133-z.
- 776 **Antonenko YN**, Pohl P, Denisov GA. Permeation of Ammonia across Bilayer Lipid Membranes Studied by
777 Ammonium Ion Selective Microelectrodes. Biophysical Journal. 1997 May; 72(5):2187-2195.
- 778 **Assentoft M**, Kaptan S, Schneider HP, Deitmer JW, de Groot BL, MacAulay N. Aquaporin 4 as a NH₃ Channel.
779 Journal of Biological Chemistry. 2016 Sep; 291(36):19184-19195. doi: 10.1074/jbc.M116.740217.
- 780 **Baba T**, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction
781 of *Escherichia coli*K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular Systems Biology.
782 2006 Jun; 2(1):2460.
- 783 **Basan M**, Zhu M, Dai X, Warren M, Sévin D, Wang YP, Hwa T. Inflating bacterial cells by increased protein
784 synthesis. Molecular Systems Biology. 2015 Oct; 11(10):836.
- 785 **Bauer S**, Ziv E. Dense Growth of Aerobic Bacteria in a Bench-Scale Fermentor. Biotechnology and Bioengineering.
786 1976; 18(1):81-94. doi: 10.1002/bit.260180107, _eprint:
787 https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107.
- 788 **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
789 of the National Academy of Sciences. 2018 May; 115(21):E4796-E4805. doi: 10.1073/pnas.1722055115.
- 790 **Booth IR**, Mitchell WJ, Hamilton WA. Quantitative Analysis of Proton-Linked Transport Systems. The Lactose
791 Permease of *Escherichia coli*. Biochemical Journal. 1979 Sep; 182(3):687-696.
- 792 **Dai X**, Zhu M, Warren M, Balakrishnan R, Okano H, Williamson JR, Fredrick K, Hwa T. Slowdown of Translational
793 Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. mBio. 2018 Feb; 9(1):281.
- 794 **Dai X**, Zhu M, Warren M, Balakrishnan R, Patsalo V, Okano H, Williamson JR, Fredrick K, Wang YP, Hwa T. Reduction
795 of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. Nature
796 Microbiology. 2016 Dec; 2(2):16231.
- 797 **Datsenko KA**, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR
798 products. Proceedings of the National Academy of Sciences. 2000 Jun; 97(12):6640-6645.
- 799 **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 Formation. Applied Microbiology and Biotechnology. 2012 Jun; 94(6):1483-1494. doi: 10.1007/s00253-012-4101-5.
- 800 **Feist AM**, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BØ. A
801 Genome-Scale Metabolic Reconstruction for *Escherichia coli* K-12 MG1655 That Accounts for 1260 ORFs and
802 Thermodynamic Information. Molecular Systems Biology. 2007 Jan; 3(1):121. doi: 10.1038/msb4100155.
- 803 **Gama-Castro S**, Salgado H, Santos-Zavaleta A, Ledezma-Tejeida D, Muñiz-Rascado L, García-Sotelo JS, Alquicira-Hernández K, Martínez-Flores I, Pannier L, Castro-Mondragón JA, Medina-Rivera A, Solano-Lira H, Bonavides-Martínez C, Pérez-Rueda E, Alquicira-Hernández S, Porrón-Sotelo L, López-Fuentes A, Hernández-Koutoucheva A, Moral-Chávez VD, Rinaldi F, et al. RegulonDB Version 9.0: High-Level Integration of Gene Regulation, Coexpression, Motif Clustering and Beyond. Nucleic Acids Research. 2016 Jan; 44(D1):D133-D143. doi: 10.1093/nar/gkv1156.
- 804 **Harris RM**, Webb DC, Howitt SM, Cox GB. Characterization of PitA and PitB from *Escherichia coli*. Journal of
805 Bacteriology. 2001 Sep; 183(17):5008-5014. doi: 10.1128/JB.183.17.5008-5014.2001.
- 806 **van Heeswijk WC**, Westerhoff HV, Boogerd FC. Nitrogen Assimilation in *Escherichia coli*: Putting Molecular
807 Data into a Systems Perspective. Microbiology and Molecular Biology Reviews. 2013 Dec; 77(4):628-695. doi:
808 10.1128/MMBR.00025-13.
- 809 **Heldal M**, Norland S, Tumyr O. X-Ray Microanalytic Method for Measurement of Dry Matter and Elemental
810 Content of Individual Bacteria. Applied and Environmental Microbiology. 1985 Nov; 50(5):1251-1257.
- 811 **Ireland WT**, Beeler SM, Flores-Bautista E, Belliveau NM, Sweredoski MJ, Moradian A, Kinney JB, Phillips R.
812 Deciphering the Regulatory Genome of *Escherichia coli*, One Hundred Promoters at a Time. bioRxiv. 2020 Jan;
813 doi: 10.1101/2020.01.18.910323.

- 823 **Jacob F**, Monod J. Genetic Regulatory Mechanisms in the Synthesis of Proteins. *Journal of Molecular Biology*.
 824 1961 Jun; 3(3):318–356. doi: 10.1016/S0022-2836(61)80072-7.
- 825 **Jun S**, Si F, Pugatch R, Scott M. Fundamental Principles in Bacterial Physiology - History, Recent Progress, and
 826 the Future with Focus on Cell Size Control: A Review. *Reports on Progress in Physics*. 2018 May; 81(5):056601.
 827 doi: 10.1088/1361-6633/aaa628.
- 828 **Karp PD**, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, Keseler IM, Krummenacker M, Midford PE,
 829 Ong Q, Ong WK, Paley SM, Subhraveti P. The BioCyc collection of microbial genomes and metabolic pathways.
 830 *Briefings in Bioinformatics*. 2019 Aug; 20(4):1085–1093.
- 831 **Keseler IM**, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, Fulcher C, Gama-Castro
 832 S, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti
 833 P, Velázquez-Ramírez DA, Weaver D, Collado-Vides J, Paulsen I, et al. The EcoCyc database: reflecting new
 834 knowledge about *Escherichia coli*K-12. *Nucleic Acids Research*. 2017 Jan; 45(D1):D543–D550.
- 835 **Khademi S**, O'Connell J, Remis J, Robles-Colmenares Y, Miercke LJW, Stroud RM. Mechanism of Ammonia
 836 Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*. 2004 Sep; 305(5690):1587–1594. doi:
 837 10.1126/science.1101952.
- 838 **Lex A**, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: visualization of intersecting sets. *IEEE Transactions
 839 on Visualization and Computer Graphics*. 2014; 20(12):1983–1992.
- 840 **Li GW**, Burkhardt D, Gross C, Weissman JS. Quantifying Absolute Protein Synthesis Rates Reveals Principles
 841 Underlying Allocation of Cellular Resources. *Cell*. 2014 Apr; 157(3):624–635. doi: 10.1016/j.cell.2014.02.033.
- 842 **Liu M**, Durfee T, Cabrera JE, Zhao K, Jin DJ, Blattner FR. Global Transcriptional Programs Reveal a Carbon Source
 843 Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*. 2005 Apr; 280(16):15921–15927. doi:
 844 10.1074/jbc.M414050200.
- 845 **Lu D**, Grayson P, Schulten K. Glycerol Conductance and Physical Asymmetry of the *Escherichia coli* Glycerol
 846 Facilitator GlpF. *Biophysical Journal*. 2003 Nov; 85(5):2977–2987. doi: 10.1016/S0006-3495(03)74718-3.
- 847 **Mikucki JA**, Pearson A, Johnston DT, Turchyn AV, Farquhar J, Schrag DP, Anbar AD, Priscu JC, Lee PA. A Contem-
 848 porary Microbially Maintained Subglacial Ferrous "Ocean". *Science*. 2009 Apr; 324(5925):397–400.
- 849 **Milo R**, Jorgensen P, Moran U, Weber G, Springer M. BioNumbers—the Database of Key Numbers in Molecular
 850 and Cell Biology. *Nucleic Acids Research*. 2010 Jan; 38(suppl_1):D750–D753. doi: 10.1093/nar/gkp889.
- 851 **Monod J**. The Phenomenon of Enzymatic Adaptation and Its Bearings on Problems of Genetics and Cellular
 852 Differentiation. *Growth Symposium*. 1947; 9:223–289.
- 853 **Monod J**. The Growth of Bacterial Cultures. *Annual Review of Microbiology*. 1949 Oct; 3(1):371–394.
- 854 **Neidhardt FC**, Ingraham JL, Schaechter M. Physiology of the Bacterial Cell - A Molecular Approach, vol. 1.
 855 Elsevier; 1991.
- 856 **Ojkic N**, Serbanescu D, Banerjee S. Surface-to-volume scaling and aspect ratio preservation in rod-shaped
 857 bacteria. *eLife*. 2019 Aug; 8:642.
- 858 **Peebo K**, Valgepea K, Maser A, Nahku R, Adamberg K, Vilu R. Proteome Reallocation in *Escherichia coli* with
 859 Increasing Specific Growth Rate. *Molecular BioSystems*. 2015; 11(4):1184–1193. doi: 10.1039/C4MB00721B.
- 860 **Phillips R**. Membranes by the Numbers. In: *Physics of Biological Membranes* Cham: Springer, Cham; 2018.p.
 861 73–105.
- 862 **Ramos S**, Kaback HR. The Relation between the Electrochemical Proton Gradient and Active Transport in
 863 *Escherichia coli* Membrane Vesicles. *Biochemistry*. 1977 Mar; 16(5):854–859. doi: 10.1021/bi00624a007.
- 864 **Rosenberg H**, Gerdes RG, Chegwidden K. Two Systems for the Uptake of Phosphate in *Escherichia coli*. *Journal
 865 of Bacteriology*. 1977 Aug; 131(2):505–511.
- 866 **Schaechter M**, Maaløe O, Kjeldgaard NO. Dependency on Medium and Temperature of Cell Size and Chemical
 867 Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*. 1958 Dec; 19(3):592–606.
- 868 **Schmidt A**, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heine-
 869 mann M. The Quantitative and Condition-Dependent *Escherichia coli* Proteome. *Nature Biotechnology*. 2016
 870 Jan; 34(1):104–110. doi: 10.1038/nbt.3418.

- 871 **Scott M**, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. Interdependence of cell growth and gene expression:
872 origins and consequences. *Science*. 2010 Nov; 330(6007):1099–1102.
- 873 **Sekowska A**, Kung HF, Danchin A. Sulfur Metabolism in *Escherichia coli* and Related Bacteria: Facts and Fiction.
874 *Journal of Molecular Microbiology and Biotechnology*. 2000; 2(2):34.
- 875 **Si F**, Le Treut G, Sauls JT, Vadia S, Levin PA, Jun S. Mechanistic Origin of Cell-Size Control and Homeostasis in
876 Bacteria. *Current Biology*. 2019 Jun; 29(11):1760–1770.e7. doi: [10.1016/j.cub.2019.04.062](https://doi.org/10.1016/j.cub.2019.04.062).
- 877 **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
878 and Predictability of Cell Size in *Escherichia coli*. *Current Biology*. 2017 May; 27(9):1278–1287.
- 879 **Sirko A**, Zatyka M, Sadowy E, Hulanicka D. Sulfate and Thiosulfate Transport in *Escherichia coli* K-12: Evidence
880 for a Functional Overlapping of Sulfate- and Thiosulfate-Binding Proteins. *Journal of Bacteriology*. 1995 Jul;
881 177(14):4134–4136. doi: [10.1128/jb.177.14.4134-4136.1995](https://doi.org/10.1128/jb.177.14.4134-4136.1995).
- 882 **Soufi B**, Krug K, Harst A, Macek B. Characterization of the *E. coli* proteome and its modifications during growth
883 and ethanol stress. *Frontiers in Microbiology*. 2015 Feb; 6:198.
- 884 **Stasi R**, Neves HI, Spira B. Phosphate Uptake by the Phosphonate Transport System PhnCDE. *BMC Microbiology*.
885 2019 Apr; 19. doi: [10.1186/s12866-019-1445-3](https://doi.org/10.1186/s12866-019-1445-3).
- 886 **Szenk M**, Dill KA, de Graff AMR. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the
887 Membrane Real Estate Hypothesis. *Cell Systems*. 2017 Aug; 5(2):95–104. doi: [10.1016/j.cels.2017.06.005](https://doi.org/10.1016/j.cels.2017.06.005).
- 888 **Taheri-Araghi S**, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, Vergassola M, Jun S. Cell-size control and
889 homeostasis in bacteria. - PubMed - NCBI. *Current Biology*. 2015 Feb; 25(3):385–391.
- 890 **Taniguchi Y**, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS. Quantifying *E. coli* proteome and
891 transcriptome with single-molecule sensitivity in single cells. *Science (New York, NY)*. 2010 Jul; 329(5991):533–
892 538.
- 893 **Taymaz-Nikerel H**, Borujeni AE, Verheijen PJT, Heijnen JJ, van Gulik WM. Genome-Derived Minimal
894 Metabolic Models for *Escherichia coli* MG1655 with Estimated in Vivo Respiratory ATP Stoichiometry.
895 *Biotechnology and Bioengineering*. 2010; 107(2):369–381. doi: [10.1002/bit.22802](https://doi.org/10.1002/bit.22802), _eprint:
896 <https://onlinelibrary.wiley.com/doi/10.1002/bit.22802>.
- 897 **Valgepea K**, Adamberg K, Seiman A, Vilu R. *Escherichia coli* Achieves Faster Growth by Increasing Catalytic and
898 Translation Rates of Proteins. *Molecular BioSystems*. 2013; 9(9):2344. doi: [10.1039/c3mb70119k](https://doi.org/10.1039/c3mb70119k).
- 899 **Virtanen P**, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser
900 W, Bright J, van der Walt SJ, Brett M, Wilson J, Jarrod Millman K, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson
901 E, Carey C, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*.
902 2020; 17:261–272. doi: <https://doi.org/10.1038/s41592-019-0686-2>.
- 903 **Volkmer B**, Heinemann M. Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate
904 Data Conversion for Systems Biology Modeling. *PLOS ONE*. 2011 Jul; 6(7):e23126.
- 905 **Willsky GR**, Bennett RL, Malamy MH. Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two
906 Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*. 1973 Feb; 113(2):529–
907 539.
- 908 **You C**, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, Wang YP, Lenz P, Yan D, Hwa T. Coordination of bacterial
909 proteome with metabolism by cyclic AMP signalling. *Nature*. 2013 Aug; 500(7462):301–306.
- 910 **Zhang L**, Jiang W, Nan J, Almqvist J, Huang Y. The *Escherichia coli* CysZ Is a pH Dependent Sulfate Transporter That
911 Can Be Inhibited by Sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2014 Jul; 1838(7):1809–1816.
912 doi: [10.1016/j.bbamem.2014.03.003](https://doi.org/10.1016/j.bbamem.2014.03.003).
- 913 **Zhang Z**, Aboulwafa M, Saier MH. Regulation of Crp Gene Expression by the Catabolite Repressor/Activator,
914 Cra, in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*. 2014; 24(3):135–141. doi:
915 [10.1159/000362722](https://doi.org/10.1159/000362722).