

Fundamental limits on the rate of bacterial cell division

³ **Nathan M. Belliveau^{†, 1}, Griffin Chure^{†, 2, 3}, Christina L. Hueschen⁴, Hernan G.
⁴ Garcia⁵, Jané Kondev⁶, Daniel S. Fisher⁷, Julie Theriot^{1, 8}, Rob Phillips^{2, 9, *}**

*For correspondence:

[†]These authors contributed equally to this work

⁵ ¹Department of Biology, University of Washington, Seattle, WA, USA; ²Division of
⁶ Biology and Biological Engineering, California Institute of Technology, Pasadena, CA,
⁷ USA; ³Department of Applied Physics, California Institute of Technology, Pasadena, CA,
⁸ USA; ⁴Department of Chemical Engineering, Stanford University, Stanford, CA, USA;
⁹ ⁵Department of Molecular Cell Biology and Department of Physics, University of
¹⁰ California Berkeley, Berkeley, CA, USA; ⁶Department of Physics, Brandeis University,
¹¹ Waltham, MA, USA; ⁷Department of Applied Physics, Stanford University, Stanford, CA,
¹² USA; ⁸Allen Institute for Cell Science, Seattle, WA, USA; ⁹Department of Physics,
¹³ California Institute of Technology, Pasadena, CA, USA; *Contributed equally

¹⁴

¹⁵ **Abstract** This will be written next

¹⁶

¹⁷ Introduction

¹⁸ The range of bacterial growth rates is enormously diverse. In natural environments, some micro-
¹⁹ bial organisms might double only once per year while in comfortable laboratory conditions, growth
²⁰ can be rapid with several divisions per hour. This six order of magnitude difference illustrates the
²¹ intimate relationship between environmental conditions and the rates at which cells convert nu-
²² trients into new cellular material – a relationship that has remained a major topic of inquiry in
²³ bacterial physiology for over a century (*Jun et al., 2018*). As was noted by Jacques Monod, “the
²⁴ study of the growth of bacterial cultures does not constitute a specialized subject or branch of re-
²⁵ search, it is the basic method of Microbiology.” Those words ring as true today as they did when
²⁶ they were written 70 years ago. Indeed, the study of bacterial growth has undergone a molecular
²⁷ resurgence since many of the key questions addressed by the pioneering efforts in the middle of
²⁸ the last century can be revisited by examining them through the lens of the increasingly refined
²⁹ molecular census that is available for bacteria such as the microbial workhorse *Escherichia coli*. Sev-
³⁰ eral of the outstanding questions that can now be studied about bacterial growth include: what
³¹ sets the fastest time scale that bacteria can divide, and how is growth rate tied to the quality of the
³² carbon source. In this paper, we address these two questions from two distinct angles. First, as
³³ a result of an array of high-quality proteome-wide measurements of the *E. coli* proteome under a
³⁴ myriad of different growth conditions, we have a census that allows us to explore how the num-
³⁵ ber of key molecular players change as a function of growth rate. This census provides a window
³⁶ onto whether the processes they mediate such as molecular transport into the cells and molecular
³⁷ synthesis within cells can run faster. Second, because of our understanding of the molecular path-
³⁸ ways responsible for many of the steps in bacterial growth, we can also make order of magnitude
³⁹ estimates to infer the copy numbers that would be needed to achieve a given growth rate. In this
⁴⁰ paper, we pass back and forth between the analysis of a variety of different proteomic datasets and
⁴¹ order-of-magnitude estimations to determine possible molecular bottlenecks that limit bacterial

42 growth and to see how the growth rate varies in different carbon sources.

43 Specifically, we leverage a combination of *E. coli* proteomic data sets collected over the past
 44 decade using either mass spectrometry (*Schmidt et al., 2016; Peebo et al., 2015; Valgepea et al.,
 45 2013*) or ribosomal profiling (*Li et al., 2014*) across 31 unique growth conditions. Broadly speaking,
 46 we entertain several classes of hypotheses as are illustrated in **Figure 1**. First, we consider potential
 47 limits on the transport of nutrients into the cell. We address this hypothesis by performing an
 48 order-of-magnitude estimate for how many carbon, phosphorous, and sulfur atoms are needed to
 49 facilitate this requirement given a 5000 second division time. As a second hypothesis, we consider
 50 the possibility that there exists a fundamental limit on how quickly the cell can generate ATP. We
 51 approach this hypothesis from two angles, considering how many ATP synthase complexes must
 52 be needed to churn out enough ATP to power protein translation followed by an estimation of
 53 how many electron transport complexes must be present to maintain the proton motive force.
 54 A third class of estimates considers the need to maintain the size and shape of the cell through
 55 the construction of new lipids for the cell membranes as well as the glycan polymers which make
 56 up the rigid peptidoglycan. Our final class of hypotheses centers on the synthesis of a variety of
 57 biomolecules. Our focus is primarily on the stages of the central dogma as we estimate the number
 58 of protein complexes needed for DNA replication, transcription, and protein translation.

59 In broad terms, we find that the majority of these estimates are in close agreement with the
 60 experimental observations, with protein copy numbers apparently well-tuned for the task of cell
 61 doubling. This allows us to systematically scratch off the hypotheses diagrammed in **Figure 1** as
 62 setting possible speed limits. Ultimately, we find that protein translation (particularly the genera-
 63 tion of new ribosomes) acts as 1) a rate limiting step for the *fastest* bacterial division, and 2) the
 64 major determinant of bacterial growth across all nutrient conditions we have considered under
 65 steady state, exponential growth. This perspective is in line with the linear correlation observed
 66 between growth rate and ribosomal content (typically quantified through the ratio of RNA to pro-
 67 tein) for fast growing cells (*Scott et al., 2010*), but suggests a more prominent role for ribosomes
 68 in setting the doubling time across all conditions of nutrient limitation. Here we again leverage the
 69 quantitative nature of this data set and present a quantitative model of the relationship between
 70 the fraction of the proteome devoted to ribosomes and the speed limit of translation, revealing
 71 a fundamental tradeoff between the translation capacity of the ribosome pool and the maximal
 72 growth rate.

73 Uptake of Nutrients

74 In order to build new cellular mass, the molecular and elemental building blocks must be scav-
 75 enged from the environment in different forms. Carbon, for example, is acquired via the transport
 76 of carbohydrates and sugar alcohols with some carbon sources receiving preferential treatment
 77 in their consumption (*Monod, 1947*). Phosphorus, sulfur, and nitrogen, on the other hand, are har-
 78 vested primarily in the forms of inorganic salts, namely phosphate, sulfate, and ammonia (*Jun et al.,
 79 2018; Assentoft et al., 2016; Stasi et al., 2019; Antonenko et al., 1997; Rosenberg et al., 1977; Will-
 80 sky et al., 1973*). All of these compounds have different permeabilities across the cell membrane
 81 (*Phillips 2018*) and most require some energetic investment either via ATP hydrolysis or through
 82 the proton electrochemical gradient to bring the material across the hydrophobic cell membrane.
 83 Given the diversity of biological transport mechanisms and the vast number of inputs needed to
 84 build a cell, we begin by considering transport of some of the most important cellular ingredients:
 85 carbon, nitrogen, oxygen, hydrogen, phosphorus, and sulfur.

86 The elemental composition of *E. coli* has received much quantitative attention over the past
 87 half century (*Neidhardt et al., 1991; Taymaz-Nikerel et al., 2010; Heldal et al., 1985; Bauer and
 88 Ziv, 1976*), providing us with a starting point for estimating the copy numbers of various trans-
 89 porters. While there is some variability in the exact elemental percentages (with different uncer-
 90 tainties), we can estimate that the dry mass of a typical *E. coli* cell is \approx 45% carbon (BNID: 100649,
 91 *Milo et al. (2010)*), \approx 15% nitrogen (BNID: 106666, *Milo et al. (2010)*), \approx 3% phosphorus (BNID:

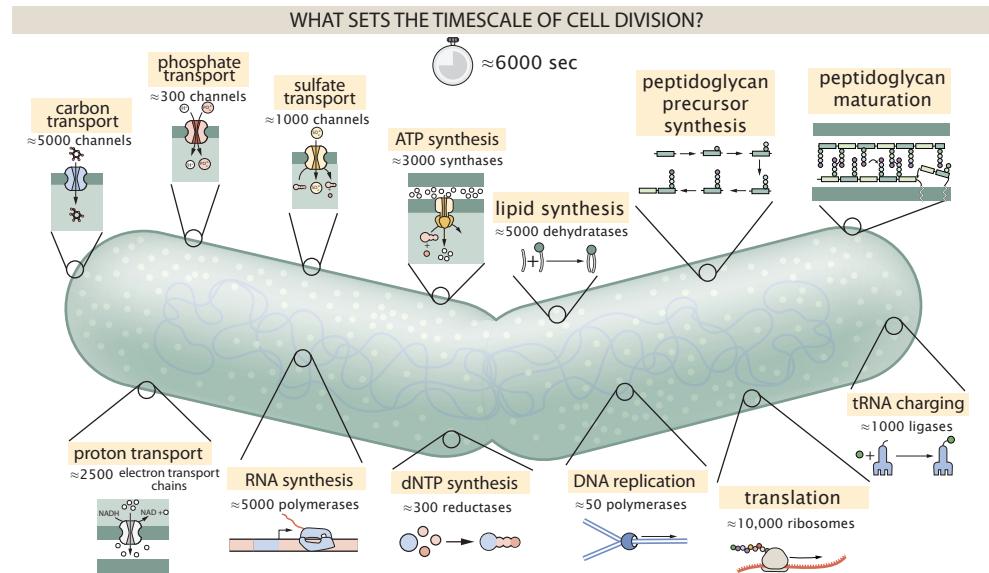


Figure 1. Transport and synthesis processes necessary for cell division. We consider an array of processes necessary for a cell to double its molecular components. Such processes include the transport of carbon across the cell membrane, the production of ATP, and fundamental processes of the central dogma namely RNA, DNA, and protein synthesis. A schematic of each synthetic or transport category is shown with an approximate measure of the complex abundance at a growth rate of 0.5 per hour. In this work, we consider a standard bacterial division time of ≈ 5000 sec.

92 100653, *Milo et al. (2010)*), and 1% sulfur (BNID: 100655, *Milo et al. (2010)*). In the coming para-
 93 graphs, we will engage in a dialogue between back-of-the-envelope estimates for the numbers of
 94 transporters needed to facilitate these chemical stoichiometries and the experimental proteomic
 95 measurements of the biological reality. Such an approach provides the opportunity to test if our
 96 biological knowledge is sufficient to understand the scale at which these complexes are produced.
 97 Specifically, we will make these estimates considering a modest doubling time of 5000 s, a growth
 98 rate of $\approx 0.5 \text{ hr}^{-1}$, the range in which the majority of the experimental measurements reside.

99 Nitrogen Transport

100 Before we begin our back-of-the-envelope estimations, we must address which elemental sources
 101 must require proteinaceous transport, meaning that the cell cannot acquire appreciable amounts
 102 simply via diffusion from the membrane. The permeability of the lipid membrane to a large num-
 103 ber of solutes has been extensively characterized over the past century. Large, polar molecular
 104 species (such as various sugar molecules, sulfate, and phosphate) have low permeabilities while
 105 small, non-polar compounds (such as oxygen, carbon dioxide, and ammonia) can readily diffuse
 106 across the membrane. Ammonia, a primary source of nitrogen in typical laboratory conditions,
 107 has a permeability on par with water ($\approx 10^5 \text{ nm/s}$, BNID:110824 *Milo et al. (2010)*). In particularly
 108 nitrogen-poor conditions, *E. coli* expresses a transporter (AmtB) which appears to aid in nitrogen
 109 assimilation, though the mechanism and kinetic details of transport is still a matter of debate (*van*
 110 *Heeswijk et al., 2013; Khademi et al., 2004*). Beyond ammonia, another plentiful source of nitrogen
 111 come in the form of glutamate, which has its own complex metabolism and scavenging pathways.
 112 However, nitrogen is plentiful in the growth conditions examined in this work, permitting us to ne-
 113 glect nitrogen transport as a potential rate limiting process in cell division in typical experimental
 114 conditions. We direct the reader to the supplemental information for a more in-depth discussion of
 115 permeabilities and a series of calculations revealing that active nitrogen transport can be neglected
 116 for the purposes of this article.

117 **Carbon Transport**

118 We begin our estimations with the most abundant element in *E. coli* by mass, carbon. Using ≈ 0.3
 119 pg as the typical *E. coli* dry mass (BNID: 103904, *Milo et al. (2010)*), we estimate that $\approx 10^{10}$ carbon
 120 atoms must be brought into the cell in order to double all of the carbon-containing molecules (*Fig-*
 121 *ure 2(A, top)*). Typical laboratory growth conditions, such as those explored in the aforementioned
 122 proteomic data sets, provide carbon as a single class of sugar such as glucose, galactose, or xylose
 123 to name a few. *E. coli* has evolved myriad mechanisms by which these sugars can be transported
 124 across the cell membrane. One such mechanism of transport is via the PTS system which is a
 125 highly modular system capable of transporting a diverse range of sugars (*Escalante et al., 2012*).
 126 The glucose-specific component of this system transports ≈ 200 glucose molecules per second per
 127 transporter (BNID: 114686, *Milo et al. (2010)*). Making the assumption that this is a typical sugar
 128 transport rate, coupled with the need to transport 10^{10} carbon atoms, we arrive at the conclusion
 129 that on the order of 1,000 transporters must be expressed in order to bring in enough carbon
 130 atoms to divide in 5000 s, diagrammed in the top panel of *Figure 2(A)*. This estimate, along with
 131 the observed average number of the PTS system carbohydrate transporters present in the pro-
 132 teomic data sets (*Schmidt et al., 2016; Peebo et al., 2015; Valgepea et al., 2013; Li et al., 2014*), is
 133 shown in *Figure 2(A)*. While we estimate 1500 transporters are needed with a 5000 s division time,
 134 we can abstract this calculation to consider any particular growth rate given knowledge of the cell
 135 density and volume as a function of growth rate and direct the reader to the SI for more informa-
 136 tion. As revealed in *Figure 2(A)*, experimental measurements exceed the estimate by several fold,
 137 illustrating that transport of carbon in to the cell is not rate limiting for cell division. Abstracting
 138 this point estimate at 5000 s to a continuum of growth rates (grey line in *Figure 2(A)*) reveals an
 139 excess of transporters at other growth rates, though in rapid growth regimes, the abundance is
 140 below our simple estimate.

141 The estimate presented in *Figure 2(A)* neglects any specifics of the regulation of carbon trans-
 142 port system and presents a data-averaged view of how many carbohydrate transporters are present
 143 on average. Using the diverse array of growth conditions explored in the proteomic data sets, we
 144 can explore how individual carbon transport systems depend on the population growth rate. In
 145 *Figure 2(B)*, we show the total number of carbohydrate transporters specific to different carbon
 146 sources. A striking observation, shown in the top-left plot of *Figure 2(B)*, is the constancy in the
 147 expression of the glucose-specific transport systems. Additionally, we note that the total number
 148 of glucose-specific transporters is tightly distributed $\approx 10^4$ per cell, the approximate number of
 149 transporters needed to sustain rapid growth of several divisions per hour. This illustrates that *E.*
 150 *coli* maintains a substantial number of complexes present for transporting glucose which is known
 151 to be the preferential carbon source (*Monod, 1947; Liu et al., 2005; Aidelberg et al., 2014*).

152 It is now understood that a large number of metabolic operons are regulated with dual-input
 153 logic gates that are only expressed when glucose concentrations are low (mediated by cyclic-AMP
 154 receptor protein CRP) and the concentration of other carbon sources are elevated (*Gama-Castro*
 155 *et al., 2016; Zhang et al., 2014b*). A famed example of such dual-input regulatory logic is in the regu-
 156 lation of the *lac* operon which is only natively activated in the absence of glucose and the presence
 157 of allolactose, an intermediate in lactose metabolism (*Jacob and Monod, 1961*), though we now
 158 know of many other such examples (*Ireland et al., 2020; Gama-Castro et al., 2016; Belliveau et al.,*
 159 *2018*). This illustrates that once glucose is depleted from the environment, cells have a means to
 160 dramatically increase the abundance of the specific transporter needed to digest the next sugar
 161 that is present. Several examples of induced expression of specific carbon-source transporters
 162 are shown in *Figure 2(B)*. Points colored in red (labeled by red text-boxes) correspond to growth
 163 conditions in which the specific carbon source (glycerol, xylose, or fructose) is present. These plots
 164 show that, in the absence of the particular carbon source, expression of the transporters is main-
 165 tained on the order of $\sim 10^2$ per cell. However, when the transport substrate is present, expression
 166 is induced and the transporters become highly-expressed. The grey lines in *Figure 2(B)* show the

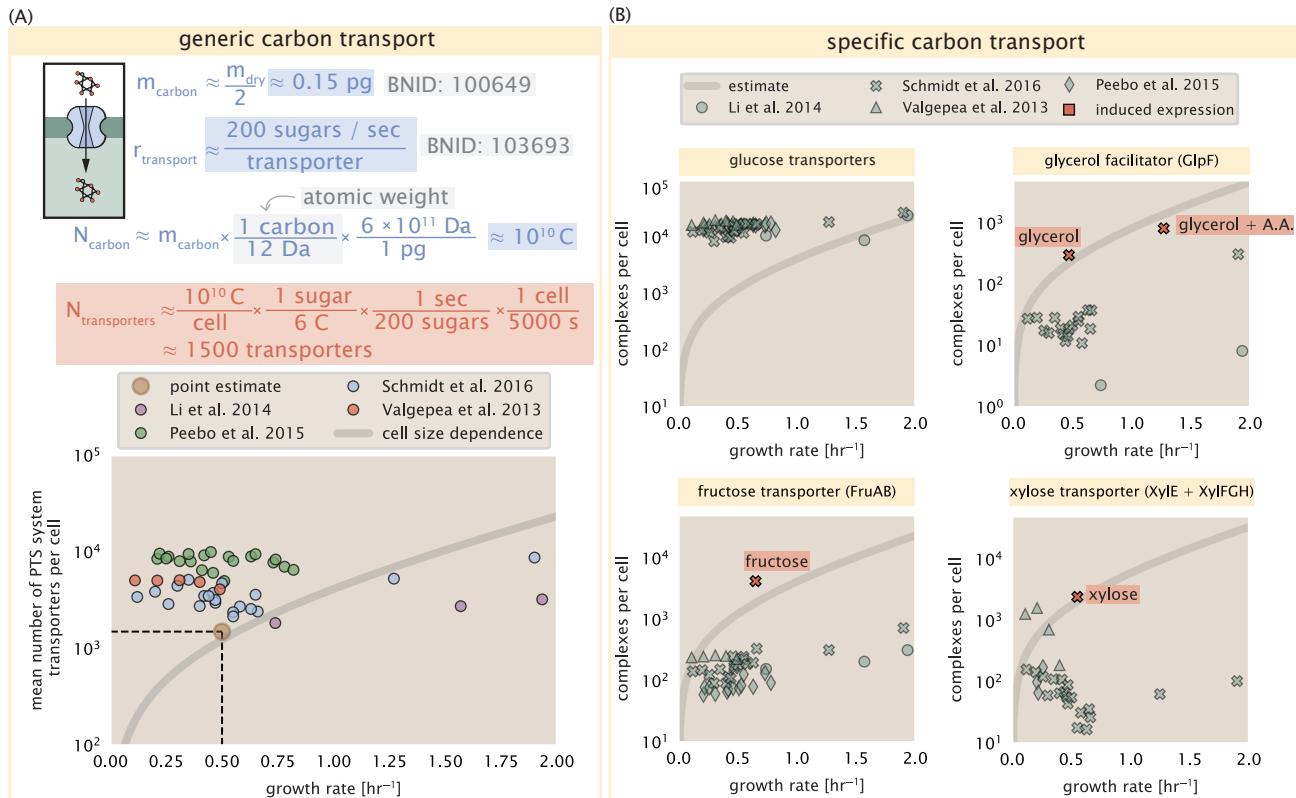


Figure 2. The abundance of carbon transport systems across growth rates. (A) A simple estimate for the minimum number of generic carbohydrate transport systems (top) assumes $\approx 10^{10}$ C are needed to complete division, each transported sugar contains ≈ 6 C, and each transporter conducts sugar 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 PTS system sugar transporters (complexes annotated with the Gene Ontology term GO:0009401) 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, *Milo et al. (2010)*), 2000 glycerol- s^{-1} (*Lu et al., 2003*), 200 fructose- s^{-1} (assumed to be similar to PtsL, BNID: 103693, *Milo et al. (2010)*), and 50 xylose- s^{-1} (assumed to be comparable to LacY, BNID: 103159, *Milo et al. (2010)*). 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.

167 estimated number of transporters needed at each growth rate to satisfy the cellular carbon re-
 168 quirement. It is notable that in all cases, the magnitude of induced expression (shown in red) falls
 169 close to the estimate, illustrating the ability of the cell to tune expression in response to changing
 170 environments. Together, this generic estimation and the specific examples of induced expression
 171 suggest that transport of carbon across the cell membrane, while critical for growth, is not the
 172 rate-limiting step of cell division.

173 **Phosphorus and Sulfur Transport**

174 We now turn our attention towards other essential elements, namely phosphorus and sulfur. Phos-
 175 phorus is critical to the cellular energy economy in the form of high-energy phosphodiester bonds
 176 making up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-translational
 177 modification of proteins and defining the polar-heads of lipids. In total, phosphorus makes up
 178 $\approx 3\%$ of the cellular dry mass which in typical experimental conditions is in the form of inorganic
 179 phosphate. The cell membrane has remarkably low permeability to this highly-charged and critical
 180 molecule, therefore requiring the expression of active transport systems. In *E. coli*, the proton elec-
 181 trochemical gradient across the inner membrane is leveraged to transport inorganic phosphate
 182 into the cell (Rosenberg et al., 1977). Proton-solute symporters are widespread in *E. coli* (Ramos
 183 and Kaback, 1977; Booth et al., 1979) and can have rapid transport rates of 50 to 100 molecules
 184 per second for sugars and other solutes (BNID: 103159; 111777, Milo et al. (2010)). As a more
 185 extreme example, the proton transporters in the F₁-F₀ ATP synthase, which leverage the proton
 186 electrochemical gradient for rotational motion, can shuttle protons across the membrane at a rate
 187 of ≈ 1000 per second (BNID: 104890; 103390, (Milo et al., 2010)). In *E. coli* the PitA phosphate trans-
 188 port system has been shown to be very tightly coupled with the proton electrochemical gradient
 189 with a 1:1 proton:phosphate stoichiometric ratio (Harris et al., 2001; Feist et al., 2007). Taking the
 190 geometric mean of the aforementioned estimates gives a plausible rate of phosphate transport
 191 on the order of 300 per second. Illustrated in **Figure 3(A)**, we can estimate that ≈ 150 phosphate
 192 transporters are necessary to maintain an $\approx 3\%$ dry mass with a 5000 s division time. This estimate
 193 is again satisfied when we examine the observed copy numbers of PitA in proteomic data sets (plot
 194 in **Figure 3(A)**). While our estimate is very much in line with the observed numbers, we emphasize
 195 that this is likely a slight overestimate of the number of transporters needed as there are other
 196 phosphorous scavenging systems, such as the ATP-dependent phosphate transporter Pst system
 197 which we have neglected.

198 Satisfied that there are a sufficient number of phosphate transporters present in the cell, we
 199 now turn to sulfur transport as another potentially rate limiting process. Similar to phosphate,
 200 sulfate is highly-charged and not particularly membrane permeable, requiring active transport.
 201 While there exists a H+/sulfate symporter in *E. coli*, it is in relatively low abundance and is not well
 202 characterized (Zhang et al., 2014a). Sulfate is predominantly acquired via the ATP-dependent ABC
 203 transporter CysUWA system which also plays an important role in selenium transport (Sekowska
 204 et al., 2000; Sirko et al., 1995). While specific kinetic details of this transport system are not readily
 205 available, generic ATP transport systems in prokaryotes transport on the order of 1 to 10 molecules
 206 per second (BNID: 109035, Milo et al. (2010)). Combining this generic transport rate, measurement
 207 of sulfur comprising 1% of dry mass, and a 5000 second division time yields an estimate of ≈ 1000
 208 CysUWA complexes per cell (**Figure 3(B)**). Once again, this estimate is in notable agreement with
 209 proteomic data sets, suggesting that there are sufficient transporters present to acquire the nec-
 210 essary sulfur. In a similar spirit of our estimate of phosphorus transport, we emphasize that this is
 211 likely an overestimate of the number of necessary transporters as we have neglected other sulfur
 212 scavenging systems that are in lower abundance.

213 **Limits on Transporter Expression**

214 So which, if any, of these processes may be rate limiting for growth? As suggested by **Figure 2**
 215 (B), induced expression can lead to an order-of-magnitude (or more) increase in the amount of

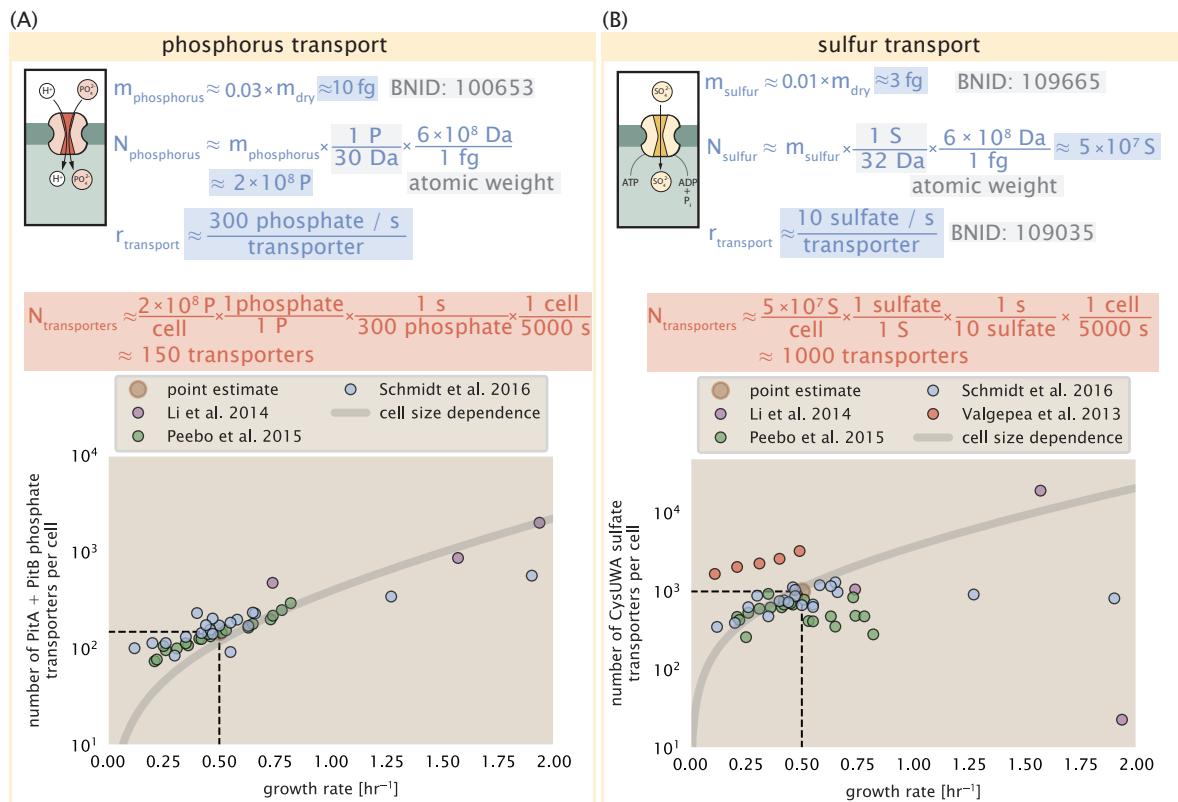


Figure 3. Estimates and measurements of phosphate and sulfate transport systems as a function of growth rate. (A) Estimate for the number of PitA phosphate transport systems needed to maintain a 3% phosphorus *E. coli* dry mass. Points in plot correspond to the total number of PitA transporters per cell.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

216 transporters needed to facilitate transport. Thus, if acquisition of nutrients was the limiting state
 217 in cell division, could expression simply be increased to accommodate faster growth? A way to
 218 approach this question is to compute the amount of space in the bacterial membrane that could
 219 be occupied by nutrient transporters. Considering a rule-of-thumb for the surface area of *E. coli* of
 220 about $6 \mu\text{m}^2$ (BNID: 101792, *Milo et al. (2010)*), we expect an areal density for 1000 transporters to
 221 be approximately 200 transporters/ μm^2 . For a typical transporter occupying about $50 \text{ nm}^2/\text{dimer}$,
 222 this amounts to about only 1 percent of the total inner membrane (*Szenk et al., 2017*). In addition,
 223 bacterial cell membranes typically have densities of 10^5 proteins/ μm^2 (*Phillips, 2018*), implying that
 224 the cell could accommodate more transporters of a variety of species if it were rate limiting. As we
 225 will see in the next section, however, occupancy of the membrane can impose other limits on the
 226 rate of energy production.

227 Energy Production

228 While the transport of nutrients is required to build new cell mass, the metabolic pathways both
 229 consume and generate energy in the form of NTPs. The high-energy phosphodiester bonds of
 230 (primarily) ATP power a variety of cellular processes that drive biological systems away from ther-
 231 modynamic equilibrium. The next set of processes we hypothesize might control the rate of cell
 232 division considers the energy budget of a dividing cell in terms of the synthesis of ATP from ADP
 233 and inorganic phosphate as well as maintenance of the electrochemical proton gradient which
 234 powers it.

235 ATP Synthesis

236 Hydrolysis of the terminal phosphodiester bond of ATP forming ADP and an inorganic phosphate is
 237 a kinetic driving force in a wide array of biochemical reactions. One such reaction is the formation
 238 of peptide bonds during translation which requires ≈ 2 ATPs for the charging of an amino acid
 239 to the tRNA and ≈ 2 ATP equivalents for the formation of the peptide bond between amino acids.
 240 Considering the ATP costs associated with error correction and post-translational modifications
 241 of proteins, we can make the approximation that each peptide bond has a net cost of ≈ 5 ATP
 242 (BNID: 107782, *Milo et al. (2010)*). In total, the energetic costs of peptide bond formation consume
 243 $\approx 80\%$ of the cells ATP budget (BNID: 107782; 106158; 101637; 111918, *Milo et al. (2010); Lynch*
and Marinov (2015); Stouthamer (1973)). The pool of ATP is produced by the F₁-F₀ ATP synthase
 245 – a membrane-bound rotary motor which under ideal conditions can yield ≈ 300 ATP per second
 246 (BNID: 114701; *Milo et al. (2010); Weber and Senior (2003)*).

247 To estimate the total number of ATP equivalents consumed during a cell cycle, we will make
 248 the approximation that there are $\approx 3 \times 10^6$ proteins per cell with an average protein length of ≈ 300
 249 peptide bonds (BNID: 115702; 108986; 104877, *Milo et al. (2010)*). Taking these values together,
 250 we estimate that the typical *E. coli* cell consumes $\approx 5 \times 10^9$ ATP per cell cycle on protein synthesis
 251 alone and $\approx 6 \times 10^9$ ATP in total. Assuming that the ATP synthases are operating at their fastest
 252 possible rate, ≈ 3000 ATP synthases are needed to keep up with the energy demands of the cell.
 253 This estimate and a comparison with the data are shown in *Figure 4* (A). Despite our assumption
 254 of maximal ATP production rate per synthase and approximation of all NTP consuming reactions
 255 being the same as ATP, we find that an estimate of a few thousand complete synthases per cell
 256 to agree well with the experimental data. Much as we did for the estimates of transporter copy
 257 number in the previous section, we can generalize this estimation to consider a continuum of
 258 growth rates rather than a point estimate of 5000 s, indicated by the gray lines in *Figure 4*, and find
 259 that this approach adequately describes the observed growth rate dependence.

260 If the direct production of ATP was a rate limiting step for growth, could the cell simply express
 261 more ATP synthase complexes? This requires us to consider several features of cellular physiology,
 262 namely the physical space on the inner membrane as well as the ability to maintain the proton
 263 chemical gradient leveraged by the synthase to drive ATP production out of equilibrium.

264 Generating the Proton Electrochemical Gradient

265 In order to produce ATP, the F₁-F₀ ATP synthase itself must consume energy. Rather than burning
 266 through its own product, this intricate macromolecular machine has evolved to exploit the elec-
 267 trochemical potential established across the inner membrane through cellular respiration. This
 268 electrochemical gradient is manifest by the pumping of protons into the intermembrane space via
 269 the electron transport chains as they reduce NADH. In *E. coli*, this potential difference is ≈ -200 mV
 270 (BNID: 102120, *Milo et al. (2010)*). A simple estimate of the inner membrane as a capacitor with a
 271 working voltage of -200 mV (as performed in the Supplemental Information) reveals that $\approx 2 \times 10^4$
 272 protons must be present in the intermembrane space.

273 However, the constant rotation of the ATP synthases would rapidly abolish this potential dif-
 274 ference if it were not being actively maintained. To undergo a complete rotation (and produce a
 275 single ATP), the F₁-F₀ ATP synthase must shuttle ≈ 4 protons across the membrane into the cytosol
 276 (BNID: 103390, *Milo et al. (2010)*). With ≈ 3000 ATP synthases each generating 300 ATP per second,
 277 the 2×10^4 protons establishing the 200 mV potential would be consumed in only a few millisec-
 278 onds. This brings us to our next estimate: how many electron transport complexes are needed to
 279 support the consumption rate of the ATP synthases?

280 The electrochemistry of the electron transport complexes of *E. coli* have been the subject of
 281 intense biochemical and biophysical study over the past half century (*Ingledew and Poole, 1984;*
282 Khademian and Imlay, 2017; Cox et al., 1970; Henkel et al., 2014). A recent work (*Szenk et al.,*
283 2017) examined the respiratory capacity of the *E. coli* electron transport complexes using struc-
 284 tural and biochemical data, revealing that each electron transport chain rapidly pumps protons
 285 into the intermembrane space at a clip of ≈ 1500 protons per second (BIND: 114704; 114687, *Milo*
286 et al. (2010)). Using our estimate of the number of ATP synthases required per cell (*Figure 4(A)*),
 287 coupled with these recent measurements, we estimate that ≈ 1000 electron transport complexes
 288 would be necessary to facilitate the $\approx 4 \times 10^6$ protons per second diet of the cellular ATP synthases.
 289 This estimate (along with a generalization to the entire range of observed growth rates) is in agree-
 290 ment with the number of complexes identified in the proteomic datasets (plot in *Figure 4(B)*). This
 291 suggests that every ATP synthase must be accompanied by ≈ 1 functional electron transport chain.
 292 Again, if this were a rate limiting process for bacterial growth, one must conclude that it is not
 293 possible for the cell to simply increase the production of both the number of electron transport
 294 chain complexes as well as ATP synthases. As both of these components only function bound to
 295 the inner membrane, we now turn our attention towards the available space in the membrane as
 296 well as surface-area-to-volume constraints.

297 Energy Production in a Crowded Membrane.

298 For each protein considered so far, the data shows that in general their numbers increase with
 299 growth rate. This is in part a consequence of the increase in cell length and width at that is common
 300 to many rod-shaped bacteria at faster growth rates (*Ojikic et al., 2019; Harris and Theriot, 2018*).
 301 For the particular case of *E. coli*, the total cellular protein and cell size increase logarithmically
 302 with growth rate (*Schaechter et al., 1958; Si et al., 2017*). Indeed, this is one reason why we have
 303 considered only a single, common growth condition across all our estimates so far. Such a scaling
 304 will require that the total number of proteins and net demand on resources also grow in proportion
 305 to the increase in cell size divided by the cell's doubling time. Recall however that each transport
 306 process, as well as the ATP production via respiration, is performed at the bacterial membrane.
 307 This means that their maximum productivity can only increase in proportion to the cell's surface
 308 area divided by the cell doubling time. This difference in scaling would vary in proportion to the
 309 surface area-to-volume (S/V) ratio.

310 While we found that there was more than sufficient membrane real estate for carbon intake in
 311 our earlier estimate, the total number of ATP synthases and electron chain transport complexes
 312 both exhibit a clear increase in copy number with growth rate, reaching in excess of 10^4 copies per
 313 cell (*Figure 4*). Here we consider the consequences of this S/V ratio scaling in more detail.

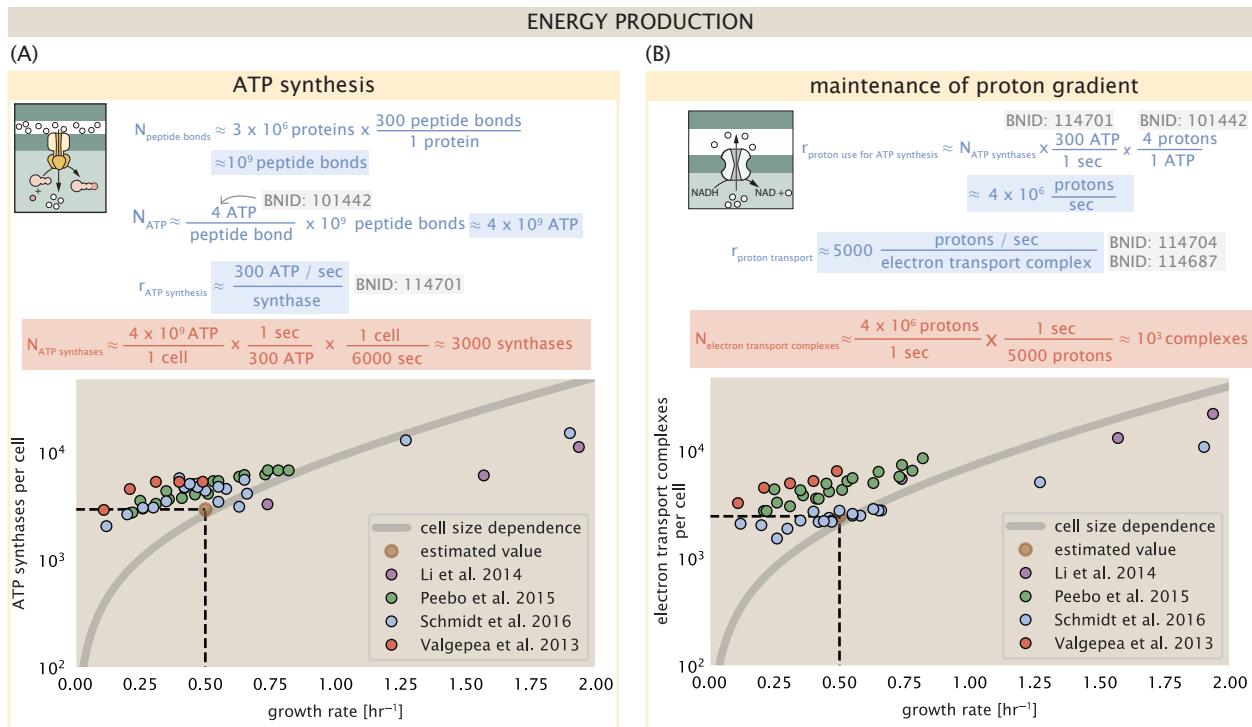


Figure 4. The abundance of F₁-F₀ ATP synthases and electron transport chain complexes as a function of growth rate. (A) Estimate of the number of F₁-F₀ ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes. Points in plot correspond to the mean number of complete F₁-F₀ ATP synthase complexes that can be formed given proteomic measurements and the subunit stoichiometry [AtpE]₁₀[AtpF]₂[AtpB][AtpC][AtpH][AtpA]₃[AtpG][AtpD]₃. (B) Estimate of the number of electron transport chain complexes needed to maintain a membrane potential of -200 mV given estimate of number of F₁-F₀ ATP synthases from (A). Points in plot correspond to the average number of complexes identified as being involved in aerobic respiration by the Gene Ontology identifier GO:0019646 that could be formed given proteomic observations. These complexes include cytochromes *bd*_I ([CydA][CydB][CydX][CydH]), *bd*_{II} ([AppC][AppB]), *bo*₃,([CyoD][CyoA][CyoB][CyoC]) and NADH:quinone oxioreductase I ([NuoA][NuoH][NuoJ][NuoK][NuoL][NuoM][NuoN][NuoB][NuoC][NuoE][NuoF][NuoG][NuoI]) and II ([Ndh]). Grey lines in both (A) and (B) correspond to the estimate procedure described, but applied to a continuum of growth rates. We direct the reader to the Supporting Information for a more thorough description of this approach.

314 In our estimate of ATP production above we found that a cell demands about 6×10^9 ATP or 10^6
 315 ATP/s. With a cell volume of roughly 1 fL, this corresponds to about 2×10^{10} ATP per fL of cell volume,
 316 in line with previous estimates (*Stouthamer and Bettenhausen, 1977; Szenk et al., 2017*). In *Figure 5*(A) we plot this ATP demand as a function of the S/V ratio in green, where we have considered
 317 a range of cell shapes from spherical to rod-shaped with an aspect ratio (length/width) equal to 4
 318 (See appendix for calculations of cell volume and surface area). In order to consider the maximum
 319 power that could be produced, we consider the amount of ATP that can be generated by a membrane
 320 filled with ATP synthase and electron transport complexes, which provides a maximal production
 321 of about 3 ATP / (nm²·s) (*Szenk et al., 2017*). This is shown in blue in *Figure 5*(A), which shows that
 322 at least for the growth rates observed, the energy demand is roughly an order of magnitude less.
 323

324 Interestingly, *Szenk et al. (2017)* also found that ATP production by respiration is less efficient
 325 than by fermentation per membrane area occupied due to the additional proteins of the electron
 326 transport chain. This suggests that even under anaerobic growth, there will be sufficient mem-
 327 brane space for ATP production in general.

328 While this serves to highlight the diminishing capacity to provide resources to grow if the cell
 329 increases in size (and its S/V decreases), the blue region in *Figure 5*(A) represents a somewhat
 330 unachievable limit since the inner membrane must also include other proteins such as those re-
 331 quired for lipid and membrane synthesis. To gain insight, we used the proteomic data to look at
 332 the distribution of proteins on the inner membrane. Here we use Gene Ontology (GO) annotations
 333 (*Ashburner et al., 2000; The Gene Ontology Consortium, 2018*) to identify all proteins embedded
 334 or peripheral to the inner membrane (GO term: 0005886). Those associated but not membrane-
 335 bound include proteins like MreB and FtsZ, that traverse the inner membrane by treadmilling and
 336 must nonetheless be considered as a vital component occupying space on the membrane. In *Fig-*
 337 *ure 5*(B), we find that the total protein mass per μm^2 is relatively constant with growth rate. Inter-
 338 estingly, when we consider the distribution of proteins grouped by their Clusters of Orthologous
 339 Groups (COG) (*Tatusov et al., 2000*), the relative abundance for those in metabolism (including ATP
 340 synthesis via respiration) is also relatively constant.

341 Function of the Central Dogma

342 Up to this point, we have considered a variety of transport and biosynthetic processes that are
 343 critical to acquiring and generating new cell mass. While there are of course many other metabolic
 344 processes we could consider and perform estimates of (such as the components of fermentative
 345 versus aerobic respiration), we now turn our focus to some of the most central processes which
 346 *must* be undertaken irrespective of the growth conditions – the processes of the central dogma.

347 DNA

348 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal
 349 plasmids up to ~ 100 kbp in length. We will focus our quantitative thinking solely on the chromo-
 350 some of *E. coli* which harbors ≈ 5000 genes and $\approx 5 \times 10^6$ base pairs. To successfully divide and
 351 produce viable progeny, this chromosome must be faithfully replicated and segregated into each
 352 nascent cell. We again rely on the near century of literature in molecular biology to provide some
 353 insight on the rates and mechanics of the replicative feat as well as the production of the required
 354 starting materials, dNTPs.

355 dNTP synthesis

356 We begin our exploration of DNA replication by examining the production of the deoxyribonucleotide
 357 triphosphates (dNTPs). The four major dNTPs (dATP, dTTP, dCTP, and dGTP) are synthesized *de*
 358 *novo* in separate pathways, requiring different building blocks. However, a critical step present in
 359 all dNTP synthesis pathways is the conversion from ribonucleotide to deoxyribonucleotide via the
 360 removal of the 3' hydroxyl group of the ribose ring (*Rudd et al., 2016*). This reaction is mediated
 361 by a class of enzymes termed ribonucleotide reductases, of which *E. coli* possesses two aerobically

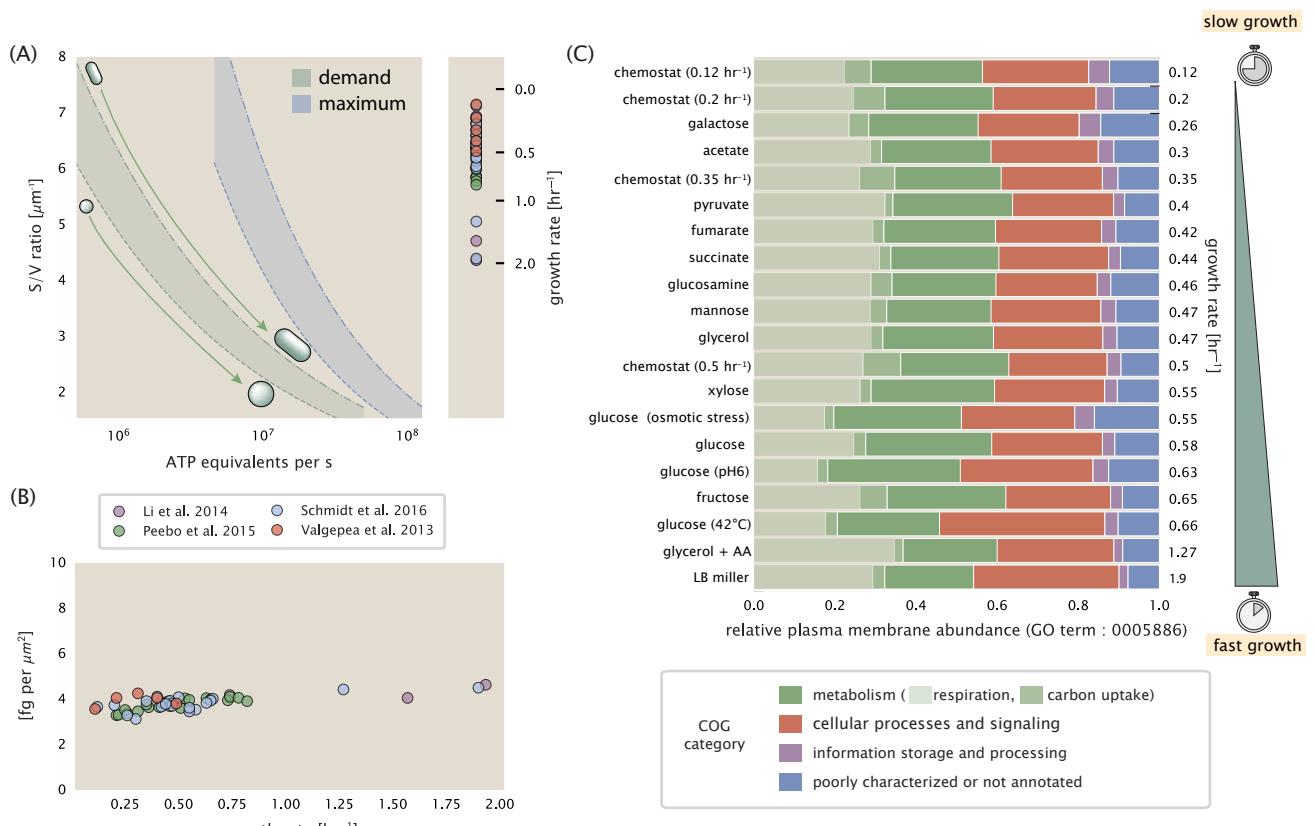


Figure 5. Influence of cell size and S/V ratio on ATP production and inner membrane composition. (A) Scaling of ATP demand and maximum ATP production as a function of S/V ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (---) reflecting a rod-shaped bacterium like *E. coli* with a typical aspect ratio (length / width) of 0.4 (Shi et al., 2018). Approximately 50% of the bacterial inner membrane is assumed to be protein, with the remainder lipid. (B) Total protein mass calculated for proteins with inner membrane annotation (GO term: 0005886). (C) Relative protein abundances by mass based on COG annotation. Metabolic proteins are further separated into respiration (F_1 - F_0 ATP synthase, NADH dehydrogenase I, succinate:quinone oxidoreductase, cytochrome bo₃ ubiquinol oxidase, cytochrome bd-I ubiquinol oxidase) and carbohydrate transport (GO term: GO:0008643). Note that the elongation factor EF-Tu can also associate with the inner membrane, but was excluded in this analysis due to its high relative abundance (roughly identical to the total mass shown in part (B)).

362 active complexes (termed I and II) and a single anaerobically active enzyme. Due to their peculiar
 363 formation of a radical intermediate, these enzymes have received much biochemical, kinetic, and
 364 structural characterization. One such work (**Ge et al., 2003**) performed a detailed *in vitro* measure-
 365 ment of the steady-state kinetic rates of these complexes, revealing a turnover rate of ≈ 10 dNTP
 366 per second.

367 Since this reaction is central to the synthesis of all dNTPs, it is reasonable to consider the abun-
 368 dence of these complexes is a measure of the total dNTP production in *E. coli*. Illustrated schemati-
 369 cally in **Figure 6** (A), we consider the fact that to replicate the cell's genome, on the order of $\approx 10^7$
 370 dNTPs must be synthesized. Assuming a production rate of 10 per second per ribonucleotide
 371 reductase complex and a cell division time of 5000 seconds, we arrive at an estimate of ≈ 200
 372 complexes needed per cell. As shown in the bottom panel of **Figure 6** (A), this estimate agrees
 373 with the experimental measurements of these complexes abundances within $\approx 1/2$ an order of
 374 magnitude. Extension of this estimate across a continuum of growth rate, including the fact that
 375 multiple chromosomes can be replicated at a given time, is shown as a grey transparent line in
 376 **Figure 6** (A). Similarly to our point estimate, this refinement agrees well with the data, accurately
 377 describing both the magnitude of the complex abundance and the dependence on growth rate.

378 Recent work has revealed that during replication, the ribonucleotide reductase complexes co-
 379 alesce to form discrete foci colocalized with the DNA replisome complex (**Sánchez-Romero et al.,**
 380 **2011**). This is particularly pronounced in conditions where growth is slow, indicating that spatial
 381 organization and regulation of the activity of the complexes plays an important role.

382 DNA Replication

383 We now turn our focus to the integration of these dNTP building blocks into the replicated chromo-
 384 some strand via the DNA polymerase. Replication is initiated at a single region of the chromosome
 385 termed the *oriC* locus at which a pair of DNA polymerases bind and begin their high-fidelity repli-
 386 cation of the genome in opposite directions. Assuming equivalence between the two replication
 387 forks, this means that the two DNA polymerase complexes (termed replisomes) meet at the mid-
 388 way point of the circular chromosome termed the *ter* locus. The kinetics of the five types of DNA
 389 polymerases (I – V) have been intensely studied, revealing that DNA polymerase III performs the
 390 high fidelity processive replication of the genome with the other "accessory" polymerases playing
 391 auxiliary roles (**Fijalkowska et al., 2012**). *In vitro* measurements have shown that DNA Polymerase
 392 III copies DNA at a rate of ≈ 600 nucleotides per second (BNID: 104120, **Milo et al. (2010)**). There-
 393 fore, to replicate a single chromosome, two replisomes (containing two DNA polymerase III each)
 394 moving at their maximal rate would copy the entire genome in ≈ 4000 s. Thus, with a division time
 395 of 5000 s (our "typical" growth rate for the purposes of this work), there is sufficient time for a pair
 396 of replisomes complexes to replicate the entire genome. However, this estimate implies that 4000
 397 s would be the upper-limit time scale for bacterial division which is at odds with the familiar 20
 398 minute (1200 s) doubling time of *E. coli* in rich medium.

399 It is well known that *E. coli* can parallelize its DNA replication such that multiple chromosomes
 400 are being replicated at once, with as many as 10 - 12 replication forks at a given time (**Bremer**
 401 **and Dennis, 2008; Si et al., 2017**). Thus, even in rapidly growing cultures, we expect only a few
 402 polymerases (≈ 10) are needed to replicate the chromosome per cell doubling. However, as shown
 403 in **Figure 6** (B), DNA polymerase III is nearly an order of magnitude more abundant. This
 404 discrepancy can be understood by considering its binding constant to DNA. DNA polymerase III is
 405 highly processive, facilitated by a strong affinity of the complex to the DNA. *In vitro* biochemical
 406 characterization has quantified the K_D of DNA polymerase III holoenzyme to single-stranded and
 407 double-stranded DNA to be 50 and 200 nM, respectively (**Ason et al., 2000**). The bottom plot in
 408 **Figure 6** (B) shows that the concentration of the DNA polymerase III across all data sets and growth
 409 conditions is within this range. Thus, while the copy number of the DNA polymerase III is in excess
 410 of the strict number required to replicate the genome, its copy number appears to vary such that its
 411 concentration is approximately equal to the dissociation constant to the DNA. While the processes

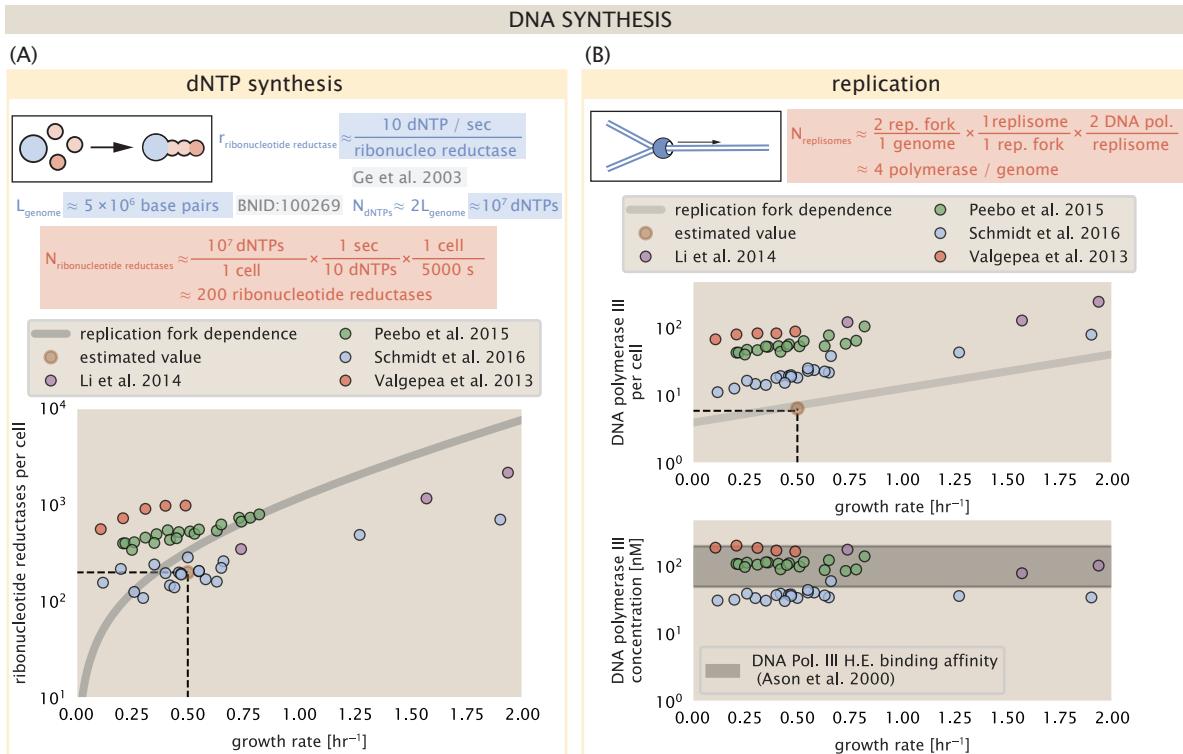


Figure 6. Complex abundance estimates for dNTP synthesis and DNA replication. (A) Estimate of the number of ribonucleotide reductase enzymes needed to facilitate the synthesis of $\approx 10^7$ dNTPs over the course of a 5000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ($[\text{NrdA}]_2[\text{NrdB}]_2$) and ribonucleotide reductase II ($[\text{NrdE}]_2[\text{NrdF}]_2$) complexes. (B) An estimate for the minimum number of DNA polymerase holoenzyme complexes needed to facilitate replication of a single genome. Points in the top plot correspond to the total number of DNA polymerase III holoenzyme complexes ($[\text{DnaE}]_3[\text{DnaQ}]_3[\text{HolE}]_3[\text{DnaX}]_5[\text{HolB}][\text{HolA}][\text{DnaN}]_4[\text{HolC}]_4[\text{HolD}]_4$) per cell. Bottom plot shows the effective concentration of DNA polymerase III holoenzyme given cell volumes calculated from the growth rate as in *Si et al. (2019)* (See Supplemental Information Section 4). Grey lines in (A) and top panel of (B) show the estimated number of complexes needed as a function of growth, the details of which are described in the Supplemental Information.

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

RNA Synthesis

With the machinery governing the replication of the genome accounted for, we now turn our attention to the next stage of the central dogma – the transcription of DNA to form RNA. We primarily consider three major groupings of RNA, namely the RNA associated with ribosomes (rRNA), the RNA encoding the amino-acid sequence of proteins (mRNA), and the RNA which links codon sequence to amino-acid identity during translation (tRNA). Despite the varied function of these RNA species, they share a commonality in that they are transcribed from DNA via the action of RNA polymerase. In the coming paragraphs, we will consider the synthesis of RNA as a rate limiting step in bacterial division by estimating how many RNA polymerases must be present to synthesize all necessary rRNA, mRNA, and tRNA.

rRNA

We begin with an estimation of the number of RNA polymerases needed to synthesize the rRNA that serve as catalytic and structural elements of the ribosome. Each ribosome contains three rRNA molecules of lengths 120, 1542, and 2904 nucleotides (BNID: 108093, *Milo et al. (2010)*), meaning each ribosome contains ≈ 4500 nucleotides. As the *E. coli* RNA polymerase transcribes DNA to RNA at a rate of ≈ 40 nucleotides per second (BNID: 101904, *Milo et al. (2010)*), it takes a single RNA polymerase ≈ 100 s to synthesize the RNA needed to form a single functional ribosome. Therefore, in a 5000 s division time, a single RNA polymerase transcribing rRNA at a time would result in only ≈ 50 functional ribosomal rRNA units – far below the observed number of $\approx 10^4$ ribosomes per cell.

Of course, there can be more than one RNA polymerase transcribing the rRNA genes at any given time. To elucidate the *maximum* number of rRNA units that can be synthesized given a single copy of each rRNA gene, we will consider a hypothesis in which the rRNA operon is completely tiled with RNA polymerase. *In vivo* measurements of the kinetics of rRNA transcription have revealed that RNA polymerases are loaded onto the promoter of an rRNA gene at a rate of ≈ 1 per second (BNID: 111997; 102362, *Milo et al. (2010)*). If RNA polymerases are being constantly loaded on to the rRNA genes at this rate, then we can assume that ≈ 1 functional rRNA unit is synthesized per second. With a 5000 second division time, this hypothesis leads to a maximal value of 5000 functional rRNA units, still undershooting the observed number of 10^4 ribosomes per cell.

E. coli has evolved a clever mechanism to surpass this kinetic limit for the rate of rRNA production. Rather than having only one copy of each rRNA gene, *E. coli* has seven copies of the operon (BIND: 100352, *Milo et al. (2010)*) four of which are localized directly adjacent to the origin of replication (*Birnbaum and Kaplan, 1971*). As fast growth also implies an increased gene dosage due to parallelized chromosomal replication, the total number of rRNA genes can be on the order of $\approx 10 - 70$ copies at moderate to fast growth rates (*Stevenson and Schmidt, 2004*). Using our standard time scale of a 5000 second division time, we can make the lower-bound estimate that the typical cell will have 7 copies of the rRNA operon. Synthesizing one functional rRNA unit per second per rRNA operon, a total of 4×10^4 rRNA units can be synthesized, comfortably above the observed number of ribosomes per cell.

How many RNA polymerases are then needed to constantly transcribe 7 copies of the rRNA genes? We approach this estimate by considering the maximum number of RNA polymerases tiled along the rRNA genes with a loading rate of 1 per second and a transcription rate of 40 nucleotides per second. Considering that a RNA polymerase has a physical footprint of approximately 40 nu-

461 cleotides (BNID: 107873, *Milo et al. (2010)*), we can expect ≈ 1 RNA polymerase per 80 nucleotides.
462 With a total length of ≈ 4500 nucleotides per operon and 7 operons per cell, the maximum number
463 of RNA polymerases that can be transcribing rRNA at any given time is ≈ 400 . As we will see in the
464 coming sections, the synthesis of rRNA is the dominant requirement of the RNA polymerase pool.

465 mRNA

466 To form a functional protein, all protein coding genes must first be transcribed from DNA to form
467 an mRNA molecule. While each protein requires an mRNA blueprint, many copies of the protein
468 can be synthesized from a single mRNA. Factors such as strength of the ribosomal binding site,
469 mRNA stability, and rare codon usage frequency dictate the number of proteins that can be made
470 from a single mRNA, with yields ranging from 10^1 to 10^4 (BNID: 104186; 100196; 106254, *Milo et al.*
471 (2010)). Computing the geometric mean of this range yields ≈ 1000 proteins synthesized per mRNA,
472 a value that agrees with experimental measurements of the number of proteins per cell ($\approx 3 \times 10^6$,
473 BNID: 100088, *Milo et al. (2010)*) and total number of mRNA per cell ($\approx 3 \times 10^3$, BNID:100064, *Milo*
474 *et al. (2010)*).

475 This estimation captures the *steady-state* mRNA copy number, meaning that at any given time,
476 there will exist approximately 3000 unique mRNA molecules. To determine the *total* number of
477 mRNA that need to be synthesized over the cell's lifetime, we must consider degradation of the
478 mRNA. In most bacteria, mRNAs are rather unstable with life times on the order of several minutes
479 (BNID: 104324; 106253; 111927; 111998, *Milo et al. (2010)*). For convenience, we assume that the
480 typical mRNA in our cell of interest has a typical lifetime of ≈ 300 seconds. Using this value, we
481 can determine the total mRNA production rate to maintain a steady-state copy number of 3000
482 mRNA per cell. While we direct the reader to the appendix for a more detailed discussion of mRNA
483 transcriptional dynamics, we state here that the total mRNA production rate must be on the order
484 of ≈ 15 mRNA per second. In *E. coli*, the average protein is ≈ 300 amino acids in length (BNID:
485 108986; *Milo et al. (2010)*), meaning that the corresponding mRNA is ≈ 900 nucleotides which we
486 will further approximate as ≈ 1000 nucleotides to account for the non-protein coding regions on
487 the 5' and 3' ends. This means that the cell must have enough RNA polymerase molecules about
488 to sustain a transcription rate of $\approx 1.5 \times 10^4$ nucleotides per second. Knowing that a single RNA
489 polymerase polymerizes RNA at a clip of 40 nucleotides per second, we arrive at a comfortable
490 estimate of ≈ 250 RNA polymerase complexes needed to satisfy the mRNA demands of the cell. It
491 is worth noting that this number is approximately half of that required to synthesize enough rRNA,
492 as we saw in the previous section. We find this to be a striking result as these 250 RNA polymerase
493 molecules are responsible for the transcription of the ≈ 4000 protein coding genes that are not
494 ribosome associated.

495 tRNA

496 The final class of RNA molecules worthy of quantitative consideration is the the pool of tRNAs
497 used during translation to map codon sequence to amino acid identity. Unlike mRNA or rRNA,
498 each individual tRNA is remarkably short, ranging from 70 to 95 nucleotides each (BNID: 109645;
499 102340, *Milo et al. (2010)*). What they lack in length, they make up for in abundance. There are
500 many measurements of the size of the *E. coli* tRNA pool, ranging from $\approx 6 \times 10^4$ (BNID:105280, *Milo*
501 *et al. (2010)*) to $\approx 4 \times 10^5$ (BNID: 108611). To test tRNA synthesis as a possible growth-rate limiting
502 stage, we will err towards a higher abundance of $\approx 4 \times 10^5$ per cell. Combining the abundance and
503 tRNA length measurements, we make the estimate that $\approx 5 \times 10^7$ nucleotides are sequestered in
504 tRNA per cell. Unlike mRNA, tRNA is remarkably stable with typical lifetimes *in vivo* on the order of
505 ≈ 48 hours (*Abelson et al., 1974; Svenningsen et al., 2017*) – well beyond the timescale of division.
506 Once again using our rule-of-thumb for the rate of transcription to be 40 nucleotides per second
507 and assuming a division time of ≈ 5000 seconds, we arrive at an estimate of ≈ 150 RNA polymerases
508 to synthesize enough tRNA. This requirement pales in comparison to the number of polymerases
509 needed to generate the rRNA and mRNA pools and can be neglected as a significant transcriptional

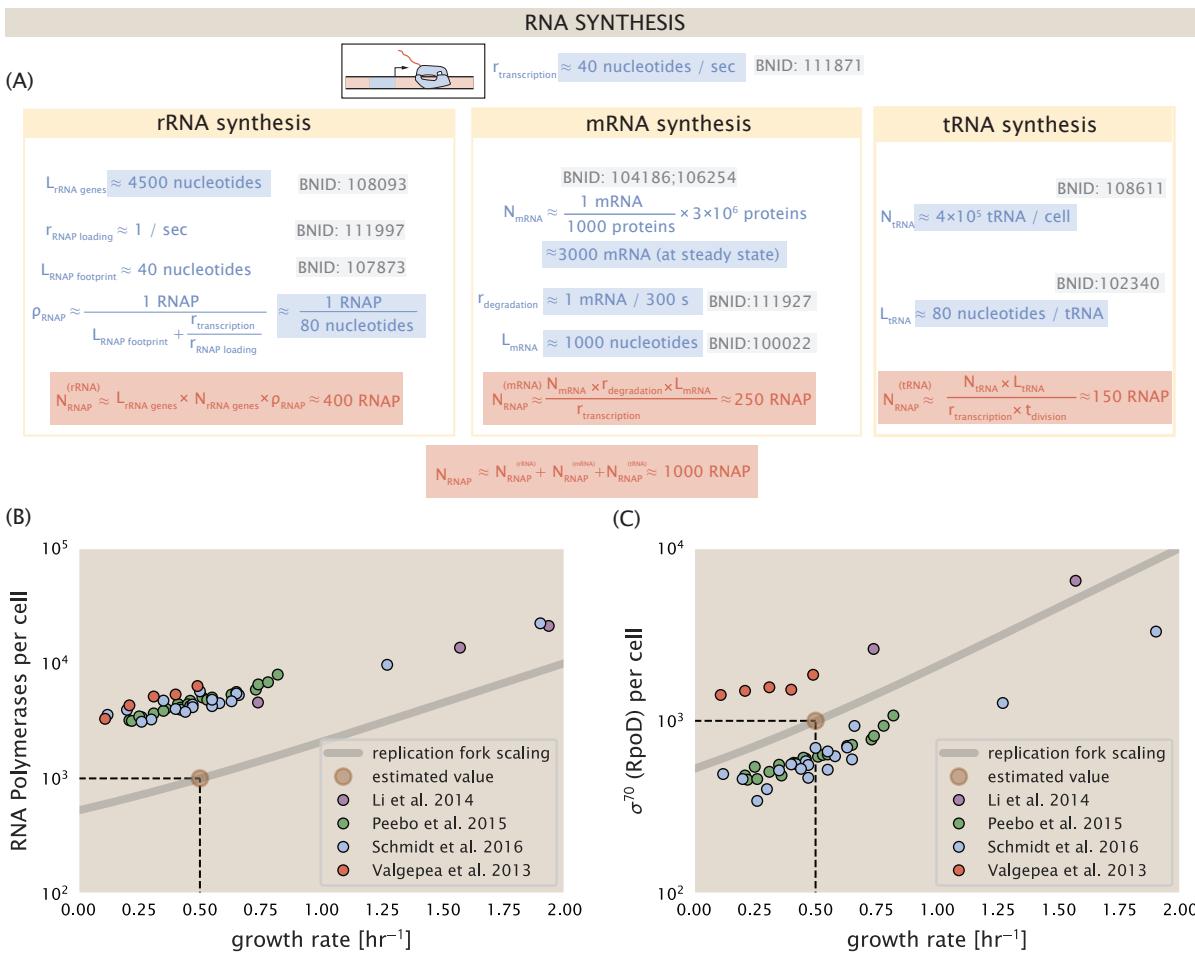


Figure 7. Estimation of the RNA polymerase demand and comparison with experimental data. (A)

Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. Bionumber Identifiers (BNIDs) are provided for key quantities used in the estimates. (B) The RNA polymerase core enzyme copy number as a function of growth rate. Colored points correspond to the average number RNA polymerase core enzymes that could be formed given a subunit stoichiometry of $[RpoA]_2[RpoC][RpoB]$. (C) The abundance of σ^{70} as a function of growth rate. Estimated value for the number of RNAP is shown in (B) and (C) as a translucent brown point at a growth rate of 0.5 hr^{-1} .

burden.

511 RNA Polymerase and σ -factor Abundance

512 These estimates, summarized in **Figure 7 (A)**, reveal that synthesis of rRNA and mRNA are the domi-
 513 nant RNA species synthesized by RNA polymerase, suggesting the need for ≈ 700 RNA polymerases
 514 per cell. As is revealed in **Figure 7 (B)**, this estimate is about an order of magnitude below the ob-
 515 served number of RNA polymerase complexes per cell ($\approx 5000 - 7000$). The disagreement between
 516 the estimated number of RNA polymerases and these observations are at least consistent with
 517 recent literature revealing that $\approx 80\%$ of RNA polymerases in *E. coli* are not transcriptionally active
 518 (*Patrick et al., 2015*). Our estimate ignores the possibility that some fraction is only nonspecifically
 519 bound to DNA, as well as the obstacles that RNA polymerase and DNA polymerase present for each
 520 other as they move along the DNA (*Finkelstein and Greene, 2013*).

521 In addition, it is also vital to consider the role of σ -factors which help RNA polymerase identify
 522 and bind to transcriptional start sites (*Browning and Busby, 2016*). Here we consider σ^{70} (RpoD)
 523 which is the dominant "general-purpose" σ -factor in *E. coli*. While initially thought of as being solely

involved in transcriptional initiation, the past two decades of single-molecule work has revealed a more multipurpose role for σ^{70} including facilitating transcriptional elongation (Kapanidis et al., 2005; Goldman et al., 2015; Perdue and Roberts, 2011; Mooney and Landick, 2003; Mooney et al., 2005). **Figure 7** (B) is suggestive of such a role as the number of σ^{70} proteins per cell is in close agreement with our estimate of the number of transcriptional complexes needed.

These estimates provide insight as to the observed magnitude of both RNA polymerase and the σ -70 factor. As we have done in the previous sections, and described in the supplemental information, we can generalize these estimates across a wide range of growth rates (grey line in **Figure 7(B)**). While there remains some disagreement in the magnitude of the copy number, this estimate appears to very adequately describe the growth rate dependence of these complexes. Furthermore, these findings illustrate that transcription cannot be the rate limiting step in bacterial division. **Figure 7** (A) reveals that the availability of RNA polymerase is not a limiting factor for cell division as the cell always has an apparent \sim 10-fold excess than needed. Furthermore, if more transcriptional activity was needed to satisfy the cellular requirements, more σ^{70} -factors could be expressed to utilize a larger fraction of the RNA polymerase pool.

539 Translation and ribosomal synthesis

540 Lastly, we turn our attention to the process of synthesizing new proteins, translation. These processes stand as good candidates for defining the growth limit as the synthesis of new proteins 541 relies on the generation of ribosomes, themselves proteinaceous molecules. As we will see in the 542 coming sections of this work, this poses a "chicken-or-the-egg" problem where the synthesis of 543 ribosomes requires ribosomes in the first place.

544 We will begin our exploration of protein translation in the same spirit as we have in previous 545 sections – we will draw order-of-magnitude estimates based on our intuition and available literature, 546 and then compare these estimates to the observed data. In doing so, we will estimate both 547 the absolute number of ribosomes necessary for replication of the proteome as well as the synthe- 548 sis of amino-acyl tRNAs. In the closing sections, we will explore the details of ribosome biogenesis 549 in granular detail, ultimately presenting a quantitative model tying ribosome abundance to the 550 concentration of amino acids as well as the state of chromosome replication.

552 tRNA synthetases

553 We begin by first estimating the number of tRNA ligases in *E. coli* needed to convert free amino- 554 acids to polypeptide chains. At a modest growth rate of \approx 5000 s, *E. coli* has roughly 3×10^6 proteins 555 per cell (BNID: 115702; Milo et al. (2010)). Assuming that the typical protein is on the order of \approx 556 300 amino acids in length (BNID: 100017; Milo et al. (2010)), we can estimate that a total of $\approx 10^9$ 557 amino acids are stitched together by peptide bonds.

558 How many tRNAs are needed to facilitate this remarkable number of amino acid delivery events 559 to the translating ribosomes? It is important to note that tRNAs are recycled after they've passed 560 through the ribosome and can be recharged with a new amino acid, ready for another round of 561 peptide bond formation. While some *in vitro* data exists on the turnover of tRNA in *E. coli* for 562 different amino acids, we can make a reasonable estimate by comparing the number of amino 563 acids to be polymerized to cell division time. Using our stopwatch of 5000 s and 10^9 amino acids, 564 we arrive at a requirement of $\approx 2 \times 10^5$ tRNA molecules. This estimate is in line with experimental 565 measurements of $\approx 3 \times 10^5$ per cell (BNID: 108611, Milo et al. (2010)), suggesting we are on the 566 right track.

567 There are many processes which go into synthesizing a tRNA and ligating it with the appropriate 568 amino acids. As we covered in the previous section, there appear to be more than enough RNA 569 polymerases per cell to synthesize the needed pool of tRNAs. Without considering the many ways 570 in which amino acids can be scavenged or synthesized *de novo*, we can explore ligation the as a 571 potential rate limiting step. The enzymes which link the correct amino acid to the tRNA, known as 572 tRNA synthetases or tRNA ligases, are incredible in their proofreading of substrates with the incor-

573 rect amino acid being ligated once out of every 10^4 to 10^5 times (BNID: 103469, *Milo et al. (2010)*).
 574 This is due in part to the consumption of energy as well as a multi-step pathway to ligation. While
 575 the rate at which tRNA is ligated is highly dependent on the identity of the amino acid, it is reason-
 576 able to state that the typical tRNA synthetase has charging rate of ≈ 20 AA per tRNA synthetase per
 577 second (BNID: 105279, *Milo et al. (2010)*).

578 Combining these estimates together, as shown schematically in *Figure 8(A)*, yields an estimate
 579 of $\approx 10^4$ tRNA synthetases per cell with a division time of 5000 s. This point estimate is in very close
 580 agreement with the observed number of synthetases (the sum of all 20 tRNA synthetases in *E. coli*).
 581 This estimation strategy seems to adequately describe the observed growth rate dependence of
 582 the tRNA synthetase copy number (shown as the grey line in *Figure 8(B)*), suggesting that the copy
 583 number scales with the cell volume.

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

590 Protein synthesis

591 With the number of tRNA synthetases accounted for, we now consider the abundance of the pro-
 592 tein synthesis machines themselves, ribosomes. Ribosomes are enormous protein/rRNA com-
 593 plexes that facilitate the peptide bond formation between amino acids in the correct sequence
 594 as defined by the coding mRNA. Before we examine the synthesis of the ribosome proteins and
 595 the limits that may place on the observed bacterial growth rates, let's consider replication of the
 596 cellular proteome.

597 As described in the previous section, *E. coli* consists of $\approx 3 \times 10^6$ proteins at a growth rate of \approx
 598 5000 s. If we again assume that each protein is composed of ≈ 300 amino acids and each amino
 599 acid is linked together by one peptide bond, we arrive at an estimate that the cellular proteome
 600 consists of $\approx 10^{10}$ peptide bonds. While the rate at which ribosomes translates is well known to
 601 have a growth rate dependence *Dai et al. (2018)* and is a topic which we discuss in detail in the
 602 coming sections. However, for the purposes of our order-of-magnitude estimate, we can make
 603 the approximation that translation occurs at a rate of ≈ 15 amino acids per second per ribosome
 604 (BNID: 100233, *Milo et al. (2010)*). Under this approximation and assuming a division time of 5000
 605 s, we can arrive at an estimate of $\approx 10^4$ ribosomes are needed to replicate the cellular proteome,
 606 shown in *Figure 8(B)*. This point estimate, while glossing over important details such as chromo-
 607 some copy number and growth-rate dependent translation rates, proves to be notably accurate
 608 when compared to the experimental observations (*Figure 8(B)*).

609 Translation as a growth-rate limiting step

610 Thus far in our work, the general back-of-the-envelope estimates have been reasonably successful
 611 in explaining what sets the scale of absolute protein copy number. In many cases, these estimates
 612 can be adapted to consider a continuum of growth rates in lieu of our single 5000 s point estimate,
 613 the details of which are described in the Supplemental Information. A recurring theme we have
 614 found is the ability of the cell to parallelize different processes, whether they be transport or the
 615 synthesize of certain biomolecules. For example, growth on different carbon sources resulted in
 616 the induced expression of particular transporters, often producing more than needed to acquire
 617 enough carbon to build new cell mass (??(B)). In examining replication of the DNA, we described
 618 how cells can replicate multiple copies of the chromosome at any given time, permitting growth
 619 rates faster than the limit at which the chromosome can be faithfully replicated. As a final example,
 620 we showed how increasing the gene dosage of the rRNA operons is necessary to produce enough
 621 rRNA to form functional ribosomes. However, when it comes to ribosome biogenesis, namely the

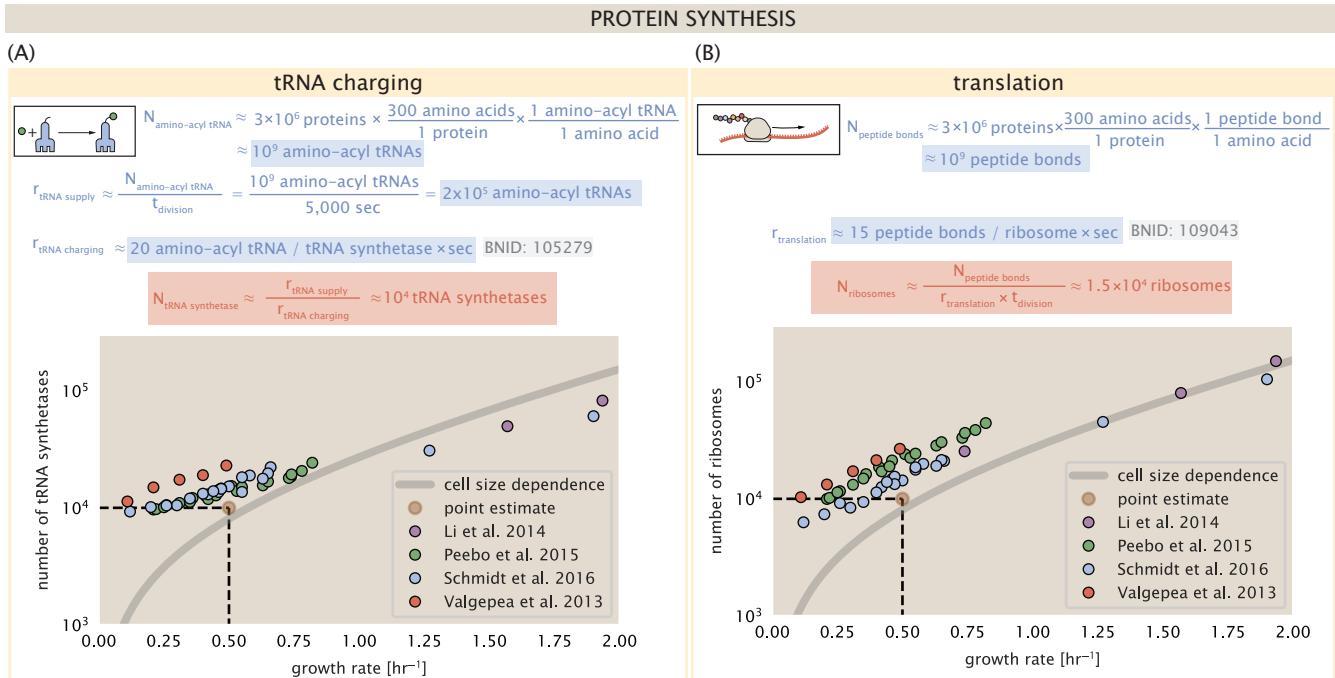


Figure 8. Estimation of the required tRNA synthetases and ribosomes. (A) Estimation for the number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted as a function of growth rate ([ArgS], [CysS], [GlnS], [Glx], [IleS], [LeuS],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

622 translation of ribosomal proteins, such parallelization is not possible, suggesting that translation
623 may be a key factor determining the cellular growth rate.

624 Understanding the allocation of resources and large-scale structure of bacterial proteome has
625 been an area of intense quantitative study over the last decade by Hwa and others (*Scott et al., 2010; Hui et al., 2015*). From the perspective of limiting growth, our earlier estimate of rRNA high-
626 lighted the necessity for multiple copies of rRNA genes in order to make enough rRNA. For *E. coli*'s
627 fastest growth rates at 2 hr⁻¹, the additional demand for rRNA is further supported by parallelized
628 DNA replication and increased rRNA gene dosage. This suggests the possibility that synthesis of
629 ribosomes might become rate limiting.

630 To gain some intuition into how translation may set the speed limit for bacterial growth, we
631 again consider the total number of peptide bonds that must be synthesized, N_{AA} . Noting that cell
632 mass grows exponentially (*Godin et al., 2010*), we can compute the number of amino acids to be
633 polymerized as

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

634 where λ is the cell growth rate in s⁻¹, r_t is the maximum translation rate in amino acids per second,
635 and R is the average ribosome copy number per cell. Knowing the number of peptide bonds to be
636 formed permits us to compute the translation-limited growth rate as

$$\lambda_{\text{translation-limited}} = \frac{r_t R}{N_{\text{AA}}}. \quad (2)$$

637 Alternatively, since N_{AA} is related to the total protein mass through the molecular weight of
638 each protein, we can also consider the growth rate in terms of the fraction of the total proteome

mass that is dedicated to ribosomal protein mass. By making the approximation that an average amino acid has a molecular weight of 110 Da (see **Figure 9(A)**), we can approximate $R/N_{AA} \approx \Phi_R/L_R$, where Φ_R is the ribosomal mass fraction and L_R is the total length in amino acids that make up a ribosome. We can then rewrite the growth rate as,

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

This is plotted as a function of ribosomal fraction Φ_R in **Figure 9(A)**, where we take $L_R \approx 7500$ aa, corresponding to the length in amino acids for all ribosomal subunits of the 50S and 30S complex (BNID: 101175, (Milo et al., 2010)).

The growth rate defined by Equation 3 reflects mass-balance under steady-state growth and has long provided a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function of growth rate (Maaløe, 1979; Scott et al., 2010). For our purposes, we note that there is a maximum growth rate of $\lambda \approx 8 \text{ hr}^{-1}$, or a doubling time just under 6 minutes (**Figure 9(B)**, dashed line). This growth rate can be viewed as an inherent speed limit due to the need for the cell to double its entire ribosomal mass. Interestingly, this limit is independent of the absolute number of ribosomes and is simply given by the time to translate an entire ribosome, L_R/r_t . As shown in **Figure 9(B)**, we can reconcile this with the observation that in order to double the average number of ribosomes, each ribosome must produce a second ribosome. This process cannot be parallelized.

Here we again return to rRNA synthesis, but here consider the maximum rRNA that can be produced at different growth rates.

[expand on.]

Relationship between cell size and growth rate.

While our analysis suggest that ribosomes set the maximum growth rate, our estimates also point to a more important role for ribosomes in setting cell size and in the maintenance of steady-state growth when considered in light of a number of recent studies. In the coming section, we consider the influence of these features on ribosomal biogenesis, beginning with cell size.

The relationship between cell size and growth rate has long been of interest in the study of bacteria, particularly following the now decades-old observation that cell volume appears to increase exponentially with growth rate; known as Schaechter's growth law (Schaechter et al., 1958; Taheri-Araghi et al., 2015). Wild-type *E. coli* growing at relatively fast growth rates, show a remarkably constant cell cycle time t_{cyc} (referring to the C and D periods of DNA replication and cell division, respectively), as shown in **Figure 12(A)** for the data reproduced from (Si et al., 2017). With a constant cell cycle time, the exponential scaling in size has long been considered a direct consequence of cells initiating replication at a constant volume per origin. However, the particular mechanism that governs this relationship, and even the question of whether the change in average cell size is truly exponential have remained under debate (Si et al., 2017; Harris and Theriot, 2018).

Given the constraint prescribed by Equation 3, the simplest strategy to increase growth rate would be to generate more ribosomes. In contrast, it is clear that a large portion of the proteome increases in absolute abundance [GC: I wonder if we need a plot of just proteome size v growth rate. I think just stating this fact is not sufficient]. There is now overwhelming experimental evidence and phenomenological description that cells maintain a linear scaling between size and the number of chromosomal origins of replication, $\langle \# \text{ ori} \rangle$, that is robust to a remarkable array of perturbations (Si et al., 2017). In **Figure 12(B)** we show this linear trend for ribosomes from each of the proteomic data sets. The number of origins $\langle \# \text{ ori} \rangle$ is determined by how often replication forks must be initiated per cell cycle to maintain steady-state growth and grow at rates faster than the chromosome replication time. This quantity is calculated as

$$\langle \# \text{ ori} \rangle \approx 2^{\tau_{cyc}/\tau}, \quad (4)$$

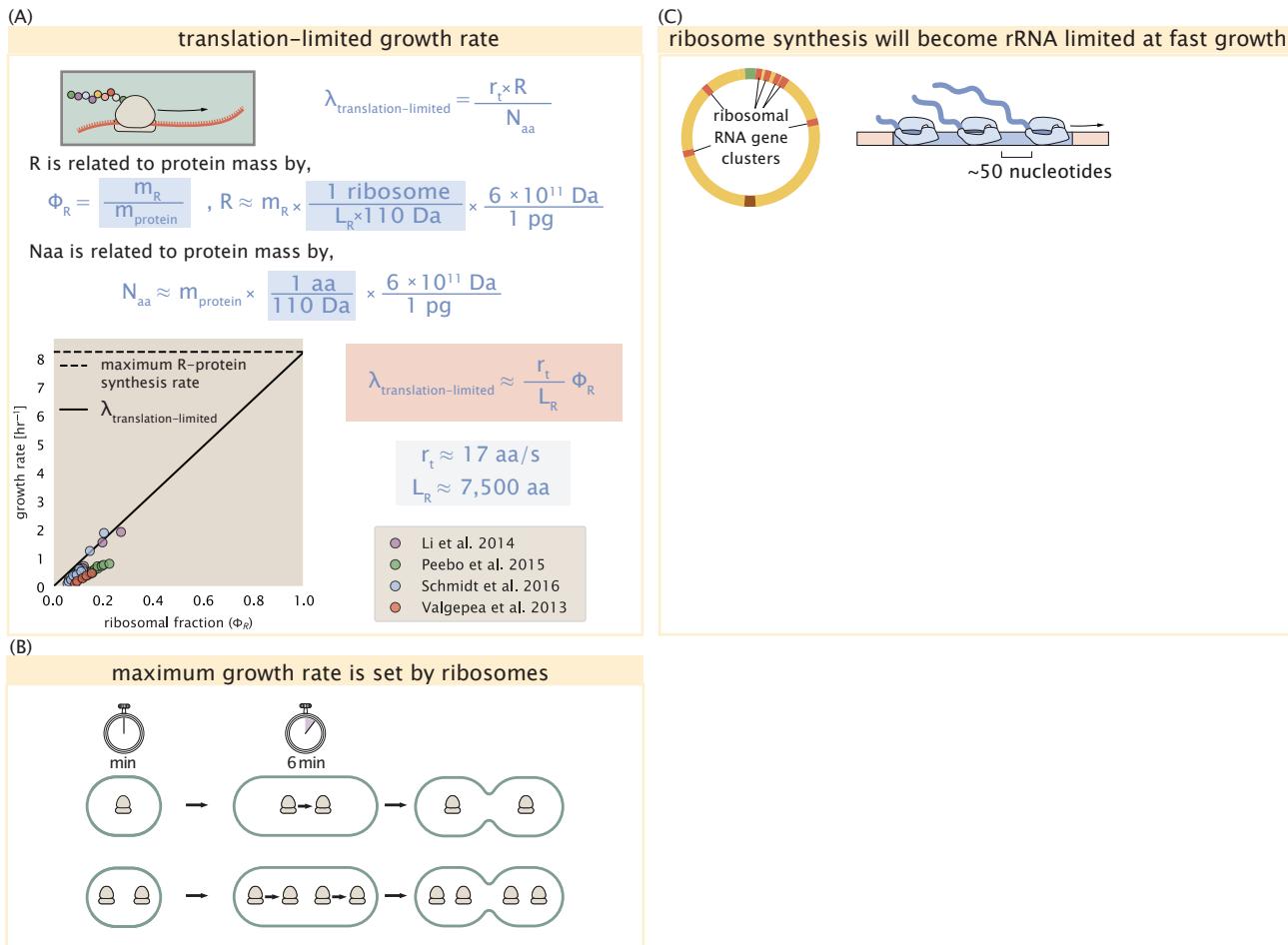


Figure 9. Translation-limited growth rate. (A) Here we consider the translation-limited growth as a function of ribosomal fraction. By mass balance, the time required to double the entire proteome (N_{AA} / r_t) sets the translation-limited growth rate, $\lambda_{\text{translation-limited}}$. Here N_{aa} is effectively the number of peptide bonds that must be translated, r_t is the translation elongation rate, and R is the number of ribosomes. This can also be re-written in terms of the ribosomal mass fraction $\Phi_R = m_R / m_{\text{protein}}$, where m_R is the total ribosomal mass and m_{protein} is the mass of all proteins in the cell. L_R refers to the summed length of the ribosome in amino acids. $\lambda_{\text{translation-limited}}$ is plotted as a function of Φ_R (solid line). (B) The dashed line in part (A) identifies a maximum growth rate that is set by the ribosome. Specifically, this growth rate corresponds to the time required to translate an entire ribosome, L_R / r_t . This is a result that is independent of the number of ribosomes in the cell as shown schematically here. (C)

685 where τ is the doubling time. Here [GC: where?] we have estimated $\langle \# \text{ ori} \rangle$ from the measurements
 686 of *Si et al. (2017)* for wild-type cells grown under nutrient limitation.

687 Through our estimates in the sections on the central dogma, it is apparent that the processes
 688 of transcription (i.e. synthesis of mRNA) and translation are unlikely limiting steps in the process
 689 of doubling the cell mass. In both cases, there is an overabundance of the requisite protein com-
 690 plexes (DNA and RNA polymerase, respectively) and there are mechanisms by which the synthesis
 691 processes can be parallelized. This argues that as DNA replication begins to parallelize, the pro-
 692 teome should only change in ways that reflect the change in DNA gene dosage and mRNA distribu-
 693 tion, and any additional aspects of gene regulation [GC: I don't think I understand the sentence. Are
 694 you just trying to say, "As replication is parallelized, gene dosage increases, which in turn means
 695 that protein expression increases"?]. The total amount of protein synthesized over a cell cycle is
 696 nevertheless determined by $\frac{r_R}{\lambda}$ [GC: I do find it a little confusing going back and forth between λ
 697 and τ . Is there a reason we can't just stick with one for these calculations? I also don't understand
 698 your usage of "nevertheless" here. The preceding sentence is not contradictory.]. Since protein
 699 accounts for most of the cellular dry mass (*Bremer and Dennis, 2008; Basan et al., 2015*), cell size
 700 will also vary in proportional t. The relationship between cell size and growth rate, however, will
 701 then depend only on how the cell varies its ribosomal fraction, as highlighted by Equation 3. [GC:
 702 In re-reading, I think this paragraph is the most confusing for me and, coincidentally, is one of
 703 the most critical in conveying that there is a direct relationship between cell size and proteome
 704 mass/composition. I can try to rewrite this paragraph if you would like, but some clarification on
 705 my earlier points would help.]

706 Exponential relationship between cell size and growth rate is set by ribosomal abundance
 707 at moderate to fast growth rates.

708 It is notable that the majority of ribosomal proteins and rRNA operons are found closer to the DNA
 709 origin. For a relatively constant cell cycle time t_{cyc} and in particular, for a constant DNA replication
 710 period t_C , parallelized DNA replication has the important consequence that it will skew absolute
 711 gene dosage and mRNA abundance in favor of genes closer to the origin (*Scholz et al., 2019*) [more
 712 cites] (*Figure 12(C)*). This raises the possibility that for moderately fast growth rates (above about
 713 0.5 hr⁻¹ [GC: why this cutoff? Seems to come out of the blue a bit.]), where these time scales are
 714 nearly constant, that the increased number of chromosomal origins can be viewed as a way for the
 715 cell to skew its ribosomal abundance. Importantly, alternative solutions such as simply increasing
 716 the expression of ribosomes for a particular $\langle \# \text{ ori} \rangle$ will not be possible if rRNA synthesis is nearly
 717 limiting. Biasing DNA dosage in this manner also provides a means to maintain levels of expression
 718 for most other proteins that are nonetheless needed for growth, as we've seen throughout our
 719 estimates.

720 We were unaware, however, of whether such the skew in gene dosage at fast growth material-
 721 izes at the proteomic level. In *Figure 12(D)* we show a running boxcar average (500 kbp window)
 722 of protein copy number as a function of each gene's transcriptional start site (*Figure 12(D)*) [GC: In
 723 rereading, I think this point would be made much more clearly if we did remake Fig. 10 D to start
 724 on the left size at oric. I know that's not the same as the 'genomic position', but I thikn it would help.
 725 Another option is a radial diagram if you want to show the symmetry.]. While the protein copy num-
 726 bers of individual proteins can vary substantially across the entire chromosome, we nonetheless
 727 observe a bias in expression under fast growth conditions (dark blue lines). The dramatic change
 728 in protein copy number near the origin is primarily due to the increase in ribosomal protein ex-
 729 pression. This trend is in contrast to slower growth conditions (yellow) where the average copy
 730 number is much more uniform across the length of the chromosome.

731 This result provides important evidence that although total protein content scales with $\langle \# \text{ ori} \rangle$,
 732 parallel DNA replication also skews the proteome in favor of genes closer to the origin and in par-
 733 ticular supports the synthesis of additional ribosomes. We can then view an increase in ribosomal
 734 fraction Φ_R (and therefore, λ) as requiring a geometric increase in total protein abundance that

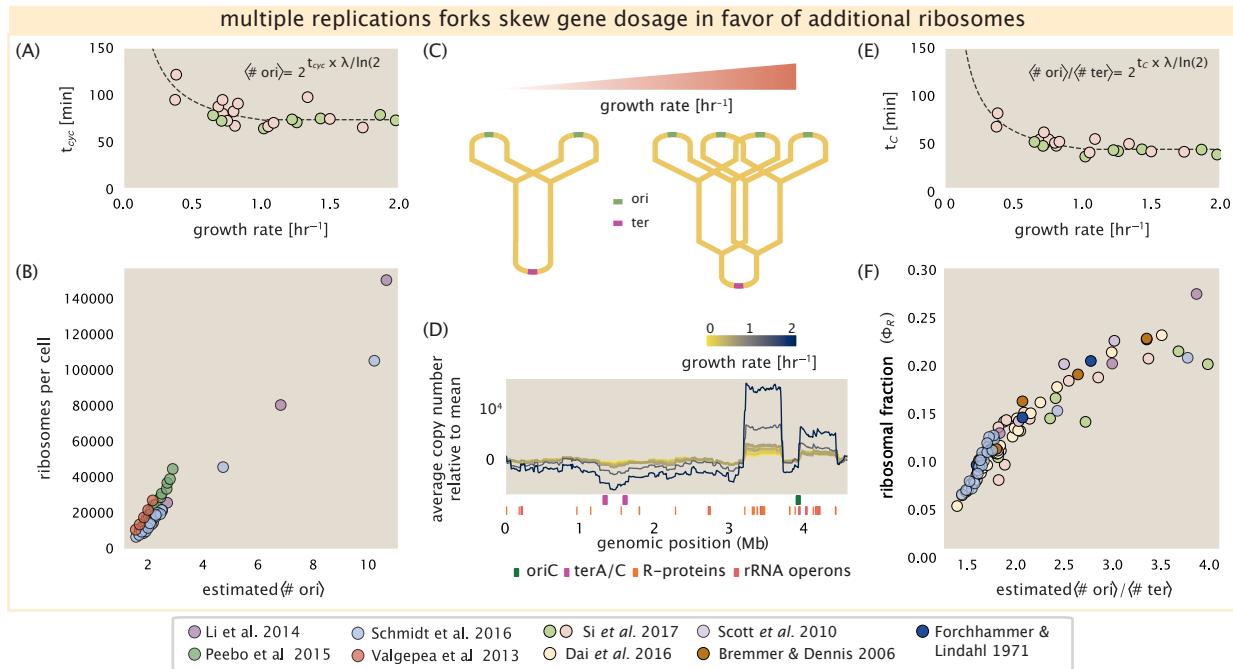


Figure 10. Multiple replication forks skew gene dosage and ribosomal content. (A) Experimental data from Si *et al.* (2017). Dashed line shows fit to the data, which were used to estimate ($\# \text{ori}$). t_{cyc} was assumed to vary in proportion to τ for doubling times great than 40 minutes, and then reach a minimum value of [fill] minutes below this (see Supplemental Appendix X for additional details). Red data points correspond to measurements in strain MG1655, while light green points are for strain NCM3722. (B) Plot of the ribosome copy number estimated from the proteomic data against the estimated ($\# \text{ori}$) [NB: change to total protein abundance?]. (C) Schematic shows the expected increase in replication forks (or number of ori regions) as *E. coli* cells grow faster. (D) A running boxcar average of protein copy number is calculated for each each growth condition considered by (Schmidt *et al.*, 2016). A 0.5 Mb averaging window was used. Protein copy numbers are reported relative to their condition-specific means in order to center all data sets. (E) Experimental data from Si *et al.* (2017) showing t_c as a function of growth rate. Dashed line shows a best fit to the data, similar to part (A). (F) Plot compares our estimate of ($\# \text{ori} / \# \text{ter}$) to the experimental measurements of ribosomal abundance. Ribosomal fraction was approximated from the RNA/protein ratios of Dai *et al.* (2016) (yellow) and Si *et al.* (2017) (light red and light green) by the conversion RNA/protein ratio $\approx \Phi_R \cdot 2.1$.

735 is proportional to $\langle \# \text{ ori} \rangle$. For *E. coli*, which initiates chromosomal replication at fixed volume per
 736 origin [cite], it then follows that the growth rate λ will only increase with an exponential increase
 737 in total protein.

738 An exponential increase in chromosomal content provides shows a diminishing benefit
 739 on growth.

740 [GC: If we want to go down this route, then I think a more conceptual diagram is absolutey critical.
 741 In fact, I think we may want a final "summary" figure for the discussion which illustrates this point. I
 742 will think about what that may look like. .] The ratio $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ can also be used to consider this
 743 skew in chromosomal content. Quantitatively, this will depend on how quickly the chromosome is
 744 replicated (i.e. its C period) relative the cell's doubling time τ and this is given by $2^{\tau_C/\tau}$. In **Figure 12(C)**
 745 we plot the measured τ_C versus τ (computed as $\tau = \log(2)/\lambda$), similarly using measurements on wild-
 746 type *E. coli* from (*Si et al., 2017*). In **Figure 12(D)** we plot the $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ ratio against ribosomal
 747 fraction across a number of recent measurements. Here we see that the ribosomal fraction doesn't
 748 increase as much at higher $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$. This may be in part because not all rRNA operons are
 749 located exactly at the origin. However, it also serves to highlight a diminishing benefit on growth
 750 rate with additional rounds of chromosomal initiation, which likely reflects difficulty in increasing
 751 the absolute number of ribosomes at fast growth.

752 **Growth in poor nutrient conditions.**

753 While the above results suggest that it is the need to increase the number of ribosomes that sets
 754 an exponential scaling in cell size, this relationship is likely to falter at slow growth rates (below
 755 about 0.5h^{-1}). In this regime ribosomal abundance R no longer reflects the cell's protein synthesis
 756 capacity. Although R still appears to scale with $\langle \# \text{ ori} \rangle$, leading to a minimum ribosomal fraction of
 757 about 0.06, additional regulatory control through the small-molecule alarmones (p)ppGpp reduces
 758 the fraction of actively translating ribosomes at slow growth (*Dai et al., 2016; Bosdriesz et al., 2015;*
 759 *Zhu and Dai, 2019*). In this section we consider the consequence of having excess ribosomes on
 760 maintaining steady-state growth in poor nutrient conditions.

761 The challenge here lies in the ability of the cell to maintain homeostasis when consumption
 762 of amino acids for protein synthesis might exceed the rate of supply if all ribosomes were ac-
 763 tively translating **Figure 13A**. Without additional regulatory control, this would prevent continuous
 764 growth, and indeed for (p)ppGpp null strains, cells only grow in minimal media if additional amino
 765 acid supplements are present. In contrast, wild-type *E. coli* are able to maintain a relatively high
 766 elongation rate even in stationary phase ($\approx 8 \text{ AA/s}$, (*Dai et al., 2016, 2018*)).

767 Mitigation of ribosome activity helps maintain homeostasis in poor nutrient conditions
 768 To better understand how regulation of ribosomes influence growth rate in this slow growth regime
 769 we consider a coarse-grained model that relates elongation rate to a limiting supply of amino acids,
 770 which for simplicity we treat as a single, effective rate-limiting species. Under such a scenario, the
 771 elongation rate can be described as simply depending on the maximum elongation rate ($r_t^{\max} \approx$
 772 17.1 aa/s , (*Dai et al., 2016, 2018*)), an effective binding constant K_d , and the limiting amino acid
 773 concentration $[AA]_{eff}$, [GC: I think we need to explain how to interpret these effective parameters.
 774 This is a section that needs to be a bit more pedagogical. I haven't touched anything here as I'm
 775 not 100% sure I can accurately summarize the parameters!]

$$r_t = r_t^{\max} \cdot \frac{1}{1 + K_d/[AA]_{eff}}. \quad (5)$$

776 For cells growing in minimal medium supplemented with glucose, the amino acid concentration
 777 is of order 100 mM (BNID: 110093, (*Milo et al., 2010; Bennett et al., 2009*)). With a growth rate of
 778 about 0.6 hr^{-1} and elongation rate of 12.5 amino acids per second (*Dai et al., 2016*), we can esti-
 779 mate an effective K_d of about 40 mM [GC: This kind of comes out of the blue.]. The maintenance

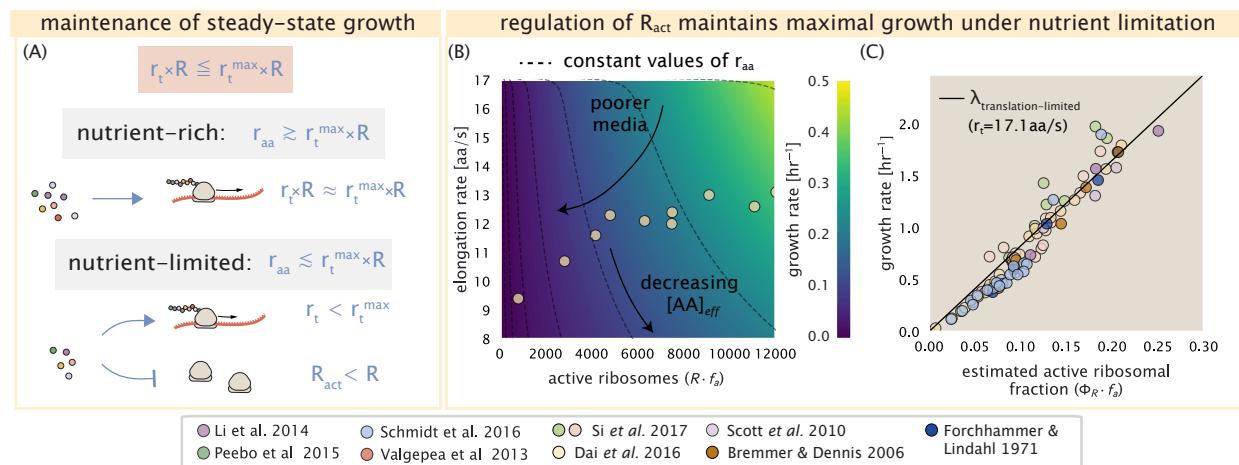


Figure 11. *E. coli* must regulate ribosomal activity in limiting nutrient conditions. (A) Schematic showing translation-specific requirements for maintenance of steady-state growth. In a nutrient rich environment, amino acid supply r_{aa} is sufficiently in excess of the demand by ribosomes translating at their maximal rate. In poorer nutrient conditions, reduced amino acid supply r_{aa} will decrease the rate of elongation. In a regime where r_{aa} is less than $r_t \cdot R$, the number of actively translating ribosomes will need to be reduced in order to maintain steady-state growth. (B) Translation elongation rate is plotted as a function of the number of actively translating ribosomes $R \cdot f_a$. Dashed lines correspond to a range of amino acid synthesis rates r_{aa} , from 10^3 to 10^6 . Growth rates are calculated according to Equation 1, assuming a constant ribosomal fraction of 8 percent. See appendix XX for additional details. (C) Experimental data from Dai *et al.* are used to estimate the fraction of actively translating ribosomes. The solid line represents the translation-limited growth rate for ribosomes elongating at 17.1 AA/s.

780 of this amino acid pool $[AA]_{eff}$ will depend on the difference between the synthesis/supply rate
 781 of amino acids r_{AA} and consumption by ribosomes $r_t R f_a$, where we use f_a to account for the pos-
 782 sible reduction of actively translating ribosomes (see Supplemental Appendix XX for a complete
 783 description of this model [GC: Maybe the complete description should just go here.]).

784 In **Figure 13(B)** we plot the growth rate and elongation rate as a function of the number of
 785 actively translating ribosomes. If we consider constant values of amino acid synthesis rate r_{AA}
 786 (dashed lines) to reflect the available parameter space for a specific growth condition, cells will
 787 grow fastest by maximizing their fraction of actively translating ribosomes. When we consider the
 788 experimental measurements from **Dai et al. (2018)** (yellow circles), which reflect growth in different
 789 nutrient conditions, we see that although cells reduce $R \times f_a$ in poorer nutrient conditions, they
 790 do so in a way that keeps $[AA]_{eff}$ relatively constant. Given our estimate for the K_D of 40 mM,
 791 we would only expect a decrease from 100 mM to about 35 mM in the slowest growth conditions.
 792 While experimental data is scarce, amino acid concentrations only decrease to about 60 mM for
 793 cells grown in minimal media + acetate ($\lambda \approx 0.3 \text{ hr}^{-1}$ in our proteomic data) (**Bennett et al., 2009**),
 794 qualitatively consistent with our expectations. One explanation for the experimental data then is
 795 that the cell is regulating ribosome activity in order to maintain a sufficient pool of amino acids
 796 for growth. Any further increase in $R \times f_a$ at constant r_{AA} would otherwise be associated with an
 797 additional drop in cellular amino acids concentrations.

798 *E. coli* maximizes its steady-state growth rate by tuning both ribosomal content and ac-
 799 tivity.

800 Using the active fraction f_a measurements across a broad range of nutrient-limited growth con-
 801 ditions from the work of **Dai et al. (2016)**, we also estimated the active fraction of ribosomal protein
 802 across our collated data (**Figure 13(C)**). Importantly, we find that across all growth conditions con-
 803 sidered, cells appear to maintain a growth rate consistent with Equation 3 with an elongation rate
 804 of $r_t \approx 17.1 \text{ aa/s}$. While somewhat counter intuitive given that ribosomes translate at almost half

805 this rate in the poorest of growth conditions, it is because cells tune $r_t \times R \times f_a$ that they're able
 806 achieve these steady-state growth rates over such a broad range of conditions.

807 Recently it was shown that growth in a (p)ppGpp null strain abolishes both the scaling in cell
 808 size and the $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ ratio. Instead, cells exhibited a high $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ closer to 4 and
 809 cell size more consistent with a fast growth state where (p)ppGpp levels are low (*Fernández-Coll*
 810 *et al.*, 2020) and high ribosomal fraction (*Zhu and Dai*, 2019). This raises the possibility that the
 811 action of (p)ppGpp is also mediating growth control and size scaling over the entire range of growth
 812 conditions. Specifically, in rich media, it is through additional rounds of DNA replication initiation
 813 that *E. coli* increases ribosomal abundance to increase growth rate, while growth in poor nutrient
 814 conditions is achieved through mitigation of ribosome activity.

815 Discussion

816 [Fill in.]

817 Maximizing growth rate requires coordination of biosynthesis at all growth rates.

818 However, the mechanism behind growth rate control has remained elusive and has only been
 819 described at a phenomenological level.

820 Here we attempt to place our observations across the proteomic data sets in the context of *E.*
 821 *coli* maximizing its steady-state growth rate across a wide array of conditions.

822 Parallel DNA replication biases gene dosage in support of ribosome synthesis.

823 *E. coli* cells grow by a so-called "adder" mechanism, whereby cells add a constant volume with
 824 each cell division (*Taheri-Araghi et al.*, 2015). In conjunction with this, additional rounds of DNA
 825 replication are triggered when cells reach a critical volume per origin of replication (*Figure 12(A)*).
 826 This leads to the classically-described exponential increase in cell size with growth rate *Schaechter*
 827 *et al.* (1958); *Si et al.* (2017, 2019). In the context of maximizing growth rate, it is notable that the
 828 majority of ribosomal proteins and rRNA operons are found closer to the DNA origin.

829 While an increase in transcription has been observed for genes closer to the origin in rapidly
 830 growing *E. coli* (*Scholz et al.*, 2019), we were unaware of such characterization at the proteomic
 831 level. In order to see whether there is a relative increase in protein expression for genes closer to
 832 the origin at faster growth, we calculated a running boxcar average (500 kbp window) of protein
 833 copy number as a function of each gene's transcriptional start site (*Figure 12(B)*). While absolute
 834 protein copy numbers can vary substantially across the chromosome, we indeed observe a bias in
 835 expression under fast growth conditions (dark blue), showing the result. The dramatic change in
 836 protein copy number near the origin is primarily due to the increase in ribosomal protein expres-
 837 sion. This trend is in contrast to slower growth conditions (yellow) where the average copy number
 838 is more uniform across the length of the chromosome.

839 If ribosomal genes (rRNA and ribosomal proteins) are being synthesized at their maximal rate
 840 according to their rRNA gene dosage and maximal transcription rate, we can make two related
 841 hypotheses about how their ribosome abundance should vary with chromosomal content. First,
 842 the ribosomal protein fraction should increase in proportion to the average ratio of DNA origins to
 843 DNA termini ($\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ ratio). This is a consequence of the skew in DNA dosage as cells grow
 844 faster. The second hypothesis is that the absolute number of ribosomes should increase with the
 845 number of DNA origins ($\langle \# \text{ ori} \rangle$), since this will reflect the total gene dosage at a particular growth
 846 condition.

847 In order to test each of these expectations we considered the experimental data from *Si et al.*
 848 (2017), which inferred these parameters for cells under nutrient-limited growth. The ratio $\langle \# \text{ ori} \rangle /$
 849 $\langle \# \text{ ter} \rangle$ depends on how quickly chromosomes are replicated relative the cell's doubling time τ and
 850 is given by $2^{\tau_C/\tau}$. Here τ_C is the time taken to replicate *E. coli*'s chromosome, referred to as the C
 851 period of cell division. In *Figure 12(C)* we plot the measured τ_C versus τ (computed as $\tau = \log(2)/\lambda$),
 852 with data points in red corresponding to *E. coli* strain MG1655, and blue to strain NCM3722. *Si*

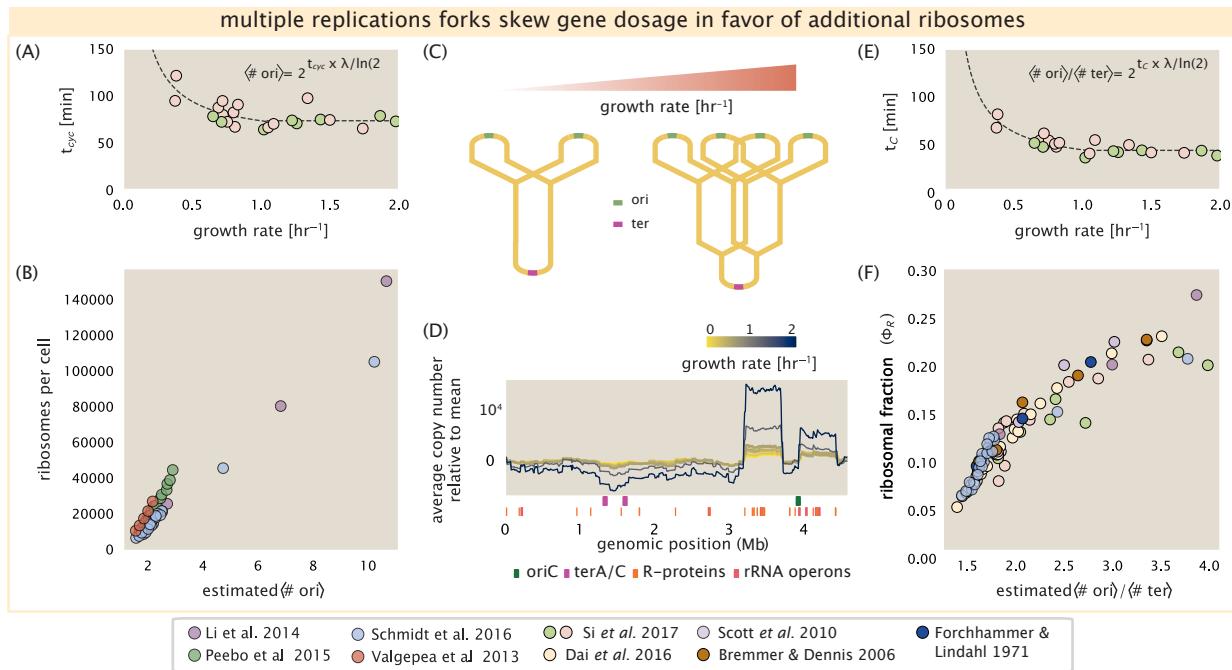


Figure 12. Multiple replication forks skew gene dosage and ribosomal content. (A) Schematic shows the expected increase in replication forks (or number of ori regions) as *E. coli* cells grow faster. (B) A running boxcar average of protein copy number is calculated for each each growth condition considered by Schmidt *et al.*. A 0.5 Mb averaging window was used. Protein copy numbers are reported relative to their condition-specific means in order to center all data sets. (C) and (E) show experimental data from Si *et al.* (2017) Solid lines show fits to the data, which were used to estimate $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ and $\langle \# \text{ori} \rangle$ [NB: to note fit equations]. Red data points correspond to measurements in strain MG1655, while light green points are for strain NCM3722. (D) Plot compares our estimate of $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$ to the experimental measurements of ribosomal abundance. Ribosomal fraction was approximated from the RNA/protein ratios of Dai *et al.* (2016) (yellow) and Si *et al.* (2017) (light red and light green) by the conversion RNA/protein ratio $\approx \Phi_R \cdot 2.1$. (F) Plot of the ribosome copy number estimated from the proteomic data against the estimated $\langle \# \text{ori} \rangle$.

853 *et al.* (2017) also measured the total RNA to protein ratio which reflects ribosomal abundance and
 854 we show that data along with other recent measurements from *Dai et al.* (2016, 2018). Indeed, we
 855 find that the ribosomal fraction increases with $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ (*Figure 12(C)*). We note a systematic
 856 difference in the relative abundances from *Peebo et al.* (2015) and *Valgepea et al.* (2013) that was
 857 inconsistent with a number of other measurements of total RNA-to-protein ratios ($\approx \Phi_R \times 2.1$ *Dai*
 858 *et al.* (2016)) and only show the data from *Schmidt et al.* (2016) and *Li et al.* (2014) for relative
 859 ribosome abundances (see supplemental section XX for a more complete discussion). For the data
 860 shown, the ribosomal fraction doesn't increase as much at higher $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$. Since several
 861 rRNA operons are actually located approximately half-way between the origin and terminus, the
 862 trend may in part be a consequence of a diminishing increase in rRNA gene dosage at higher $\langle \#$
 863 $\text{ori} \rangle / \langle \# \text{ ter} \rangle$ ratios.

864 We can similarly estimate $\langle \# \text{ ori} \rangle$, which depends on how often replication forks are initiated
 865 per cell cycle. This is given by the number of overlapping cell cycles, $2^{\tau_{\text{cyc}}/\tau}$, where τ_{cyc} refers to
 866 the total time of chromosome replication and cell division. *Figure 12(E)* shows the associated data
 867 from *Si et al.* (2019), which we use to estimate $\langle \# \text{ ori} \rangle$ for each growth condition of the proteomic
 868 data. In agreement with our expectations, we find that ribosome copy number increases with the
 869 estimated $\langle \# \text{ ori} \rangle$ (*Figure 12(F)*).

870 While it is difficult to distinguish between causality and correlation, the data is consistent with
 871 the need for cells to increase their effective rRNA gene dosage in order to grow according to the
 872 constraint set by Equation 2. These results may also shed some light on the notable increase
 873 in ribosomal content that is observed when sublethal doses of antibiotics (*Scott et al.*, 2010; *Dai*
 874 *et al.*, 2016). Specifically, if rRNA synthesis is rate limiting, and nutrient conditions largely dictate the
 875 extent of overlapping DNA replication cycles, than addition of antibiotic will lengthen the doubling
 876 time and allow an increased rRNA synthesis relative to the rate of cell division. In Supplemental
 877 Section XX, we consider this further using additional data from *Si et al.* (2017).

878 Regulation of translating ribosomes helps maintain maximal growth according to nutrient-
 879 availability.

880 While the above observations show how *E. coli* can vary its ribosomal content to increase growth
 881 rate, it also presents a challenge in the limit of poorer nutrient conditions. Recall from Equation 3
 882 that ribosomal content should decrease to zero as growth decreases to zero. While bacteria tend to
 883 decrease their ribosomal abundance in poorer nutrient conditions, they do so only to some fixed,
 884 non-zero amount (*Scott et al.*, 2010; *Liebermeister et al.*, 2014). Here we find a minimal ribosomal
 885 fraction of ≈ 0.06 in the slowest growth conditions. From the perspective of a bacterium dealing
 886 with uncertain nutrient conditions, there is likely a benefit for the cell to maintain some relative
 887 fraction of ribosomes to support rapid growth as nutrient conditions improve.

888 The challenge however, lies in the cell's ability to maintain growth when ribosomes are in excess
 889 of the rate that nutrients can be harvested and amino acids synthesized for consumption
 890 (*Figure 13A*). In the limit of poor growth conditions, ribosomes would consume their amino acid
 891 supply and be unable to maintain steady-state growth. In reality, *E. coli* is still able to maintain a
 892 relatively high elongation rate even in stationary phase ($\approx 8 \text{ AA/s}$, (*Dai et al.*, 2016, 2018)). A explanation
 893 for this is that the cell further regulates its biological activity in conditions of stress and
 894 nutrient-limitation; in particular through the small-molecule alarmones (p)ppGpp (*Harris and The-*
 895 *riot*, 2018). In (p)ppGpp null strains, cells are unable to grow in nutrient-poor media. Indeed, these
 896 small molecules play a role in controlling biosynthesis rates throughout the central dogma [NB
 897 citations]. Here we explore this further in the context of growth by maximizing protein synthesis.

898 We consider slow growth conditions (λ less than 0.5 hr^{-1}) by assuming that the decrease in
 899 elongation rate is due to a limiting supply of amino acids and a need for the cell to maintain ex-
 900 cess nutrients for cellular homeostasis under steady-state growth. There is some experimental
 901 support showing that in poorer nutrient growth conditions, cells have lower amino acids concen-
 902 trations (*Bennett et al.*, 2009). We proceed by coarse graining the cell's amino acid supply as an

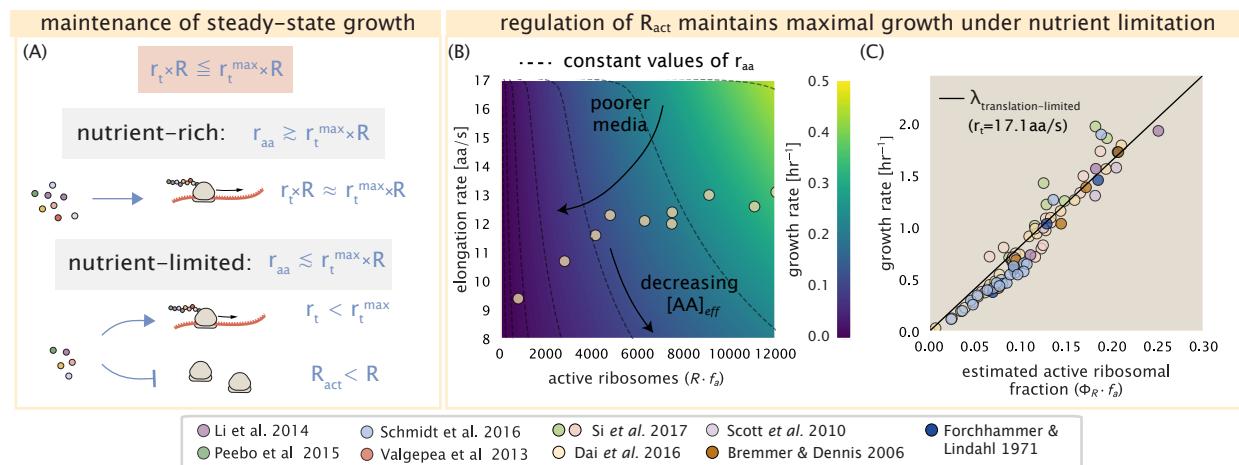


Figure 13. *E. coli* must regulate ribosomal activity in limiting nutrient conditions. (A) Schematic showing translation-specific requirements for maintenance of steady-state growth. In a nutrient rich environment, amino acid supply r_{aa} is sufficiently in excess of the demand by ribosomes translating at their maximal rate. In poorer nutrient conditions, reduced amino acid supply r_{aa} will decrease the rate of elongation. In a regime where r_{aa} is less than $r_t \cdot R$, the number of actively translating ribosomes will need to be reduced in order to maintain steady-state growth. (B) Translation elongation rate is plotted as a function of the number of actively translating ribosomes $R \cdot f_a$. Dashed lines correspond to a range of amino acid synthesis rates r_{aa} , from 10^3 to 10^6 . Growth rates are calculated according to Equation 1, assuming a constant ribosomal fraction of 8 percent. See appendix XX for additional details. (C) Experimental data from Dai et al. are used to estimate the fraction of actively translating ribosomes. The solid line represents the translation-limited growth rate for ribosomes elongating at 17.1 AA/s.

single, effective rate-limiting species (see Supplmental Section XX for a more complete discussion). Under such a scenario, the elongation rate can described as simply depending on the maximum elongation rate (≈ 17.1 aa/s, (Dai et al., 2016, 2018)), an effective K_d , and the limiting amino acid concentration $[AA]_{eff}$. Specifically, the elongation rate is given by,

$$r_t = r_t^{\max} \cdot \frac{1}{1 + K_d/[AA]_{eff}}. \quad (6)$$

For cells growing in minimal media + glucose, the amino acid concentration is of order 100 mM (BNID: 110093, (Milo et al., 2010; Bennett et al., 2009)). With a growth rate of about 0.6 hr^{-1} and elongation rate of 12.5 aa per second (Dai et al., 2016), we can estimate an effective K_d of about 40 mM. Ultimately the steady state amino acid concentration will depend on the difference between the supply of amino acids r_{aa} and consumption by ribosomes $r_t \cdot R \cdot f_a$, where f_a accounts for the possible reduction of actively translating ribosomes.

In Figure 13B we consider how the maximal growth rate and elongation rates vary as a function of the number of actively translating ribosomes in this slow growth regime (see Supplmental Section XX for a complete description of this model). If we consider r_{AA} to be reflective of a specific growth condition, by considering lines of constant r_{AA} , we find that cells grow fastest by maximizing their fraction of actively translating ribosomes. When we consider the experimental measurements from Dai et al. (2018), we see that although cells indeed reduce $R \times f_a$, they do so in a way that keeps $[AA]_{eff}$ relatively constant. Given our estimate for the K_d of 40 mM, we would only expect a decrease from 100 mM to about 35 mM in the slowest growth conditions. While experimental data is limited, amino acid concentrations only decrease to about 60 mM for cells grown in minimal media + acetate ($\lambda = 0.3$ hr^{-1} in our proteomic data; value obtained from Bennett et al. (2009)), qualitatively consistent with our expectations.

Given the quantitative data from Dai et al. (2018), which determined f_a across the entire range of growth rates across our data, we next estimated the active fraction of ribosomal protein. As

926 shown in **Figure 13(C)**, we find that cells grow at a rate near the expected translation maximum
927 expected from Equation 1, using the maximum elongation rate of $r_e = 17.1$ aa per second. This is in
928 contrast to the reality that ribosomes are translating at almost half this rate in the poorest growth
929 conditions. This highlights that there are alternative ways to grow according to the translated-
930 limited growth rate that is expected based with ribosomes translating at their maximal elongation
931 rate. Specifically, it is by adjusting $r_e \times R \times f_a$ to match maximal growth rate set by Equation 2, through
932 the parameters $r_{tmax} \times R'$, that cells are able to maximize their growth rate under steady-state.

933 Global regulatory control across central dogma may provide an explanation for the ro-
934 bust scaling laws in *E. coli*.

935 A number of recent papers further highlight the possibility that (p)ppGpp may even provide a causal
936 explanation for the scaling laws in *E. coli*. In the context of ribosomal activity, increased levels of
937 (p)ppGpp are associated with lower ribosomal content, and at slow growth appear to help reduce the
938 fraction of actively translating ribosomes (*Dai et al., 2016, 2018*). Titration of the cellular (p)ppGpp
939 concentrations (up or down) can invoke similar proteomic changes reminiscent of those observed
940 under nutrient limitation (*Zhu and Dai, 2019*). In light of the limiting dependence of ribosome copy
941 number on chromosomal gene dosage, it was recently shown that growth in a (p)ppGpp null strain
942 abolishes both the scaling in cell size and the $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ ratio. Instead, cells exhibited a high
943 $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$ closer to 4 and cell size more consistent with a fast growth state where (p)ppGpp
944 levels are low (*Fernández-Coll et al., 2020*.)]

945 [NB, expand on to consider how activity of RNAP and other aspects(?) may follow a similar
946 behaviour and are under related control mechanisms.]

947 References

- 948 Abelson, H., Johnson, L., Penman, S., and Green, H. (1974). Changes in RNA in relation to growth of the fibroblast:
949 II. The lifetime of mRNA, rRNA, and tRNA in resting and growing cells. *Cell*, 1(4):161–165.
- 950 Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose
951 sugars in *Escherichia coli*. *BMC Systems Biology*, 8(1):133.
- 952 Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes
953 studied by ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 954 Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S.,
955 Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E.,
956 Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*,
957 25(1):25–29.
- 958 Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B.
959 (2000). A Model for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends:
960 DNA Triggers A Change In Binding Specificity of the γ Complex Clamp Loader. *Journal of Biological Chemistry*,
961 275(4):3006–3015.
- 962 Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4
963 as a NH₃ Channel. *Journal of Biological Chemistry*, 291(36):19184–19195.
- 964 Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by
965 increased protein synthesis. *Molecular Systems Biology*, 11(10):836.
- 966 Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and
967 Bioengineering*, 18(1):81–94. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.260180107>.
- 968 Belliveau, N. M., Barnes, S. L., Ireland, W. T., Jones, D. L., Sweredoski, M. J., Moradian, A., Hess, S., Kinney, J. B.,
969 and Phillips, R. (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional
970 regulation in bacteria. *Proceedings of the National Academy of Sciences*, 115(21):E4796–E4805.
- 971 Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009). Absolute
972 metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical
973 Biology*, 5(8):593–599.
- 974 Birnbaum, L. S. and Kaplan, S. (1971). Localization of a Portion of the Ribosomal RNA Genes in *Escherichia coli*.
975 *Proceedings of the National Academy of Sciences*, 68(5):925–929.
- 976 Booth, I. R., Mitchell, W. J., and Hamilton, W. A. (1979). Quantitative analysis of proton-linked transport systems.
977 The lactose permease of *Escherichia coli*. *Biochemical Journal*, 182(3):687–696.
- 978 Bosdriesz, E., Molenaar, D., Teusink, B., and Bruggeman, F. J. (2015). How fast-growing bacteria robustly tune
979 their ribosome concentration to approximate growth-rate maximization. *The FEBS Journal*, 282(10):2029–
980 2044.
- 981 Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at
982 Different Exponential Growth Rates. *EcoSal Plus*, 3(1).
- 983 Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria.
984 *Nature Reviews Microbiology*, 14(10):638–650.
- 985 Cox, G. B., Newton, N. A., Gibson, F., Snoswell, A. M., and Hamilton, J. A. (1970). The function of ubiquinone in
986 *Escherichia coli*. *Biochemical Journal*, 117(3):551–562.
- 987 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018).
988 Slowdown of Translational Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*,
989 9(1):281.
- 990 Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P.,
991 and Hwa, T. (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates
992 during slow growth. *Nature Microbiology*, 2(2):16231.
- 993 Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli*
994 phosphoenolpyruvate–carbohydrate phosphotransferase system: Peculiarities of regulation and impact on
995 growth and product formation. *Applied Microbiology and Biotechnology*, 94(6):1483–1494.

- 996 Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis,
 997 V., and Palsson, B. Ø. (2007). A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that
 998 accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology*, 3(1):121.
- 999 Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M.,
 1000 and Dunny, G. M. (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA
 1001 Synthesis Independent of Growth Rates. *mBio*, 11(2):45.
- 1002 Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A multi-DNA
 1003 polymerase affair. *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 1004 Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*, 42(1):241–
 1005 263.
- 1006 Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J. S.,
 1007 Alquicira-Hernández, K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-
 1008 Lira, H., Bonavides-Martínez, C., Pérez-Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes,
 1009 A., Hernández-Koutoucheva, A., Moral-Chávez, V. D., Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB
 1010 version 9.0: High-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic
 1011 Acids Research*, 44(D1):D133–D143.
- 1012 Ge, J., Yu, G., Ator, M. A., and Stubbe, J. (2003). Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli* Class
 1013 I Ribonucleotide Reductase. *Biochemistry*, 42(34):10071–10083.
- 1014 Godin, M., Delgado, F. F., Son, S., Grover, W. H., Bryan, A. K., Tzur, A., Jorgensen, P., Payer, K., Grossman, A. D.,
 1015 Kirschner, M. W., and Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature
 1016 Methods*, 7(5):387–390.
- 1017 Goldman, S. R., Nair, N. U., Wells, C. D., Nickels, B. E., and Hochschild, A. (2015). The primary σ factor in *Es-
 1018 cherichia coli* can access the transcription elongation complex from solution *in vivo*. *eLife*, 4:e10514.
- 1019 Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogen-
 1020 esis. *Trends in microbiology*, 26(10):815–832.
- 1021 Harris, R. M., Webb, D. C., Howitt, S. M., and Cox, G. B. (2001). Characterization of PitA and PitB from *Escherichia
 1022 coli*. *Journal of Bacteriology*, 183(17):5008–5014.
- 1023 Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and
 1024 elemental content of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 1025 Henkel, S. G., Beek, A. T., Steinsiek, S., Stagge, S., Bettenbrock, K., de Mattos, M. J. T., Sauter, T., Sawodny, O.,
 1026 and Ederer, M. (2014). Basic Regulatory Principles of *Escherichia coli*'s Electron Transport Chain for Varying
 1027 Oxygen Conditions. *PLoS ONE*, 9(9):e107640.
- 1028 Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015).
 1029 Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular
 1030 Systems Biology*, 11(2):e784–e784.
- 1031 Ingledew, W. J. and Poole, R. K. (1984). The respiratory chains of *Escherichia coli*. *Microbiological Reviews*,
 1032 48(3):222–271.
- 1033 Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B.,
 1034 and Phillips, R. (2020). Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a
 1035 time. *bioRxiv*.
- 1036 Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular
 1037 Biology*, 3(3):318–356.
- 1038 Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in *Caulobacter crescentus*.
 1039 *The EMBO journal*, 20(17):4952–4963.
- 1040 Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent
 1041 progress, and the future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- 1042 Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonjia, E., Mekler, V.,
 1043 Ebright, R. H., and Weiss, S. (2005). Retention of Transcription Initiation Factor Σ 70 in Transcription Elonga-
 1044 tion: Single-Molecule Analysis. *Molecular Cell*, 20(3):347–356.

- 1045 Khademi, S., O'Connell, J., Remis, J., Robles-Colmenares, Y., Miercke, L.J. W., and Stroud, R. M. (2004). Mechanism
1046 of Ammonia Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*, 305(5690):1587–1594.
- 1047 Khademian, M. and Imlay, J. A. (2017). *Escherichia coli* cytochrome c peroxidase is a respiratory oxidase that
1048 enables the use of hydrogen peroxide as a terminal electron acceptor. *Proceedings of the National Academy
1049 of Sciences*, 114(33):E6922–E6931.
- 1050 Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates
1051 reveals principles underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 1052 Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein
1053 investment in cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 1054 Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs
1055 Reveal a Carbon Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–
1056 15927.
- 1057 Lu, D., Grayson, P., and Schulten, K. (2003). Glycerol Conductance and Physical Asymmetry of the *Escherichia
1058 coli* Glycerol Facilitator GlpF. *Biophysical Journal*, 85(5):2977–2987.
- 1059 Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of
1060 Sciences*, 112(51):15690–15695.
- 1061 Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On. Gene
1062 Expression*. Springer.
- 1063 Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key num-
1064 bers in molecular and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753.
- 1065 Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and
1066 cellular differentiation. *Growth Symposium*, 9:223–289.
- 1067 Mooney, R. A., Darst, S. A., and Landick, R. (2005). Sigma and RNA Polymerase: An On-Again, Off-Again Rela-
1068 tionship? *Molecular Cell*, 20(3):335–345.
- 1069 Mooney, R. A. and Landick, R. (2003). Tethering Σ70 to RNA polymerase reveals high in vivo activity of σ factors
1070 and Σ70-dependent pausing at promoter-distal locations. *Genes & Development*, 17(22):2839–2851.
- 1071 Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*,
1072 volume 1. Elsevier.
- 1073 Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in
1074 rod-shaped bacteria. *eLife*, 8:642.
- 1075 Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*.
1076 *Biochimie*, 119:80–91.
- 1077 Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Es-
1078 cherichia coli* with increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 1079 Perdue, S. A. and Roberts, J. W. (2011). σ⁷⁰-dependent Transcription Pausing in *Escherichia coli*. *Journal of
1080 Molecular Biology*, 412(5):782–792.
- 1081 Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer,
1082 Cham, Cham.
- 1083 Ramos, S. and Kaback, H. R. (1977). The relation between the electrochemical proton gradient and active trans-
1084 port in *Escherichia coli* membrane vesicles. *Biochemistry*, 16(5):854–859.
- 1085 Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in *Escherichia
1086 coli*. *Journal of Bacteriology*, 131(2):505–511.
- 1087 Rudd, S. G., Valerie, N. C. K., and Helleday, T. (2016). Pathways controlling dNTP pools to maintain genome
1088 stability. *DNA Repair*, 44:193–204.
- 1089 Sánchez-Romero, M. A., Molina, F., and Jiménez-Sánchez, A. (2011). Organization of ribonucleoside diphosphate
1090 reductase during multifork chromosome replication in *Escherichia coli*. *Microbiology*, 157(8):2220–2225.

- ¹⁰⁹¹ Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size
¹⁰⁹² and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- ¹⁰⁹³ Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold,
¹⁰⁹⁴ R., and Heinemann, M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature
¹⁰⁹⁵ Biotechnology*, 34(1):104–110.
- ¹⁰⁹⁶ Scholz, S. A., Diao, R., Wolfe, M. B., Fivenson, E. M., Lin, X. N., and Freddolino, P. L. (2019). High-Resolution
¹⁰⁹⁷ Mapping of the *Escherichia coli* Chromosome Reveals Positions of High and Low Transcription. *Cell Systems*,
¹⁰⁹⁸ 8(3):212–225.e9.
- ¹⁰⁹⁹ Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and
¹¹⁰⁰ gene expression: origins and consequences. *Science*, 330(6007):1099–1102.
- ¹¹⁰¹ Sekowska, A., Kung, H.-F., and Danchin, A. (2000). Sulfur Metabolism in *Escherichia coli* and Related Bacteria:
¹¹⁰² Facts and Fiction. *Journal of Molecular Microbiology and Biotechnology*, 2(2):34.
- ¹¹⁰³ Shi, H., Bratton, B. P., Gitai, Z., and Huang, K. C. (2018). How to Build a Bacterial Cell: MreB as the Foreman of
¹¹⁰⁴ *E. coli* Construction. *Cell*, 172(6):1294–1305.
- ¹¹⁰⁵ Si, F., Le Treut, G., Sauls, J. T., Vadia, S., Levin, P. A., and Jun, S. (2019). Mechanistic Origin of Cell-Size Control
¹¹⁰⁶ and Homeostasis in Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- ¹¹⁰⁷ Si, F., Li, D., Cox, S. E., Sauls, J. T., Azizi, O., Sou, C., Schwartz, A. B., Erickstad, M. J., Jun, Y., Li, X., and Jun, S. (2017).
¹¹⁰⁸ Invariance of Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- ¹¹⁰⁹ Sirko, A., Zatyka, M., Sadowy, E., and Hulanicka, D. (1995). Sulfate and thiosulfate transport in *Escherichia coli* K-
¹¹¹⁰ 12: Evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins. *Journal of Bacteriology*,
¹¹¹¹ 177(14):4134–4136.
- ¹¹¹² Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE.
¹¹¹³ *BMC Microbiology*, 19.
- ¹¹¹⁴ Stevenson, B. S. and Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in *Escherichia
¹¹¹⁵ coli*. *Applied and Environmental Microbiology*, 70(11):6670–6677.
- ¹¹¹⁶ Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell
¹¹¹⁷ material. *Antonie van Leeuwenhoek*, 39(1):545–565.
- ¹¹¹⁸ Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant
¹¹¹⁹ of *Escherichia coli*. *Archives of Microbiology*, 113(3):185–189.
- ¹¹²⁰ Svenningsen, S. L., Kongstad, M., Stenum, T. S. n., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017). Transfer RNA
¹¹²¹ is highly unstable during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Research*, 45(2):793–804.
- ¹¹²² Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism?
¹¹²³ Testing the Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- ¹¹²⁴ Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015).
¹¹²⁵ Cell-size control and homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- ¹¹²⁶ Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale
¹¹²⁷ analysis of protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- ¹¹²⁸ Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010).
¹¹²⁹ Genome-derived minimal metabolic models for *Escherichia coli* mg1655 with estimated in vivo
¹¹³⁰ respiratory ATP stoichiometry. *Biotechnology and Bioengineering*, 107(2):369–381. _eprint:
¹¹³¹ <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- ¹¹³² The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic
¹¹³³ Acids Research*, 47(D1):D330–D338.
- ¹¹³⁴ Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing
¹¹³⁵ catalytic and translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- ¹¹³⁶ van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen Assimilation in *Escherichia coli*: Putting
¹¹³⁷ Molecular Data into a Systems Perspective. *Microbiology and Molecular Biology Reviews*, 77(4):628–695.

- 1138 Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1FO-ATP synthase. *FEBS Letters*,
1139 545(1):61–70.
- 1140 Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement
1141 of Two Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–
1142 539.
- 1143 Zhang, L., Jiang, W., Nan, J., Almqvist, J., and Huang, Y. (2014a). The *Escherichia coli* CysZ is a pH dependent sulfate
1144 transporter that can be inhibited by sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(7):1809–
1145 1816.
- 1146 Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014b). Regulation of crp gene expression by the catabolite repres-
1147 sor/activator, cra, in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 1148 Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource
1149 allocation in *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.