

# Fundamental limits on the rate of bacterial cell division

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

\*For correspondence:

<sup>†</sup>These authors contributed equally to this work

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

<sup>14</sup> 

---

## <sup>15</sup> Abstract

<sup>16</sup> 

---

## <sup>17</sup> Introduction

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

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

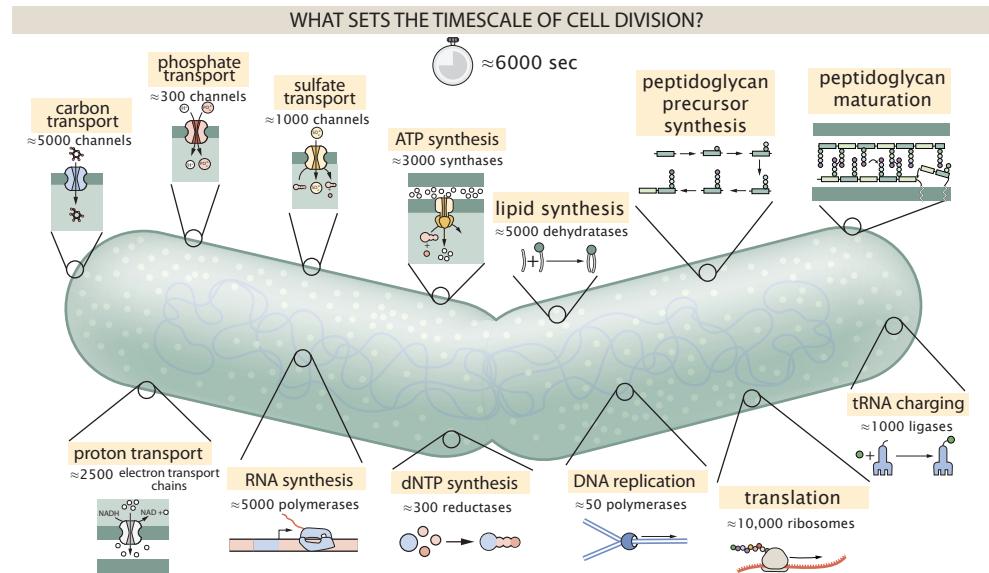
43 Specifically, we leverage a combination of *E. coli* proteomic data sets collected over the past  
 44 decade using either mass spectrometry (???) or ribosomal profiling (?) across 31 unique growth  
 45 conditions. Broadly speaking, we entertain several classes of hypotheses as are illustrated in ??.  
 46 First, we consider potential limits on the transport of nutrients into the cell. We address this hy-  
 47 pothesis by performing an order-of-magnitude estimate for how many carbon, phosphorous, and  
 48 sulfur atoms are needed to facilitate this requirement given a 5000 second division time. As a sec-  
 49 ond hypothesis, we consider the possibility that there exists a fundamental limit on how quickly  
 50 the cell can generate ATP. We approach this hypothesis from two angles, considering how many  
 51 ATP synthase complexes must be needed to churn out enough ATP to power protein translation  
 52 followed by an estimation of how many electron transport complexes must be present to main-  
 53 tain the proton motive force. A third class of estimates considers the need to maintain the size  
 54 and shape of the cell through the construction of new lipids for the cell membranes as well as the  
 55 glycan polymers which make up the rigid peptidoglycan. Our final class of hypotheses centers on  
 56 the synthesis of a variety of biomolecules. Our focus is primarily on the stages of the central dogma  
 57 as we estimate the number of protein complexes needed for DNA replication, transcription, and  
 58 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 ?? as setting  
 62 possible speed limits. Ultimately, we find that protein translation (particularly the generation of  
 63 new ribosomes) acts as 1) a rate limiting step for the *fastest* bacterial division, and 2) the major  
 64 determinant of bacterial growth across all nutrient conditions we have considered under steady  
 65 state, exponential growth. This perspective is in line with the linear correlation observed between  
 66 growth rate and ribosomal content (typically quantified through the ratio of RNA to protein) for fast  
 67 growing cells (?), but suggests a more prominent role for ribosomes in setting the doubling time  
 68 across all conditions of nutrient limitation. Here we again leverage the quantitative nature of this  
 69 data set and present a quantitative model of the relationship between the fraction of the proteome  
 70 devoted to ribosomes and the speed limit of translation, revealing a fundamental tradeoff between  
 71 the translation capacity of the ribosome pool and the maximal growth rate.

## 72 Uptake of Nutrients

73 In order to build new cellular mass, the molecular and elemental building blocks must be scav-  
 74 enged from the environment in different forms. Carbon, for example, is acquired via the transport  
 75 of carbohydrates and sugar alcohols with some carbon sources receiving preferential treatment  
 76 in their consumption (?). Phosphorus, sulfur, and nitrogen, on the other hand, are harvested pri-  
 77 marily in the forms of inorganic salts, namely phosphate, sulfate, and ammonia (??????). All of  
 78 these compounds have different permeabilities across the cell membrane and most require some  
 79 energetic investment either via ATP hydrolysis or through the proton electrochemical gradient to  
 80 bring the material across the hydrophobic cell membrane. Given the diversity of biological trans-  
 81 port mechanisms and the vast number of inputs needed to build a cell, we begin by considering  
 82 transport of some of the most important cellular ingredients: carbon, nitrogen, oxygen, hydrogen,  
 83 phosphorus, and sulfur.

84 The elemental composition of *E. coli* has received much quantitative attention over the past  
 85 half century (????), providing us with a starting point for estimating the copy numbers of various  
 86 transporters. While there is some variability in the exact elemental percentages (with different un-  
 87 certainties), we can estimate that the dry mass of a typical *E. coli* cell is  $\approx$  45% carbon (BNID: 100649,  
 88 ?),  $\approx$  15% nitrogen (BNID: 106666, ?),  $\approx$  3% phosphorus (BNID: 100653, ?), and 1% sulfur (BNID:  
 89 100655, ?). In the coming paragraphs, we will engage in a dialogue between back-of-the-envelope  
 90 estimates for the numbers of transporters needed to facilitate these chemical stoichiometries and  
 91 the experimental proteomic measurements of the biological reality. Such an approach provides



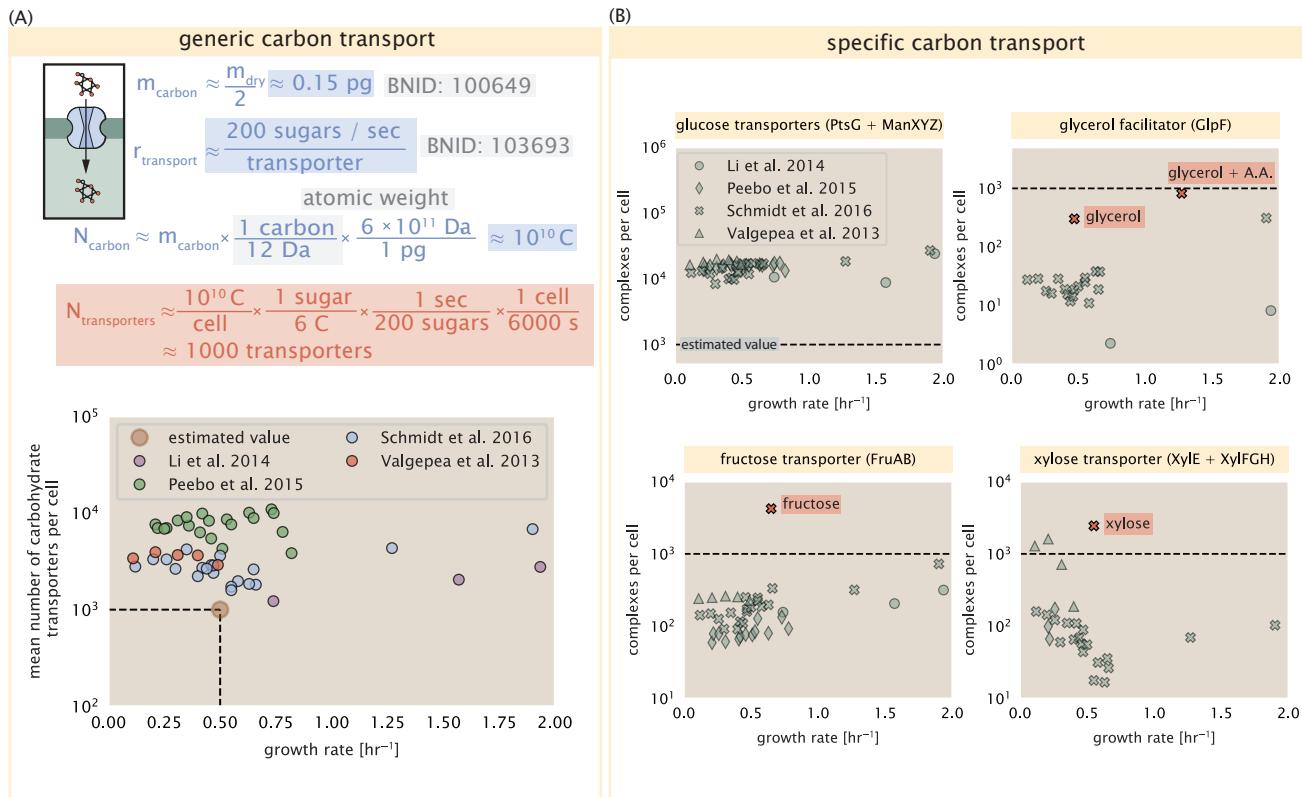
**Figure 1. Transport and synthesis processes necessary for cell division.** We consider an array of processes necessary for a cell to double its molecular components. Such processes include the transport of carbon across the cell membrane, the production of ATP, and fundamental processes of the central 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  $\approx 5000$  sec.

the opportunity to test if our biological knowledge is sufficient to understand the scale at which these complexes are produced. Specifically, we will make these estimates considering a modest doubling time of 5000 s, a growth rate of  $\approx 0.5 \text{ hr}^{-1}$ , the range in which the majority of the experimental measurements reside.

### Carbon Transport

We begin with the most abundant element by mass, carbon. Using  $\approx 0.3 \text{ pg}$  as the typical *E. coli* dry mass (BNID: 103904, ?), we estimate that  $\approx 10^{10}$  carbon atoms must be brought into the cell in order to double all of the carbon-containing molecules (??(A, top)). Typical laboratory growth conditions, such as those explored in the aforementioned proteomic data sets, provide carbon as a single class of sugar such as glucose, galactose, or xylose to name a few. *E. coli* has evolved myriad mechanisms by which these sugars can be transported across the cell membrane. One such mechanism of transport is via the PTS system which is a highly modular system capable of transporting a diverse range of sugars (?). The glucose-specific component of this system transports  $\approx 200$  glucose molecules per second per channel (BNID: 114686, ?). Making the assumption that this is a typical sugar transport rate, coupled with the need to transport  $10^{10}$  carbon atoms, we arrive at the conclusion that on the order of 1,000 transporters must be expressed in order to bring in enough carbon atoms to divide in 6,000 s, diagrammed in the top panel of ??(A). This estimate, along with the observed average number of carbohydrate transporters present in the proteomic data sets (????), is shown in ??(A). While we estimate 1,000 transporters are needed, the data reveals that at a division time of  $\approx 5000$  s there is nearly a ten-fold excess of transporters. Furthermore, the data illustrates that the average number of carbohydrate transporters present is largely-growth rate independent.

The estimate presented in ??(A) neglects any specifics of the regulation of carbon transport system and presents a data-averaged view of how many carbohydrate transporters are present on average. Using the diverse array of growth conditions explored in the proteomic data sets, we can explore how individual carbon transport systems depend on the population growth rate. In ??(B),



**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  $\approx 6$  C, and each transporter conducts sugar molecules at a rate of  $\approx 200$  per second. Bottom plot shows the estimated number of transporters needed at a growth rate of  $\approx 0.5$  per hr (light-brown point and dashed lines). Colored points correspond to the mean number of carbohydrate transporters 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. Red points and red-highlighted text indicate conditions in which the only source of carbon in the growth medium induces expression of the transport system.

118 we show the total number of carbohydrate transporters specific to different carbon sources. A striking  
 119 observation, shown in the top-left plot of ??(B), is the constancy in the expression of the glucose-  
 120 specific transport systems (the PtsG enzyme of the PTS system and the glucose-transporting ManXYZ  
 121 complex). Additionally, we note that the total number of glucose-specific transporters is tightly dis-  
 122 tributed  $\approx 10^4$  per cell, an order of magnitude beyond the estimate shown in ??(A). This illustrates  
 123 that *E. coli* maintains a substantial number of complexes present for transporting glucose which is  
 124 known to be the preferential carbon source (???).

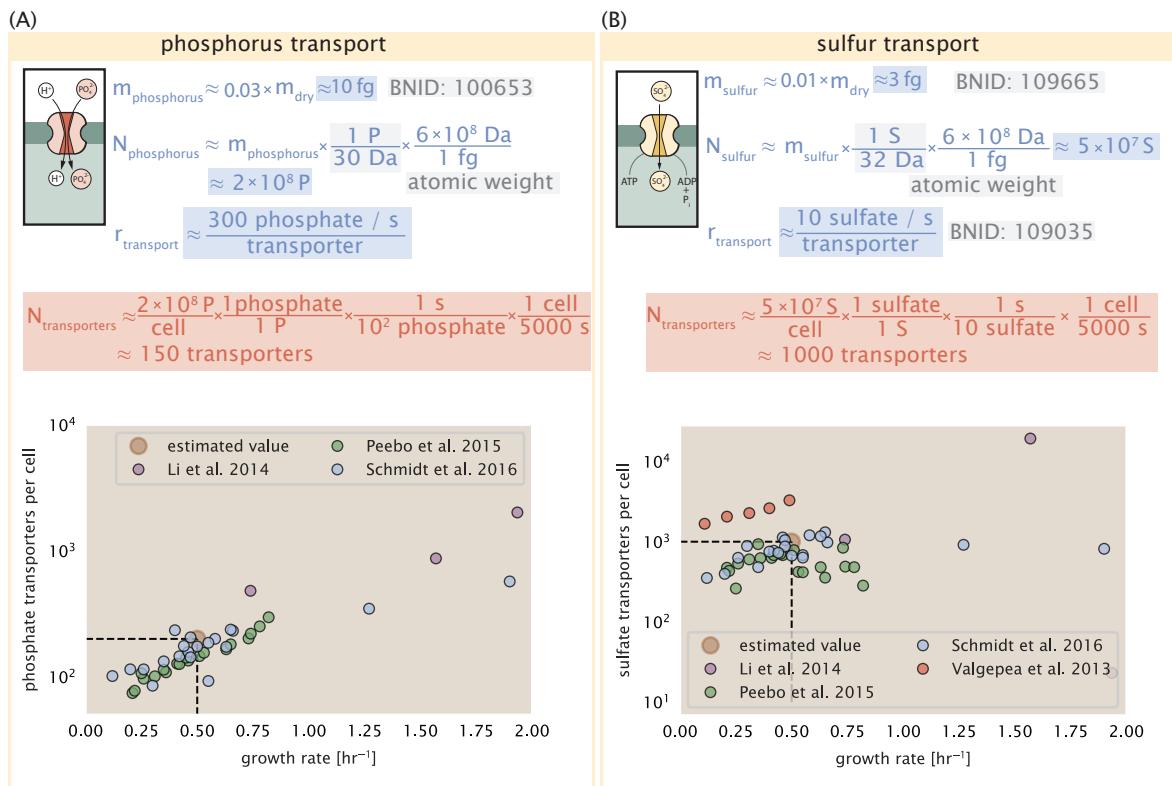
125 It is now understood that a large number of metabolic operons are regulated with dual-input  
 126 logic gates that are only expressed when glucose concentrations are low (mediated by cyclic-AMP  
 127 receptor protein CRP) and the concentration of other carbon sources are elevated (??). A famed  
 128 example of such dual-input regulatory logic is in the regulation of the *lac* operon which is only  
 129 natively activated in the absence of glucose and the presence of allolactose, an intermediate in  
 130 lactose metabolism (?), though we now know of many other such examples (???). This illustrates  
 131 that once glucose is depleted from the environment, cells have a means to dramatically increase  
 132 the abundance of the specific transporter needed to digest the next sugar that is present. Several  
 133 examples of induced expression of specific carbon-source transporters are shown in ??(B). Points  
 134 colored in red (labeled by red text-boxes) correspond to growth conditions in which the specific  
 135 carbon source (glycerol, xylose, or fructose) is present. These plots show that, in the absence of the  
 136 particular carbon source, expression of the transporters is maintained on the order of  $\sim 10^2$  per cell.  
 137 However, when induced, the transporters become highly-expressed and are present on the order  
 138 of  $\sim 10^4$  per cell, which exceeds the generic estimate given in ??(A). Together, this generic estimation  
 139 and the specific examples of induced expression suggest that transport of carbon across the cell  
 140 membrane, while critical for growth, is not the rate-limiting step of cell division.

141 In the context of speeding up growth, one additional limitation is the fact that the cell's inner  
 142 membrane is occupied by a multitude of other membrane proteins. Considering a rule-of-thumb  
 143 for the surface area of *E. coli* of about  $6 \mu\text{m}^2$  (BNID: 101792, ?), we expect an areal density for 1,000  
 144 transporters to be approximately 200 transporters/ $\mu\text{m}^2$ . For a glucose transporter occupying about  
 145  $50 \text{ nm}^2/\text{dimer}$ , this amounts to about only 1 percent of the total inner membrane (?). In addition,  
 146 bacterial cell membranes typically have densities of  $10^5$  proteins// $\mu\text{m}^2$  (?), implying that the cell  
 147 could accommodate more transporters if it were rate limiting.

#### 148 **Phosphorus and Sulfur Transport**

149 We now turn our attention towards other essential elements, namely phosphorus and sulfur. Phos-  
 150 phorus is critical to the cellular energy economy in the form of high-energy phosphodiester bonds  
 151 making up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-translational  
 152 modification of proteins and defining the polar-heads of lipids. In total, phosphorus makes up  
 153  $\approx 3\%$  of the cellular dry mass which in typical experimental conditions is in the form of inorganic  
 154 phosphate. The cell membrane has remarkably low permeability to this highly-charged and critical  
 155 molecule, therefore requiring the expression of active transport systems. In *E. coli*, the proton elec-  
 156 trochemical gradient across the inner membrane is leveraged to transport inorganic phosphate  
 157 into the cell (?). Proton-solute symporters are widespread in *E. coli* (??) and can have rapid trans-  
 158 port rates of 50 molecules per second for sugars and other solutes (BNID: 103159; 111777, ?).  
 159 In *E. coli* the PitA phosphate transport system has been shown to very tightly coupled with the  
 160 proton electrochemical gradient with a 1:1 proton:phosphate stoichiometric ratio (??). Illustrated  
 161 in ??(A), we can estimate that  $\approx 300$  phosphate transporters are necessary to maintain an  $\approx 3\%$   
 162 dry mass with a 6,000 s division time. This estimate is again satisfied when we examine the ob-  
 163 served copy numbers of PitA in proteomic data sets (plot in ??(A)). While our estimate is very much  
 164 in line with the observed numbers, we emphasize that this is likely a slight over estimate of the  
 165 number of transporters needed as there are other phosphorous scavenging systems, such as the  
 166 ATP-dependent phosphate transporter Pst system which we have neglected.

167 Satisfied that there are a sufficient number of phosphate transporters present in the cell, we



**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]<sub>2</sub>[CysU][CysW][Sbp/CysP].

168 now turn to sulfur transport as another potentially rate limiting process. Similar to phosphate, sul-  
 169 fide is highly-charged and not particularly membrane permeable, requiring active transport. While  
 170 there exists a H<sup>+</sup>/sulfate symporter in *E. coli*, it is in relatively low abundance and is not well char-  
 171 acterized (?). Sulfate is predominantly acquired via the ATP-dependent ABC transporter CysUWA  
 172 system which also plays an important role in selenium transport (??). While specific kinetic details  
 173 of this transport system are not readily available, generic ATP transport systems in prokaryotes  
 174 transport on the order of 1 to 10 molecules per second (BNID: 109035, ?). Combining this generic  
 175 transport rate, measurement of sulfur comprising 1% of dry mass, and a 6,000 second division  
 176 time yields an estimate of  $\approx 1000$  CysUWA complexes per cell (??(B)). Once again, this estimate is  
 177 in notable agreement with proteomic data sets, suggesting that there are sufficient transporters  
 178 present to acquire the necessary sulfur. In a similar spirit of our estimate of phosphorus transport,  
 179 we emphasize that this is likely an overestimate of the number of necessary transporters as we  
 180 have neglected other sulfur scavenging systems that are in lower abundance.

### 181 Nitrogen Transport

182 Finally, we turn to nitrogen transport as the last remaining transport system highlighted in ??.  
 183 Unlike phosphate, sulfate, and various sugar molecules, nitrogen in the form of ammonia can  
 184 readily diffuse across the cell membrane and has a permeability on par with water ( $\approx 10^5$  nm/s,  
 185 BNID:110824 ?). In particularly nitrogen-poor conditions, *E. coli* expresses a transporter (AmtB)  
 186 which appears to aid in nitrogen assimilation, though the mechanism and kinetic details of trans-  
 187 port is still a matter of debate (??). Beyond ammonia, another plentiful source of nitrogen come

188 in the form of glutamate, which has its own complex metabolism and scavenging pathways. How-  
 189 ever, nitrogen is plentiful in the growth conditions examined in this work, permitting us to neglect  
 190 nitrogen transport as a potential rate limiting process in cell division.

### 191 Function of the Central Dogma

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

### 197 DNA

198 Most bacteria (including *E. coli*) harbor a single, circular chromosome and can have extra-chromosomal  
 199 plasmids up to ~ 100 kbp in length. We will focus our quantitative thinking solely on the chromo-  
 200 some of *E. coli* which harbors ≈ 5000 genes and ≈ 5 × 10<sup>6</sup> base pairs. To successfully divide and  
 201 produce viable progeny, this chromosome must be faithfully replicated and segregated into each  
 202 nascent cell. We again rely on the near century of literature in molecular biology to provide some  
 203 insight on the rates and mechanics of the replicative feat as well as the production of the required  
 204 starting materials, dNTPs.

### 205 dNTP synthesis

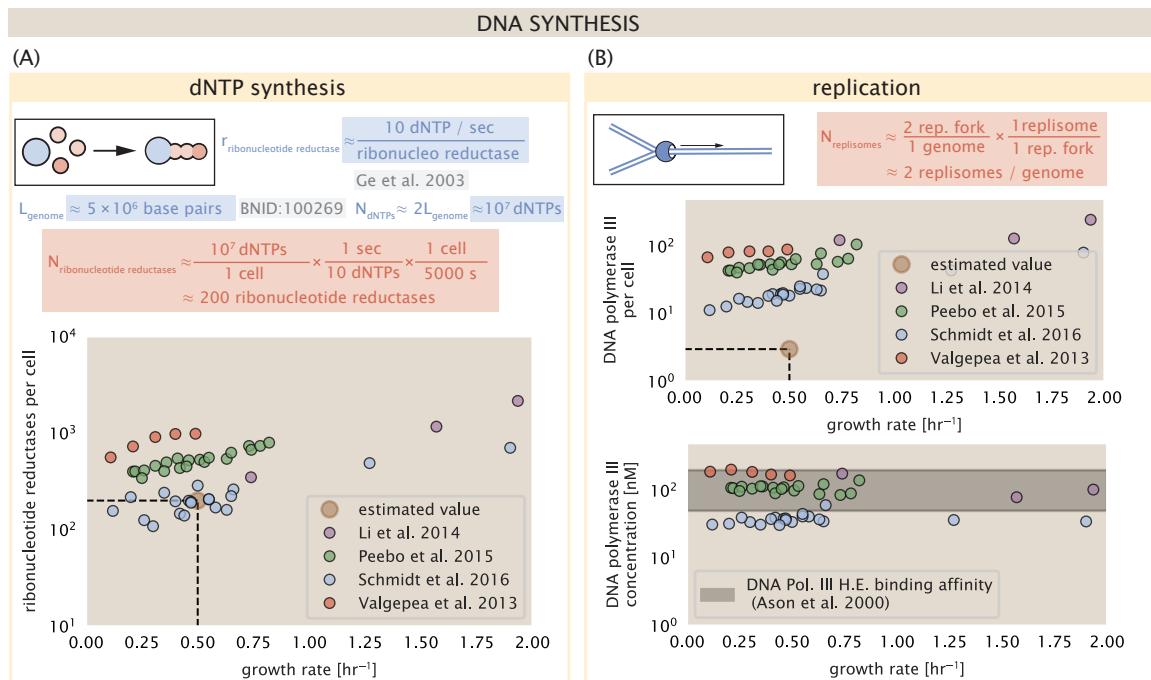
206 We begin our exploration DNA replication by examining the production of the deoxyribonucleotide  
 207 triphosphates (dNTPs). The four major dNTPs (dATP, dTTP, dCTP, and dGTP) are synthesized *de*  
 208 *novo* in separate pathways, requiring different building blocks. However, a critical step present  
 209 in all dNTP synthesis pathways is the conversion from ribonucleotide to deoxyribonucleotide via  
 210 the removal of the 3' hydroxyl group of the ribose ring (?). This reaction is mediated by a class of  
 211 enzymes termed ribonucleotide reductases, of which *E. coli* possesses two aerobically active com-  
 212 plexes (termed I and II) and a single anaerobically active enzyme. Due to their peculiar formation  
 213 of a radical intermediate, these enzymes have received much biochemical, kinetic, and structural  
 214 characterization. One such work (?) performed a detailed *in vitro* measurement of the steady-state  
 215 kinetic rates of these complexes, revealing a turnover rate of ≈ 10 dNTP per second.

216 Since this reaction is central to the synthesis of all dNTPs, it is reasonable to consider the abun-  
 217 dance of these complexes is a measure of the total dNTP production in *E. coli*. Illustrated schemati-  
 218 cally in ?? (A), we consider the fact that to replicate the cell's genome, on the order of ≈ 10<sup>7</sup> dNTPs  
 219 must be synthesized. Assuming a production rate of 10 per second per ribonucleotide reductase  
 220 complex and a cell division time of 6000 seconds, we arrive at an estimate of ≈ 150 complexes  
 221 needed per cell. As shown in the bottom panel of ?? (A), this estimate agrees with the experimen-  
 222 tal measurements of these complexes abundances within ≈ 1/2 an order of magnitude.

223 Recent work has revealed that during replication, the ribonucleotide reductase complexes co-  
 224 alesce to form discrete foci colocalized with the DNA replisome complex (?). This is particularly  
 225 pronounced in conditions where growth is slow, indicating that spatial organization and regula-  
 226 tion of the activity of the complexes plays an important role.

### 227 DNA Replication

228 We now turn our focus to the integration of these dNTP building blocks into the replicated chromo-  
 229 some strand via the DNA polymerase. Replication is initiated at a single region of the chromosome  
 230 termed the *oriC* locus at which a pair of DNA polymerases bind and begin their high-fidelity repli-  
 231 cation of the genome in opposite directions. Assuming equivalence between the two replication  
 232 forks, this means that the two DNA polymerase complexes (termed replisomes) meet at the mid-  
 233 way point of the circular chromosome termed the *ter* locus. The kinetics of the five types of DNA  
 234 polymerases (I – V) have been intensely studied, revealing that DNA polymerase III performs the



**Figure 4. 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 6000 second generation time. Points in the plot correspond to the total number of ribonucleotide reductase I ( $[NrdA]_2[NrdB]_2$ ) and ribonucleotide reductase II ( $[NrdE]_2[NrdF]_2$ ) 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 ( $[DnaE]_3[DnaQ]_3[HolE]_3[DnaX]_5[HolB][HolA][DnaN]_4[HolC]_4[HolD]_4$ ) per cell. Bottom plot shows the effective concentration of DNA polymerase III holoenzyme given cell volumes calculated from the growth rate as in ? (See Appendix Section 4).

235 high fidelity processive replication of the genome with the other "accessory" polymerases playing  
 236 auxiliary roles (?). *In vitro* measurements have shown that DNA Polymerase III copies DNA at a rate  
 237 of  $\approx$  600 nucleotides per second (BNID: 104120, ?). Therefore, to replicate a single chromosome,  
 238 two DNA polymerases replicating at their maximal rate would copy their entire genome in  $\approx$  4000  
 239 s. Thus, with a division time of 5000 s (our "typical" growth rate for the purposes of this work),  
 240 there is sufficient time for a pair of DNA polymerase III complexes to replicate the entire genome.  
 241 However, this estimate implies that 4000 s would be the upper-limit time scale for bacterial division  
 242 which is at odds with the familiar 20 minute (1200 s) doubling time of *E. coli* in rich medium.

243 It is well known that *E. coli* can parallelize its DNA replication such that multiple chromosomes  
 244 are being replicated at once, with as many as 10 - 12 replication forks at a given time (??). Thus,  
 245 even in rapidly growing cultures, we expect only a few polymerases ( $\approx$  10) are needed to replicate  
 246 the chromosome per cell doubling. However, as shown in ??(B), DNA polymerase III is nearly an  
 247 nearly an order of magnitude more abundant. This discrepancy can be understood by considering  
 248 its binding constant to DNA. DNA polymerase III is highly processive, facilitated by a strong affinity  
 249 of the complex to the DNA. *In vitro* biochemical characterization has quantified the  $K_D$  of DNA  
 250 polymerase III holoenzyme to single-stranded and double-stranded DNA to be 50 and 200 nM,  
 251 respectively (?). The bottom plot in ?? (B) shows that the concentration of the DNA polymerase  
 252 III across all data sets and growth conditions is within this range. Thus, while the copy number  
 253 of the DNA polymerase III is in excess of the strict number required to replicate the genome, its  
 254 copy number appears to vary such that its concentration is approximately equal to the dissociation  
 255 constant to the DNA. While the processes regulating the initiation of DNA replication are complex  
 256 and involve more than just the holoenzyme, these data indicate that the kinetics of replication  
 257 rather than the explicit copy number of the DNA polymerase III holoenzyme is the more relevant  
 258 feature of DNA replication to consider. In light of this, the data in ??(B) suggests that for bacteria  
 259 like *E. coli*, DNA replication does no that represent a rate-limiting step in cell division. However,  
 260 it is worth noting that for bacterium like *C. crescentus* whose chromosomal replication is initiated  
 261 only once per cell cycle (?), the time to double their chromosome likely represents an upper limit  
 262 to their growth rate.

### 263 RNA Synthesis

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

### 273 rRNA

274 We begin with an estimation of the number of RNA polymerases needed to synthesize the rRNA  
 275 that serve as catalytic and structural elements of the ribosome. Each ribosome contains three rRNA  
 276 molecules of lengths 120, 1542, and 2904 nucleotides (BNID: 108093, ?), meaning each ribosome  
 277 contains  $\approx$  4500 nucleotides. As the *E. coli* RNA polymerase transcribes DNA to RNA at a rate of  $\approx$   
 278 40 nucleotides per second (BNID: 101904, ?), it takes a single RNA polymerase  $\approx$  100 s to synthesize  
 279 the RNA needed to form a single functional ribosome. Therefore, in a 5000 s division time, a single  
 280 RNA polymerase transcribing rRNA at a time would result in only  $\approx$  50 functional ribosomal rRNA  
 281 units – far below the observed number of  $\approx$   $10^4$  ribosomes per cell.

282 Of course, there can be more than one RNA polymerase transcribing the rRNA genes at any  
 283 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  $\approx 1$  per second (BNID: 111997; 102362, ?). If RNA polymerases are being constantly loaded on to the rRNA genes at this rate, then we can assume that  $\approx 1$  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, ?) four of which are localized directly adjacent to the origin of replication (?). 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 (?). 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  $4 \times 10^4$  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 nucleotides (BNID: 107873, ?), we can expect  $\approx 1$  RNA polymerase per 80 nucleotides. With a total length of  $\approx 4500$  nucleotides per operon and 7 operons per cell, the maximum number of RNA polymerases that can be transcribing rRNA at any given time is  $\approx 400$ . As we will see in the coming sections, the synthesis of rRNA is the dominant requirement of the RNA polymerase pool.

### 308 mRNA

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

This estimation captures the *steady-state* mRNA copy number, meaning that at any given time, there will exist approximately 3000 unique mRNA molecules. To determine the *total* number of mRNA that need to be synthesized over the cell's lifetime, we must consider degradation of the mRNA. In most bacteria, mRNAs are rather unstable with life times on the order of several minutes (BNID: 104324; 106253; 111927; 111998, ?). For convenience, we assume that the typical mRNA in our cell of interest has a typical lifetime of  $\approx 300$  seconds. Using this value, we can determine the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell. While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional dynamics, we state here that the total mRNA production rate must be on the order of  $\approx 15$  mRNA per second. In *E. coli*, the average protein is  $\approx 300$  amino acids in length (BNID: 108986; ?), meaning that the corresponding mRNA is  $\approx 900$  nucleotides which we will further approximate as  $\approx 1000$  nucleotides to account for the non-protein coding regions on the 5' and 3' ends. This means that the cell must have enough RNA polymerase molecules about to sustain a transcription rate of  $\approx 1.5 \times 10^4$  nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA at a clip of 40 nucleotides per second, we arrive at a comfortable estimate of  $\approx 250$  RNA polymerase complexes needed to satisfy the mRNA demands of the cell. It is worth noting that this number is approximately half of that required to synthesize enough rRNA, as we saw in the previous section.

**334** We find this to be a striking result as these 250 RNA polymerase molecules are responsible for the  
**335** transcription of the  $\approx$  4000 protein coding genes that are not ribosome associated.

### **336** tRNA

**337** The final class of RNA molecules worthy of quantitative consideration is the pool of tRNAs  
**338** used during translation to map codon sequence to amino acid identity. Unlike mRNA or rRNA,  
**339** each individual tRNA is remarkably short, ranging from 70 to 95 nucleotides each (BNID: 109645;  
**340** 102340, ?). What they lack in length, they make up for in abundance. There are approximately  
**341**  $\approx$  3000 tRNA molecules present for each of the 20 amino acids (BNID: 105280, ?), although the  
**342** precise copy number is dependent on the identity of the ligated amino acid. Using these values,  
**343** we make the estimate that  $\approx 5 \times 10^6$  nucleotides are sequestered in tRNA per cell. Unlike mRNA,  
**344** tRNA is remarkably stable with typical lifetimes *in vivo* on the order of  $\approx$  48 hours (??) – well beyond  
**345** the timescale of division. Once again using our rule-of-thumb for the rate of transcription to be 40  
**346** nucleotides per second and assuming a division time of  $\approx$  5000 seconds, we arrive at an estimate  
**347** of  $\approx$  20 RNA polymerases to synthesize enough tRNA. This requirement pales in comparison to the  
**348** number of polymerases needed to generate the rRNA and mRNA pools and can be neglected as a  
**349** significant transcriptional burden.

### **350** RNA Polymerase and $\sigma$ -factor Abundance

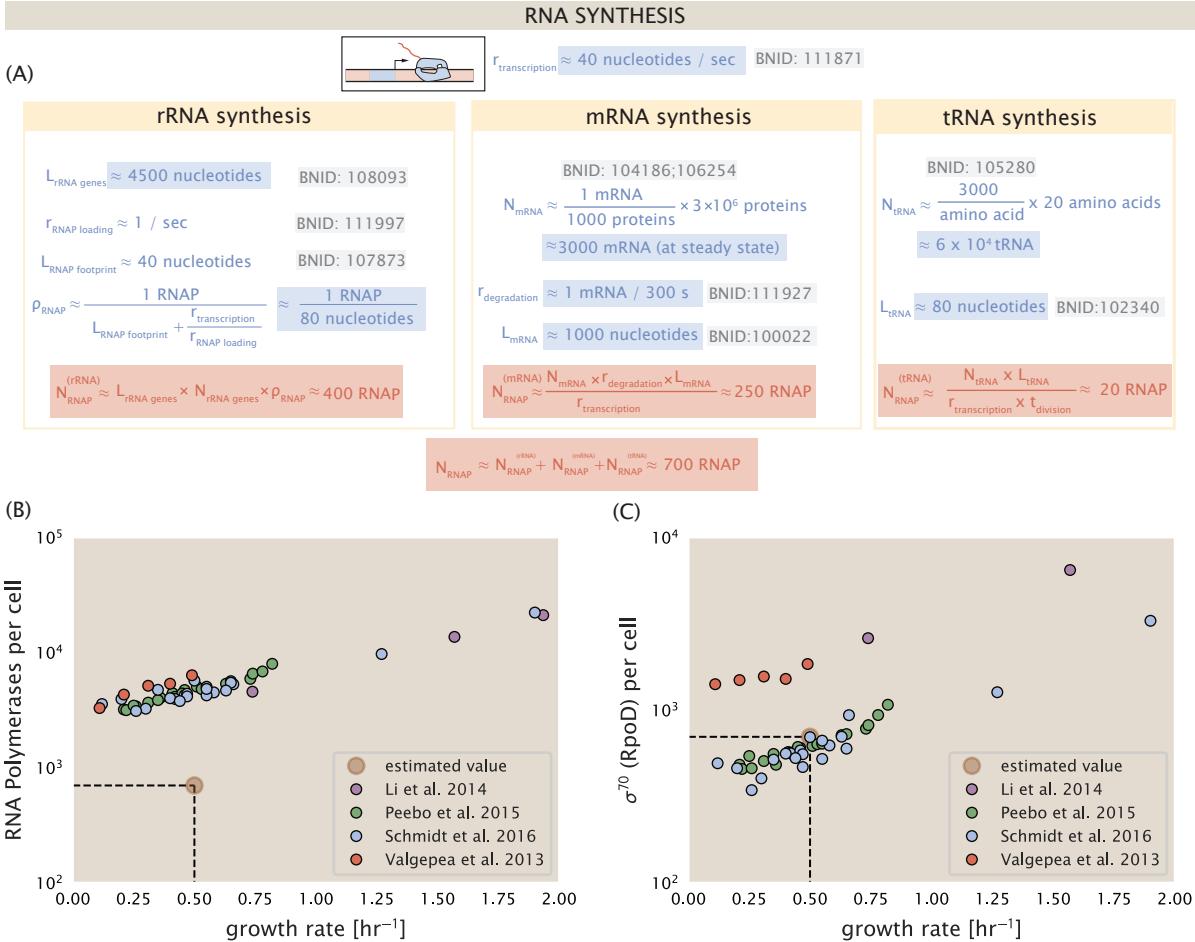
**351** These estimates, summarized in ?? (A), reveal that synthesis of rRNA and mRNA are the dominant  
**352** RNA species synthesized by RNA polymerase, suggesting the need for  $\approx$  700 RNA polymerases  
**353** per cell. As is revealed in ?? (B), this estimate is about an order of magnitude below the observed  
**354** number of RNA polymerase complexes per cell ( $\approx$  5000 - 7000). The disagreement between the  
**355** estimated number of RNA polymerases and these observations are at least consistent with recent  
**356** literature revealing that  $\approx$  80 % of RNA polymerases in *E. coli* are not transcriptionally active (?).  
**357** Our estimate ignores the possibility that some fraction is only nonspecifically bound to DNA, as  
**358** well as the obstacles that RNA polymerase and DNA polymerase present for each other as they  
**359** move along the DNA (?).

**360** In addition, it is also vital to consider the role of  $\sigma$ -factors which help RNA polymerase identify  
**361** and bind to transcriptional start sites (?). Here we consider  $\sigma^{70}$  (RpoD) which is the dominant  
**362** "general-purpose"  $\sigma$ -factor in *E. coli*. While initially thought of as being solely involved in transcriptional  
**363** initiation, the past two decades of single-molecule work has revealed a more multipurpose  
**364** role for  $\sigma^{70}$  including facilitating transcriptional elongation (?????). ?? (B) is suggestive of such a role  
**365** as the number of  $\sigma^{70}$  proteins per cell is in close agreement with our estimate of the number of  
**366** transcriptional complexes needed.

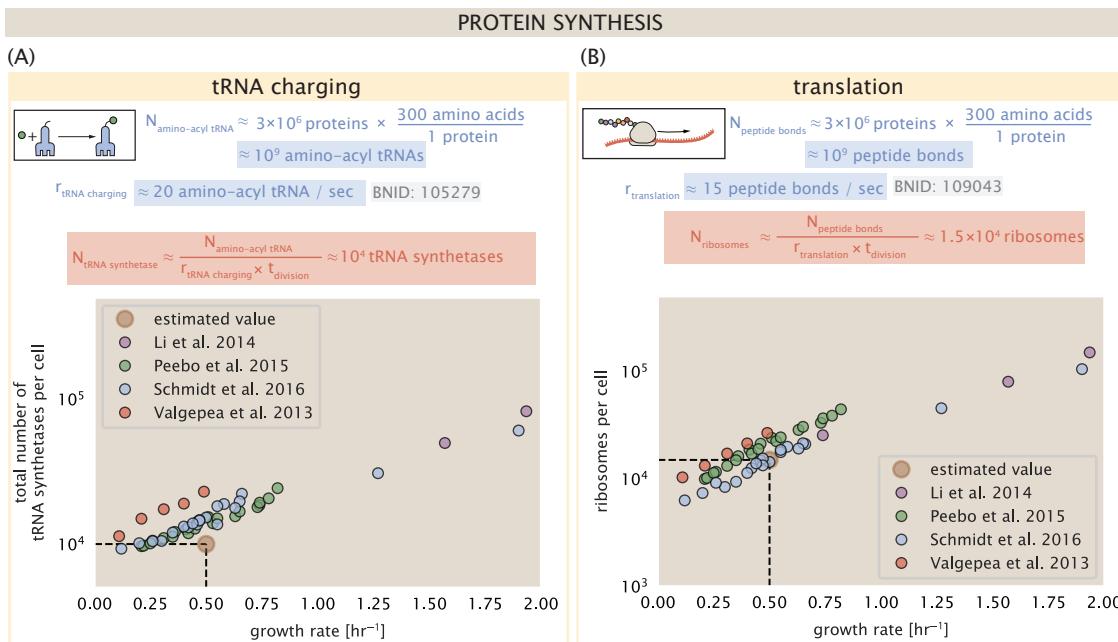
**367** While these estimates and comparison with experimental data reveal an interesting dynamic  
**368** at play between the transcriptional demand and copy numbers of the corresponding machinery,  
**369** these findings illustrate that transcription cannot be the rate limiting step in bacterial division. ??  
**370** (A) reveals that the availability of RNA polymerase is not a limiting factor for cell division as the cell  
**371** always has an apparent  $\sim$  10-fold excess than needed. Furthermore, if more transcriptional activity  
**372** was needed to satisfy the cellular requirements, more  $\sigma^{70}$ -factors could be expressed to utilize a  
**373** larger fraction of the RNA polymerase pool.

### **374** Protein synthesis

**375** Lastly, we turn our attention to the process of translation. So far our estimates have led to protein  
**376** copy numbers that are consistent with the proteomic data, or even in excess of what might be  
**377** needed for each task under limiting growth conditions. Even in our example of *E. coli* grown under  
**378** different carbohydrate sources (??(B)), cells can utilize alternative carbon sources by inducing the  
**379** expression of additional membrane transporters and enzymes. Optimal resource allocation and  
**380** the role of ribosomal proteins have been an area of intense quantitative study over the last decade



**Figure 5. 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  $\sigma^{70}$  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  $0.5 \text{ hr}^{-1}$ .



**Figure 6. 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], [GltX], [IleS], [LeuS], [ValS], [AlaS]<sub>2</sub>, [AsnS]<sub>2</sub>, [AspS]<sub>2</sub>, [TyrS]<sub>2</sub>, [TrpS]<sub>2</sub>, [ThrS]<sub>2</sub>, [SerS]<sub>2</sub>, [ProS]<sub>2</sub>, [PheS]<sub>2</sub>[PheT]<sub>2</sub>, [MetG]<sub>2</sub>, [LysS]<sub>2</sub>, [HisS]<sub>2</sub>, [GlyS]<sub>2</sub>[GlyQ]<sub>2</sub>). (B) Estimation for the number of ribosomes required to synthesize all proteins in the cell. 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<sup>-1</sup>.

381 by Hwa and others (??). From the perspective of limiting growth, our earlier estimate of rRNA high-  
 382 lighted the necessity for multiple copies of rRNA genes in order to make enough rRNA, suggesting  
 383 the possibility that synthesis of ribosomes might be rate limiting. While the transcriptional demand  
 384 for the ribosomal proteins is substantially lower than rRNA genes, since many proteins can be trans-  
 385 lated from relatively fewer mRNA, other ribosomal proteins like the translation elongation factor  
 386 EF-Tu also present a substantial burden. For EF-Tu in particular, it is the most highly expressed  
 387 protein in *E. coli* and is expressed by multiple genes on the chromosome, tufA and tufB.

388 We begin by first estimating the number of tRNA synthetases and ribosomes required for a  
 389 doubling time of 5000 seconds. *E. coli* has roughly  $3 \times 10^6$  proteins per cell, which for an average  
 390 protein of 300 aa, amounts to the formation of  $\approx 10^9$  peptide bonds. This also corresponds to  
 391 the number of amino-acyl tRNA that are used by ribosomes, with the pool of tRNA continuously  
 392 recharging new amino acids by tRNA synthetases. At a rate of charging of about 20 amino-acyl  
 393 tRNA per second (BNID: 105279, ?), we find that cells have more than sufficient tRNA synthetases  
 394 to meet the demand of ribosomes during protein synthesis (??(A)). If we consider an elongation  
 395 rate of  $\approx 15$  peptide bonds per second (BNID: 114271, ??), the formation of  $\approx 10^9$  peptide bonds  
 396 would require  $1.5 \times 10^4$  ribosomes at a growth rate of  $0.5 \text{ hr}^{-1}$ . This is indeed consistent with the  
 397 experimental data shown in ??(B).

398 [NB: How about moving this estimates paragraph and associated ?? to SI after all?]

399 We can begin to gain some intuition into how translation might limit growth by noting that the  
 400 total number of peptide bonds generated as the cell doubles,  $N_{aa}$ , which we used in our calculation  
 401 above, will be given by,  $\tau \cdot r_t \cdot R$ . Here,  $\tau$  refers to the doubling time of the cell under steady-state  
 402 growth,  $r_t$  is the maximum translation elongation rate, and  $R$  is the average number of ribosomes  
 403 per cell. With the growth rate related to the cell doubling time by  $\lambda = \ln(2)/\tau$ , we can write the

404 translation-limited growth rate as,

$$\lambda_{\text{translation-limited}} = \frac{\ln(2) \cdot r_t \cdot R}{N_{aa}}. \quad (1)$$

405 Alternatively, since  $N_{aa}$  is related to the total protein mass through the molecular weight of each  
 406 protein, we can also consider the growth rate in terms of ribosomal mass fraction. By making the  
 407 approximation that an average amino acid has a molecular weight of 110 Da (see ??(A)), we can  
 408 rewrite the growth rate as,

$$\lambda_{\text{translation-limited}} \approx \frac{\ln(2) \cdot r_t}{L_R} \Phi_R, \quad (2)$$

409 where  $L_R$  is the total length in amino acids that make up a ribosome, and  $\Phi_R$  is the ribosomal mass  
 410 fraction. This is plotted as a function of ribosomal fraction  $\Phi_R$  in ??(A), where we take  $L_R \approx 7459$  aa,  
 411 corresponding to the length in amino acids for all ribosomal subunits of the 50S and 30S complex.  
 412 This formulation assumes that the cell can transcribe the required amount of rRNA, which appears  
 413 reasonable for *E. coli* under the allowing us to consider the inherent limit on growth set by the  
 414 ribosome.

415 The growth rate defined by Equation ?? reflects mass-balance under steady-state growth and  
 416 has long provided a rationalization to the apparent linear increase in *E. coli*'s ribosomal content  
 417 as a function of growth rate (??). For our purposes, there are several important consequences  
 418 of this trend. Perhaps the first thing to notice is that there is a maximum growth rate at about  
 419  $\lambda \approx 6\text{hr}^{-1}$ , or doubling time of about 7 minutes (dashed line). This growth rate can be viewed as an  
 420 inherent maximum growth rate due to the need for the cell to double the cell's entire ribosomal  
 421 mass. Interestingly, this limit is independent of the absolute number of ribosomes, but rather is  
 422 simply given by time to translate an entire ribosome,  $L_R/r_t$ . As shown in ??(B), we can reconcile  
 423 this with the observation that in order to double the average number of ribosomes, each ribosome  
 424 must produce a second ribosome. This is a process that cannot be parallelized.

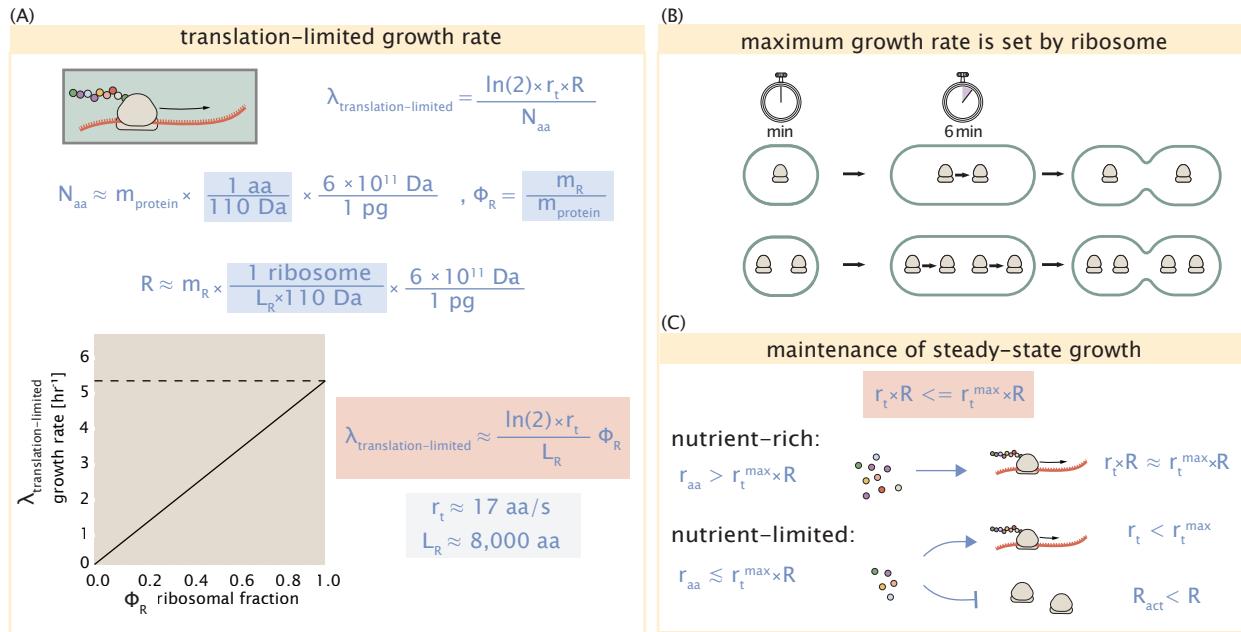
425 For reasonable values of  $\Phi_R$ , between about 0.1 - 0.3 (?), the maximum growth rate is in line  
 426 with experimentally reported growth rates around  $0.5 - 2\text{ hr}^{-1}$ . Here we are implicitly assuming that  
 427 translation proceeds randomly, without preference between ribosomal or non-ribosomal mRNA,  
 428 which appears reasonable. Importantly, in order for a cell to scale this growth limit set by  $\Phi_R$ , cells  
 429 must increase their ribosomal abundance. This can be achieved by either synthesizing more ribo-  
 430 somes or reducing the fraction of non-ribosomal proteins. Reduction of non-ribosomal proteins is  
 431 not straight forward since, as we have found throughout our estimates, doubling a cell requires a  
 432 substantial number of other enzymes and transporters. Increasing the absolute ribosomal abun-  
 433 dance is limited by the number of rRNA operons.

434 While it is common for bacteria to decrease their ribosomal abundance in poorer nutrient conditions  
 435 ??, this does not decrease to zero. From the perspective of a bacterium dealing with uncertain  
 436 nutrient conditions, there is likely a benefit for the cell to maintain some relative fraction of ribo-  
 437 somes to support rapid growth as nutrient conditions improve. However, if we consider a scenario  
 438 where nutrient conditions become poorer and poorer, there will be a regime where ribosomes are  
 439 in excess of the nutrient supply. If the cell is to maintain steady-state growth, it will need to attenu-  
 440 ate its translational activity since ribosomes would otherwise exhaust their supply of amino acids  
 441 and bring cell growth to a halt (??(C)). In the next section we will consider this more specifically for  
 442 *E. coli*, which has been shown to maintain a relatively high elongation rate even in stationary phase  
 443 ( $\approx 8\text{ aa/s}$ , ?) where cell growth is minimal.

444 [NB: I'm considering moving this paragraph near the end of the next section].

#### 445 **Multiple replication forks bias ribosome abundance.**

446 *E. coli* cells grow by an adder mechanism, whereby cells add a constant volume with each cell di-  
 447 vision (?). In conjunction with this, additional rounds of DNA replication are triggered when cells



**Figure 7. 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 ( $N_{\text{aa}} / r_t$ ) sets the translation-limited growth rate,  $\lambda_{\text{translation-limited}}$ . Here  $N_{\text{aa}}$  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  $m_{\text{protein}}$  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  $\Phi_R$  (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) Schematic showing translation-specific requirements for maintenance of steady-state growth. In a nutrient rich environment, amino acid supply  $r_{\text{aa}}$  is sufficiently in excess of demand by ribosomes translating at their maximal rate. In poorer nutrient conditions, reduced amino acid supply  $r_{\text{aa}}$  will decrease the rate of elongation. In a regime where  $r_{\text{aa}} < r_t \cdot R$ , the number of actively translating ribosomes will need to be reduced in order to maintain steady-state growth.

448 reach a critical volume per origin of replication (??(A)). This leads to the classically-described ex-  
 449 ponential increase in cell size with growth rate ???. In the context of maximizing growth rate, it  
 450 is notable that the majority of ribosomal proteins and rRNA operons are found closer to the DNA  
 451 origin. Given the need to increase to total gene dosage of rRNA operons at faster growth rates, and  
 452 the intimate relationship between ribosomal content and growth rate we considered above, this  
 453 raises the possibility that the observed size scaling and increase in chromosomal content might  
 454 simply be a means for the cell to tune biosynthesis according to its physiological state.

455 While an increase in transcription has been observed for genes near the origin in rapidly grow-  
 456 ing *E. coli* (?), we were unaware of such characterization at the proteomic level. In order to test  
 457 whether there is a relative increase in protein expression for genes closer to the origin, we calcu-  
 458 lated a running boxcar average of protein copy number as a function of each gene's transcriptional  
 459 start site. While absolute protein copy numbers can vary substantially across the chromosome, we  
 460 indeed observe a bias in expression under fast growth conditions (??(B), showing the result using  
 461 a 0.5 kb averaging window). The dramatic change in protein copy number near the origin mainly  
 462 reflects the increase in ribosomal protein expression. This trend is in contrast to slower growth con-  
 463 ditions where the average copy number is more uniform across the length of the chromosome.

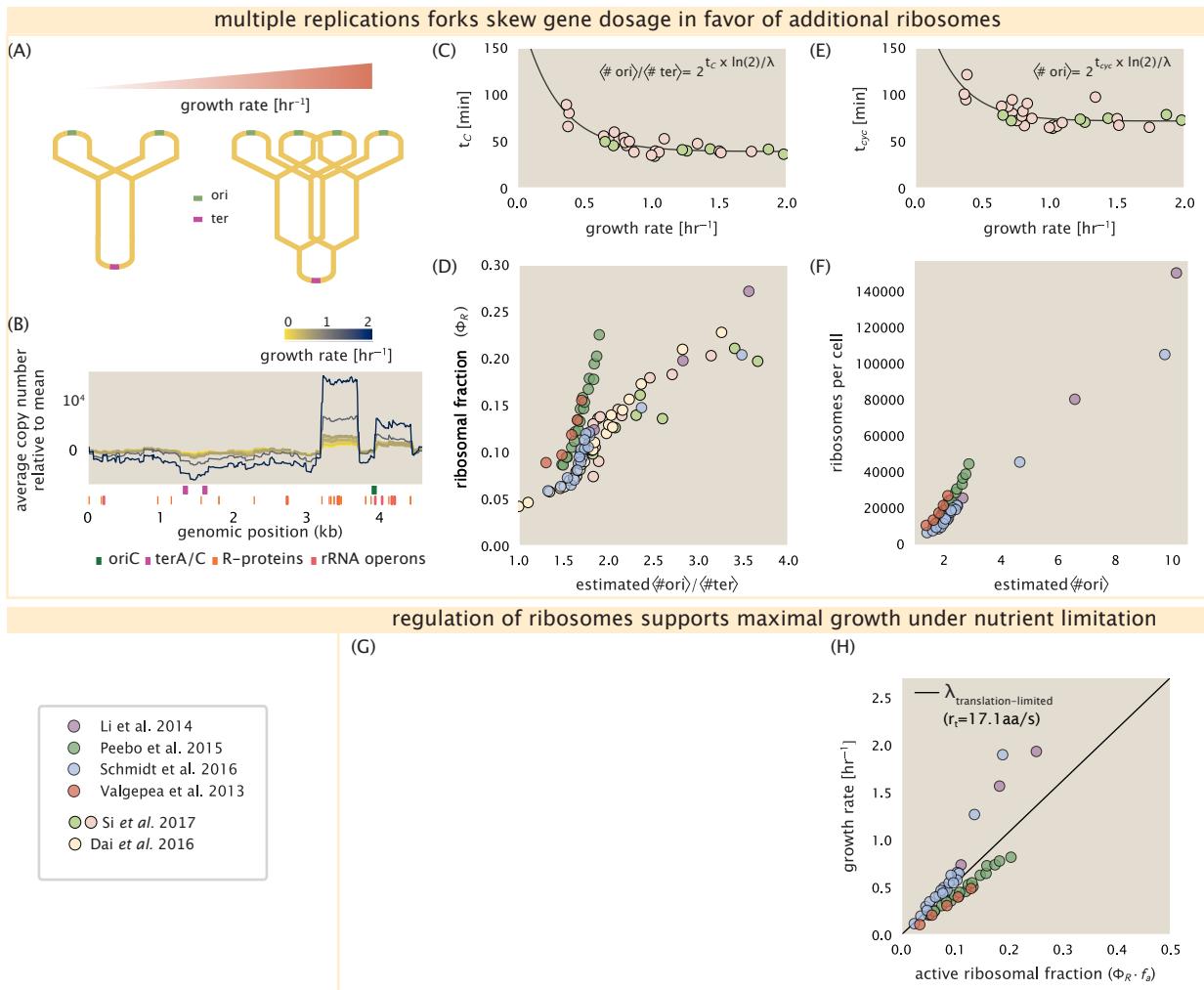
464 If ribosomal genes (rRNA and ribosomal proteins) are being synthesized according to their avail-  
 465 able gene dosage we can make two related hypotheses about how their abundance should vary  
 466 with chromosomal content. The first is that the ribosomal protein fraction should increase in pro-  
 467 portion to the average ratio of DNA origins to DNA termini ( $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$  ratio). This is a conse-  
 468 quence of the skew in DNA dosage as cells grow faster. The second is that the absolute number  
 469 of ribosomes should increase in proportion to the number of DNA origins ( $\langle \# \text{ ori} \rangle$ ), since this will  
 470 reflect the total gene dosage at a particular growth condition.

471 In order to test each of these expectations we considered the experimental data from Si *et*  
 472 *al.* (2017), which inferred these parameters for cells under nutrient-limited growth.  $\langle \# \text{ ori} \rangle / \langle \#$   
 $\text{ter} \rangle$  ratio depends on how quickly chromosomes are replicated relative the cell's doubling time  $\tau$   
 474 and is given by  $2^{\tau_C/\tau}$ . Here  $\tau_C$  is the time taken to replicate *E. coli*'s chromosome, referred to as  
 475 the C period of cell division. In ??(C) we plot  $\tau_C$  versus  $\tau$  that were measured, with data points in  
 476 red corresponding to *E. coli* strain MG1655, and blue to strain NCM3722. In their work they also  
 477 measured the total RNA to protein ratio which reflects ribosomal abundance and we show that data  
 478 along with other recent measurements from Dai *et al.*. Indeed we find that the ribosomal fraction  
 479 increases with  $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$  (??(C)). Across our different proteomic data sets there also appeared  
 480 two distinct trends. To consider the possibility that this may reflect systematic differences in how  
 481 the data was generated, we also considered recent measurements of total RNA to protein ratio  
 482 across the growth rates considered, which provide an alternative measure of ribosomal abundance  
 483 (RNA to protein ratio  $\approx \Phi_R \times 2.1$  ?). While these showed a similar correlation, they were most  
 484 consistent with the proteomic data from Schmidt *et al.* (2016) and Li *et al.* (2014).

485 We can similarly estimate  $\langle \# \text{ ori} \rangle$ , which depends on how often replication forks are initiated  
 486 per cell cycle. This is given by the number of overlapping cell cycles,  $2^{\tau_{\text{cycle}}/\tau}$ , where  $\tau_{\text{cycle}}$  refers to  
 487 the total time of chromosome replication and cell division. ??(E) shows the associated data from  
 488 Si *et al.*, which we use to estimate  $\langle \# \text{ ori} \rangle$  for each growth condition of the proteomic data. In  
 489 agreement with our expectations, we find a strong correlation between the ribosome copy number  
 490 and estimated  $\langle \# \text{ ori} \rangle$  (??(F)).

491 [NB: to do. 1) slow growth regime, 2) putting it all together ; cells appear to grow near the  
 492 translation-limited rate ( $r_t = 17 \text{ aa/s}$ ) across all growth conditions. Need to provide some rational-  
 493 ization for points above line. Maybe it's the interpretation of  $L_R$ , or the reality that a ribosome  
 494 complex is more complex than the simple picture of a 50S + 30S subunit considered here. In any  
 495 case, in the fast growth regime, this amounts to differences of minutes. ]

496 [NB: to incorporate. Titration of the cellular ppGpp concentration invoked similar proteomic  
 497 changes to those observed under nutrient limitation (?). In light of our hypothesis that such changes  
 498 to the proteome are intimately linked to the details of DNA replication, it was recently shown that



**Figure 8. 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 growth condition considered by Schmidt *et al.*. A 0.5 kb 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  $\text{RNA/protein ratio} \approx \Phi_R \cdot 2.1$ . (F) plots the ribosome copy number estimated from the proteomic data against our estimate of  $\langle \# \text{ori} \rangle$ . (G) [in progress], (H) 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.

499 both the  $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$  and cell size lost their growth rate dependent scaling in a ppGpp null strain.  
500 Rather, cells exhibit a  $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$  closer to 4 and cell size more consistent with a fast growth  
501 state (?). This supports the possibility that in addition to coordinating ribosome activity, (p)ppGpp  
502 signaling may be acting to coordinate other cellular processes in accordance with nutrient condi-  
503 tions and biosynthetic demand. From this perspective, the increase in the rate of DNA initiation and  
504 associated increase in cell size may be viewed as a way for the cell to vary its proteomic composition  
505 and biosynthetic capacity according to its available nutrient conditions. ]