

# Fundamental limits on the rate of bacterial cell division

**Nathan M. Belliveau<sup>†, 1</sup>, Griffin Chure<sup>†, 2, 3</sup>, Christina L. Hueschen<sup>4</sup>, Hernan G. 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>1</sup>Department of Biology, University of Washington, Seattle, WA, USA; <sup>2</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA; <sup>3</sup>Department of Applied Physics, California Institute of Technology, Pasadena, CA, USA; <sup>4</sup>Department of Chemical Engineering, Stanford University, Stanford, CA, USA; <sup>5</sup>Department of Molecular Cell Biology and Department of Physics, University of California Berkeley, Berkeley, CA, USA; <sup>6</sup>Department of Physics, Brandeis University, Waltham, MA, USA; <sup>7</sup>Department of Applied Physics, Stanford University, Stanford, CA, USA; <sup>8</sup>Allen Institute for Cell Science, Seattle, WA, USA; <sup>9</sup>Department of Physics, California Institute of Technology, Pasadena, CA, USA; \*Contributed equally

## Abstract

## Introduction

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

42 the growth rate varies in different carbon sources.

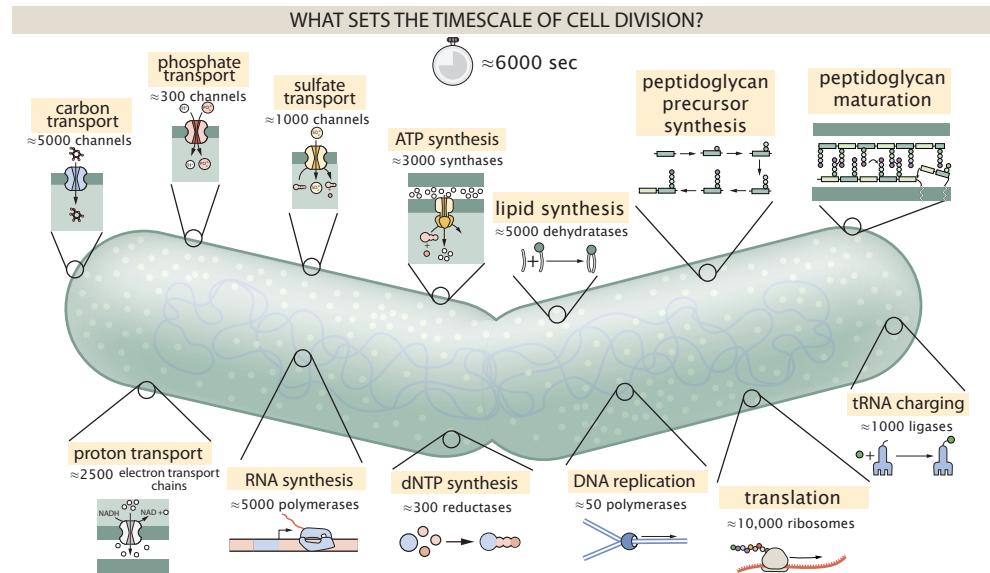
43 Specifically, we leverage a combination of *E. coli* proteomic data sets collected over the past  
 44 decade using either mass spectrometry (48; 40; 61) or ribosomal profiling (31) across 31 unique  
 45 growth conditions. Broadly speaking, we entertain several classes of hypotheses as are illustrated  
 46 in **Figure 1**. First, we consider potential limits on the transport of nutrients into the cell. We address  
 47 this hypothesis by performing an order-of-magnitude estimate for how many carbon, phosphorous,  
 48 and sulfur atoms are needed to facilitate this requirement given a 5000 second division time. As a  
 49 second 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 maintain  
 53 the proton motive force. A third class of estimates considers the need to maintain the size and  
 54 shape of the cell through the construction of new lipids for the cell membranes as well as the glycan  
 55 polymers which make up the rigid peptidoglycan. Our final class of hypotheses centers on the  
 56 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 **Figure 1** as  
 62 setting possible speed limits. Ultimately, we find that protein translation (particularly the generation  
 63 of 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 (50), 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 scavenged  
 74 from the environment in different forms. Carbon, for example, is acquired via the transport of  
 75 carbohydrates and sugar alcohols with some carbon sources receiving preferential treatment in their  
 76 consumption (35). Phosphorus, sulfur, and nitrogen, on the other hand, are harvested primarily in  
 77 the forms of inorganic salts, namely phosphate, sulfate, and ammonia (28; 6; 55; 4; 44; 63). 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 transport  
 81 mechanisms and the vast number of inputs needed to build a cell, we begin by considering transport  
 82 of some of the most important cellular ingredients: carbon, nitrogen, oxygen, hydrogen, phosphorus,  
 83 and sulfur.

84 The elemental composition of *E. coli* has received much quantitative attention over the past half  
 85 century (38; 60; 23; 7), 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  
 87 uncertainties), we can estimate that the dry mass of a typical *E. coli* cell is  $\approx$  45% carbon (BNID:  
 88 100649, (34)),  $\approx$  15% nitrogen (BNID: 106666, (34)),  $\approx$  3% phosphorus (BNID: 100653, (34)), and  
 89 1% sulfur (BNID: 100655, (34)). In the coming paragraphs, we will engage in a dialogue between  
 90 back-of-the-envelope estimates for the numbers of transporters needed to facilitate these chemical  
 91 stoichiometries and the experimental proteomic measurements of the biological reality. Such an



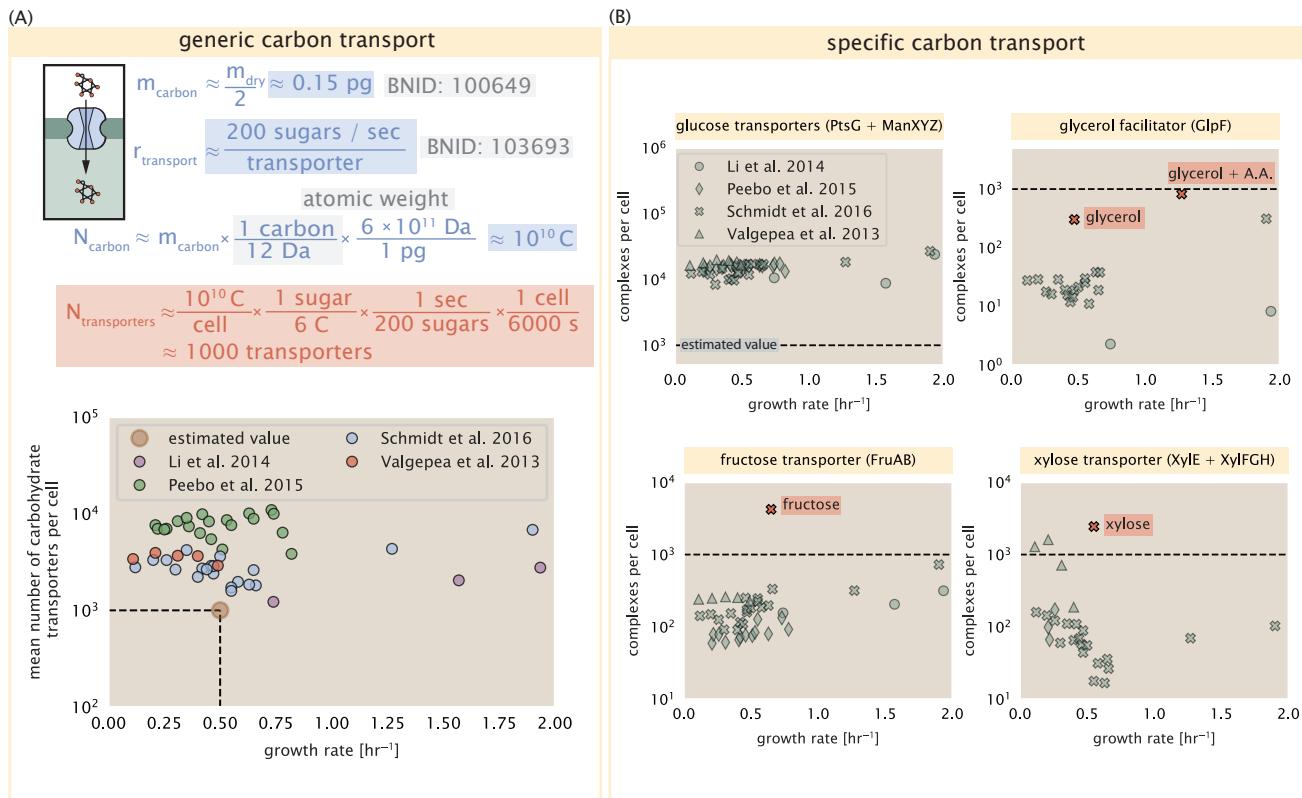
**Figure 1. Transport and synthesis processes necessary for cell division.** We consider an array of processes necessary for a cell to double its molecular components. Such processes include the transport of carbon across the cell membrane, the production of ATP, and fundamental processes of the central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

92 approach provides the opportunity to test if our biological knowledge is sufficient to understand the  
 93 scale at which these complexes are produced. Specifically, we will make these estimates considering  
 94 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  
 95 experimental measurements reside.

### 96 Carbon Transport

97 We begin with the most abundant element by mass, carbon. Using  $\approx 0.3 \text{ pg}$  as the typical *E. coli*  
 98 dry mass (BNID: 103904, (34)), we estimate that  $\approx 10^{10}$  carbon atoms must be brought into the  
 99 cell in order to double all of the carbon-containing molecules (**Figure 2(A, top)**). Typical laboratory  
 100 growth conditions, such as those explored in the aforementioned proteomic data sets, provide  
 101 carbon as a single class of sugar such as glucose, galactose, or xylose to name a few. *E. coli* has  
 102 evolved myriad mechanisms by which these sugars can be transported across the cell membrane.  
 103 One such mechanism of transport is via the PTS system which is a highly modular system capable  
 104 of transporting a diverse range of sugars (14). The glucose-specific component of this system  
 105 transports  $\approx 200$  glucose molecules per second per channel (BNID: 114686, (34)). Making the  
 106 assumption that this is a typical sugar transport rate, coupled with the need to transport  $10^{10}$  carbon  
 107 atoms, we arrive at the conclusion that on the order of 1,000 transporters must be expressed  
 108 in order to bring in enough carbon atoms to divide in 6,000 s, diagrammed in the top panel of  
 109 **Figure 2(A)**. This estimate, along with the observed average number of carbohydrate transporters  
 110 present in the proteomic data sets (48; 40; 61; 31), is shown in **Figure 2(A)**. While we estimate 1,000  
 111 transporters are needed, the data reveals that at a division time of  $\approx 5000$  s there is nearly a ten-fold  
 112 excess of transporters. Furthermore, the data illustrates that the average number of carbohydrate  
 113 transporters present is largely-growth rate independent.

114 The estimate presented in **Figure 2(A)** neglects any specifics of the regulation of carbon transport  
 115 system and presents a data-averaged view of how many carbohydrate transporters are present  
 116 on average. Using the diverse array of growth conditions explored in the proteomic data sets, we  
 117 can explore how individual carbon transport systems depend on the population growth rate. In



**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.

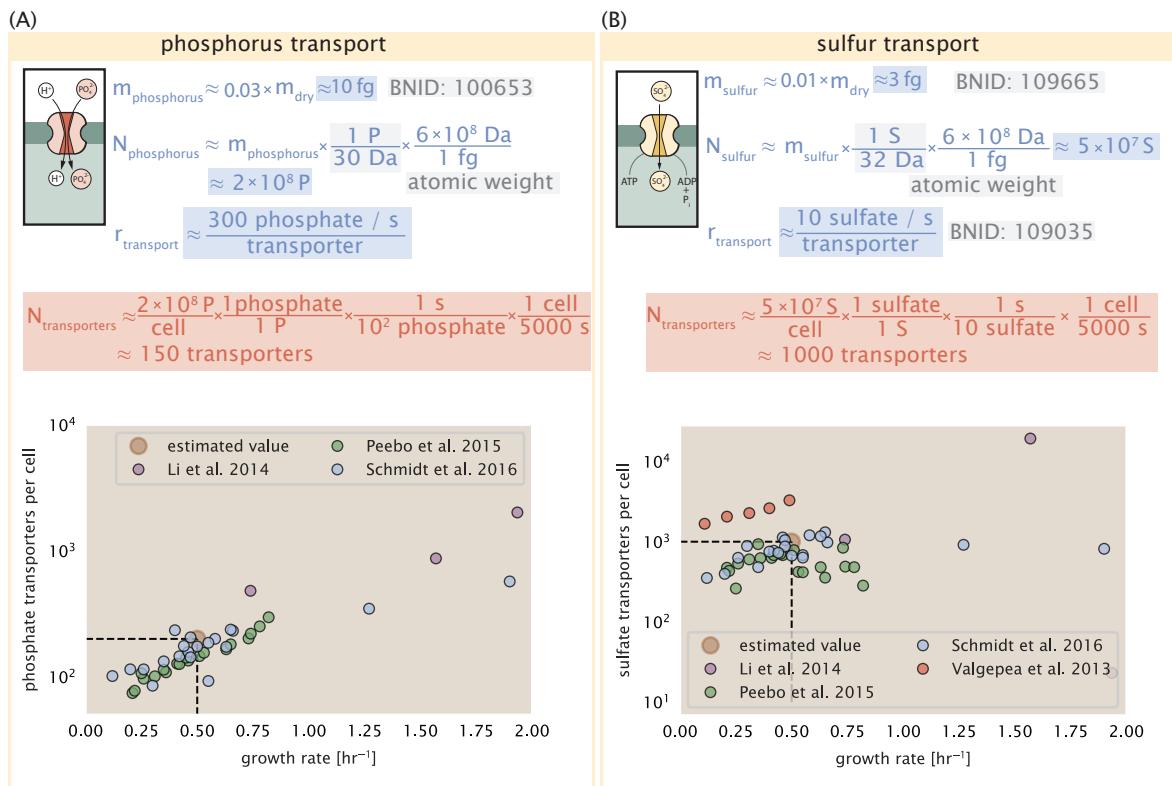
118 **Figure 2(B)**, we show the total number of carbohydrate transporters specific to different carbon  
 119 sources. A striking observation, shown in the top-left plot of **Figure 2(B)**, is the constancy in the  
 120 expression of the glucose-specific transport systems (the PtsG enzyme of the PTS system and the  
 121 glucose-transporting ManXYZ complex). Additionally, we note that the total number of glucose-  
 122 specific transporters is tightly distributed  $\approx 10^4$  per cell, an order of magnitude beyond the estimate  
 123 shown in **Figure 2(A)**. This illustrates that *E. coli* maintains a substantial number of complexes  
 124 present for transporting glucose which is known to be the preferential carbon source (35; 33; 3).

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 (19; 65). 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 (26), though we now know of many other such examples (25; 19; 8). This  
 131 illustrates that once glucose is depleted from the environment, cells have a means to dramatically  
 132 increase the abundance of the specific transporter needed to digest the next sugar that is present.  
 133 Several examples of induced expression of specific carbon-source transporters are shown in  
 134 **Figure 2(B)**. Points colored in red (labeled by red text-boxes) correspond to growth conditions in  
 135 which the specific carbon source (glycerol, xylose, or fructose) is present. These plots show that, in  
 136 the absence of the particular carbon source, expression of the transporters is maintained on the  
 137 order of  $\sim 10^2$  per cell. However, when induced, the transporters become highly-expressed and  
 138 are present on the order of  $\sim 10^4$  per cell, which exceeds the generic estimation given in **Figure 2(A)**.  
 139 Together, this generic estimation and the specific examples of induced expression suggest that  
 140 transport of carbon across the cell membrane, while critical for growth, is not the rate-limiting step  
 141 of cell division.

142 In the context of speeding up growth, one additional limitation is the fact that the cell's inner  
 143 membrane is occupied by a multitude of other membrane proteins. Considering a rule-of-thumb  
 144 for the surface area of *E. coli* of about  $6 \mu\text{m}^2$  (BNID: 101792, (34)), we expect an areal density for  
 145 1,000 transporters to be approximately 200 transporters/ $\mu\text{m}^2$ . For a glucose transporter occupying  
 146 about  $50 \text{ nm}^2/\text{dimer}$ , this amounts to about only 1 percent of the total inner membrane (58). In  
 147 addition, bacterial cell membranes typically have densities of  $10^5$  proteins// $\mu\text{m}^2$  (42), implying that  
 148 the cell could accommodate more transporters if it were rate limiting.

### 149 Phosphorus and Sulfur Transport

150 We now turn our attention towards other essential elements, namely phosphorus and sulfur.  
 151 Phosphorus is critical to the cellular energy economy in the form of high-energy phosphodiester  
 152 bonds making up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-  
 153 translational modification of proteins and defining the polar-heads of lipids. In total, phosphorus  
 154 makes up  $\approx 3\%$  of the cellular dry mass which in typical experimental conditions is in the form of  
 155 inorganic phosphate. The cell membrane has remarkably low permeability to this highly-charged  
 156 and critical molecule, therefore requiring the expression of active transport systems. In *E. coli*, the  
 157 proton electrochemical gradient across the inner membrane is leveraged to transport inorganic  
 158 phosphate into the cell (44). Proton-solute symporters are widespread in *E. coli* (43; 10) and can  
 159 have rapid transport rates of 50 molecules per second for sugars and other solutes (BNID: 103159;  
 160 111777, (34)). In *E. coli* the PitA phosphate transport system has been shown to very tightly coupled  
 161 with the proton electrochemical gradient with a 1:1 proton:phosphate stoichiometric ratio (22; 15).  
 162 Illustrated in **Figure 3(A)**, we can estimate that  $\approx 300$  phosphate transporters are necessary to  
 163 maintain an  $\approx 3\%$  dry mass with a 6,000 s division time. This estimate is again satisfied when we  
 164 examine the observed copy numbers of PitA in proteomic data sets (plot in **Figure 3(A)**). While our  
 165 estimate is very much in line with the observed numbers, we emphasize that this is likely a slight  
 166 over estimate of the number of transporters needed as there are other phosphorous scavenging  
 167 systems, such as the ATP-dependent phosphate transporter Pst system which we have neglected.



**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     Satisfied that there are a sufficient number of phosphate transporters present in the cell, we now  
169     turn to sulfur transport as another potentially rate limiting process. Similar to phosphate, sulfate is  
170     highly-charged and not particularly membrane permeable, requiring active transport. While there  
171     exists a H<sup>+</sup>/sulfate symporter in *E. coli*, it is in relatively low abundance and is not well characterized  
172     (64). Sulfate is predominantly acquired via the ATP-dependent ABC transporter CysUWA system  
173     which also plays an important role in selenium transport (51; 54). While specific kinetic details  
174     of this transport system are not readily available, generic ATP transport systems in prokaryotes  
175     transport on the order of 1 to 10 molecules per second (BNID: 109035, (34)). Combining this generic  
176     transport rate, measurement of sulfur comprising 1% of dry mass, and a 6,000 second division time  
177     yields an estimate of  $\approx 1000$  CysUWA complexes per cell (**Figure 3(B)**). Once again, this estimate is  
178     in notable agreement with proteomic data sets, suggesting that there are sufficient transporters  
179     present to acquire the necessary sulfur. In a similar spirit of our estimate of phosphorus transport,  
180     we emphasize that this is likely an overestimate of the number of necessary transporters as we  
181     have neglected other sulfur scavenging systems that are in lower abundance.

### 182     Nitrogen Transport

183     Finally, we turn to nitrogen transport as the last remaining transport system highlighted in **Figure 1**.  
184     Unlike phosphate, sulfate, and various sugar molecules, nitrogen in the form of ammonia can  
185     readily diffuse across the cell membrane and has a permeability on par with water ( $\approx 10^5$  nm/s,  
186     BNID:110824 (34)). In particularly nitrogen-poor conditions, *E. coli* expresses a transporter (AmtB)  
187     which appears to aid in nitrogen assimilation, though the mechanism and kinetic details of transport  
188     is still a matter of debate (62; 30). Beyond ammonia, another plentiful source of nitrogen come in

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

## 192 Function of the Central Dogma

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

## 198 DNA

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

## 206 dNTP synthesis

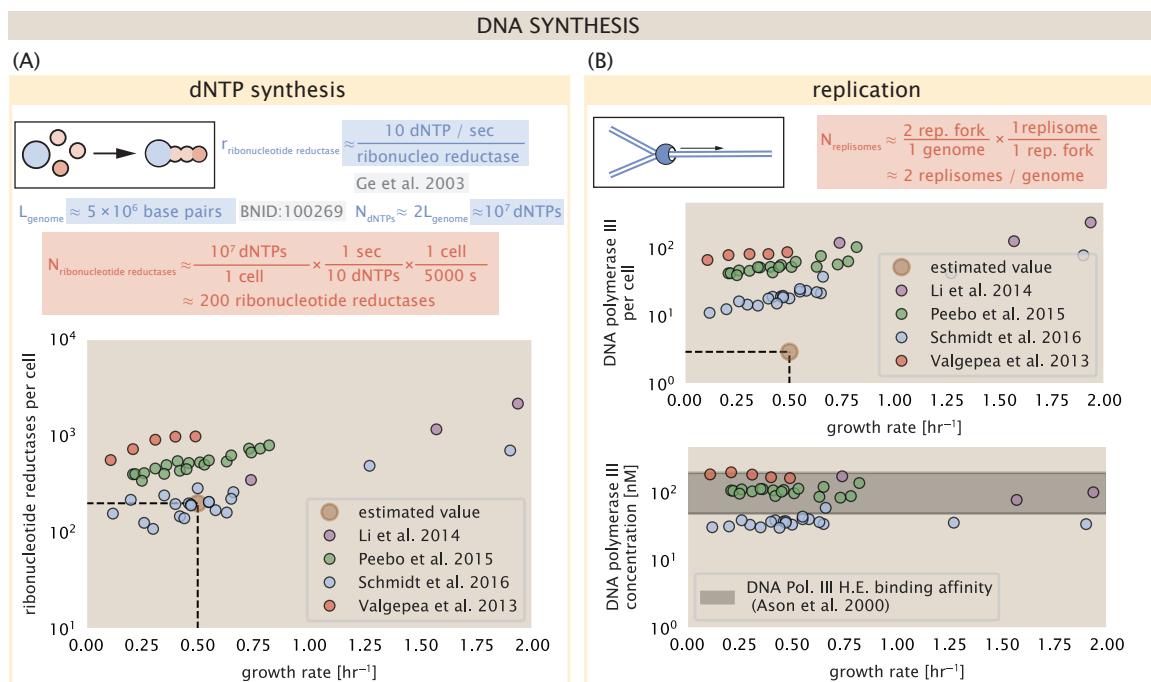
207 We begin our exploration DNA replication by examining the production of the deoxyribonucleotide  
 208 triphosphates (dNTPs). The four major dNTPs (dATP, dTTP, dCTP, and dGTP) are synthesized *de novo*  
 209 in separate pathways, requiring different building blocks. However, a critical step present in all dNTP  
 210 synthesis pathways is the conversion from ribonucleotide to deoxyribonucleotide via the removal  
 211 of the 3' hydroxyl group of the ribose ring (45). This reaction is mediated by a class of enzymes  
 212 termed ribonucleotide reductases, of which *E. coli* possesses two aerobically active complexes  
 213 (termed I and II) and a single anaerobically active enzyme. Due to their peculiar formation of  
 214 a radical intermediate, these enzymes have received much biochemical, kinetic, and structural  
 215 characterization. One such work (20) performed a detailed *in vitro* measurement of the steady-state  
 216 kinetic rates of these complexes, revealing a turnover rate of ≈ 10 dNTP per second.

217 Since this reaction is central to the synthesis of all dNTPs, it is reasonable to consider the  
 218 abundance of these complexes is a measure of the total dNTP production in *E. coli*. Illustrated  
 219 schematically in **Figure 4** (A), we consider the fact that to replicate the cell's genome, on the order of  
 220 ≈ 10<sup>7</sup> dNTPs must be synthesized. Assuming a production rate of 10 per second per ribonucleotide  
 221 reductase complex and a cell division time of 6000 seconds, we arrive at an estimate of ≈ 150  
 222 complexes needed per cell. As shown in the bottom panel of **Figure 4** (A), this estimate agrees  
 223 with the experimental measurements of these complexes abundances within ≈ 1/2 an order of  
 224 magnitude.

225 Recent work has revealed that during replication, the ribonucleotide reductase complexes  
 226 coalesce to form discrete foci colocalized with the DNA replisome complex (46). This is particularly  
 227 pronounced in conditions where growth is slow, indicating that spatial organization and regulation  
 228 of the activity of the complexes plays an important role.

## 229 DNA Replication

230 We now turn our focus to the integration of these dNTP building blocks into the replicated chromo-  
 231 some strand via the DNA polymerase. Replication is initiated at a single region of the chromosome  
 232 termed the *oriC* locus at which a pair of DNA polymerases bind and begin their high-fidelity replica-  
 233 tion of the genome in opposite directions. Assuming equivalence between the two replication forks,  
 234 this means that the two DNA polymerase complexes (termed replisomes) meet at the midway point  
 235 of the circular chromosome termed the *ter* locus. The kinetics of the five types of DNA polymerases



**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 (52) (See Appendix Section 4).

(I – V) have been intensely studied, revealing that DNA polymerase III performs the high fidelity processive replication of the genome with the other "accessory" polymerases playing auxiliary roles (17). *In vitro* measurements have shown that DNA Polymerase III copies DNA at a rate of  $\approx$  600 nucleotides per second (BNID: 104120, (34)). Therefore, to replicate a single chromosome, two DNA polymerases replicating at their maximal rate would copy their entire genome in  $\approx$  4000 s. Thus, with a division time of 5000 s (our "typical" growth rate for the purposes of this work), there is sufficient time for a pair of DNA polymerase III complexes to replicate the entire genome. However, this estimate implies that 4000 s would be the upper-limit time scale for bacterial division which is at odds with the familiar 20 minute (1200 s) doubling time of *E. coli* in rich medium.

It is well known that *E. coli* can parallelize its DNA replication such that multiple chromosomes are being replicated at once, with as many as 10 - 12 replication forks at a given time (11; 53). Thus, even in rapidly growing cultures, we expect only a few polymerases ( $\approx$  10) are needed to replicate the chromosome per cell doubling. However, as shown in **Figure 4(B)**, DNA polymerase III is nearly an order of magnitude more abundant. This discrepancy can be understood by considering its binding constant to DNA. DNA polymerase III is highly processive, facilitated by a strong affinity of the complex to the DNA. *In vitro* biochemical characterization has quantified the  $K_D$  of DNA polymerase III holoenzyme to single-stranded and double-stranded DNA to be 50 and 200 nM, respectively (5). The bottom plot in **Figure 4 (B)** shows that the concentration of the DNA polymerase III across all data sets and growth conditions is within this range. Thus, while the copy number of the DNA polymerase III is in excess of the strict number required to replicate the genome, its copy number appears to vary such that its concentration is approximately equal to the dissociation constant to the DNA. While the processes regulating the initiation of DNA replication are complex and involve more than just the holoenzyme, these data indicate that the kinetics of replication rather than the explicit copy number of the DNA polymerase III holoenzyme is the more relevant feature of DNA replication to consider. In light of this, the data in **Figure 4(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 (27), the time to double their chromosome likely represents an upper limit to their growth rate.

### 265 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.

### 275 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, (34)), meaning each ribosome contains  $\approx$  4500 nucleotides. As the *E. coli* RNA polymerase transcribes DNA to RNA at a rate of  $\approx$  40 nucleotides per second (BNID: 101904, (34)), it takes a single RNA polymerase  $\approx$  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  $\approx$  50 functional ribosomal rRNA units – far below the observed number of  $\approx$   $10^4$  ribosomes per cell.

284 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  $\approx 1$  per second (BNID: 111997; 102362, (34)). 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, (34)) four of which are localized directly adjacent to the origin of replication (9). 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 (56). 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, (34)), 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.

### 310 mRNA

311 To form a functional protein, all protein coding genes must first be transcribed from DNA to form  
 312 an mRNA molecule. While each protein requires an mRNA blueprint, many copies of the protein can  
 313 be synthesized from a single mRNA. Factors such as strength of the ribosomal binding site, mRNA  
 314 stability, and rare codon usage frequency dictate the number of proteins that can be made from a  
 315 single mRNA, with yields ranging from  $10^1$  to  $10^4$  (BNID: 104186; 100196; 106254, (34)). Computing  
 316 the geometric mean of this range yields  $\approx 1000$  proteins synthesized per mRNA, a value that agrees  
 317 with experimental measurements of the number of proteins per cell ( $\approx 3 \times 10^6$ , BNID: 100088, (34))  
 318 and total number of mRNA per cell ( $\approx 3 \times 10^3$ , BNID: 100064, (34)).

319 This estimation captures the *steady-state* mRNA copy number, meaning that at any given time,  
 320 there will exist approximately 3000 unique mRNA molecules. To determine the *total* number of  
 321 mRNA that need to be synthesized over the cell's lifetime, we must consider degradation of the  
 322 mRNA. In most bacteria, mRNAs are rather unstable with life times on the order of several minutes  
 323 (BNID: 104324; 106253; 111927; 111998, (34)). For convenience, we assume that the typical mRNA  
 324 in our cell of interest has a typical lifetime of  $\approx 300$  seconds. Using this value, we can determine  
 325 the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell.  
 326 While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional  
 327 dynamics, we state here that the total mRNA production rate must be on the order of  $\approx 15$  mRNA  
 328 per second. In *E. coli*, the average protein is  $\approx 300$  amino acids in length (BNID: 108986; (34)),  
 329 meaning that the corresponding mRNA is  $\approx 900$  nucleotides which we will further approximate as  $\approx$   
 330 1000 nucleotides to account for the non-protein coding regions on the 5' and 3' ends. This means  
 331 that the cell must have enough RNA polymerase molecules about to sustain a transcription rate of  
 332  $\approx 1.5 \times 10^4$  nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA at a  
 333 clip of 40 nucleotides per second, we arrive at a comfortable estimate of  $\approx 250$  RNA polymerase  
 334 complexes needed to satisfy the mRNA demands of the cell. It is worth noting that this number is

335 approximately half of that required to synthesize enough rRNA, as we saw in the previous section.  
 336 We find this to be a striking result as these 250 RNA polymerase molecules are responsible for the  
 337 transcription of the  $\approx 4000$  protein coding genes that are not ribosome associated.

### 338 tRNA

339 The final class of RNA molecules worthy of quantitative consideration is the the pool of tRNAs used  
 340 during translation to map codon sequence to amino acid identity. Unlike mRNA or rRNA, each  
 341 individual tRNA is remarkably short, ranging from 70 to 95 nucleotides each (BNID: 109645; 102340,  
 342 (34)). What they lack in length, they make up for in abundance. There are approximately  $\approx 3000$   
 343 tRNA molecules present for each of the 20 amino acids (BNID: 105280, (34)), although the precise  
 344 copy number is dependent on the identity of the ligated amino acid. Using these values, we make  
 345 the estimate that  $\approx 5 \times 10^6$  nucleotides are sequestered in tRNA per cell. Unlike mRNA, tRNA is  
 346 remarkably stable with typical lifetimes *in vivo* on the order of  $\approx 48$  hours (2; 57) – well beyond the  
 347 timescale of division. Once again using our rule-of-thumb for the rate of transcription to be 40  
 348 nucleotides per second and assuming a division time of  $\approx 5000$  seconds, we arrive at an estimate  
 349 of  $\approx 20$  RNA polymerases to synthesize enough tRNA. This requirement pales in comparison to the  
 350 number of polymerases needed to generate the rRNA and mRNA pools and can be neglected as a  
 351 significant transcriptional burden.

### 352 RNA Polymerase and $\sigma$ -factor Abundance

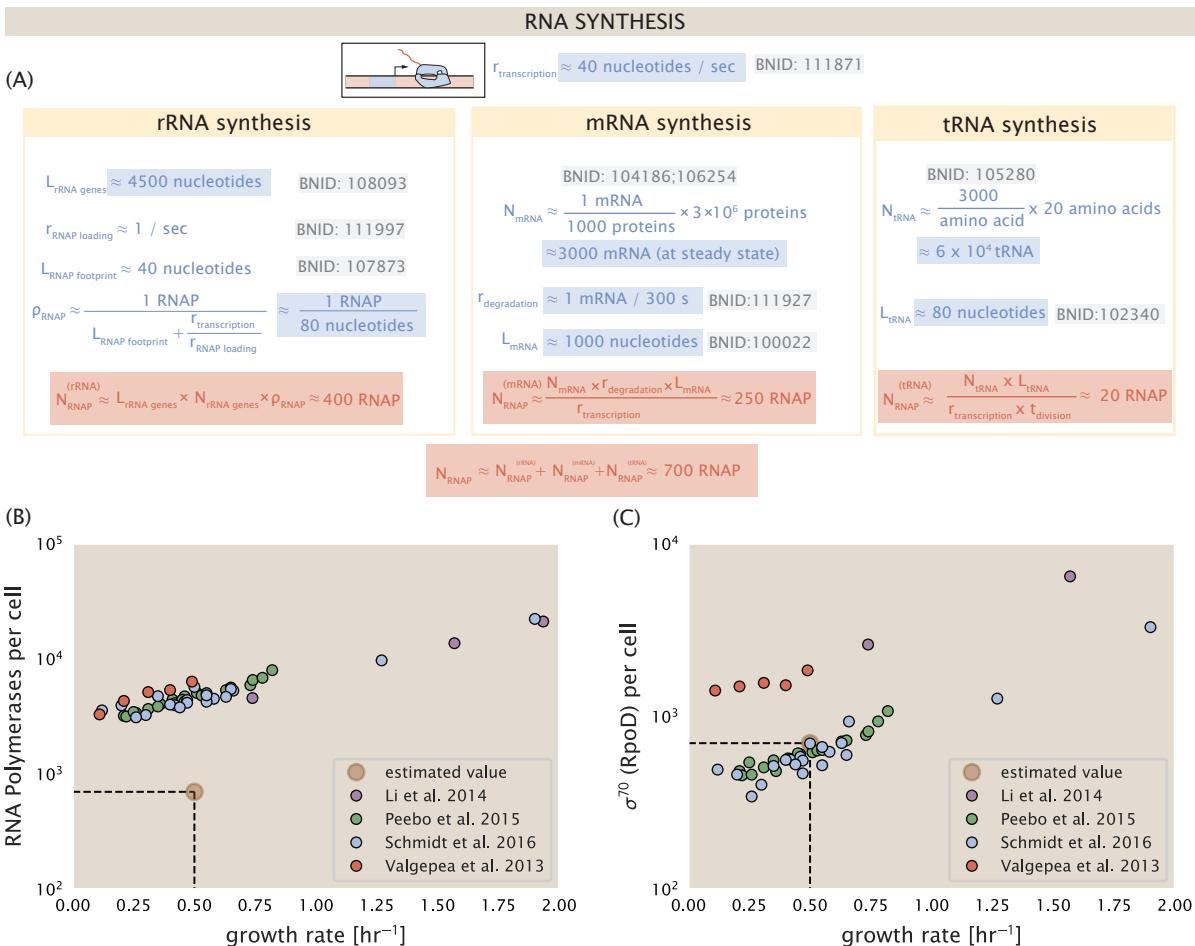
353 These estimates, summarized in **Figure 5** (A), reveal that synthesis of rRNA and mRNA are the  
 354 dominant RNA species synthesized by RNA polymerase, suggesting the need for  $\approx 700$  RNA poly-  
 355 merases per cell. As is revealed in **Figure 5** (B), this estimate is about an order of magnitude below  
 356 the observed number of RNA polymerase complexes per cell ( $\approx 5000$  - 7000). The disagreement  
 357 between the estimated number of RNA polymerases and these observations are at least consistent  
 358 with recent literature revealing that  $\approx 80$  % of RNA polymerases in *E. coli* are not transcriptionally  
 359 active (39). Our estimate ignores the possibility that some fraction is only nonspecifically bound to  
 360 DNA, as well as the obstacles that RNA polymerase and DNA polymerase present for each other as  
 361 they move along the DNA (18).

362 In addition, it is also vital to consider the role of  $\sigma$ -factors which help RNA polymerase identify and  
 363 bind to transcriptional start sites (12). Here we consider  $\sigma^{70}$  (RpoD) which is the dominant "general-  
 364 purpose"  $\sigma$ -factor in *E. coli*. While initially thought of as being solely involved in transcriptional  
 365 initiation, the past two decades of single-molecule work has revealed a more multipurpose role for  
 366  $\sigma^{70}$  including facilitating transcriptional elongation (29; 21; 41; 37; 36). **Figure 5** (B) is suggestive of  
 367 such a role as the number of  $\sigma^{70}$  proteins per cell is in close agreement with our estimate of the  
 368 number of transcriptional complexes needed.

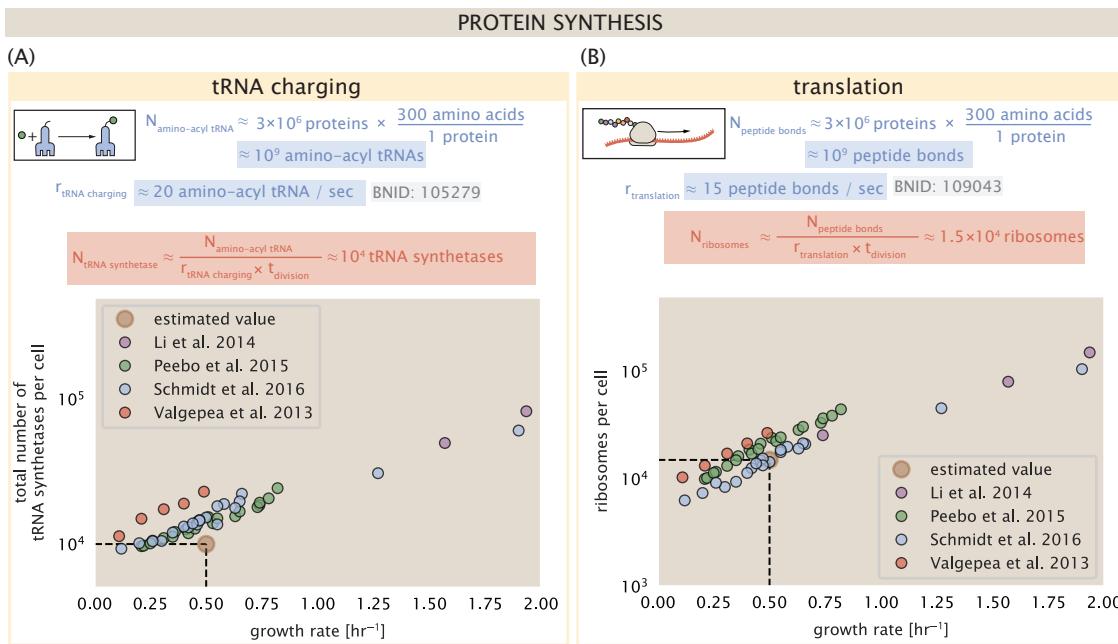
369 While these estimates and comparison with experimental data reveal an interesting dynamic at  
 370 play between the transcriptional demand and copy numbers of the corresponding machinery, these  
 371 findings illustrate that transcription cannot be the rate limiting step in bacterial division. **Figure 5**  
 372 (A) reveals that the availability of RNA polymerase is not a limiting factor for cell division as the cell  
 373 always has an apparent  $\sim 10$ -fold excess than needed. Furthermore, if more transcriptional activity  
 374 was needed to satisfy the cellular requirements, more  $\sigma^{70}$ -factors could be expressed to utilize a  
 375 larger fraction of the RNA polymerase pool.

### 376 Protein synthesis

377 Lastly, we turn our attention to the process of translation. So far our estimates have led to protein  
 378 copy numbers that are consistent with the proteomic data, or even in excess of what might be  
 379 needed for each task under limiting growth conditions. Even in our example of *E. coli* grown under  
 380 different carbohydrate sources (**Figure 2(B)**), cells can utilize alternative carbon sources by inducing  
 381 the expression of additional membrane transporters and enzymes. Optimal resource allocation  
 382 and the role of ribosomal proteins have been an area of intense quantitative study over the last



**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 ( $[{\text{ArgS}}, {\text{CysS}}, {\text{GlnS}}, {\text{GlxS}}, [{\text{IleS}}, {\text{LeuS}}, {\text{ValS}}, {\text{AlaS}}]_2, [{\text{AsnS}}, {\text{AspS}}]_2, [{\text{TyrS}}, {\text{TrpS}}]_2, [{\text{ThrS}}]_2, [{\text{SerS}}]_2, [{\text{ProS}}]_2, [{\text{PheS}}]_2[{\text{PheT}}]_2, [{\text{MetG}}]_2, [{\text{lysS}}]_2, [{\text{HisS}}]_2, [{\text{GlyS}}]_2[{\text{GlyQ}}]_2]$ ). (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  $0.5 \text{ hr}^{-1}$ .

383 decade by Hwa and others (50; 24). From the perspective of limiting growth, our earlier estimate  
 384 of rRNA highlighted the necessity for multiple copies of rRNA genes in order to make enough  
 385 rRNA, suggesting the possibility that synthesis of ribosomes might be rate limiting. While the  
 386 transcriptional demand for the ribosomal proteins is substantially lower than rRNA genes, since  
 387 many proteins can be translated from relatively fewer mRNA, other ribosomal proteins like the  
 388 translation elongation factor EF-Tu also present a substantial burden. For EF-Tu in particular, it is  
 389 the most highly expressed protein in *E. coli* and is expressed by multiple genes on the chromosome,  
 390 *tufA* and *tufB*.

391 We begin by first estimating the number of tRNA synthetases and ribosomes required for a  
 392 doubling time of 5000 seconds. *E. coli* has roughly  $3 \times 10^6$  proteins per cell, which for an average  
 393 protein of 300 aa, amounts to the formation of  $\approx 10^9$  peptide bonds. This also corresponds to  
 394 the number of amino-acyl tRNA that are used by ribosomes, with the pool of tRNA continuously  
 395 recharging new amino acids by tRNA synthetases. At a rate of charging of about 20 amino-acyl tRNA  
 396 per second (BNID: 105279, (34)), we find that cells have more than sufficient tRNA synthetases to  
 397 meet the demand of ribosomes during protein synthesis (Figure 6(A)). If we consider an elongation  
 398 rate of  $\approx 15$  peptide bonds per second (BNID: 114271, (34; 13)), the formation of  $\approx 10^9$  peptide  
 399 bonds would require  $1.5 \times 10^4$  ribosomes at a growth rate of  $0.5 \text{ hr}^{-1}$ . This is indeed consistent with  
 400 the experimental data shown in Figure 6(B).

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

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

407 translation-limited growth rate as,

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

408 Alternatively, since  $N_{aa}$  is related to the total protein mass through the molecular weight of each  
 409 protein, we can also consider the growth rate in terms of ribosomal mass fraction. By making the  
 410 approximation that an average amino acid has a molecular weight of 110 Da (see **Figure 7(A)**), we  
 411 can rewrite the growth rate as,

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

412 where  $L_R$  is the total length in amino acids that make up a ribosome, and  $\Phi_R$  is the ribosomal  
 413 mass fraction. This is plotted as a function of ribosomal fraction  $\Phi_R$  in **Figure 7(A)**, where we take  
 414  $L_R \approx 7459$  aa, corresponding to the length in amino acids for all ribosomal subunits of the 50S and  
 415 30S complex. This formulation assumes that the cell can transcribe the required amount of rRNA,  
 416 which appears reasonable for *E. coli* under the allowing us to consider the inherent limit on growth  
 417 set by the ribosome.

418 The growth rate defined by Equation 2 reflects mass-balance under steady-state growth and  
 419 has long provided a rationalization to the apparent linear increase in *E. coli*'s ribosomal content as  
 420 a function of growth rate (Gol; 50). For our purposes, there are several important consequences  
 421 of this trend. Perhaps the first thing to notice is that there is a maximum growth rate at about  
 422  $\lambda \approx 6\text{hr}^{-1}$ , or doubling time of about 7 minutes (dashed line). This growth rate can be viewed as an  
 423 inherent maximum growth rate due to the need for the cell to double the cell's entire ribosomal  
 424 mass. Interestingly, this limit is independent of the absolute number of ribosomes, but rather  
 425 is simply given by time to translate an entire ribosome,  $L_R/r_t$ . As shown in **Figure 7(B)**, we can  
 426 reconcile this with the observation that in order to double the average number of ribosomes, each  
 427 ribosome must produce a second ribosome. This is a process that cannot be parallelized.

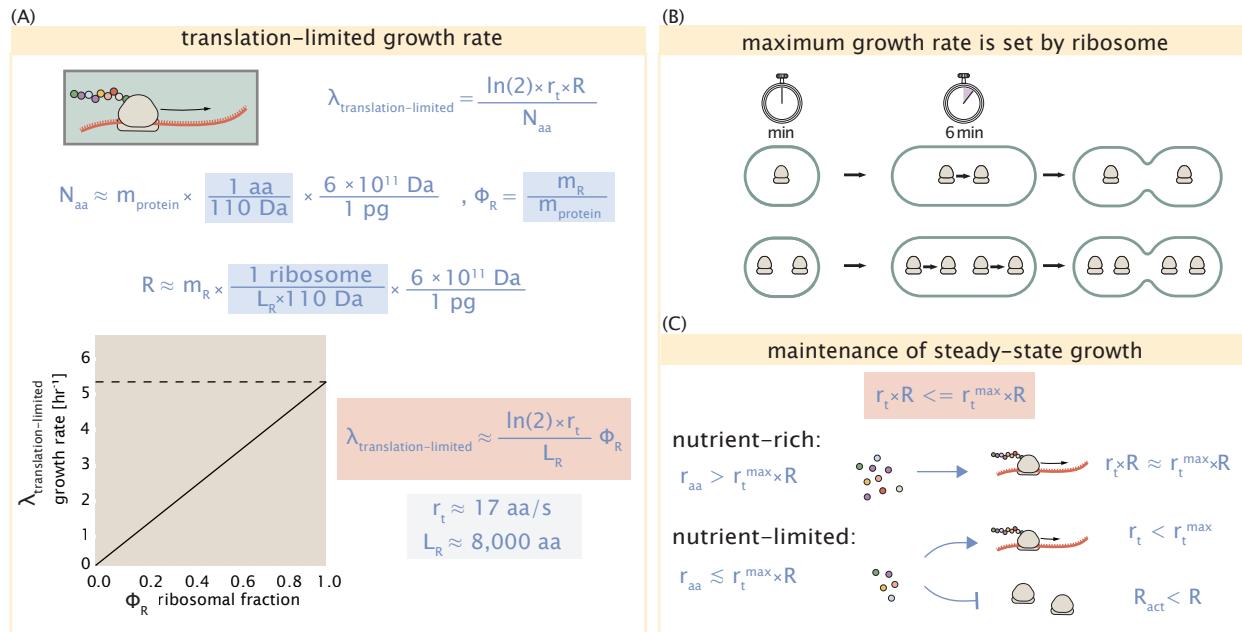
428 For reasonable values of  $\Phi_R$ , between about 0.1 - 0.3 (50), the maximum growth rate is in line  
 429 with experimentally reported growth rates around 0.5 - 2  $\text{hr}^{-1}$ . Here we are implicitly assuming  
 430 that translation proceeds randomly, without preference between ribosomal or non-ribosomal  
 431 mRNA, which appears reasonable. Importantly, in order for a cell to scale this growth limit set by  
 432  $\Phi_R$ , cells *must* increase their ribosomal abundance. This can be achieved by either synthesizing  
 433 more ribosomes or reducing the fraction of non-ribosomal proteins. Reduction of non-ribosomal  
 434 proteins is not straight forward since, as we have found throughout our estimates, doubling a  
 435 cell requires a substantial number of other enzymes and transporters. Increasing the absolute  
 436 ribosomal abundance is constrained by production of rRNA.

437 While it is common for bacteria to decrease their ribosomal abundance in poorer nutrient  
 438 conditions (50; 32), this does not decrease to zero. From the perspective of a bacterium dealing with  
 439 uncertain nutrient conditions, there is likely a benefit for the cell to maintain some relative fraction  
 440 of ribosomes to support rapid growth as nutrient conditions improve. However, if we consider  
 441 a scenario where nutrient conditions become poorer and poorer, there will be a regime where  
 442 ribosomes are in excess of the nutrient supply. If the cell is to maintain steady-state growth, it will  
 443 need to attenuate its translational activity since ribosomes would otherwise exhaust their supply of  
 444 amino acids and bring cell growth to a halt (**Figure 7(C)**). In the next section we will consider this  
 445 more specifically for *E. coli*, which has been shown to maintain a relatively high elongation rate even  
 446 in stationary phase ( $\approx 8$  aa/s, (13)) where cell growth is minimal.

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

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

449 *E. coli* cells grow by an adder mechanism, whereby cells add a constant volume with each cell  
 450 division (59). In conjunction with this, additional rounds of DNA replication are triggered when cells  
 451 reach a critical volume per origin of replication (**Figure 8(A)**). This leads to the classically-described



**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  $r_{\text{aa}}$  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.

452 exponential increase in cell size with growth rate (47; 53; 52). In the context of maximizing growth  
 453 rate, it is notable that the majority of ribosomal proteins and rRNA operons are found closer to the  
 454 DNA origin. Given the need to increase total gene dosage of rRNA operons at faster growth rates,  
 455 this raises the possibility that the observed size scaling and increase in chromosomal equivalents  
 456 might simply be a means for the cell to tune biosynthesis according to their physiological state.

457 While an increase in transcription has been observed for genes near the origin in rapidly growing  
 458 *E. coli* (49), we were unaware of such characterization at the proteomic level. In order to test whether  
 459 there is a relative increase in protein expression for genes closer to the origin, we calculated a  
 460 running boxcar average of protein copy number as a function of each gene's transcriptional start site.  
 461 While absolute protein copy numbers can vary substantially across the chromosome, we indeed  
 462 observe a bias in expression under fast growth conditions (**Figure 8(B)**, showing the result using  
 463 a 0.5 kb averaging window). The dramatic change in protein copy number near the origin mainly  
 464 reflects the increase in ribosomal protein expression. This trend is in contrast to slower growth  
 465 conditions where the average copy number is more uniform across the length of the chromosome.

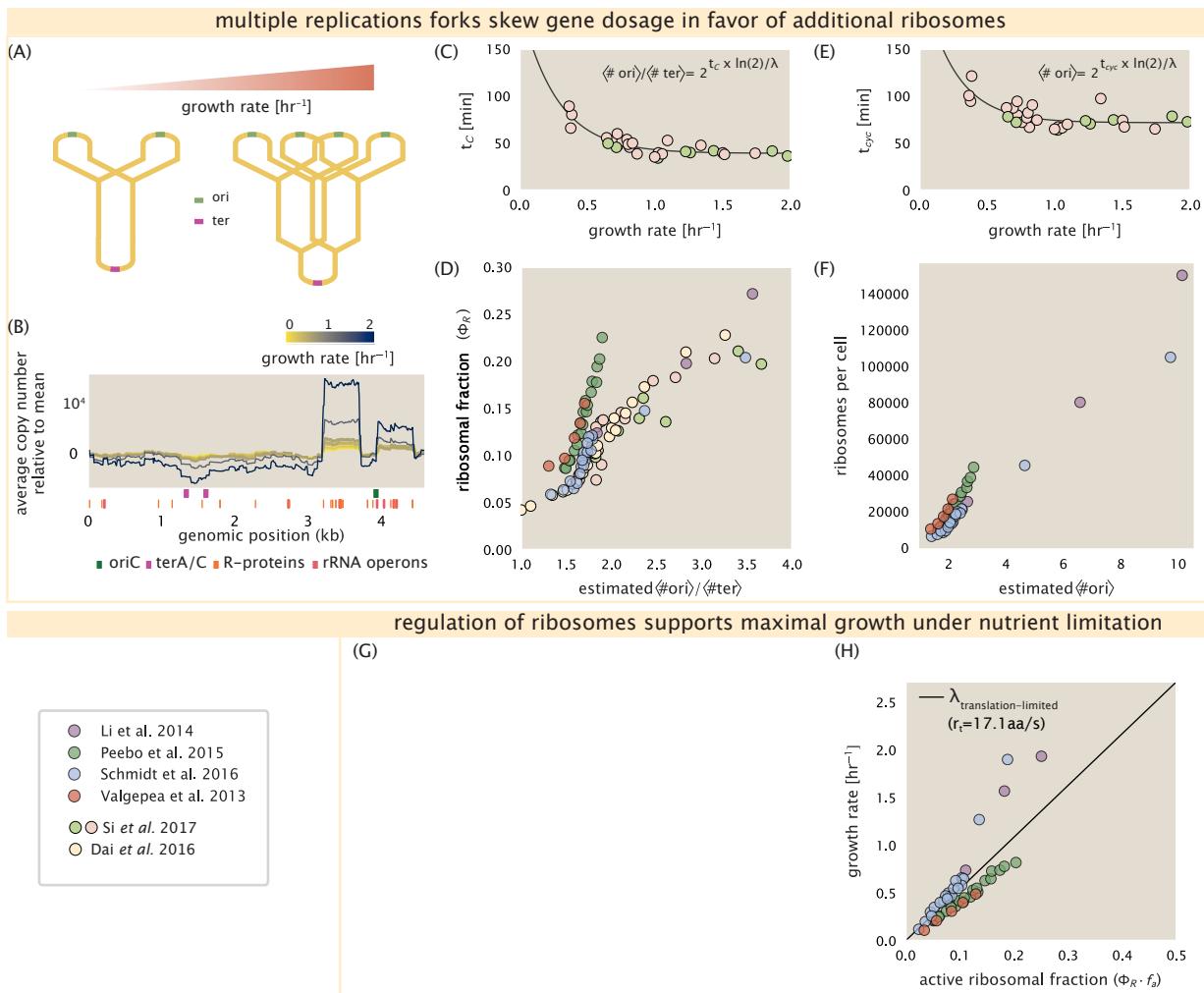
466 If ribosomal genes (rRNA and ribosomal proteins) are being synthesized according to their  
 467 available gene dosage we can make two related hypotheses about how their abundance should  
 468 vary with chromosomal content. The first is that the ribosomal protein fraction should increase  
 469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

473 In order to test each of these expectations we considered the experimental data from Si *et 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  
 474  $\tau$  and is given by  $2^{\tau_C/\tau}$ . Here  $\tau_C$  is the time taken to replicate *E. coli*'s chromosome, referred to  
 475 as the C period of cell division. In **Figure 8(C)** we plot  $\tau_C$  versus  $\tau$  that were measured, with data  
 476 points in red corresponding to *E. coli* strain MG1655, and blue to strain NCM3722. In their work  
 477 they also measured the total RNA to protein ratio which reflects ribosomal abundance and we  
 478 show that data along with other recent measurements from Dai *et al.*. Indeed we find that the  
 479 ribosomal fraction increases with  $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$  (**Figure 8(C)**). Across our different proteomic  
 480 data sets there also appeared two distinct trends. To consider the possibility that this may reflect  
 481 systematic differences in how the data was generated, we also considered recent measurements of  
 482 total RNA to protein ratio across the growth rates considered, which provide an alternative measure  
 483 of ribosomal abundance (RNA to protein ratio  $\approx \Phi_R \times 2.1$  (13)). While these showed a similar  
 484 correlation, they were most consistent with the proteomic data from Schmidt *et al.* (2016) and Li *et  
 485 al.* (2014).

486 We can similarly estimate  $\langle \# \text{ ori} \rangle$ , which depends on how often replication forks are initiated  
 487 per cell cycle. This is given by the number of overlapping cell cycles,  $2^{\tau_{\text{cyc}}/\tau}$ , where  $\tau_{\text{cyc}}$  refers to  
 488 the total time of chromosome replication and cell division. **Figure 8(E)** shows the associated data  
 489 from Si *et al.*, which we use to estimate  $\langle \# \text{ ori} \rangle$  for each growth condition of the proteomic data. In  
 490 agreement with our expectations, we find a strong correlation between the ribosome copy number  
 491 and estimated  $\langle \# \text{ ori} \rangle$  (**Figure 8(F)**).

492 [NB: to do. 1) slow growth regime, 2) putting it all together ; cells appear to grow near the  
 493 translation-limited rate ( $r_i = 17 \text{ aa/s}$ ) across all growth conditions. Need to provide some rationalization  
 494 for points above line.]

495 [NB: Titration of the cellular ppGpp concentration invoked similar proteomic changes to those  
 496 observed under nutrient limitation (66). In light of our hypothesis that such changes to the proteome  
 497 are intimately linked to the details of DNA replication, it was recently shown that both the  $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$   
 498 and cell size lost their growth rate dependent scaling in a ppGpp null strain. Rather, cells exhibit  
 499 a  $\langle \# \text{ ori} \rangle / \langle \# \text{ ter} \rangle$  closer to 4 and cell size more consistent with a fast growth state (16). This supports  
 500 the possibility that in addition to coordinating ribosome activity, (p)ppGpp signaling may be acting  
 501 502



**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  $\langle \# \text{ori} \rangle / \langle \# \text{ter} \rangle$  [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) 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.

503 to coordinate other cellular processes in accordance with nutrient conditions and biosynthetic  
 504 demand. From this perspective, the increase in the rate of DNA initiation and associated increase in  
 505 cell size may be viewed as a way for the cell to vary its proteomic composition and biosynthetic  
 506 capacity according to its available nutrient conditions. ]

## 507 References

- 508 [Gol] *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On.*
- 509 [2] Abelson, H., Johnson, L., Penman, S., and Green, H. (1974). Changes in RNA in relation to growth of the  
 510 fibroblast: II. The lifetime of mRNA, rRNA, and tRNA in resting and growing cells. *Cell*, 1(4):161–165.
- 511 [3] Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose  
 512 sugars in *Escherichia coli*. *BMC Systems Biology*, 8(1):133.
- 513 [4] Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes  
 514 studied by ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- 515 [5] Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B.  
 516 (2000). A Model for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends:  
 517 DNA Triggers A Change In Binding Specificity of the  $\gamma$  Complex Clamp Loader. *Journal of Biological Chemistry*,  
 518 275(4):3006–3015.
- 519 [6] Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin  
 520 4 as a NH<sub>3</sub> Channel. *Journal of Biological Chemistry*, 291(36):19184–19195.
- 521 [7] Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and  
 522 Bioengineering*, 18(1):81–94. \_eprint: <https://onlinelibrary.wiley.com/doi/10.1002/bit.260180107>.
- 523 [8] Belliveau, N. M., Barnes, S. L., Ireland, W. T., Jones, D. L., Sweredoski, M. J., Moradian, A., Hess, S., Kinney,  
 524 J. B., and Phillips, R. (2018). Systematic approach for dissecting the molecular mechanisms of transcriptional  
 525 regulation in bacteria. *Proceedings of the National Academy of Sciences*, 115(21):E4796–E4805.
- 526 [9] Birnbaum, L. S. and Kaplan, S. (1971). Localization of a Portion of the Ribosomal RNA Genes in *Escherichia  
 527 coli*. *Proceedings of the National Academy of Sciences*, 68(5):925–929.
- 528 [10] Booth, I. R., Mitchell, W. J., and Hamilton, W. A. (1979). Quantitative analysis of proton-linked transport  
 529 systems. The lactose permease of *Escherichia coli*. *Biochemical Journal*, 182(3):687–696.
- 530 [11] Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell  
 531 at Different Exponential Growth Rates. *EcoSal Plus*, 3(1).
- 532 [12] Browning, D. F. and Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria.  
 533 *Nature Reviews Microbiology*, 14(10):638–650.
- 534 [13] Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang,  
 535 Y.-P., and Hwa, T. (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation  
 536 rates during slow growth. *Nature Microbiology*, 2(2):16231.
- 537 [14] Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia  
 538 coli* phosphoenolpyruvate–carbohydrate phosphotransferase system: Peculiarities of regulation and impact  
 539 on growth and product formation. *Applied Microbiology and Biotechnology*, 94(6):1483–1494.
- 540 [15] Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis,  
 541 V., and Palsson, B. Ø. (2007). A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that  
 542 accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology*, 3(1):121.
- 543 [16] Fernández-Coll, L., Maciąg-Dorszynska, M., Tailor, K., Vadia, S., Levin, P. A., Szalewska-Palasz, A., Cashel, M.,  
 544 and Dunny, G. M. (2020). The Absence of (p)ppGpp Renders Initiation of *Escherichia coli* Chromosomal DNA  
 545 Synthesis Independent of Growth Rates. *mBio*, 11(2):45.
- 546 [17] Fijalkowska, I. J., Schaaper, R. M., and Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: A  
 547 multi-DNA polymerase affair. *FEMS Microbiology Reviews*, 36(6):1105–1121.
- 548 [18] Finkelstein, I. J. and Greene, E. C. (2013). Molecular Traffic Jams on DNA. *Annual Review of Biophysics*,  
 549 42(1):241–263.

- 550 [19] Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo,  
 551 J. S., Alquicira-Hernández, K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J. A., Medina-Rivera, A., Solano-  
 552 Lira, H., Bonavides-Martínez, C., Pérez-Rueda, E., Alquicira-Hernández, S., Porrón-Sotelo, L., López-Fuentes,  
 553 A., Hernández-Koutoucheva, A., Moral-Chávez, V. D., Rinaldi, F., and Collado-Vides, J. (2016). RegulonDB  
 554 version 9.0: High-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic  
 555 Acids Research*, 44(D1):D133–D143.
- 556 [20] Ge, J., Yu, G., Ator, M. A., and Stubbe, J. (2003). Pre-Steady-State and Steady-State Kinetic Analysis of *E. coli*  
 557 Class I Ribonucleotide Reductase. *Biochemistry*, 42(34):10071–10083.
- 558 [21] Goldman, S. R., Nair, N. U., Wells, C. D., Nickels, B. E., and Hochschild, A. (2015). The primary  $\sigma$  factor in  
 559 *Escherichia coli* can access the transcription elongation complex from solution *in vivo*. *eLife*, 4:e10514.
- 560 [22] Harris, R. M., Webb, D. C., Howitt, S. M., and Cox, G. B. (2001). Characterization of PitA and PitB from  
 561 *Escherichia coli*. *Journal of Bacteriology*, 183(17):5008–5014.
- 562 [23] Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter  
 563 and elemental content of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- 564 [24] Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015).  
 565 Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular  
 566 Systems Biology*, 11(2):e784–e784.
- 567 [25] Ireland, W. T., Beeler, S. M., Flores-Bautista, E., Belliveau, N. M., Sweredoski, M. J., Moradian, A., Kinney, J. B.,  
 568 and Phillips, R. (2020). Deciphering the regulatory genome of *Escherichia coli*, one hundred promoters at a  
 569 time. *bioRxiv*.
- 570 [26] Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of  
 571 Molecular Biology*, 3(3):318–356.
- 572 [27] Jensen, R. B., Wang, S. C., and Shapiro, L. (2001). A moving DNA replication factory in Caulobacter crescentus.  
 573 *The EMBO journal*, 20(17):4952–4963.
- 574 [28] Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history,  
 575 recent progress, and the future with focus on cell size control: A review. *Reports on Progress in Physics*,  
 576 81(5):056601.
- 577 [29] Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonja, E., Mekler,  
 578 V., Ebright, R. H., and Weiss, S. (2005). Retention of Transcription Initiation Factor  $\Sigma$ 70 in Transcription  
 579 Elongation: Single-Molecule Analysis. *Molecular Cell*, 20(3):347–356.
- 580 [30] Khademi, S., O'Connell, J., Remis, J., Robles-Colmenares, Y., Miercke, L. J. W., and Stroud, R. M. (2004).  
 581 Mechanism of Ammonia Transport by Amt/MEP/Rh: Structure of AmtB at 1.35 Å. *Science*, 305(5690):1587–  
 582 1594.
- 583 [31] Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates  
 584 reveals principles underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- 585 [32] Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of  
 586 protein investment in cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- 587 [33] Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs  
 588 Reveal a Carbon Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–  
 589 15927.
- 590 [34] Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key  
 591 numbers in molecular and cell biology. *Nucleic Acids Research*, 38(suppl\_1):D750–D753.
- 592 [35] Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and  
 593 cellular differentiation. *Growth Symposium*, 9:223–289.
- 594 [36] Mooney, R. A., Darst, S. A., and Landick, R. (2005). Sigma and RNA Polymerase: An On-Again, Off-Again  
 595 Relationship? *Molecular Cell*, 20(3):335–345.
- 596 [37] Mooney, R. A. and Landick, R. (2003). Tethering  $\Sigma$ 70 to RNA polymerase reveals high *in vivo* activity of  $\sigma$   
 597 factors and  $\Sigma$ 70-dependent pausing at promoter-distal locations. *Genes & Development*, 17(22):2839–2851.

- 598 [38] Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- 600 [39] Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- 602 [40] Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in  
603 *Escherichia coli* with increasing specific growth rate. *Molecular BioSystems*, 11(4):1184–1193.
- 604 [41] Perdue, S. A. and Roberts, J. W. (2011).  $\sigma^{70}$ -dependent Transcription Pausing in *Escherichia coli*. *Journal of Molecular Biology*, 412(5):782–792.
- 606 [42] Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer,  
607 Cham, Cham.
- 608 [43] Ramos, S. and Kaback, H. R. (1977). The relation between the electrochemical proton gradient and active  
609 transport in *Escherichia coli* membrane vesicles. *Biochemistry*, 16(5):854–859.
- 610 [44] Rosenberg, H., Gerdes, R. G., and Chegwidden, K. (1977). Two systems for the uptake of phosphate in  
611 *Escherichia coli*. *Journal of Bacteriology*, 131(2):505–511.
- 612 [45] Rudd, S. G., Valerie, N. C. K., and Helleday, T. (2016). Pathways controlling dNTP pools to maintain genome  
613 stability. *DNA Repair*, 44:193–204.
- 614 [46] Sánchez-Romero, M. A., Molina, F., and Jiménez-Sánchez, A. (2011). Organization of ribonucleoside  
615 diphosphate reductase during multifork chromosome replication in *Escherichia coli*. *Microbiology*, 157(8):2220–  
616 2225.
- 617 [47] Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature  
618 of Cell Size and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*,  
619 19(3):592–606.
- 620 [48] Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold,  
621 R., and Heinemann, M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- 623 [49] Scholz, S. A., Diao, R., Wolfe, M. B., Fivenson, E. M., Lin, X. N., and Freddolino, P. L. (2019). High-Resolution  
624 Mapping of the *Escherichia coli* Chromosome Reveals Positions of High and Low Transcription. *Cell Systems*,  
625 8(3):212–225.e9.
- 626 [50] Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth  
627 and gene expression: origins and consequences. *Science*, 330(6007):1099–1102.
- 628 [51] Sekowska, A., Kung, H.-F., and Danchin, A. (2000). Sulfur Metabolism in *Escherichia coli* and Related Bacteria:  
629 Facts and Fiction. *Journal of Molecular Microbiology and Biotechnology*, 2(2):34.
- 630 [52] Si, F., Le Treut, G., Sauls, J. T., Vadia, S., Levin, P. A., and Jun, S. (2019). Mechanistic Origin of Cell-Size Control  
631 and Homeostasis in Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- 632 [53] 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,  
633 S. (2017). Invariance of Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*,  
634 27(9):1278–1287.
- 635 [54] Sirko, A., Zatyka, M., Sadowy, E., and Hulanicka, D. (1995). Sulfate and thiosulfate transport in *Escherichia coli* K-12: Evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins. *Journal of Bacteriology*, 177(14):4134–4136.
- 638 [55] Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE.  
639 *BMC Microbiology*, 19.
- 640 [56] Stevenson, B. S. and Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in  
641 *Escherichia coli*. *Applied and Environmental Microbiology*, 70(11):6670–6677.
- 642 [57] Svenningsen, S. L., Kongstad, M., Stenum, T. S. n., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017).  
643 Transfer RNA is highly unstable during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Research*,  
644 45(2):793–804.

- 645 [58] Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism?  
646 Testing the Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- 647 [59] Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015).  
648 Cell-size control and homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- 649 [60] Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M.  
650 (2010). Genome-derived minimal metabolic models for *Escherichia coli* MG1655 with estimated  
651 in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*, 107(2):369–381. \_eprint:  
652 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- 653 [61] Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by  
654 increasing catalytic and translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- 655 [62] van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen Assimilation in *Escherichia coli*:  
656 Putting Molecular Data into a Systems Perspective. *Microbiology and Molecular Biology Reviews*, 77(4):628–695.
- 657 [63] Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*:  
658 Involvement of Two Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*,  
659 113(2):529–539.
- 660 [64] Zhang, L., Jiang, W., Nan, J., Almqvist, J., and Huang, Y. (2014a). The *Escherichia coli* CysZ is a pH dependent  
661 sulfate transporter that can be inhibited by sulfite. *Biochimica et Biophysica Acta (BBA) - Biomembranes*,  
662 1838(7):1809–1816.
- 663 [65] Zhang, Z., Aboulwafa, M., and Saier, M. H. (2014b). Regulation of crp gene expression by the catabolite  
664 repressor/activator, cra, in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology*, 24(3):135–141.
- 665 [66] Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal  
666 resource allocation in *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.