

# Fundamental limits on the rate of bacterial growth and their influence on proteomic composition

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

<sup>1</sup>Department of Biology, Howard Hughes Medical Institute, University of Washington, Seattle, WA; <sup>2</sup>Department of Applied Physics, California Institute of Technology, Pasadena, CA, USA; <sup>3</sup>Department of Chemical Engineering, Stanford University, Stanford, CA, USA; <sup>4</sup>Department of Molecular Cell Biology and Department of Physics, University of California Berkeley, Berkeley, CA, USA; <sup>5</sup>Department of Physics, Brandeis University, Waltham, MA, USA; <sup>6</sup>Department of Applied Physics, Stanford University, Stanford, CA, USA; <sup>7</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA; <sup>8</sup>Department of Physics, California Institute of Technology, Pasadena, CA, USA; \*Co-corresponding authors. Address correspondence to phillips@pboc.caltech.edu and jtheriot@uw.edu; †These authors contributed equally to this work

---

## Abstract

Recent years have seen an experimental deluge interrogating the relationship between bacterial growth rate, cell size, and protein content. However, we still lack a rigorous understanding of what sets the scale of these quantities and when protein abundances should depend on growth rate. Here, we seek to quantitatively understand this relationship across a collection of *Escherichia coli* proteomic data covering  $\approx 4000$  proteins and 36 growth rates. We estimate the basic requirements and physical constraints on steady-state growth by considering key processes in cellular growth. In contrast to perspectives that specific processes limit growth rate or dictate cell size, our analysis suggests that cells are predominantly tuned for the task of cell doubling across a continuum of growth rates. Importantly, a theoretical inability to parallelize ribosomal synthesis places a firm limit on the achievable growth rate that is reflected under moderate to fast growth rates. We expand on this assessment by exploring a model of proteomic regulation as a function of nutrient supply, revealing a mechanism that ties size and growth rate to proteomic content.

---

## Introduction

The observed range of bacterial growth rates is enormously diverse. In natural environments, some microbial organisms may double only once per year (*Mikucki et al., 2009*) while in comfortable laboratory conditions, growth can be rapid with several divisions per hour (*Schaechter et al., 1958*). This six order-of-magnitude difference in time scales of growth encompasses different microbial species and lifestyles, yet even for a single species such as *Escherichia coli*, the growth rate can be modulated over a large scale by tuning the type and amount of nutrients in the growth medium (*Liu et al., 2005*). This remarkable plasticity in growth rate illustrates the intimate relationship between environmental conditions and the rates at which cells convert nutrients into new cellular material – a relationship that has remained a major topic of inquiry in bacterial physiology for over a century (*Jun et al., 2018*).

A key discovery in bacterial physiology of the past 70 years was the identification of bacterial "growth laws" (*Schaechter et al., 1958*); empirical relationships that relate the bacterial growth rate to the protein and RNA composition of the intracellular milieu in a number of different species. Over the past decade, a flurry of work (*Molenaar et al., 2009; Scott et al., 2010; Klumpp and Hwa, 2014; Basan et al., 2015; Dai et al., 2016; Erickson et al., 2017*) has examined these growth laws at a quantitative level, developing a series of phenomenological models from which the growth laws naturally emerge. In parallel, a "molecular revolution" in biology has yielded an increasingly refined molecular census of the cell, particularly for bacteria such as the microbial workhorse *E. coli* (*Schmidt et al., 2016; Davidi et al., 2016*). In light of the now expansive trove of quantitative biological data, we can revisit several

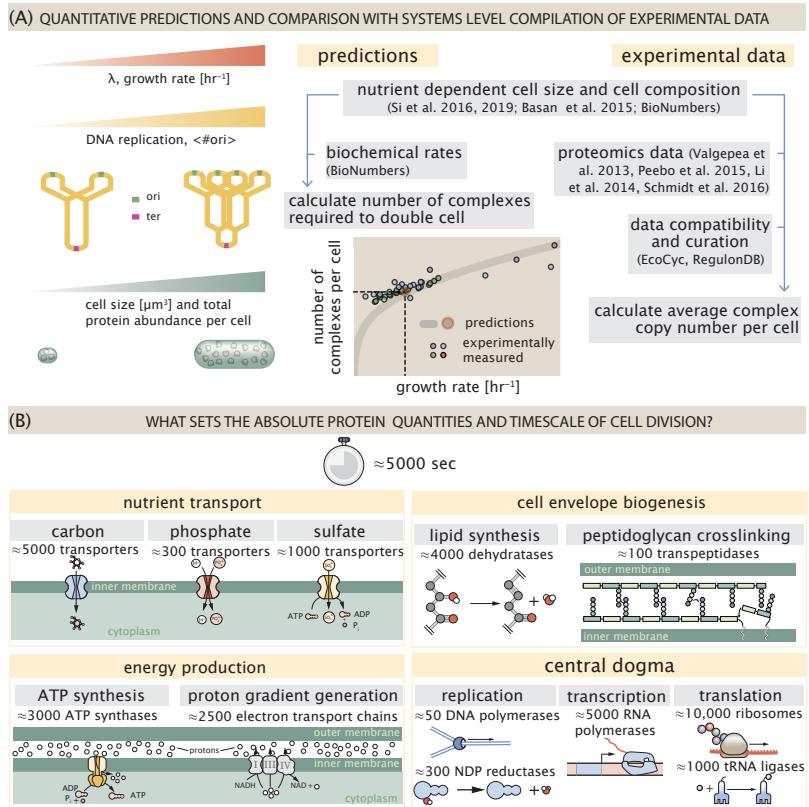
of the evergreen questions about bacterial growth and physiology that were originally raised by microbiologists in the middle of the 20th century. Specifically, what biological processes are the primary determinants for how quickly bacterial cells can grow and reproduce? Why do cells modulate the absolute numbers and relative ratios of their molecular constituents in response to changes in growth rate or nutrient availability? These questions remain under intense inquiry and have implicated processes ranging from ribosomal biogenesis and transcription to cell wall and lipid synthesis as key determinants of growth rate and cell size (*Bremer and Dennis, 2008; Scott et al., 2010; Si et al., 2017; Vadia et al., 2017; Harris and Theriot, 2018; Büke et al., 2020; Zhang et al., 2020*).

In this work, we consider these two questions in *E. coli* by considering both the biosynthetic capacity of key cellular processes – meaning, the minimal number of enzymes needed to synthesize one cell’s worth of a particular biomolecule given the observed doubling time – as well as the physical constraints given the now well-characterized change in cell size as a function of growth rate (*Taheri-Araghi et al., 2015; Si et al., 2017; Basan et al., 2015*). As a result of an array of high-quality proteome-wide measurements under diverse growth conditions, we have generated a census that allows us to explore how the number of key molecular players change as a function of growth rate. Here, we have assembled a singular data set of protein copy numbers using measurements collected over the past decade via mass spectrometry (*Schmidt et al., 2016; Peebo et al., 2015; Valgepea et al., 2013*) or ribosomal profiling (*Li et al., 2014*) of the composition of the *E. coli* proteome across a gamut of growth rates. Due to notable changes in both cell size and cellular composition as a function of growth rate (*Bremer and Dennis, 2008; Taheri-Araghi et al., 2015*), as well as differences in normalization and standardization schemes used in each experimental work, substantial care was taken to ensure consistency on a per cellular basis [*Figure 1(A)*; see the Appendix for a detailed analysis and additional discussion]. To our knowledge, this compiled and curated dataset represents the most comprehensive view to date of the *E. coli* proteome, covering  $\approx 4000$  proteins and 36 unique growth rates, with the observed abundance of any given protein being directly comparable between data sets and across growth rates. This allows us to interrogate the *E. coli* specific physiology underlying the observed abundances while minimizing the effects of experimental noise as  $\approx 75\%$  of the proteins are observed in at least two separate datasets.

By compiling molecular turnover rate measurements for many of the fundamental processes associated with bacterial growth, we make quantitative order-of-magnitude estimates of key cellular processes in nutrient transport, cell envelope biogenesis, energy generation, and the central dogma [schematized in *Figure 1(B)*] to determine whether our current understanding of the kinetics of these processes are sufficient to explain the magnitude of the observed protein copy numbers across conditions. The census, combined with these estimates, provide a window into the question of whether the rates of central processes such as energy generation or DNA synthesis vary systematically as a function of cell growth rate by altering protein copy number, and in particular, whether any of these processes may be limiting growth.

For the majority of the processes considered, we find that the protein copy numbers are tuned for the task of cell doubling across a continuum of growth rates. This suggests that most processes must be operating near their maximal biosynthetic capacity, particularly under moderate to fast growth rates, with cellular protein abundances increasing at faster growth rates to support the more rapid cell doublings. This observation contrasts with the perspective that there exists any single process that is the arbiter of growth rate. Rather, it is through the coordinated increase in protein abundance across these disparate processes that *E. coli* is able to increase its growth rate as nutrient conditions improve. This hypothesis is bolstered by the observation that, when we consider the change in cell size and the diminishing surface area to volume ratio at faster growth rates, there still appears to be sufficient space in the cell membrane for the key proteins required for energy production and nutrient uptake.

Given the observed importance of parallelization, where the synthesis of additional proteins provides increased biosynthetic capacity of each cellular process, our analysis also provides insight into the well-characterized dependence of growth rate on ribosomal mass fraction. Here, a theoretical inability to parallelize ribosomal protein synthesis places a firm upper bound on the achievable growth rate that is observed at moderate to fast growth rates, where ribosomes appear maximally utilized. Importantly, the strict dependence between the maximal growth rate and ribosomal mass fraction coincides with the regime where the growth laws appear most valid (*Amir, 2017; Scott et al., 2010*). This enables us to suggest that the long-observed correlation between growth rate and cell size (*Schaechter et al., 1958; Si et al., 2017*) can be simply attributed to a required increase in absolute number of ribosomes per cell under conditions supporting extremely rapid growth. To better understand how the observed alterations in absolute protein abundances (and in particular, changes in ribosome copy number) influence growth rate across different nutrient conditions, we consider a minimal model of cellular growth. Our conclusions from these analyses provide important insight into how *E. coli* regulates growth across conditions of differing nutrient availability and identifies fundamental constraints in bacterial growth more broadly.



**Figure 1. Quantitative predictions of required protein abundances across key transport and synthesis processes necessary for cell division.** (A) The growth rate dependent changes in bacterial size and composition provided a basis to both predict the required protein abundances required to double a cell, as well as compile and compare proteomic measurements on a per cell basis across recent data sets from *Schmidt et al. (2016)*; *Li et al. (2014)*; *Peebo et al. (2015)*; *Valgepea et al. (2013)*. Predictions rely on the wealth of molecular turnover rate measurements and additional data tabulated on the BioNumbers Database [[biornumbers.hms.harvard.edu](http://biornumbers.hms.harvard.edu), *Milo et al. (2010)*]. (B) We consider an array of processes necessary for a cell to double its molecular components, broadly grouped into four classes. These categories are nutrient transport across the cell membrane, cell envelope biogenesis, energy production (namely, ATP synthesis), and processes associated with the central dogma. Numbers shown are the approximate number of complexes of each type observed at a growth rate of 0.5 hr $^{-1}$ , or a cell doubling time of  $\approx 5000$  s.

## Nutrient Transport

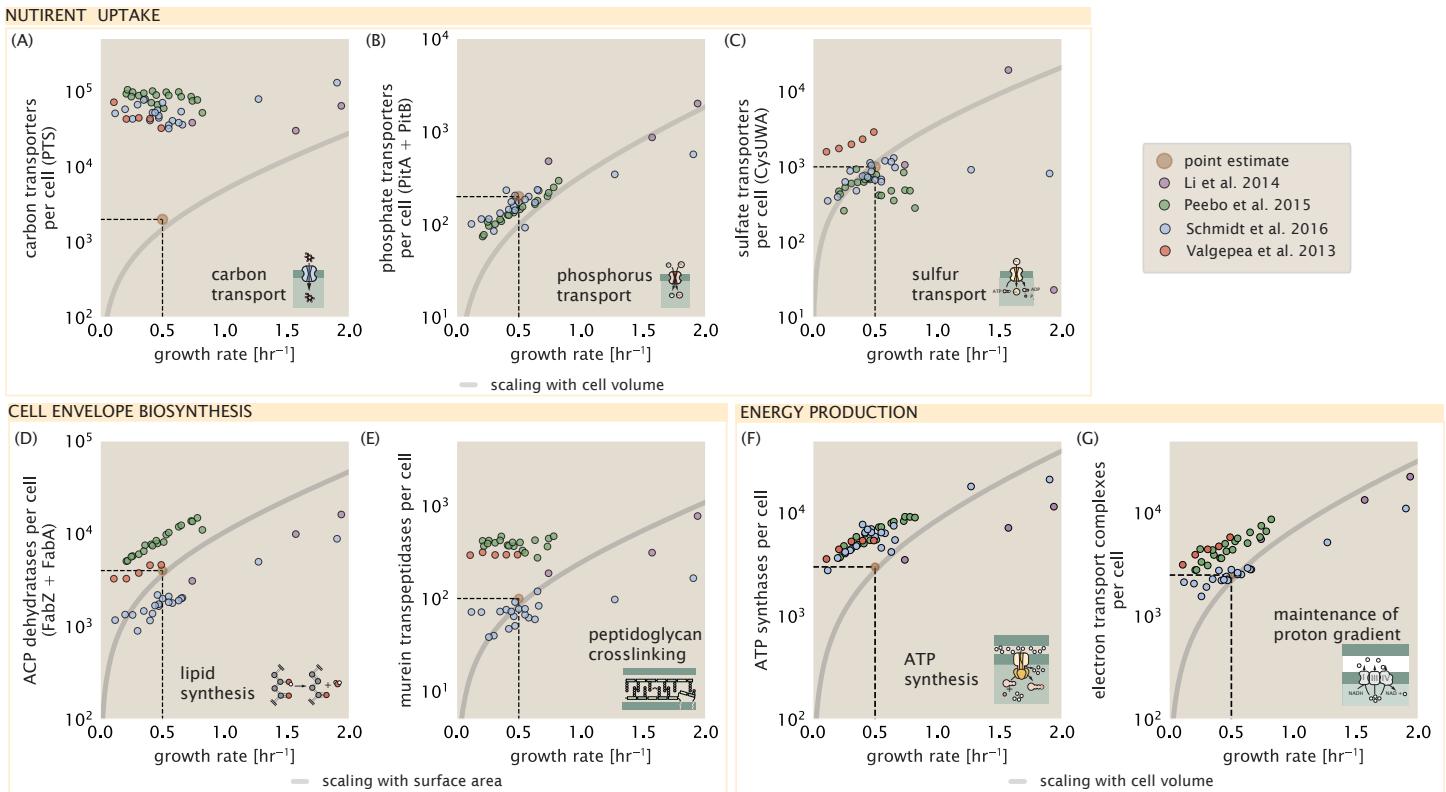
Throughout our estimates we will consider an archetypal growth rate of  $\approx 0.5 \text{ hr}^{-1}$  corresponding to a doubling time of  $\approx 5000$  seconds, as the data sets examined here heavily sample this growth regime. We will also consider how these values will vary at other growth rates due to changes in cell size, surface area, and chromosome copy number (*Taheri-Araghi et al., 2015; Harris and Theriot, 2018*) (described further in the Appendix Section "Extending Estimates to a Continuum of Growth Rates"). Here we begin by considering the critical transport processes diagrammed in *Figure 1(B)*.

In order to build new cellular mass, the molecular and elemental building blocks must be scavenged from the environment in different forms. Carbon, for example, is acquired via the transport of carbohydrates and sugar alcohols with some carbon sources receiving preferential treatment in their consumption (*Monod, 1947*). Phosphorus, sulfur, and nitrogen, on the other hand, are harvested primarily in the forms of inorganic salts, namely phosphate, sulfate, and ammonium/ammonia (*Jun et al., 2018; Assentoft et al., 2016; Stasi et al., 2019; Antonenko et al., 1997; Rosenberg et al., 1977; Willsky et al., 1973*). All of these compounds have different membrane permeabilities (*Phillips, 2018*) and most require some energetic investment either via ATP hydrolysis or through the proton electrochemical gradient to bring the material across the hydrophobic cell membrane.

The elemental composition of *E. coli* has received much quantitative attention over the past half century (*Neidhardt et al., 1991; Taymaz-Nikerei et al., 2010; Heldal et al., 1985; Bauer and Ziv, 1976*), providing us with a starting point for estimating how many atoms of each element must be scavenged from the environment:  $\approx 50\%$  carbon (BNID: 100649; see ?? for explanation of BNID references),  $\approx 15\%$  nitrogen (BNID: 106666),  $\approx 3\%$  phosphorus (BNID: 100653), and  $1\%$  sulfur (BNID: 100655) with the remainder being attributable to oxygen, hydrogen, and various transition metals. Here we estimate the abundance and growth rate dependence of a variety of transporters responsible for carbon uptake, and provide more extensive investigation of the other critical elements in the Appendix Section "Additional Estimates of Fundamental Biological Processes". Using  $\approx 0.3 \text{ pg}$  as the typical *E. coli* dry mass at a growth rate of  $\approx 0.5 \text{ hr}^{-1}$  (BNID: 103904), coupled with the approximation that  $\approx 50\%$  of this mass is carbon, we estimate that  $\approx 1 \times 10^{10}$  carbon atoms must be brought into the cell in order to double all of the carbon-containing molecules.

Typical laboratory growth conditions provide carbon as a single class of sugar (such as glucose, galactose, or xylose) often transported cross the cell membrane by a transporter complex specific to that particular sugar. One such mechanism of transport is via the PTS system, which is a highly modular system capable of transporting a diverse range of sugars with high specificity (*Escalante et al., 2012*). The glucose-specific component of this system transports  $\approx 200$  glucose molecules ( $\approx 1200$  carbon atoms) per second per transporter (BNID: 114686). Making the assumption that this is a typical sugar transport rate for the PTS system, coupled with the need to transport  $\approx 1 \times 10^{10}$  carbon atoms, we then expect on the order of  $\approx 2000$  transporters must be expressed per cell in order to bring in enough carbon atoms. We find, however, that the experimental measurements exceed this by several fold [*Figure 2(A)*], implying that the cell is capable of transporting more carbon atoms than strictly needed for biosynthesis. This constancy in the expression appears to be specific to glucose transporters, which are known to be the preferential carbon source (*Monod, 1947; Liu et al., 2005; Adelberg et al., 2014*), and stands in contrast with other species of transporters for glycerol, xylose, or fructose which we find match the required transporter abundances according to the achieved doubling time (adjusting for the specific carbon source in terms of number of carbon atoms per molecule and the rate of transport for the particular transporter species) [Supp2]. This also contrasts with our observations for uptake of phosphorus and sulfur, which turn out to align well with our expectations across different growth conditions [*Figure 2(B,C)*] and discussed further in the Appendix Section "Additional Estimates of Fundamental Biological Processes"].

If acquisition of nutrients was acting as a bottleneck on cellular growth, the growth rate could be theoretically increased simply by expressing more transporters, but is this feasible at a physiological level? A way to approach this question is to compute the amount of space in the bacterial membrane that could be occupied by nutrient transporters. Considering a rule-of-thumb for the surface area of *E. coli* of about  $5 \mu\text{m}^2$  (BNID: 101792), we expect an areal density for 2000 transporters to be approximately a few hundred transporters per  $\mu\text{m}^2$ . For a typical transporter occupying about  $50 \text{ nm}^2$ , this amounts to about only  $\approx 1\%$  of the total inner membrane surface area (*Szenk et al., 2017*). In contrast, bacterial cell membranes typically have densities of  $\approx 1 \times 10^5$  proteins/ $\mu\text{m}^2$  (*Phillips, 2018*), with roughly 60 % of the surface area occupied by protein (BNID: 100078), implying that the cell could easily accommodate more transporters. There are, however, additional constraints on the space that can be devoted to nutrient uptake due to occupancy by proteins involved in processes like cell wall synthesis and energy production, and we will consider this further in the coming sections.



**Figure 2. Key processes required for nutrient uptake, cell wall biogenesis and energy synthesis during growth.** (A) Estimate for the minimum number of generic carbohydrate transport systems. Colored points correspond to the mean number of complexes involved in carbohydrate import (complexes annotated with the Gene Ontology terms GO:0009401 and GO:0098704) for different growth conditions across different published datasets. (B) Number of PitA phosphate transport systems needed to maintain a 3% phosphorus dry mass. (C) Number of CysUWA complexes necessary to maintain a 1% sulfur *E. coli* dry mass and the experimentally observed complex copy numbers using the transporter stoichiometry  $[CysA]_2[CysU][CysW][Sbp/CysP]$ . (D) Number of ACP dehydratases necessary to form functional phospholipids, which is assumed to be a rate-limiting step on lipid synthesis, and the experimentally observed complex copy numbers using the stoichiometries  $[FabA]_2$  and  $[FabZ]_2$ . (E) Number of peptidoglycan transpeptidases needed to complete maturation of the peptidoglycan and experimental measurements of the transpeptidase complexes, following the stoichiometries  $[MrcA]_2$ ,  $[MrcB]_2$ ,  $[MrdA]_1$ , and  $[MrdB]_1$ . (F) Number of  $F_1-F_0$  ATP synthase complexes needed to accommodate peptide bond formation and other NTP dependent processes and experimental measurements following the stoichiometry  $[AtpE]_{10}[AtpF]_2[AtpB][AtpC][AtpH][AtpA]_3[AtpG][AtpD]_3$ . (G) Number of electron transport chain complexes needed to maintain a membrane potential of  $-200$  mV. Points in plot correspond to the average number of complexes identified as being involved in aerobic respiration by the Gene Ontology identifier GO:0019646. These complexes include cytochromes  $bd1$  ( $[CydA][CydB][CydX][CydH]$ ),  $bdII$  ( $[AppC][AppB]$ ),  $bo_3$  ( $[CyoD][CyoA][CyoB][CyoC]$ ) and NADH:quinone oxidoreductase I ( $[NuoA][NuoH][NuoJ][NuoK][NuoL][NuoM][NuoN][NuoB][NuoC][NuoE][NuoF][NuoG][NuoI]$ ) and II ( $[Ndh]$ ). Dashed lines indicate order of magnitude estimate needed at a growth rate of  $\approx 0.5$  per hr (light-brown point), while the gray line accounts for the growth rate dependence changes in cell size and doubling time.

## Cell Envelope Biogenesis

In contrast to nutrient transporters, which support the synthesis of biomolecules throughout the cell and therefore need to scale with the cell size, here we must consider the synthesis of components that will need to scale with the surface area of the cell. *E. coli* is a rod-shaped bacterium with a remarkably robust length-to-width aspect ratio of  $\approx 4:1$  (Harris and Theriot, 2018; Ojkic et al., 2019). In this section, we consider the synthesis of lipids as well as the complexes involved in assembling the peptidoglycan scaffold that makes up the cell envelope.

The membranes of *E. coli* are composed of a variety of different lipids, each of which are unique in their structures and biosynthetic pathways (Sohlenkamp and Geiger, 2016). Recently, a combination of stochastic kinetic modeling (Ruppe and Fox, 2018) and *in vitro* kinetic measurements (Ranganathan et al., 2012; Yu et al., 2011) has revealed remarkably slow steps in the fatty acid synthesis pathways which may serve as the rate limiting reactions for making new membrane fatty acids (that become components of a variety of membrane lipids) in *E. coli*. One such step is the removal of hydroxyl groups from the fatty-acid chain by ACP dehydratase that leads to the formation of carbon-carbon double bonds. This reaction, catalyzed by proteins FabZ and FabA (Yu et al., 2011), has been estimated to have kinetic turnover rates of  $\approx 1$  dehydration per second per enzyme (Ruppe and Fox, 2018). Thus, given this rate and the need to synthesize  $\approx 2 \times 10^7$  lipids over 5000 seconds, one can estimate that a typical cell requires  $\approx 4000$  ACP dehydratases. This is in reasonable agreement with the experimentally observed copy numbers of FabZ and FabA (Figure 2(D)).

The exquisite control of bacteria over their cell shape is due primarily to a stiff, several nanometer thick meshwork of polymerized disaccharides that makes up the cell wall termed the peptidoglycan. The formation of the peptidoglycan is an intricate process involving many macromolecular players (Shi et al., 2018; Morgenstein et al., 2015), whose coordinated action synthesizes the individual subunits and integrates them into the peptidoglycan network that maintains cell shape and integrity even in the face of large-scale chemical and osmotic perturbations (Harris and Theriot, 2018; Shi et al., 2018). Due to the extensive degree of chemical crosslinks between glycan strands, the entire peptidoglycan is a single molecule comprising  $\approx 3\%$  of the cellular dry mass (BNID: 1019360), making it the most massive molecule in *E. coli*. The polymerized unit of the peptidoglycan is a N-acetylglucosamine and N-acetylmuramic acid disaccharide, of which the former is functionalized with a short pentapeptide. With a mass of  $\approx 1000$  Da, this unit, which we refer to as a murein subunit, is polymerized to form long strands in the periplasm which are then attached to each other via their peptide linkers. Together, these quantities provide an estimate of  $\approx 5 \times 10^6$  murein subunits per cell.

There are various steps which one could consider *a priori* to be a limiting process in the synthesis of peptidoglycan, including the biosynthesis steps that occur in the cytoplasm, the transglycosylation reaction which adds new subunits to the glycan strands, and the formation of the peptide crosslinks between strands (Shi et al., 2018; Morgenstein et al., 2015; Lovering et al., 2012; Barreteau et al., 2008). Despite the extensive mechanistic characterization of these components, quantitative characterization of the individual reaction rates along the entire kinetic pathway remain scarce and make identification of any particularly slow steps difficult. However, recent measurements for the crosslinking machinery [transpeptidases, Catherwood et al. (2020)] of the peptidoglycan, which provides lateral structural integrity to the peptidoglycan shell, have found the turnover of transpeptidases to be rather slow ( $\approx 2$  crosslinking reactions per second). As the primary mechanism of subunit integration occurs by a complex with both transglycosylation and transpeptidation activities (Shi et al., 2018) we therefore consider the transpeptidation reaction as a reasonable candidate for a rate-limiting step in growth as it is vital for cell size and shape homeostasis. We estimate that on the order of  $\approx 100$  transpeptidases per cell are needed for complete maturation of the peptidoglycan, given a division time of  $\approx 5000$  seconds; a value that is comparable to experimental observations [Figure 2(E)]. Expanding this estimate to account for the changing mass of the peptidoglycan as a function of growth rate [grey line in Figure 2(E)] predicts an order-of-magnitude increase in the abundance of the transpeptidases when the grow rate is increased by a factor of four. Here, however, the measured complex abundances across the different proteomic data sets show systematic disagreements and obfuscates any significant dependence on growth rate.

While these processes represent only a small portion of proteins devoted to cell envelope biogenesis, we find it unlikely that they limit cellular growth in general. The relative amount of mass required for lipid and peptidoglycan components will decrease at faster growth rates due to a decrease in the cell's surface area to volume ratio (Ojkic et al., 2019). Furthermore, despite the slow catalytic rate of fatty acid synthesis and transpeptidation, there appear to be sufficient protein abundance to support growth. For FabZ and FabA in lipid synthesis, experimental data and recent computational modeling has shown that the rate of fatty-acid synthesis can be drastically increased by increasing the concentration of FabZ (Yu et al., 2011; Ruppe and Fox, 2018). With a proteome size of  $\approx 3 \times 10^6$  proteins, a hypothetical 10-fold increase in expression from 4000 to 40,000 ACP dehydratases would result in a paltry  $\approx 1\%$  increase in the size of the proteome.

## Energy Production

Cells consume and generate energy predominantly in the form of nucleoside triphosphates (NTPs). The high-energy phosphodiester bonds of (primarily) ATP power a variety of cellular processes that drive biological systems away from thermodynamic equilibrium. We therefore turn to the synthesis of ATP as another process that may limit growth, which will also require us to consider the maintenance of the electrochemical proton gradient that powers it.

Hydrolysis of the terminal phosphodiester bond of ATP into ADP (or alternatively GTP and GDP) and an inorganic phosphate provides the thermodynamic driving force in a wide array of biochemical reactions. One such reaction is the formation of peptide bonds during translation, which requires  $\approx 2$  ATPs for the charging of an amino acid to the tRNA and  $\approx 2$  GTPs for the formation of each peptide bond. Assuming the ATP costs associated with error correction and post-translational modifications of proteins are negligible, we can make the approximation that each peptide bond has a net cost of  $\approx 4$  ATP (BNID: 101442). Formation of GTP from ATP is achieved via the action of nucleoside diphosphate kinase, which catalyzes this reaction without an energy investment (*Lascu and Gonin, 2000*). We therefore consider all NTP requirements of the cell to be functionally equivalent to being exclusively ATP. In total, the energetic costs of peptide bond formation consumes  $\approx 80\%$  of the cells ATP budget [BNID: 107782; 106158; 101637; 111918, *Lynch and Marinov (2015); Stouthamer (1973)*] and is primarily produced by the F<sub>1</sub>-F<sub>0</sub> ATP synthase – a membrane-bound rotary motor which under ideal conditions can yield  $\approx 300$  ATP per second [BNID: 114701; *Weber and Senior (2003)*].

To estimate the total number of ATP equivalents consumed during a cell cycle, we make the approximation that there are  $\approx 3 \times 10^6$  proteins per cell with an average protein length of  $\approx 300$  peptide bonds (BNID: 115702; 108986; 104877). With  $\approx 4$  ATP equivalents per peptide bond, we find that the typical *E. coli* cell consumes  $\approx 5 \times 10^9$  ATP per cell cycle on protein synthesis alone. Assuming that each ATP synthases operates at its maximal speed,  $\approx 3000$  ATP synthases are needed to keep up with the energy demands of the cell. This estimate is comparable with the experimental observations, shown in *Figure 2(F)*. Since this estimate assumes all ATP is synthesized via ATP synthase and neglects synthesis via fermentative metabolism, this may explain why at the fastest growth rates ( $\approx 2$  hr<sup>-1</sup>), our continuum estimate predicts more synthase than is experimentally observed (data points below the gray line in *Figure 2(F)* at fast growth rates). Here, *E. coli* enters a type of overflow metabolism where non-respiratory routes for ATP synthesis become more pronounced and provide the remaining ATP demand (*Molenaar et al., 2009; Zhuang et al., 2011; Szenk et al., 2017*).

In order to produce ATP, the F<sub>1</sub>-F<sub>0</sub> ATP synthase itself must consume energy. Rather than burning through its own product (and violating thermodynamics), this intricate macromolecular machine has evolved to exploit the electrochemical potential established across the inner membrane through cellular respiration. This electrochemical gradient is manifest by the pumping of protons into the intermembrane space via the electron transport chains as they reduce NADH. In *E. coli*, this potential difference is  $\approx -200$  mV (BNID: 102120). However, each rotation of an ATP synthase shuttles  $\approx 4$  protons into the cytosol (BNID: 103390). With a few thousand ATP synthases producing ATP at their maximal rate, the potential difference would be rapidly abolished in a few milliseconds if it were not being actively maintained.

A recent work (*Szenk et al., 2017*) examined the respiratory capacity of the *E. coli* electron transport complexes using structural and biochemical data, revealing that each electron transport chain rapidly pumps protons into the intermembrane space at a rate of  $\approx 1500$  protons per second (BNID: 114704; 114687). Using our estimate of the number of ATP synthases required per cell, coupled with these recent measurements, we estimate that  $\approx 3000$  electron transport complexes would be necessary to facilitate the  $\approx 5 \times 10^6$  protons per second diet of the cellular ATP synthases. This estimate is in agreement with the number of complexes identified in the proteomic datasets [plot in *Figure 2(G)*]. This suggests that every ATP synthase must be accompanied by  $\approx 1$  functional electron transport chain.

## Limits on Biosynthesis in a Crowded Cell

Our estimates thus far have focused on the biochemistry at the periphery of the cell, with the processes of nutrient transport, cell envelope biogenesis, and energy generation all requiring space to perform their biological functions. The cell's surface area, however, does not scale as rapidly as cell size (*Harris and Theriot, 2018*) and there will be diminishing space available at the periphery to support the proteomic requirements at faster growth rates. It is therefore necessary to consider the consequences of a changing surface area to volume ratio in our effort to identify limitations on growth. Here we use our analysis of ATP production to better understand this constraint.

In our estimate of ATP production above we found that a cell demands about  $5 \times 10^9$  ATP per cell cycle or  $\approx 1 \times 10^6$  ATP/s. With a cell volume of roughly 1 fL (BNID: 100004), this corresponds to about  $2 \times 10^{10}$  ATP per fL of cell volume, in line with previous estimates (*Stouthamer and Bettenhausen, 1977; Szenk et al., 2017*). In *Figure 3(A)* we plot this ATP demand as a function of the surface area to volume ratio in green, where we have considered a range of cell shapes from spherical to rod-shaped with an aspect

ratio (length/width) equal to 4. In order to consider the maximum ATP that could be produced, we consider the amount of ATP that can be generated by a membrane filled with ATP synthase and electron transport complexes and a maximal production rate of about 3 ATP / ( $\text{nm}^2 \cdot \text{s}$ ) (Szenk *et al.*, 2017). This is shown in blue in *Figure 3(A)*, which shows that at least for the growth rates observed (right column in plot), the energy demand is roughly an order of magnitude less. Interestingly, Szenk *et al.* (2017) found that ATP production by respiration is less efficient than by fermentation on a per membrane area basis, due to the additional proteins of the electron transport chain. This suggests that, even under anaerobic growth, cells will have sufficient membrane space for ATP production.

Importantly, this analysis highlights that there will indeed be a maximum attainable cell size due to the limited capacity to provide resources as the cell increases in size. The maximum energy production in *Figure 3(A)*, however, does represent a somewhat unachievable limit since the inner membrane also includes other proteins like those we've considered for nutrient transport and cell wall biogenesis. To better understand the overall proteomic makeup of the inner membrane, we therefore used Gene Ontology (GO) annotations (Ashburner *et al.*, 2000; The Gene Ontology Consortium, 2018) to identify all proteins embedded or peripheral to the inner membrane (GO term: 0005886). Those associated but not membrane-bound include proteins like MreB and FtsZ that must nonetheless be considered as a vital component occupying space on the membrane. In *Figure 3(B)*, we find that the total protein mass per  $\mu\text{m}^2$  is nearly constant across growth rates. Interestingly, when we consider the distribution of proteins grouped by their Clusters of Orthologous Groups (COG) (Tatusov *et al.*, 2000), the relative abundance of each category is nearly constant across growth rates. This suggests that no one process (energy production, nutrient uptake, etc.) is dominating even at fast growth rates [*Figure 3(C)*] and in line with our supposition that each of the processes we've considered so far are not fundamentally limiting the maximum growth rate. In contrast, when we apply such an analysis to cytosolic proteins (GO term: 0005829), we observe a clear change in the proteomic composition [??(A, B)]. In particular, with increasing growth rates there is a substantial increase in the relative protein mass associated with 'information storage and processing'. This category includes proteins such as DNA polymerase, RNA polymerase, and ribosomes that are associated with the processes of the central dogma, whose increase is predominantly at the expense of 'metabolic' proteins as shown in ??(C). This notable anticorrelation shown in ??(D) is consistent with previous reports (Scott *et al.*, 2010; Hui *et al.*, 2015; Zhu and Dai, 2019). In the next section we therefore turn our attention to the processes of the central dogma.

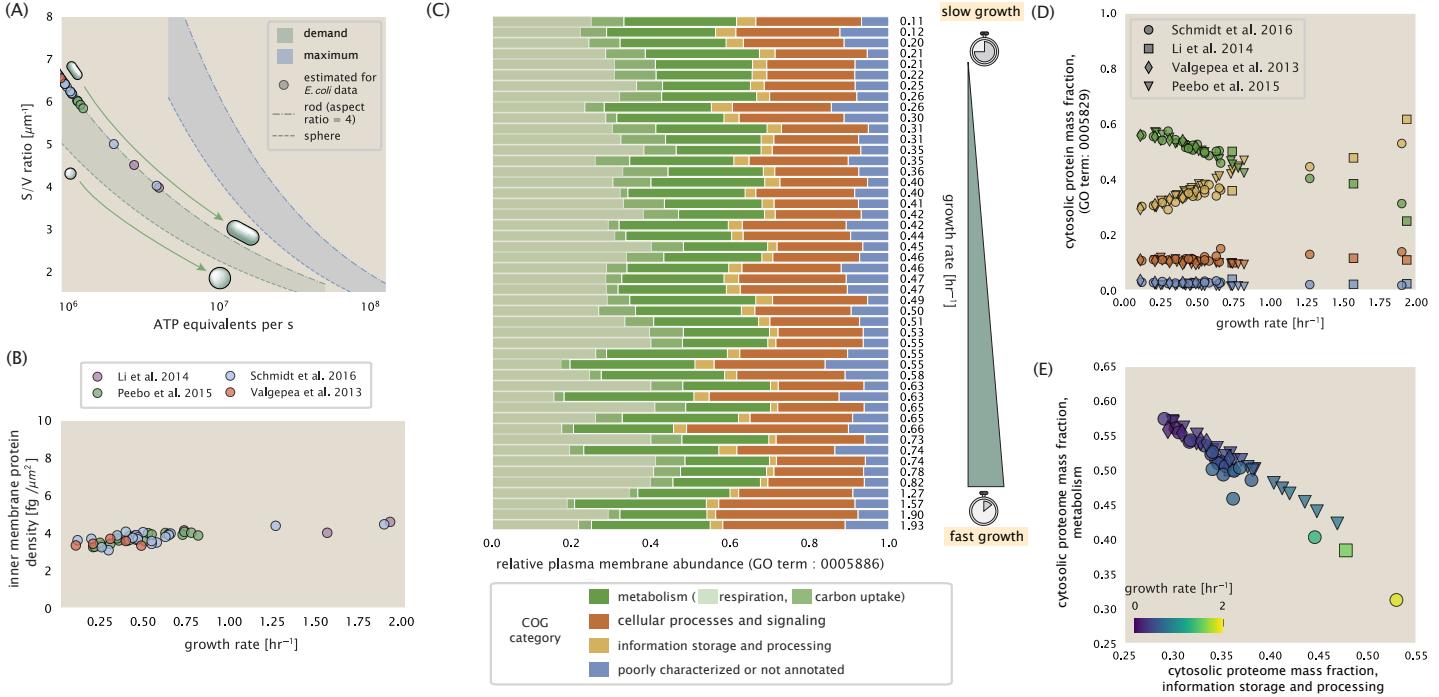
## Processes of the Central Dogma

Up to this point, we have considered a variety of transport and biosynthetic processes that are critical to acquiring and generating new cell mass. We now turn our focus to some of the most important processes which *must* be undertaken irrespective of the growth conditions – those of the central dogma.

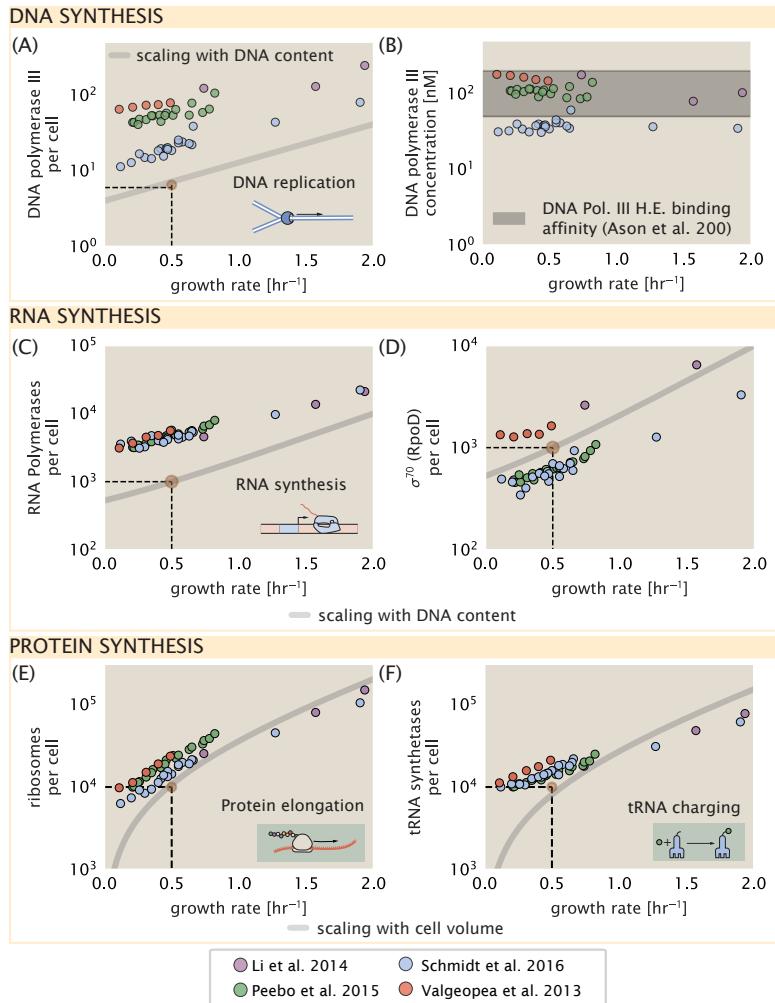
To successfully divide and produce viable progeny, the DNA must be faithfully replicated and segregated into each nascent cell. In rapidly growing cultures, bacteria like *E. coli* can initiate as many as 10 - 12 replication forks at a given time (Bremer and Dennis, 2008; Si *et al.*, 2017), suggesting only  $\approx 10$  are needed. However, as shown in *Figure 4(A)*, DNA polymerase III is nearly an order of magnitude more abundant while still maintaining a predicted growth rate dependence. This discrepancy can be understood by considering its binding constant to DNA. *In vitro* 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 (Ason *et al.*, 2000) *Figure 4(B)* (discussed further, along with the synthesis of dNTP building blocks in Appendix Section "Additional Process of the Central Dogma").

We now turn our attention to the transcription of DNA to form RNA. Here we focus on the synthesis of rRNA, which make up the majority of RNA in the cell, and discuss the synthesis of mRNA and tRNA further in the Appendix Section "Additional Process of the Central Dogma". rRNA serves as the catalytic and structural component of the ribosome, comprising approximately 2/3 of the total ribosomal mass, and is decorated with  $\approx 50$  ribosomal proteins. Each ribosome contains three rRNA molecules of lengths 120, 1542, and 2904 nucleotides (BNID: 108093), meaning each ribosome contains  $\approx 4500$  nucleotides overall. *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 constantly loaded at this rate, then we can assume that  $\approx 1$  functional rRNA unit is synthesized per second per rRNA operon. At a growth rate of  $\approx 0.5 \text{ hr}^{-1}$ , the average cell has  $\approx 1$  copy of its chromosome and therefore approximately  $\approx 7$  copies of the rRNA operons, producing  $\approx 7$  rRNA units per second. With a 5000 second division time, this means the cell is able to generate around  $3 \times 10^4$  functional rRNA units, comparable within an order of magnitude to the number of ribosomes per cell.

How many RNA polymerases are then needed to constantly transcribe the required rRNA? If one polymerase is loaded once every second on average (BNID: 111997), and the transcription rate is  $\approx 40$  nucleotides per second (BNID: 101094), then the typical spacing between polymerases will be  $\approx 40$  nucleotides. With a total length of  $\approx 4500$  nucleotides per operon and 7 operons per cell,



**Figure 3. Influence of cell size and surface area to volume ratio on ATP production and inner membrane composition.** (A) Scaling of ATP demand and maximum ATP production through respiration as a function of surface area to volume ratio. Cell volumes of 0.5 fL to 50 fL were considered, with the dashed (—) line corresponding to a sphere and the dash-dot line (---) reflecting a rod-shaped bacterium like *E. coli* with a typical aspect ratio (length / width) of 4 (Shi et al., 2018). The ATP demand is calculated as  $10^6 \text{ ATP}/(\mu\text{m}^3 \text{ s})$ , while the maximum ATP production rate is taken to be 3 ATP / ( $\text{nm}^2 \cdot \text{s}$ ) (Szenk et al., 2017), with calculations of *E. coli* volume and surface area detailed in Appendix Section "Estimation of Cell Size and Surface Area". In this calculation, 50% of the bacterial inner membrane is assumed to be protein, with the remainder lipid. (B) Total protein mass per  $\mu\text{m}^2$  calculated for proteins with inner membrane annotation (GO term: 0005886). (C) Relative protein abundances are grouped by their COG annotations ('metabolic', 'cellular processes and signaling', 'information storage and processing', and 'poorly characterized or not annotated') for the data from Schmidt et al. (2016). Metabolic proteins are further separated into respiration ( $F_1$ - $F_0$  ATP synthase, NADH dehydrogenase I, succinate:quinone oxidoreductase, cytochrome  $b_0$  $_3$  ubiquinol oxidase, cytochrome bd-I ubiquinol oxidase) and carbohydrate transport (GO term: GO:0008643). Note that the elongation factor EF-Tu can also associate with the inner membrane, but was excluded in this analysis due to its high relative abundance (roughly identical to the summed protein shown in part (B)). (D) Relative cytosolic protein abundances (GO term: 0005886), grouped by their COG annotations, are plotted as a function of growth rate. (E) The relative cytosolic protein abundances (GO term: 0005886) associated with the 'information storage and processing' and 'metabolic' COG categories are plotted against each other and highlight the larger mass fraction devoted to 'information storage and processing' at faster growth rates.



**Figure 4. Processes of the central dogma.** (A) The minimum number of DNA polymerase holoenzyme complexes needed to facilitate replication of the genome. Points 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. (B) The effective concentration of DNA polymerase III holoenzyme (See Appendix Section "Estimation of Cell Size and Surface Area" for calculation of cell size). Shaded region corresponds to the range of  $K_D$  values measured by [Ason et al. \(2000\)](#), from 50 and 200 nM. (C) The number of RNA polymerase core enzymes, with measurements corresponding to the average number given a subunit stoichiometry of  $[RpoA]_2[RpoC][RpoB]$ . (D) The abundance of  $\sigma^{70}$  as a function of growth rate along with the same prediction from (C). (E) Number of ribosomes required to synthesize  $10^9$  peptide bonds with an elongation rate of 15 peptide bonds per second. (F) Number of tRNA synthetases that will supply the required amino acid demand. The sum of all tRNA synthetases copy numbers are plotted ( $[ArgS], [CysS], [GlnS], [Glx], [IleS], [LeuS], [ValS], [AlaS]_2, [AsnS]_2, [AspS]_2, [TyrS]_2, [TrpS]_2, [ThrS]_2, [SerS]_2, [ProS]_2, [PheS]_2[PheT]_2, [MetG]_2, [LysS]_2, [HisS]_2, [GlyS]_2[GlyQ]_2$ ). Dashed lines indicate order of magnitude estimate needed at a growth rate of  $\approx 0.5$  per hr (light-brown point), while the gray line accounts for the growth rate dependence changes in cell size and doubling time.

the number of RNA polymerases transcribing rRNA at any given time is then  $\approx 1000$  per cell. We also find that cells require on the order of another  $\approx 400$  RNAP for the synthesis of mRNA and tRNA, predicting a total of  $\approx 1500$  RNAP to satisfy its transcriptional demands. As is revealed in ??(B), this estimate is about an order of magnitude below the observed number of RNA polymerase complexes per cell ( $\approx 5000 - 7000$ ). Consistent with this, a majority of RNAP is known to be non-specifically bound to DNA during its search for promoters from which to begin transcription (*Klumpp and Hwa, 2008; Patrick et al., 2015*). In *Figure 4(D)*, we find that the predicted RNA polymerase copy number indeed is more comparable with the abundance of  $\sigma$ -70 (RpoD), the primary workhorse sigma factor for transcription in *E. coli*. We can conclude that the observed RNA polymerase abundances are generally sufficient for what appears needed for growth.

We conclude our dialogue between back-of-the-envelope estimates and comparison with the proteomic data by examining the final process in the central dogma – translation. In doing so, we will begin with an estimate of the number of ribosomes needed to replicate the cellular proteome. While the rate at which ribosomes translate is well known to depend on the growth rate [*Dai et al. (2018)*, a phenomenon we consider later in this work] we begin by making the approximation that translation occurs at a modest rate of  $\approx 15$  amino acids per second per ribosome (BNID: 100233). Under this approximation and our previous estimate of  $10^9$  peptide bonds per cell at a growth rate of  $0.5 \text{ hr}^{-1}$ , we can easily arrive at an estimate of  $\approx 10^4$  ribosomes needed per cell to replicate the entire protein mass *Figure 4(E)* and proves comparable to the experimental observations. While the ribosome is responsible for the formation of peptide bonds, we do not diminish the importance of charging tRNAs with their appropriate amino acid, a process with occurs with remarkable fidelity. In *Figure 4(F)* we show our estimate for the required number of tRNA synthetases, which show similar accord with the experimental data and is discussed further in the Appendix Section "Additional Estimates of Fundamental Biological Processes."

Having completed our circuit through key processes of cellular growth outlined in *Figure 1*, we can now take stock of our understanding of the observed growth rate dependence and abundances of various protein complexes. We note that, broadly speaking, these simple estimates have been reasonably successful in quantitatively describing the observations in the proteomic data, and importantly suggests that the proteome is tuned in composition and absolute abundance to match the growth rate requirements without any one process representing a singular bottleneck or rate limiting step in division. However, in our effort to identify key limitations on growth, there are two notable observations worthy of additional emphasize.

The first is a recurring theme throughout the estimates investigated here, which is that any inherent biochemical rate limitation can be overcome by expressing more proteins. We can view this as a parallelization of each biosynthesis task, which helps explain why bacteria tend to increase their protein content and cell size as growth rate increases (*Ojikic et al., 2019*). The second, and ultimately the most significant in defining the cellular growth rate, is that the synthesis of ribosomal proteins presents a special case where parallelization is *not* possible and thereby imposes a limit on the fastest possible growth rate. Each ribosome has  $\approx 7500$  amino acids across all of its protein components which must be strung together as peptide bonds through the action of another ribosome. Once again using a modest elongation rate of  $\approx 15$  amino acids per second, we arrive at an estimate of  $\approx 500$  seconds or  $\approx 7$  minutes to replicate a single ribosome. This limit, as remarked upon by others (*Dill et al., 2011*), serves as a hard theoretical boundary for how quickly a bacterium like *E. coli* can replicate.

## Maximum Growth Rate is Determined by the Rate of Ribosomal Protein and rRNA Synthesis

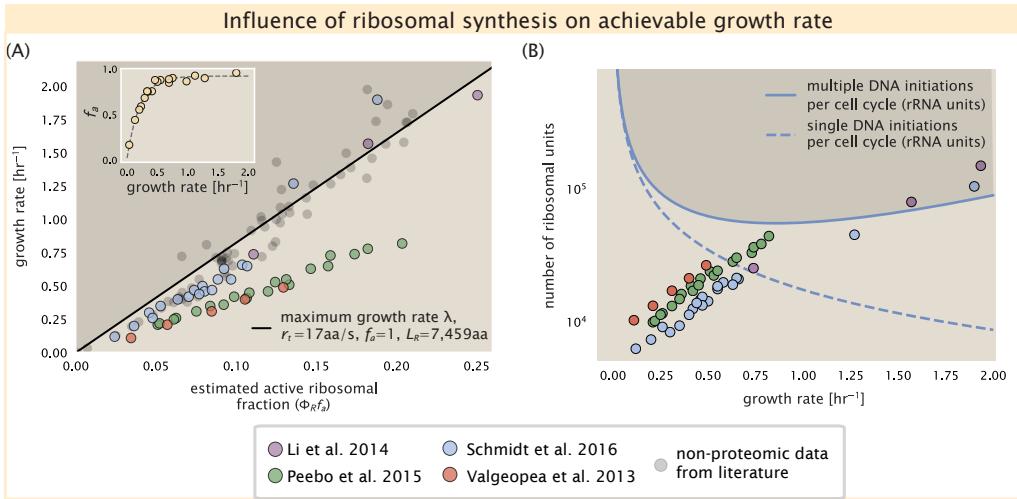
The 7 minute speed limit assumes all proteins in the cell are ribosomes. In order to connect this to the experimental data (and physiological reality more broadly), we first need to relax this assumption and determine a translation-limited growth rate. Here, we will assume that the cell is composed of  $N_{\text{pep}}$  peptide bonds and  $R$  ribosomes, whose precise values will depend on the growth rate  $\lambda$ . The protein subunits of each ribosomal protein sum to a total of  $\approx 7500$  amino acids as noted earlier, which we denote by  $L_R$ . With an average mass of an amino acid of  $m_{\text{AA}} \approx 110 \text{ Da}$  (BNID: 104877), the total ribosomal mass fraction  $\Phi_R$  is given by

$$\Phi_R = \frac{m_{\text{ribosomes}}}{m_{\text{proteome}}} \approx \frac{m_{\text{AA}} \times R \times L_R}{m_{\text{AA}} \times N_{\text{pep}}} = \frac{R \times L_R}{N_{\text{pep}}}. \quad (1)$$

For exponentially growing cells (*Godin et al., 2010*), the rate of cellular growth will be related to the rate of protein synthesis via

$$\lambda N_{\text{pep}} = r_t \times R \times f_a, \quad (2)$$

where  $r_t$  is the translation rate. Here, we've introduced a multiplicative factor  $f_a$  which represents the fraction of the ribosomes that are actively translating. This term allows us to account for immature or non-functional ribosomes or active sequestration of ribosomes through the action of the secondary messenger alarmone (p)ppGpp in poorer nutrient conditions (*Hauryliuk et al., 2015*).



**Figure 5. Limitations on ribosomal protein synthesis and growth rate.** (A) Translation-limited growth as a function of the ribosomal fraction. The dashed line corresponds to the maximum rate of ribosomal protein synthesis ( $\approx 7$  min). Translation-limited growth rate as a function of the actively translating ribosomal fraction. The actively translating ribosomal fraction is calculated using the estimated values of  $f_a$  from [Dai et al. \(2016\)](#) (shown in inset; see Section "Calculation of active ribosomal fraction" for additional detail). Shaded region defines boundary due to constraint set on growth rate by [Equation 3](#). The solid line is calculated for an elongation rate of 17 aa per second. Gray data points show additional measurements based on measurements of cellular RNA to protein ratio, with  $\Phi_R \approx$  the cellular RNA to protein ratio divided by 2.1 ([Dai et al., 2016](#)) and come from [Forchhammer and Lindahl \(1971\)](#); [Bremer and Dennis \(2008\)](#); [Scott et al. \(2010\)](#); [Dai et al. \(2016\)](#); [Si et al. \(2017\)](#). (B) Maximum number of rRNA units that can be synthesized as a function of growth rate. Solid curve corresponding to the rRNA copy number is calculated by multiplying the number of rRNA operons by the estimated number of (# ori) at each growth rate. The quantity (# ori) was calculated using [Equation 4](#) and the measurements from [Si et al. \(2017\)](#). The dashed line shows the maximal number of functional rRNA units produced from a single chromosomal initiation per cell cycle. Shaded region defines boundary due to maximal rRNA synthesis.

Combining [Equation 1](#) and [Equation 2](#) results in an expression for a translation-limited growth rate, which is given by

$$\lambda_{\text{translation-limited}} = \frac{r_i \times \Phi_R \times f_a}{L_R}. \quad (3)$$

This result, derived in a similar manner by others ([Dennis et al., 2004](#); [Klumpp et al., 2013](#)), reflects mass-balance under steady state growth and has long provided a rationalization of the apparent linear increase in *E. coli*'s ribosomal content as a function of growth rate ([Maaløe, 1979](#); [Dennis et al., 2004](#); [Scott et al., 2010](#); [Dai et al., 2016](#)). The left-hand panel of [Figure 5\(A\)](#) shows this growth rate plotted as a function of the ribosomal mass fraction. In the regime where all ribosomes are active ( $f_a = 1$ ) and the entire proteome is composed of ribosomal proteins ( $\Phi_R = 1$ ), indeed, we arrive at the maximum theoretical growth rate of  $r_i/L_R$ , and  $\approx 7$  min for *E. coli*.

Connecting [Equation 3](#) to the proteomic data, however, requires knowledge of  $f_a$  at each growth rate as proteomic measurements only provide a measure of  $\Phi_R$ . While commonly considered constant with growth rate ([Klumpp et al., 2013](#); [Bosdriesz et al., 2015](#); [Kostinski and Reuveni, 2020](#)), [Dai et al. \(2016\)](#) recently inferred  $f_a$  as a function of the growth rate ([Figure 5\(A\)](#), right-hand panel, inset), revealing that  $f_a \approx 1$  at growth rates above 0.75 hr<sup>-1</sup> but drops dramatically at slower growth rates. Using these data, we inferred the approximate active fraction (see the Appendix Section "Calculation of active ribosomal fraction") at each growth rate and used this to compute  $\Phi_R \times f_a$  ([Figure 5\(A\)](#), colored points in right-hand panel). Importantly, these data skirt the translation-limited growth rate determined using [Equation 3](#), where we have taken  $r_i$  to be the maximal elongation rate of 17 amino acids per second measured by [Dai et al. \(2016\)](#). There is a notable discrepancy between the data collected in [Schmidt et al. \(2016\)](#); [Li et al. \(2014\)](#) and that collected from [Valgepea et al. \(2013\)](#); [Peebo et al. \(2015\)](#). When compared to other measurements (non-proteomic based) of the active ribosome mass fraction based on measurements of total RNA to total protein mass ratios ([Figure 5\(B\)](#)), grey points in right-hand panel and further detailed in [Figure 5-???](#), the data from [Valgepea et al. \(2013\)](#) and [Peebo et al. \(2015\)](#) are notably different, suggesting there may be a systematic bias in these two sets of measurements.

Even under idealized experimental conditions, however, *E. coli* rarely exhibits growth rates above 2 hr<sup>-1</sup> ([Bremer and Dennis, 2008](#)), which is still well-below the synthesis rate of a single ribosome, and below the maximum growth rates reported for several other bacteria ([Roller et al., 2016](#)). While we have considered potential limits imposed by translation of ribosomal proteins, we must also consider potential limiting regimes specific to the synthesis of rRNA. Due to multiple initiations of chromosomal replication per cell doubling, the effective number of rRNA operons increases with growth rate and will do so in proportion to the average number

of chromosomal origins per cell,  $\langle \# \text{ ori} \rangle$ . This latter parameter is set by how often replication must be initiated in order to keep up with the cell doubling time  $\tau$ , whose time may be shorter than the cell cycle time  $\tau_{\text{cyc}}$  (referring to the time from replication initiation to cell division) (Dennis et al., 2004; Ho and Amir, 2015). This is quantified by

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

where the doubling time  $\tau$  is related to the growth rate by  $\tau = \log(2)/\lambda$ . As the rRNA operons are predominantly located close to the origin of replication (BNID: 100352), we make the simplifying assumption that the number of rRNA operons will be directly proportional to  $\langle \# \text{ ori} \rangle$ . We used the experimental measurements of  $\tau_{\text{cyc}}$  and  $\tau$  [Figure 5–??(B)] to calculate  $\langle \# \text{ ori} \rangle$  with Equation 4 as a function of growth rate. For growth rates above about  $0.5 \text{ hr}^{-1}$ ,  $t_{\text{cyc}}$  is approximately constant at about 70 minutes, implying an exponential increase in  $\langle \# \text{ ori} \rangle$  and the rRNA operon copy number for growth rates above  $0.5 \text{ hr}^{-1}$ .

Returning to our rule-of-thumb that one functional rRNA unit is produced per second per transcribing operon, we can estimate the maximum number of ribosomes that could be made as a function of growth rate (Figure 5(B), blue curve). This provides a useful reference alongside the proteomic measurements, particularly in the regime of fast growth. For growth rates above about  $1 \text{ hr}^{-1}$  in particular, we find that cells will need to transcribe rRNA near their maximal rate. The convergence between the maximum rRNA production and measured ribosome copy number shows that rRNA synthesis may begin to present a bottleneck at the fastest growth rates in *E. coli* due to the still-limited copies of rRNA genes. This suggests that the exponential increase in rRNA gene copy number would be insufficient to support further scaling of ribosomal copy number without an increase in the number of rRNA gene copies on the chromosome.

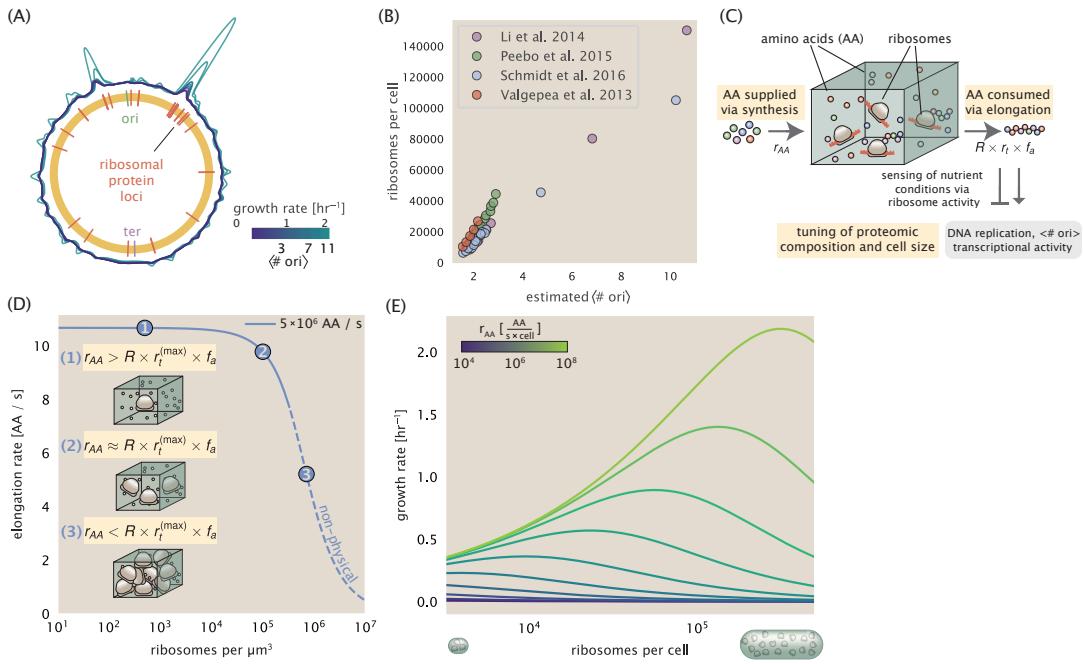
## Nutrient-Mediated Growth Rate Control Dictates both Ribosomal Content and Cell Size

While the preceding section highlights a dominant role for ribosomes in setting the achievable growth rate, our analysis thus far has also shown how the proteomic content and cell size will need to change in response to variable growth conditions and growth rate. The specific mechanism of growth rate control under nutrient limitation that leads to the observed exponential scaling in cell size in *E. coli* and other bacteria, however, has remained unclear (Si et al., 2017) [cite others]. Here we consider the consequences of the translation-limiting growth rate on cell size and its underlying mechanism.

Under translation-limited growth conditions ( $> \approx 0.5 \text{ hr}^{-1}$ ), cells can only increase growth rate by increasing ribosome content. While the simple addition of more ribosomes is likely constrained by macromolecular crowding (Delarue et al., 2018; Soler-Bistué et al., 2020), cells appear to bias their protein synthesis specifically in favor of ribosomes as they increase their size. To see this, we have calculated the position-dependent protein expression across the chromosome by a running Gaussian average of protein copy number (20 kbp st. dev. averaging window) based on each gene's transcriptional start site [Figure 6(B)]. These were median-subtracted to account for the increasing total protein abundance with  $\langle \# \text{ ori} \rangle$ . In particular, we find that the deviations in protein expression with  $\langle \# \text{ ori} \rangle$  are largely restricted to regions of ribosomal protein genes. This result suggests that  $\Phi_R$  is also being tuned in proportion to  $\langle \# \text{ ori} \rangle$  under nutrient-limited growth. Importantly, it is through this additional dependence on  $\Phi_R$ , combined with the exponential increase in  $\langle \# \text{ ori} \rangle$  that was noted in the previous section, that *E. coli* exhibits an exponential increase in cell size with growth rate. Interestingly, this also predicts that the size scaling relationship is not necessarily expected for growth rates below about  $\approx 0.5 \text{ hr}^{-1}$  where ribosomes are not maximally utilized, which indeed seems to be the case [SI single-cell sizes from Si].

In *E. coli*, an accumulation of de-acylated tRNAs at the ribosome's A-site and a strong increase in (p)ppGpp synthesis activity by the enzyme RelA (Hauryliuk et al., 2015). While conventionally associated with a dynamic global response to changes in nutrient conditions through the stringent response, (p)ppGpp increasingly appears to play a role in both the control of the active ribosomal fraction and cell size homeostasis under steady-state nutrient-limited growth (Dai et al., 2016; Zhu and Dai, 2019; Büke et al., 2020; Vadia et al., 2017; Parker et al., 2020). It is now well-documented that *E. coli* cells add a constant volume per origin of replication, which is robust to a remarkable array of cellular perturbations (Si et al., 2017). In addition, there is increasing evidence that (p)ppGpp acts to inhibit the initiation of protein synthesis. From the proteomic measurements, we find that cells also tune their ribosome copy number in direct proportion to  $\langle \# \text{ ori} \rangle$  (??(A)). With cellular (p)ppGpp levels following directly from the state of ribosomal activity, cells appear to use ribosome activity as a direct proxy for nutrient availability, and hence are able to better match cell size (and  $\langle \# \text{ ori} \rangle$ ) and ribosomal content according to the required biosynthetic capacity for growth [Figure 6(C)].

To consider this quantitatively, we assume that cells modulate their proteome (total number of peptide bonds  $N_{\text{pep}}$ , number of ribosomes  $R$ , and ribosomal fraction  $\Phi_R$ ) to better maximize their rate of peptide elongation  $r_e$ . The elongation rate  $r_e$  will depend on how quickly ribosomes can match codons with an amino-acyl tRNA, along with the subsequent steps of peptide bond formation and translocation. This ultimately depends on the cellular concentration of amino acids, which we treat as a single effective species,



**Figure 6. Coordination of cell size and proteomic composition via ribosome activity.** (A) A running Gaussian average (20 kbp st. dev.) of protein copy number is calculated for each growth condition considered by (Schmidt et al., 2016) based on each gene's transcriptional start site. Since total protein abundance increases with growth rate, protein copy numbers are median-subtracted to allow comparison between growth conditions. (# ori) are estimated using the data in (A) and Equation 4. (B) Plot of the ribosome copy number estimated from the proteomic data against the estimated (# ori) (see Appendix Section "Estimation of (# ori)/ (# ter) and (# ori) for additional details"). (C) We consider a unit volume of cellular material composed of amino acids (colored spheres) provided at a supply rate  $r_{AA}$ . These amino acids are polymerized by a pool of ribosomes (brown blobs) at a rate  $r_t \times R \times f_a$ , where  $r_t$  is the elongation rate,  $R$  is the ribosome copy number in the unit volume, and  $f_a$  is the fraction of those ribosomes actively translating. In addition to determining total protein synthesis rate, the nutrient status is gauged by any accumulation of de-acylated tRNAs and synthesis of the secondary messenger (p)ppGpp, which ultimately determine (# ori), cell size, and proteomic composition. (D) The observed elongation rate is plotted as a function of the number of ribosomes. The three points correspond to three regimes of ribosome copy numbers and are shown schematically on the left-hand side. The region of the curve shown as dashed lines represents a non-physical copy number, but is shown for illustrative purposes. This curve was generated using an amino acid supply rate of  $5 \times 10^6 \text{ AA/s}$ , a maximal elongation rate of  $17.1 \text{ AA/s}$ ,  $f_a = 1$ , and a unit cell volume of  $1 \text{ fL}$ . See Appendix Section "Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control" for additional model details. (E) The cellular growth rate is plotted as a function of total cellular ribosome copy number for different cellular amino acid supply rates, with blue and green curves corresponding to low and high supply rates, respectively. As the ribosome copy number is increased, so too is the cell size and total protein abundance  $N_{\text{pep}}$ .

$[\text{AA}]_{\text{eff}}$  (Bosdriesz et al., 2015). Having found that cells do not appear in the synthesis of tRNA or GTP, we determine the the rate of peptide elongation  $r_t$  and achievable growth rate as simply depending on the supply of amino acids (and, therefore, also amino-acyl tRNAs), through a parameter  $r_{AA}$  in units of AA per second, and the rate of amino acid consumption by protein synthesis ( $r_t \times R \times f_a$ ). In Figure 6(D), we illustrate how the elongation rate will depend on the ribosomal copy number for constant  $r_{AA}$ , and further described in the Appendix Section "Derivation of Minimal Model for Nutrient-Mediated Growth Rate Control".

To relate elongation rate to growth rate, we constrain the set of parameters based on our available proteomic measurements; namely, we restrict the values of  $R$ ,  $N_{\text{pep}}$ , and cell size to those associated with the amalgamated proteomic data (described in the Appendix Section "Estimation of Total Protein Content per Cell"). We then consider how changes in the nutrient conditions, through the parameter  $r_{AA}$ , influence the maximum growth rate as determined by Equation 3. Figure 6(E) shows how the growth rate depends on the rate of amino acid supply  $r_{AA}$  as a function of the cellular ribosome copy number. A feature immediately apparent is the presence of a maximal growth rate increases with increasing  $r_{AA}$ . Importantly, however, there is an optimum set of  $R$ ,  $N_{\text{pep}}$ , and cell size that are strictly dependent on the value of  $r_{AA}$ . This shows that increasing the ribosomal concentration beyond the cell's metabolic capacity will have the adverse consequence of depleting the supply of amino acids and lead to a concomitant decrease in the elongation rate  $r_t$  [Figure 6(B)] and growth rate. This helps us understand that while it is important for cells to increase their ribosomal content and cell size in order to increase growth rate, cells will better maximize their achievable growth rate by tuning these parameters according to the available nutrient conditions, since this is ultimately what allows cells to reach the peak for each

curve shown in [Figure 6\(E\)](#).

Also of note is the growth rate trends observed at low values of  $r_{AA}$  [purple and blue lines in [Figure 6\(E\)](#)], representative of growth in nutrient-poor media. In these conditions, there no longer exists a peak in the maximum growth rate, at least within the range of physiologically-relevant ribosome copy numbers considered here. This is the regime, associated with slower growth rates, where cells limit their pool of actively translating ribosomes by decreasing  $f_a$  ([Figure 5\(A\)](#), right-hand panel, inset), likely due to having excess ribosomes relative to the cell's metabolic capacity. By reducing the fraction of actively translating ribosomes, we find that cells instead appear to be prioritizing their pool of available amino acids  $[AA]_{eff}$  in order to increase their translation elongation rate. Consistent with this hypothesis and our model, while inhibition of translation with chloramphenicol further reduces the fraction of actively translating ribosomes  $f_a$ , it results in an increase in the elongation rate ([Dai et al., 2016](#)).

## Discussion

Continued experimental and technological improvements have led to a treasure trove of quantitative biological data ([Hui et al., 2015](#); [Schmidt et al., 2016](#); [Si et al., 2017](#); [Gallagher et al., 2020](#); [Peebo et al., 2015](#); [Valgepea et al., 2013](#)), and an ever advancing molecular view and mechanistic understanding of the constituents that support bacterial growth ([Taheri-Araghi et al., 2015](#); [Morgenstein et al., 2015](#); [Si et al., 2019](#); [Karr et al., 2012](#); [Kostinski and Reuveni, 2020](#); [Macklin et al., 2020](#)). In this work we have compiled and curated what we believe to be the state-of-the-art knowledge on proteomic copy number across a broad range of growth conditions in *E. coli*. Beyond compilation, we have taken a detailed approach in ensuring that the absolute protein abundances reported are directly comparable across growth rates and data sets, allowing us to make assertions about the physiology of *E. coli* rather than chalking up discrepancies with our simple estimates to experimental noise and systematic errors. We have made this data accessible through a [GitHub repository](#), and an [interactive figure](#) that allows exploration of specific protein and protein complex copy numbers.

Through a series of order-of-magnitude estimates that traverse key steps in the bacterial cell cycle, this proteomic data has been a resource to guide our understanding of two key questions: what biological processes limit the absolute speed limit of bacterial growth, and how do cells alter their molecular constituents as a function of changes in growth rate or nutrient availability? While not exhaustive, our series of estimates provide insight on the scales of macromolecular complex abundance across four classes of cellular processes – the transport of nutrients, the production of energy, the synthesis of the membrane and cell wall, and the numerous steps of the central dogma.

In general, the copy numbers of the complexes involved in these processes were in reasonable agreement with our order-of-magnitude estimates. Since many of these estimates represent soft lower-bound quantities, this suggests that cells do not express proteins grossly in excess of what is needed for a particular growth rate. Rather, cells maintain protein abundances that while nearly rate-limiting, are nevertheless sufficient for the require biosynthetic capacity given available nutrient conditions and the observed doubling time. Several exceptions, however, also highlight the dichotomy between a proteome that appears to "optimize" expression according to growth rate and one that must be able to quickly adapt to environments of different nutritional quality. Take, for example, the expression of carbon transporters. Shown in ??(B), we find that cells always express a similar number of glucose transporters irrespective of growth condition. At the same time, it is interesting to note that many of the alternative carbon transporters are still expressed in low but non-zero numbers ( $\approx 10-100$  copies per cell) across growth conditions. This may relate to the regulatory configuration for many of these operons, which require the presence of a metabolite signal in order for alternative carbon utilization operons to be induced ([Monod, 1949](#); [Laxhuber et al., 2020](#)). Furthermore, upon induction, these transporters are expressed and present in abundances in close agreement with a simple estimate.

Of the processes illustrated in [Figure 1](#), we arrive at a perspective where the different processes of bacterial growth all must be carefully coordinated to support rapid growth, but where ribosomal abundance sets a firm upper-limit on the achievable growth rate. This is in some sense unsurprising given the long-held observation that *E. coli* and many other organisms vary their ribosomal abundance as a function of growth conditions and growth rate ([Scott et al., 2010](#); [Metzl-Raz et al., 2017](#)). However, through our dialogue with the proteomic data, two additional key points emerge. The first relates to our question of what process sets the absolute speed limit of bacterial growth. While a cell can parallelize many of its processes simply by increasing the abundance of specific proteins or firing multiple rounds of DNA replication, this is not so for synthesis of ribosomes [[Figure 5\(A\)](#)]. The translation time for each ribosome [ $\approx 7$  min, [Dill et al. \(2011\)](#)] places an inherent limit on the growth rate that can only be surpassed if the cell were to increase their polypeptide elongation rate, or if they could reduce the total protein and rRNA mass of the ribosome. The second point relates to the long-observed correlations between growth rate and cell size ([Schaechter et al., 1958](#); [Si et al., 2017](#)), and between growth rate and ribosomal mass fraction. While both trends have sparked tremendous curiosity and driven substantial amounts of research in their own regards, these relationships are themselves intertwined. In particular, it is the need for cells to increase their absolute number of ribosomes under conditions of rapid growth that require cells to also grow in size.

While the generation of new ribosomes plays a dominant role in growth rate control, there exist other physical limits to the function of cellular processes. One of the key motivations for considering energy production was the physical constraints on total volume and surface area as cells vary their size (*Harris and Theriot, 2018; Ojkic et al., 2019*). As *E. coli* get larger at faster growth rates, an additional constraint begins to arise in energy production and nutrient uptake due to the relative decrease in total surface area, where ATP is predominantly produced (*Szenk et al., 2017*). Specifically, the cell interior requires an amount of energy that scales cubically with cell size, but the available surface area only grows quadratically [Figure 3(A)]. While this threshold does not appear to be met for *E. coli* cells growing at 2 hr<sup>-1</sup> or less, it highlights an additional constraint on growth given the apparent need to increase cell size in order to grow faster. This limit is relevant even to eukaryotic organisms, whose mitochondria exhibit convoluted membrane structures that nevertheless remain bacteria-sized organelles (*Guo et al., 2018*). In the context of bacterial growth and energy production more generally, we have mainly limited our analysis to the aerobic growth conditions associated with the proteomic data and further consideration will be needed for anaerobic growth.

This work is by no means meant to be a complete dissection of bacterial growth rate control, and there are many aspects of the bacterial proteome and growth that we neglected to consider. For example, other recent work (*Liebermeister et al., 2014; Hui et al., 2015; Schmidt et al., 2016*) has explored how the proteome is structured. In the work of *Hui et al. (2015)*, the authors coarse-grained the proteome into six discrete categories being related to either translation, catabolism, anabolism, and others related to signaling and core metabolism. The relative mass fraction of the proteome occupied by each sector could be modulated by external application of drugs or simply by changing the nutritional content of the medium. While we have explored how the quantities of individual complexes are related to cell growth, we acknowledge that higher-order interactions between groups of complexes or metabolic networks at a systems-level may reveal additional insights into how these growth-rate dependences are achieved. Furthermore, while we anticipate the conclusions summarized here are applicable to a wide collection of bacteria with similar lifestyles as *E. coli*, other bacteria and archaea may have evolved other strategies that were not considered. Further experiments with the level of rigor now possible in *E. coli* will need to be performed in a variety of microbial organisms to learn more about how regulation of proteomic composition and growth rate control has evolved over the past 3.5 billion years.

## Methods

### Data Analysis and Availability

All proteomic measurements come from the experimental work of *Schmidt et al. (2016); Peebo et al. (2015); Valgepea et al. (2013)* (mass spectrometry) and *Li et al. (2014)* (ribosomal profiling). Data curation and analysis was done programmatically in Python, and compiled data and analysis files are accessible through a [GitHub repository](#) (DOI:10.5281/zenodo.4091457) associated with this paper as well as on the associated [paper website](#). Additionally, we provide two interactive figures that allow for [rapid exploration of the compiled data sets](#) as well as [exploration of the parameter space of the minimal model](#).

### Acknowledgements

We thank Matthias Heinemann, Alexander Schmidt, and Gene-Wei Li for additional input regarding their data. We also thank all members of the Phillips, Theriot, Kondev, Garcia labs, as well as Ron Milo and Terry Hwa for useful discussions. We thank Suzannah M. Beeler, Jonas Cremer, Avi Flamholz, Soichi Hirokawa, and Manuel Razo-Mejia for reading and providing comments on drafts of this manuscript. R.P. is supported by La Fondation Pierre-Gilles de Gennes, the Rosen Center at Caltech, and the NIH 1R35 GM118043 (MIRA). J.A.T. is supported by the Howard Hughes Medical Institute, and NIH Grant R37-AI036929. N.M.B is a HHMI Fellow of The Jane Coffin Childs Memorial Fund. H.G.G. is supported by the Burroughs Wellcome Fund Career Award at the Scientific Interface, the Sloan Research Foundation, the Human Frontiers Science Program, the Searle Scholars Program, the Shurl & Kay Curci Foundation, the Hellman Foundation, the NIH Director's New Innovator Award (DP2 OD024541-01), and an NSF CAREER Award (1652236). D.S.F. is supported by an NSF award (PHY-1607606) and the NIH (NIH R01-AI13699201).

### Competing Interests

The authors declare no competing interests.

### References

- Aidelberg, G., Towbin, B. D., Rothschild, D., Dekel, E., Bren, A., and Alon, U. (2014). Hierarchy of non-glucose sugars in *Escherichia coli*. *BMC Systems Biology*, 8(1):133.
- Amir, A. (2017). Is cell size a spandrel? *eLife*, 6:18261.

- Antonenko, Y. N., Pohl, P., and Denisov, G. A. (1997). Permeation of ammonia across bilayer lipid membranes studied by ammonium ion selective microelectrodes. *Biophysical Journal*, 72(5):2187–2195.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- Ason, B., Bertram, J. G., Hingorani, M. M., Beechem, J. M., O'Donnell, M., Goodman, M. F., and Bloom, L. B. (2000). A Model for *Escherichia coli* DNA Polymerase III Holoenzyme Assembly at Primer/Template Ends: DNA Triggers A Change In Binding Specificity of the  $\gamma$  Complex Clamp Loader. *Journal of Biological Chemistry*, 275(4):3006–3015.
- Assentoft, M., Kaptan, S., Schneider, H.-P., Deitmer, J. W., de Groot, B. L., and MacAulay, N. (2016). Aquaporin 4 as a NH<sub>3</sub> Channel. *Journal of Biological Chemistry*, 291(36):19184–19195.
- Barreteau, H., Kovač, A., Boniface, A., Sova, M., Gobec, S., and Blanot, D. (2008). Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, 32(2):168–207.
- Basan, M., Zhu, M., Dai, X., Warren, M., Sévin, D., Wang, Y.-P., and Hwa, T. (2015). Inflating bacterial cells by increased protein synthesis. *Molecular Systems Biology*, 11(10):836.
- Bauer, S. and Ziv, E. (1976). Dense growth of aerobic bacteria in a bench-scale fermentor. *Biotechnology and Bioengineering*, 18(1):81–94. \_eprint: <https://onlinelibrary.wiley.com/doi/10.1002/bit.260180107>.
- Bosdriesz, E., Molenaar, D., Teusink, B., and Bruggeman, F. J. (2015). How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *The FEBS Journal*, 282(10):2029–2044.
- Bremer, H. and Dennis, P. P. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus*, 3(1).
- Büke, F., Grilli, J., Lagomarsino, M. C., Bokinsky, G., and Tans, S. (2020). ppGpp is a bacterial cell size regulator. *bioRxiv*, 266:2020.06.16.154187.
- Catherwood, A. C., Lloyd, A. J., Tod, J. A., Chauhan, S., Slade, S. E., Walkowiak, G. P., Galley, N. F., Punekar, A. S., Smart, K., Rea, D., Evans, N. D., Chappell, M. J., Roper, D. I., and Dowson, C. G. (2020). Substrate and Stereochemical Control of Peptidoglycan Cross-Linking by Transpeptidation by *Escherichia coli* PBP1B. *Journal of the American Chemical Society*, 142(11):5034–5048.
- Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Okano, H., Williamson, J. R., Fredrick, K., and Hwa, T. (2018). Slowdown of Translational Elongation in *Escherichia coli* under Hyperosmotic Stress. - PubMed - NCBI. *mBio*, 9(1):281.
- Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., Williamson, J. R., Fredrick, K., Wang, Y.-P., and Hwa, T. (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nature Microbiology*, 2(2):16231.
- Davidi, D., Noor, E., Liebermeister, W., Bar-Even, A., Flamholz, A., Tummler, K., Barenholz, U., Goldenfeld, M., Shlomi, T., and Milo, R. (2016). Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro kcat measurements. *Proceedings of the National Academy of Sciences*, 113(12):3401–3406.
- Delarue, M., Brittingham, G. P., Pfeffer, S., Surovtsev, I. V., Pinglay, S., Kennedy, K. J., Schaffer, M., Gutierrez, J. I., Sang, D., Poterewicz, G., Chung, J. K., Plitzko, J. M., Groves, J. T., Jacobs-Wagner, C., Engel, B. D., and Holt, L. J. (2018). mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. *Cell*, 174(2):338–349.e20.
- Dennis, P. P., Ehrenberg, M., and Bremer, H. (2004). Control of rRNA Synthesis in *Escherichia coli*: a Systems Biology Approach. *Microbiology and Molecular Biology Reviews*, 68(4):639–668.
- Dill, K. A., Ghosh, K., and Schmit, J. D. (2011). Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*, 108(44):17876–17882.
- Erickson, D. W., Schink, S. J., Patsalo, V., Williamson, J. R., Gerland, U., and Hwa, T. (2017). A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature*, 551(7678):119–123.
- Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: Peculiarities of regulation and impact on growth and product formation. *Applied Microbiology and Biotechnology*, 94(6):1483–1494.
- Forchhammer, J. and Lindahl, L. (1971). Growth rate of polypeptide chains as a function of the cell growth rate in a mutant of *Escherichia coli* 15. *Journal of Molecular Biology*, 55(3):563–568.
- Gallagher, L. A., Bailey, J., and Manoil, C. (2020). Ranking essential bacterial processes by speed of mutant death. *Proceedings of the National Academy of Sciences*.

- Godin, M., Delgado, F. F., Son, S., Grover, W. H., Bryan, A. K., Tzur, A., Jorgensen, P., Payer, K., Grossman, A. D., Kirschner, M. W., and Manalis, S. R. (2010). Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390.
- Guo, Y., Li, D., Zhang, S., Yang, Y., Liu, J.-J., Wang, X., Liu, C., Milkie, D. E., Moore, R. P., Tulu, U. S., Kiehart, D. P., Hu, J., Lippincott-Schwartz, J., Betzig, E., and Li, D. (2018). Visualizing Intracellular Organelle and Cytoskeletal Interactions at Nanoscale Resolution on Millisecond Timescales. *Cell*, 175(5):1430–1442.e17.
- Harris, L. K. and Theriot, J. A. (2018). Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis. *Trends in Microbiology*, 26(10):815–832.
- Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.
- Heldal, M., Norland, S., and Tumyr, O. (1985). X-ray microanalytic method for measurement of dry matter and elemental content of individual bacteria. *Applied and Environmental Microbiology*, 50(5):1251–1257.
- Ho, P.-Y. and Amir, A. (2015). Simultaneous regulation of cell size and chromosome replication in bacteria. *Frontiers in Microbiology*, 6.
- Hui, S., Silverman, J. M., Chen, S. S., Erickson, D. W., Basan, M., Wang, J., Hwa, T., and Williamson, J. R. (2015). Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2):e784–e784.
- Jun, S., Si, F., Pugatch, R., and Scott, M. (2018). Fundamental principles in bacterial physiology - history, recent progress, and the future with focus on cell size control: A review. *Reports on Progress in Physics*, 81(5):056601.
- Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., Assad-Garcia, N., Glass, J. I., and Covert, M. W. (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150(2):389–401.
- Klumpp, S. and Hwa, T. (2008). Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proceedings of the National Academy of Sciences*, 105(51):20245–20250.
- Klumpp, S. and Hwa, T. (2014). Bacterial growth: Global effects on gene expression, growth feedback and proteome partition. *Current Opinion in Biotechnology*, 28:96–102.
- Klumpp, S., Scott, M., Pedersen, S., and Hwa, T. (2013). Molecular crowding limits translation and cell growth. *Proceedings of the National Academy of Sciences*, 110(42):16754–16759.
- Kostinski, S. and Reuveni, S. (2020). Ribosome Composition Maximizes Cellular Growth Rates in *E. coli*. *Physical Review Letters*, 125(2):028103.
- Lascu, I. and Gonin, P. (2000). The Catalytic Mechanism of Nucleoside Diphosphate Kinases. *Journal of Bioenergetics and Biomembranes*, 32(3):237–246.
- Laxhuber, K. S., Morrison, M. J., Chure, G., Belliveau, N. M., Strandkvist, C., Naughton, K. L., and Phillips, R. (2020). Theoretical investigation of a genetic switch for metabolic adaptation. *PLOS ONE*, 15(5):e0226453.
- Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell*, 157(3):624–635.
- Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual account of protein investment in cellular functions. *Proceedings of the National Academy of Sciences*, 111(23):8488–8493.
- Liu, M., Durfee, T., Cabrera, J. E., Zhao, K., Jin, D. J., and Blattner, F. R. (2005). Global Transcriptional Programs Reveal a Carbon Source Foraging Strategy by *Escherichia coli*. *Journal of Biological Chemistry*, 280(16):15921–15927.
- Lovering, A. L., Safadi, S. S., and Strynadka, N. C. (2012). Structural Perspective of Peptidoglycan Biosynthesis and Assembly. *Annual Review of Biochemistry*, 81(1):451–478.
- Lynch, M. and Marinov, G. K. (2015). The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences*, 112(51):15690–15695.
- Maaløe, O. (1979). *Regulation of the Protein-Synthesizing Machinery—Ribosomes, tRNA, Factors, and So On*. Gene Expression. Springer.
- Macklin, D. N., Ahn-Horst, T. A., Choi, H., Ruggero, N. A., Carrera, J., Mason, J. C., Sun, G., Agmon, E., DeFelice, M. M., Maayan, I., Lane, K., Spangler, R. K., Gillies, T. E., Paull, M. L., Akhter, S., Bray, S. R., Weaver, D. S., Keseler, I. M., Karp, P. D., Morrison, J. H., and Covert, M. W. (2020). Simultaneous cross-evaluation of heterogeneous *E. coli* datasets via mechanistic simulation. *Science*, 369(6502).
- Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife*, 6:e03528.
- Mikucki, J. A., Pearson, A., Johnston, D. T., Turchyn, A. V., Farquhar, J., Schrag, D. P., Anbar, A. D., Priscu, J. C., and Lee, P. A. (2009). A Contemporary Microbially Maintained Subglacial Ferrous "Ocean". *Science*, 324(5925):397–400.

- Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Research*, 38(suppl\_1):D750–D753.
- Molenaar, D., van Berlo, R., de Ridder, D., and Teusink, B. (2009). Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular Systems Biology*, 5(1):323.
- Monod, J. (1947). The phenomenon of enzymatic adaptation and its bearings on problems of genetics and cellular differentiation. *Growth Symposium*, 9:223–289.
- Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394.
- Morgenstein, R. M., Bratton, B. P., Nguyen, J. P., Ouzounov, N., Shaevitz, J. W., and Gitai, Z. (2015). RodZ links MreB to cell wall synthesis to mediate MreB rotation and robust morphogenesis. *Proceedings of the National Academy of Sciences*, 112(40):12510–12515.
- Neidhardt, F. C., Ingraham, J., and Schaechter, M. (1991). *Physiology of the Bacterial Cell - A Molecular Approach*, volume 1. Elsevier.
- Ojkic, N., Serbanescu, D., and Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria. *eLife*, 8:642.
- Parker, D. J., Lalanne, J.-B., Kimura, S., Johnson, G. E., Waldor, M. K., and Li, G.-W. (2020). Growth-Optimized Aminoacyl-tRNA Synthetase Levels Prevent Maximal tRNA Charging. *Cell Systems*, 11(2):121–130.e6.
- Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.
- Peebo, K., Valgepea, K., Maser, A., Nahku, R., Adamberg, K., and Vilu, R. (2015). Proteome reallocation in *Escherichia coli* with increasing specific growth rate. *Molecular Biosystems*, 11(4):1184–1193.
- Phillips, R. (2018). Membranes by the Numbers. In *Physics of Biological Membranes*, pages 73–105. Springer, Cham.
- Ranganathan, S., Tee, T. W., Chowdhury, A., Zomorodi, A. R., Yoon, J. M., Fu, Y., Shanks, J. V., and Maranas, C. D. (2012). An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering*, 14(6):687–704.
- Roller, B. R. K., Stoddard, S. F., and Schmidt, T. M. (2016). Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature microbiology*, 1(11):1–7.
- Rosenberg, H., Gerdes, R. G., and Chegwidde, K. (1977). Two systems for the uptake of phosphate in *Escherichia coli*. *Journal of Bacteriology*, 131(2):505–511.
- Ruppe, A. and Fox, J. M. (2018). Analysis of Interdependent Kinetic Controls of Fatty Acid Synthases. *ACS Catalysis*, 8(12):11722–11734.
- Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *Microbiology*, 19(3):592–606.
- Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann, M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nature Biotechnology*, 34(1):104–110.
- Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. *Science*, 330(6007):1099–1102.
- Shi, H., Bratton, B. P., Gitai, Z., and Huang, K. C. (2018). How to Build a Bacterial Cell: MreB as the Foreman of *E. coli* Construction. *Cell*, 172(6):1294–1305.
- Si, F., Le Treut, G., Sauls, J. T., Vadia, S., Levin, P. A., and Jun, S. (2019). Mechanistic Origin of Cell-Size Control and Homeostasis in Bacteria. *Current Biology*, 29(11):1760–1770.e7.
- Si, F., Li, D., Cox, S. E., Sauls, J. T., Azizi, O., Sou, C., Schwartz, A. B., Erickstad, M. J., Jun, Y., Li, X., and Jun, S. (2017). Invariance of Initiation Mass and Predictability of Cell Size in *Escherichia coli*. *Current Biology*, 27(9):1278–1287.
- Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews*, 40(1):133–159.
- Soler-Bistué, A., Aguilar-Pierlé, S., García-Garcé, M., Val, M.-E., Sismeiro, O., Varet, H., Sieira, R., Krin, E., Skovgaard, O., Comerci, D. J., Eduardo P. C. Rocha, and Mazel, D. (2020). Macromolecular crowding links ribosomal protein gene dosage to growth rate in *Vibrio cholerae*. *BMC Biology*, 18(1):1–18.
- Stasi, R., Neves, H. I., and Spira, B. (2019). Phosphate uptake by the phosphonate transport system PhnCDE. *BMC Microbiology*, 19.
- Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van Leeuwenhoek*, 39(1):545–565.
- Stouthamer, A. H. and Bettenhausen, C. W. (1977). A continuous culture study of an ATPase-negative mutant of *Escherichia coli*. *Archives of Microbiology*, 113(3):185–189.

- Szenk, M., Dill, K. A., and de Graff, A. M. R. (2017). Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104.
- Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Cell-size control and homeostasis in bacteria. - PubMed - NCBI. *Current Biology*, 25(3):385–391.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, 28(1):33–36.
- Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J. T., Heijnen, J. J., and van Gulik, W. M. (2010). Genome-derived minimal metabolic models for *Escherichia coli* MG1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering*, 107(2):369–381. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.22802>.
- The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1):D330–D338.
- Vadia, S., Tse, J. L., Lucena, R., Yang, Z., Kellogg, D. R., Wang, J. D., and Levin, P. A. (2017). Fatty Acid Availability Sets Cell Envelope Capacity and Dictates Microbial Cell Size. *Current Biology*, 27(12):1757–1767.e5.
- Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. *Molecular BioSystems*, 9(9):2344.
- Weber, J. and Senior, A. E. (2003). ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Letters*, 545(1):61–70.
- Willsky, G. R., Bennett, R. L., and Malamy, M. H. (1973). Inorganic Phosphate Transport in *Escherichia coli*: Involvement of Two Genes Which Play a Role in Alkaline Phosphatase Regulation. *Journal of Bacteriology*, 113(2):529–539.
- Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 108(46):18643–18648.
- Zhang, Q., Brambilla, E., Li, R., Shi, H., Lagomarsino, M. C., Sclavi, B., and Bulman, Z. (2020). A Decrease in Transcription Capacity Limits Growth Rate upon Translation Inhibition. *mSystems*, 5(5):46.
- Zhu, M. and Dai, X. (2019). Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in *Escherichia coli*. *Nucleic Acids Research*, 47(9):4684–4693.
- Zhuang, K., Vemuri, G. N., and Mahadevan, R. (2011). Economics of membrane occupancy and respiro-fermentation. *Molecular Systems Biology*, 7(1):500.