Point Estimate vs Continuum Estimates

Thus far in the project, Nathan and I have been going about the estimates by choosing a particular growth rate λ (typically $\lambda = 0.5\,\mathrm{hr}^{-1}$, $t_{\mathrm{double}} = 5000\,\mathrm{s}$) and using some rules-of-thumb to compute how many protein complexes involved in a given biological process should be present. In general, this approach has worked well and our estimates have landed in the ball park of the observed copy numbers, ultimately allowing us to proclaim that we know what "sets the scale".

However, I (Griffin) have been trying to abstract these point estimates to a continuum of growth rates ($\lambda \in [0, 2]\,\mathrm{hr}^{-1}$). My reasoning for doing so is primarily due to some comments we received during Manu's group meeting as well as comments I received from friends who have said "shouldn't you be predicting the slope?" Furthermore, I (Griffin) think one can make the argument that there are *two* scales that we must be able to quantitatively define in order to "understand" the process. The first scale is the absolute copy numbers of the complexes (like we have done for the point estimates). The second scale is the dependence on the growth rate. In many cases, I believe this boils down to either a dependence on the cell volume (or cell size, depending on your favorite nomenclature) *or* a dependence on the chromosome copy number. These two "natural variables" have both qualitative and quantitative differences.

As an example, let's walk through an estimate using both the point estimate and the continuum estimate approaches. While we can do this for any process, I've chosen to consider the dynamics of phosphate transporters.

1 Phosphate Transporters – Background

Phosphorus is critical to the cellular energy economy in the form of high-energy phosphodiester bonds making up DNA, RNA, and the NTP energy pool as well as playing a critical role in the post-translational modification of proteins and defining the polar-heads of lipids. In total, phosphorus makes up ≈3% of the cellular dry mass which in typical experimental conditions is in the form of inorganic phosphate. The cell membrane has remarkably low permeability to this highly-charged and critical molecule, therefore requiring the expression of active transport systems. In *E. coli*, the proton electrochemical gradient across the inner membrane is leveraged to transport inorganic phosphate into the cell [1]. Proton-solute symporters are widespread in *E. coli* [2, 3] and can have rapid transport rates of 50 molecules per second for sugars and other solutes (BNID: 103159; 111777, [4]). Proton transporters, such as those in the F_1-F_0 ATP synthase, leverage the proton chemical gradient to drive conformational changes in the protein complex and can shuttle ≈ 1000 protons per second (BNID: 104890; 103390, [4]). In *E. coli* the PitA phosphate transport system has been shown to be very tightly coupled with the proton electrochemical gradient with a 1:1 proton:phosphate stoichiometric ratio [5, 6]. While the precise rate per transporter is not well known for the *E. coli* Pit system, we can state that it is likely in between the aforementioned rates for sugar transport and ATP synthesis, on the order of ≈ 300 phosphates transported per second.

1.1 Point Estimate

Using the fact that the cellular dry mass is ≈ 3% phosphorus and transport through the Pit system of *E. coli* is ≈ 300 phosphates per second, we can immediately compute a point estimate for the number of transporters present per cell.

We begin with considering a modest *E. coli* doubling time of ≈ 5000s. At this growth rate, we can approximate the cell as having a total volume of ≈ 1 fL which corresponds to a mass of ≈ 1 pg. Assuming that ≈ 70% of the *E. coli* cell mass is water, and ≈ 3% of the dry mass is phosphorus, we arrive at total phosphorus mass fraction of

$$\theta_P \approx 0.3 \times 0.03 \approx 0.01. \tag{1}$$

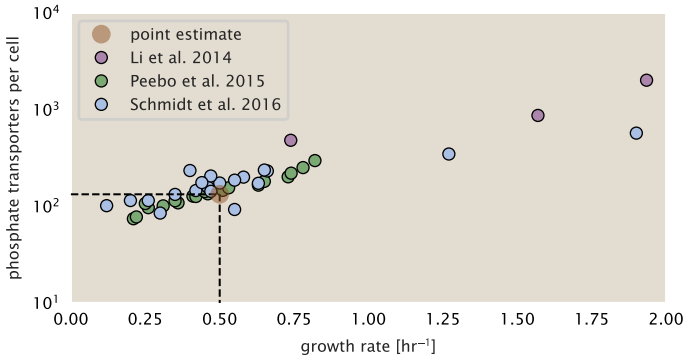

Figure 1: **Point estimate for the number of phosphate transporters and comparison with data.** Point estimate as outlined in Eq. 3 is shown in brown. Experimental meausrements of the total number of PitA and PitB phosphate transporters for each data set are shown as colored points. Vertical dashed line indicates a doubling time of 5000 seconds and horizontal dashed line indicates a copy number of ≈ 150 per cell.

Given that the molecular weight of phosphorus is 30 Da, the total number of phosphorus atoms per cell can be calculated as

$$N_{\text{phosphorus}} \approx \frac{\theta_{\text{P}} \times 1\,\text{pg}}{30\,\text{Da}} \times \frac{6 \times 10^{11}\,\text{Da}}{1\,\text{pg}} \approx 2 \times 10^{8}\,\text{P}. \tag{2}$$

We can now assume that all phosphorus atoms enter the cell via phosphate and make the further approximation that this transport comes exclusively through the Pit system. Using our estimate of ≈ 300 phosphates per second as the transport rate via the Pit system transporters, we arrive at an estimate for the total number of transporters to be

$$N_{\text{transporters}} \approx \frac{2 \times 10^{8}\,\text{P}}{1\,\text{cell}} \times \frac{1\,\text{s} \cdot \text{transporter}}{300\,\text{phosphate}} \frac{1\,\text{cell}}{5000\,\text{s}} \approx 150\,\text{transporters}. \tag{3}$$

Figure 1 shows a comparison of this estimate the experimentally measured values. While our estimate is very much in line with the observed numbers, we emphasize that this is likely a slight over estimate of the number of transporters needed as there are other phosphorous scavenging systems, such as the ATP-dependent phosphate transporter Pst system which we have neglected.

1.2 Continuum Estimate

Rather than choosing an arbitrary doubling time and rule-of-thumb for the cell mass, we can draw upon the detailed literature of *E. coli* growth to estimate the number of transporters needed across a continuum of growth rates, assuming that the entirety of the transported phosphate ions are used in the generation of new cell mass or energy production.

Recent single-cell measurements have extensively characterized the time-dependent growth of *E. coli* in terms of dimensions, cell mass, and cell volume. A combination of recent works by Fangwei Si *et al.* [7, 8] have revealed that cell volume V is exponential with growth rate λ and follows an approximate scaling of

$$V(\lambda) \approx \frac{e^{\lambda}}{2}. \tag{4}$$

Using this expression, a growth-rate independent buoyant density of $\rho \approx 1.1\,\text{pg}/\mu\text{m}^3$ (BNID: 103875, [4]), a growth-rate independent elemental composition of $\theta_P \approx 0.01$ of the total cell mass, and the atomic weight $m_P = 30Da$, the

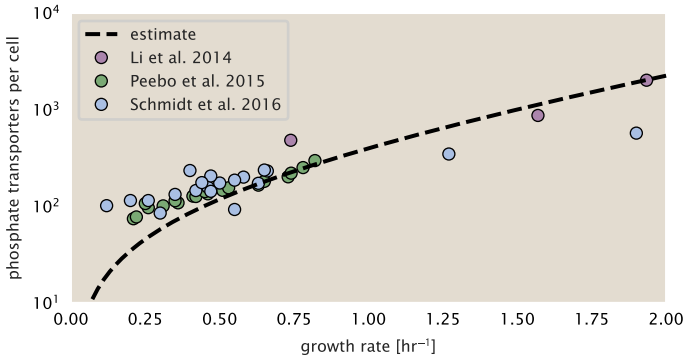

Figure 2: **Estimated number of phosphate tranposrters needed as a function of growth rate.** Dashed line shows the predicted scaling as prescribed by Eq. 6. Colored points correspond to the total number of PitA and PitB transporters present in each dataset.

total number of phosphorus atoms needed as a function of growth rate $N_{\mathrm{P}}(\lambda)$ can be computed as

$$N_{\mathrm{P}}(\lambda) \approx \frac{\theta_{\mathrm{P}} \rho V(\lambda)}{m_P}. \tag{5}$$

Assuming that the rate of phosphate transport $r = 300\,\mathrm{PO}_4^{2-} \cdot s^{-1} \cdot$ transporter is independent of the growth rate, we can now compute the estimated number of transporters per cell as a function of growth rate as

$$N_{\mathrm{transporters}} \approx \frac{\theta_{\mathrm{P}} \rho V(\lambda)}{m_{\mathrm{P}} r \left(\frac{\log 2}{\lambda}\right)}, \tag{6}$$

where the factor $\log 2/\lambda$ computes the doubling time t_{double}

$$t_{\mathrm{double}} = \log 2/\lambda. \tag{7}$$

Figure 2 shows the result of Eq. 6 plotted as a function of the growth rate and compares it with the experimental measurements. Overall, we find that this simple estimate based on cell scaling is sufficient to describe both the absolute number of the phosphate transporters as well as the dependence on the growth rate. We note that this simple scaling argument significantly underestimates the number of phosphate transporters at very slow growth rates ($\lambda < 0.2\,\mathrm{hr}^{-1}$). This may be due to the fact that at very slow growth rates, the cellular physiology prioritizes other metabolic functions over he generation of new cell mass (to anthropomorphize, the cell focuses on survival rather then reproduction).

1.3 Generalization to Other Transport Processes

The scaling argument given in Eq. 6 can be cast in very general terms for an array of transported materials. For element X the number of transporters can be computed as

$$N_{\mathrm{X}} \approx \frac{\theta_{\mathrm{X}} \rho V(\lambda)}{m_{\mathrm{X}} r_{\mathrm{X}} \left(\frac{\log 2}{\lambda}\right)}, \tag{8}$$

where the rate r_{X} must be in units of atoms per unit time. Figure 3 shows this approach applied to glucose, phosphorus, and sulfur transport.

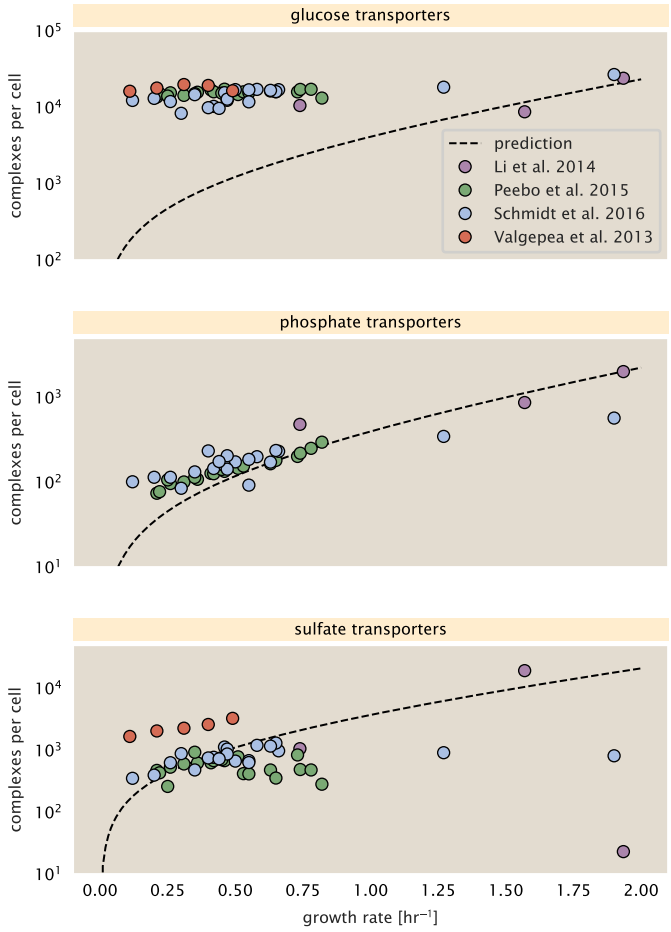

Figure 3: **Growth-rate dependent scaling of transport systems.** The estimated growth rate dependence for the transport of glucose, phosphorus, and sulfate is shown as dashed lines from top to bottom, respectively. Experimental measurements are shown as colored points.

2 Arguments and Compromise

Nathan and I have been discussing this for quite some time and I think we each have mid-grade level chilis about which way to proceed. I am in favor of the continuum estimate approach and Nathan is partial to using only the point estimates.

Here's my paraphrased version of Nathan's argument (with Nathan's approval) for using the point estimate only:

1. We want to provide some rationale just for the observed scale, and not necessarily the dependence. We consider the dependence for translation as we know some more details about how the parameters change.
2. It's not fair to make predictions at very slow growth rates because we only consider the processes necessary to make new cell mass. At slow growth rates, maybe this isn't totally fair.
3. The scaling will look the same (qualitatively) for every category where the cell size / cell mass is the natural variable.
4. Adding the exponential scaling for cell volume is very *E. coli* specific and undercuts the "fundamental limits for bacterial division" which is the main point of the paper.

Here's my argument for estimating across all growth rates.

1. While some of these parameters may change (i.e. transport rates), our rules-of-thumb for *E. coli* come from different growth regimes. We implicitly make the assumption that the rate of transport (for example) is growth rate independent.
2. The process for drawing predictions across growth rates is the exact same as using a standard time of 5000 s, just using different starting values for the cell mass and division time.
3. Maybe we can't say much at exceedingly slow growth rates, but most of the *E. coli* specific data we are comparing to lies in the regime of "moderate" growth rates and is probably fair to consider in this manner. A way around this is maybe to not show the scaling below $0.1\ \mathrm{hr}^{-1}$?
4. The processes we consider in the paper can be thought of as fundamental limits, but all of our estimates and arguments do ultimately come from *E. coli* specific biology. For example, we aren't saying anything about extracellular electron transport which ends up being a critical mode of respiration for a lot of non *E. coli* flavored bugs.

One compromise is that we proceed with the point estimates as we have currently been doing for the main text and have the continuum approach for each case in the SI. For the main-text figures, we can include both, an example of which is shown in Fig. 4

References

[1] H. Rosenberg, R. G. Gerdes, and K. Chegwidden. Two systems for the uptake of phosphate in Escherichia coli. *Journal of Bacteriology*, 131(2):505–511, August 1977.

[2] Sofia Ramos and H. Ronald Kaback. The relation between the electrochemical proton gradient and active transport in Escherichia coli membrane vesicles. *Biochemistry*, 16(5):854–859, March 1977.

[3] I R Booth, W J Mitchell, and W A Hamilton. Quantitative analysis of proton-linked transport systems. The lactose permease of Escherichia coli. *Biochemical Journal*, 182(3):687–696, September 1979.

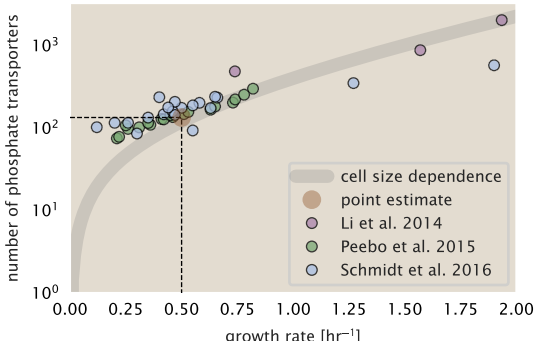

Figure 4: **An example of a compromise plot showing both the point estimate and the cell-size scaling argument.** This shows what the phosphate transporter plot for the main-text would look like under a compromise where the focus is still on the point estimate, but we also include the predicted scaling with growth rate and direct the reader to the SI for more information.

[4] Ron Milo, Paul Jorgensen, Uri Moran, Griffin Weber, and Michael Springer. BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Research*, 38(suppl_1):D750–D753, January 2010.

[5] Robyn M. Harris, Dianne C. Webb, Susan M. Howitt, and Graeme B. Cox. Characterization of PitA and PitB from *Escherichia coli*. *Journal of Bacteriology*, 183(17):5008–5014, September 2001.

[6] Adam M Feist, Christopher S Henry, Jennifer L Reed, Markus Krummenacker, Andrew R Joyce, Peter D Karp, Linda J Broadbelt, Vassily Hatzimanikatis, and Bernhard Ø Palsson. A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology*, 3(1):121, January 2007.

[7] Fangwei Si, Dongyang Li, Sarah E Cox, John T Sauls, Omid Azizi, Cindy Sou, Amy B Schwartz, Michael J Erickstad, Yonggun Jun, Xintian Li, and Suckjoon Jun. Invariance of Initiation Mass and Predictability of Cell Size in Escherichia coli. *Current Biology*, 27(9):1278–1287, May 2017.

[8] Fangwei Si, Guillaume Le Treut, John T. Sauls, Stephen Vadia, Petra Anne Levin, and Suckjoon Jun. Mechanistic Origin of Cell-Size Control and Homeostasis in Bacteria. *Current Biology*, 29(11):1760–1770.e7, June 2019.