Supplemental material for: Fundamental limits on the rate of bacterial cell division

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

***For correspondence:**

*These authors contributed equally to this work

[1]Department of Biology, University of Washington, Seattle, WA, USA; [2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA; [3]Department of Applied Physics, California Institute of Technology, Pasadena, CA, USA; [4]Department of Chemical Engineering, Stanford University, Stanford, CA, USA; [5]Department of Molecular Cell Biology and Department of Physics, University of California Berkeley, Berkeley, CA, USA; [6]Department of Physics, Brandeis University, Waltham, MA, USA; [7]Department of Applied Physics, Stanford University, Stanford, CA, USA; [8]Allen Institute for Cell Science, Seattle, WA, USA; [9]Department of Physics, California Institute of Technology, Pasadena, CA, USA; [†]Address correspondence to phillips@pboc.caltech.edu; [*]Contributed equally

Average protein expression across the chromosome.

In Figure 7(B) of the main text we plotted the average protein copy number along *E. coli*'s chromosome using a boxcar averaging (i.e. running average) window of 0.5 Mb. This means that at each position on the chromosome, proteins with a transcription start site that were +/- 0.25 Mb from that position were included in the calculated average. For *E. coli*, position 0 bp does not correspond the location of the origin and we keep to this convention, using the reported positional information from EcoCyc. Since the chromosome is circular, when calculating the average at positions near the end positions (i.e. near either 0 bp or 4.6 Mb) we include the copy numbers on opposing ends. For example, calculating the average for position 0 bp would include the range from 4.1 Mb to 0.25 Mb). Here we provide some additional analysis to show how the absolute copy numbers compare across growth conditions, as well as the effect of the specific averaging window size.

In the main text we centered each data set accoding to the mean average in order to compare the relative changes in copy number along the length of the chromosome in each data set. In reality, there is also a correlation between the total genomic content and protein copy number, which increases at faster growth. This is shown in Figure 1(A), where we plot the boxcar average from each growth condition without rescaling each about their mean values.

One of the challenges in interpreting this analysis is that the protein copy numbers for a small subset of proteins vary dramatically as a function of growth rate. This is particularly true for ribosomal proteins. In order to check whether the result is due simply due to the change in ribosomal copy number, we repeated the analysis with all ribosome proteins, and the translation elongation factor EF-Tu removed (Figure 1(B)). Indeed we still see a skew in protein abundance, with higher overall expression at the origin.

The other important parameter in this analysis is the size of the averaging window, which we took at 0.5 Mb. In Figure 2 we show the results when using averaging window sizes of 0.05 Mb, 0.25 Mb, 0.5 Mb, 1 Mb, and 2 Mb. Aside from the smallest window size of 0.05 Mb, the analysis seems to show a similar result, with proteins near the origin showing highest expression. For the window size of 0.05 Mb, the copy numbers become much noisier due to the large differences in protein

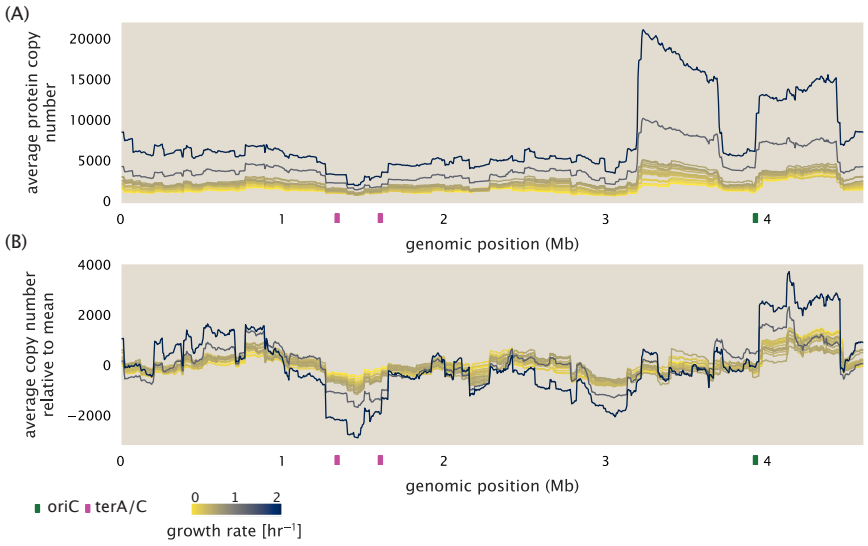

Figure 1. Position-dependent protein expression at different growth rates. (A) Protein copy number is reported along the length of the chromosome using a boxcar averaging, with window size of 0.5 Mb. (B)The boxcar average protein copy number, shifted relative to the mean value for each growth conditions, is show for a window size of 0.5 Mb. In this plot, all ribosomal proteins and elongation factor EF-Tu were excluded in the analysis.

copy number that are observed irrespective of the specific growth rate.

Estimation of ⟨#ori⟩ / ⟨#ter⟩ and ⟨#ori⟩.

In the main text we consider the possibility that cells are varying their chromosomal content in order to increase ribosomal content, which will otherwise become limiting due to the maximal rRNA production per rRNA gene copy. While it is difficult to show a causal relationship, under such a scenario we nonetheless expect the ribosomal contant to vary with the particular state of DNA replication. In particular, we can expect that the ⟨#ori⟩ / ⟨# ter⟩ ratio will reflect the skew in gene dosage for genes near the origin and therefore should relate to the relative abundance of ribosomes that are present closer to the origin. ⟨#ori⟩ is also useful to consider since it will reflect the total DNA content in the cell and therefore should relate to the total ribosomal abundance.

In order to estimate the ⟨#ori⟩ / ⟨# ter⟩ ratio, and ⟨#ori⟩, we made use of the measurements from Si *et al.* (2017). We consider their measurements of DNA replication time (t_C, 'C' period of cell division), total cell cycle time (t_{cyc}, 'C' + 'D' period of cell division), and doubling time τ from wild-type *E. coli* growing across a range of growth conditions.

We begin by considering ⟨#ori⟩. If the cell cycle time takes longer than the time of cell division, the cell will need to initiate DNA replication more often than its rate of division, $2^{\lambda t} = 2^{ln(2)\cdot t/\tau}$. Cells will need to do in in proportion to the ratio $\lambda_{cyc}/\lambda = t_{cyc}/\tau$, and the average number of origins per cell is then given by $2^{t_{cyc}/\tau}$. The average number of termini will in contrast depend on the lag time between DNA replication and cell division, t_D, with ⟨#ori⟩ / ⟨# ter⟩ ratio = $2^{t_{cyc}/\tau - t_D/\tau} = 2^{t_C/\tau}$.

In Figure 4(A) and (B) we plot the measured t_C and t_{cyc} values versus the doubling time from Si *et al.* data. The authors estimated t_C by marker frequency analysis using qPCR, while $t_{cyc} = t_C + t_D$ were inferred from t_C and τ. In the plots we see that both t_C and t_{cyc} reach a minimum at around 40 and 75 minutes, respectively. For a C period of 40 minutes, this would correspond to a maximum rate of elongation of about 1,000 bp/sec. Since we lacked a specific model to describe how each of these parameters vary with growth condition, we assumed that they were linearly dependent on the doubling time. For each parameter, t_C and t_{cyc}, we split them up into two domains corresponding to poorer nutrient conditions and rich nutrient conditions ($\tau \approx 55$ minutes). The fit lines are shown as solid black lines. In Figure 4(C) and (D) we also show t_C and t_{cyc} as a function of growth rate λ along with our piecewise linear fits, which match the plots in the main text.

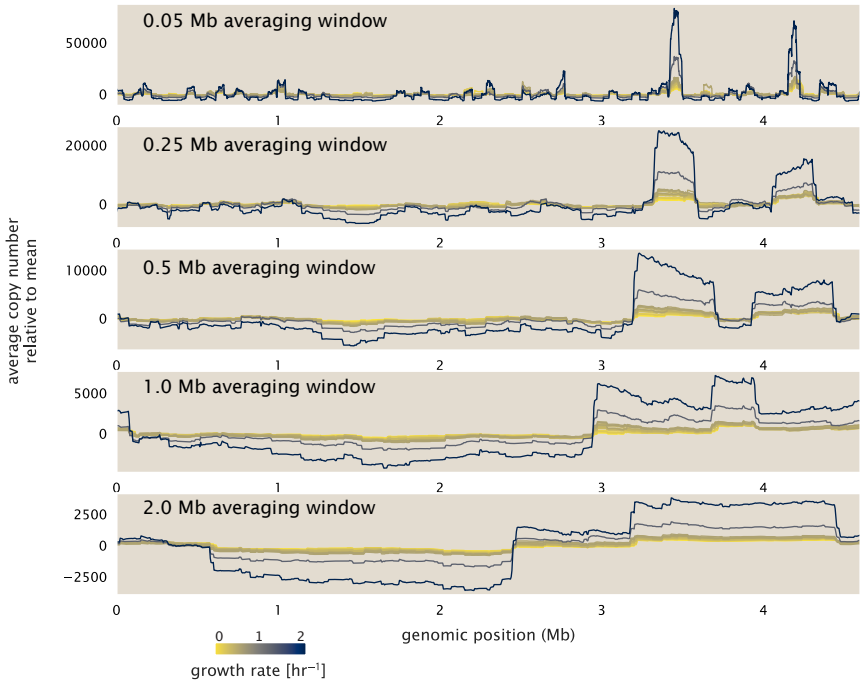

Figure 2. Position-dependent protein expression at different growth rates. (A) Protein copy number is reported along the length of the chromosome using a boxcar averaging. Here we consider different averaging window sizes: 0.05 Mb, 0.25 Mb, 0.5 Mb, 1.0 Mb, and 2 Mb.

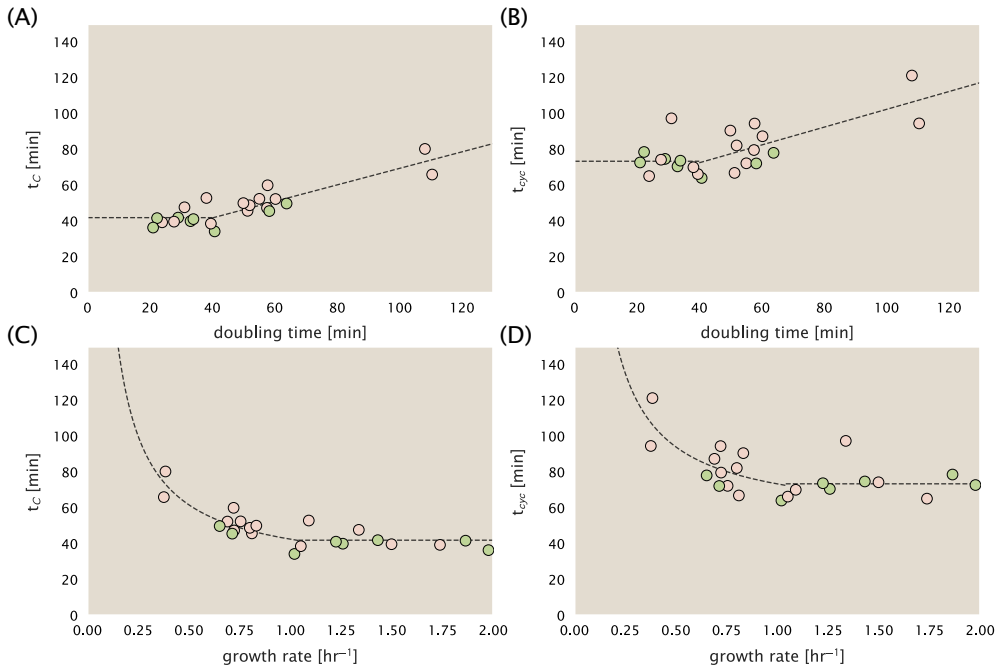

Figure 3. Estimation of ⟨#ori⟩ / ⟨# ter⟩ and ⟨#ori⟩ using data from Si *et al.* (2017). (A) and (B) plot the reported t_C and t_{cyc} as a function of cell doubling time τ, respectively. The dashed lines show a piecewise fit to the data. For short doubling times (rich media), t_C and t_{cyc} are assumed constant. At the transition, taken to occur at 40 minutes, the dashed line corresponds to an assumed proportional increase in each parameter as a function of the doubling time. (C) and (D) plot the same data as in (A) and (B), but as a function of growth rate, given by $\lambda = ln(2)/\tau$.

Hypothesis for increase in ribosomal abundance in the presence of chloramphenicol.

In the main text we note that the observed increase in ribosomal abundance upon addition of non-lethal concentrations of chloramphenicol may in part be a consequence of limiting rRNA production. Specifically, the proposal assumes that RNA polymerase are producing rRNA at their maximal rate (i.e. maximal packing of RNA polymerase on each rRNA operon). By sequestering ribosomes there will be a decrease in protein synthesis rate (i.e. lower $r_t \cdot \boldsymbol{R}$) and a corresponding increase in doubling time. Qualitatively then, we may then expect that that more rRNA (and therefore more ribosomes) may be produced for a specific growth condition and longer doubling times Figure 4(A).

Here we consider data from Si *et al.* (2017) where cells were grown in the presence of sub-lethal levels of chloramphenicol. In Figure 4(B) we plot measured RNA-to-protein ratios as a function of ⟨#ori⟩ / ⟨# ter⟩ (calculated using their reported values of τ_C and τ). While the data is relatively noisy, we do see that increasing concentrations of chloramphenicol is associated with an increased RNA-to-protein ratios and this appears roughly independent of the particular ⟨#ori⟩ / ⟨# ter⟩ ratio.

One challenge in interpreting the data is that the ⟨#ori⟩ / ⟨# ter⟩ ratio for a specific growth condition tends to decrease with increasing concentrations of chloramphenicol (indicated by marker type). Since the ⟨#ori⟩ / ⟨# ter⟩ ratio is defined by the ratio $\tau_C/\ \tau$, this is likely a reflection of chloramphenicol slowing down protein production relative to the rate of DNA replication (though, both τ_C and τ increase with added chloramphenicol).

Lastly, using the reported cell size data that was also available, we also consider how total ribosome copy number varies with growth condition and chloramphenicol. Here, as a first approaximation we assume that the total protein per cell will is proportional to cell size (with total protein ≈ cell volume x 1.1 g/ml x 30% dry mass x 55% protein). We then estimate the number of ribosomes by multiplying the protein mass by our estimate of the ribosomal fraction. Consistent with the apparent generality in how growth relates to cell size (size ∝ ⟨#ori⟩) (***Si et al., 2017***), the number of ribosomes per cell collapse onto a roughly linear trend with respect to the ⟨#ori⟩ (Figure 4(C)).

That each of the chloramphenicol curves do not collapse onto a single line when normalized relative to ⟨#ori⟩ (Figure 4(D)) may be a reflection of biosynthetic rates increasing overall in richer media. For protein translation specifically, the rate of translation increases in both nutrient-limitation (***Scott et al., 2010***), and with increasing concentrations of chloramphenicol (***Dai et al., 2016***) for poorer media (up to maximum of about 17 aa per second). It may be that other processes, and production of rRNA in particular may also slow down in poorer media.

References

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.

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.

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.

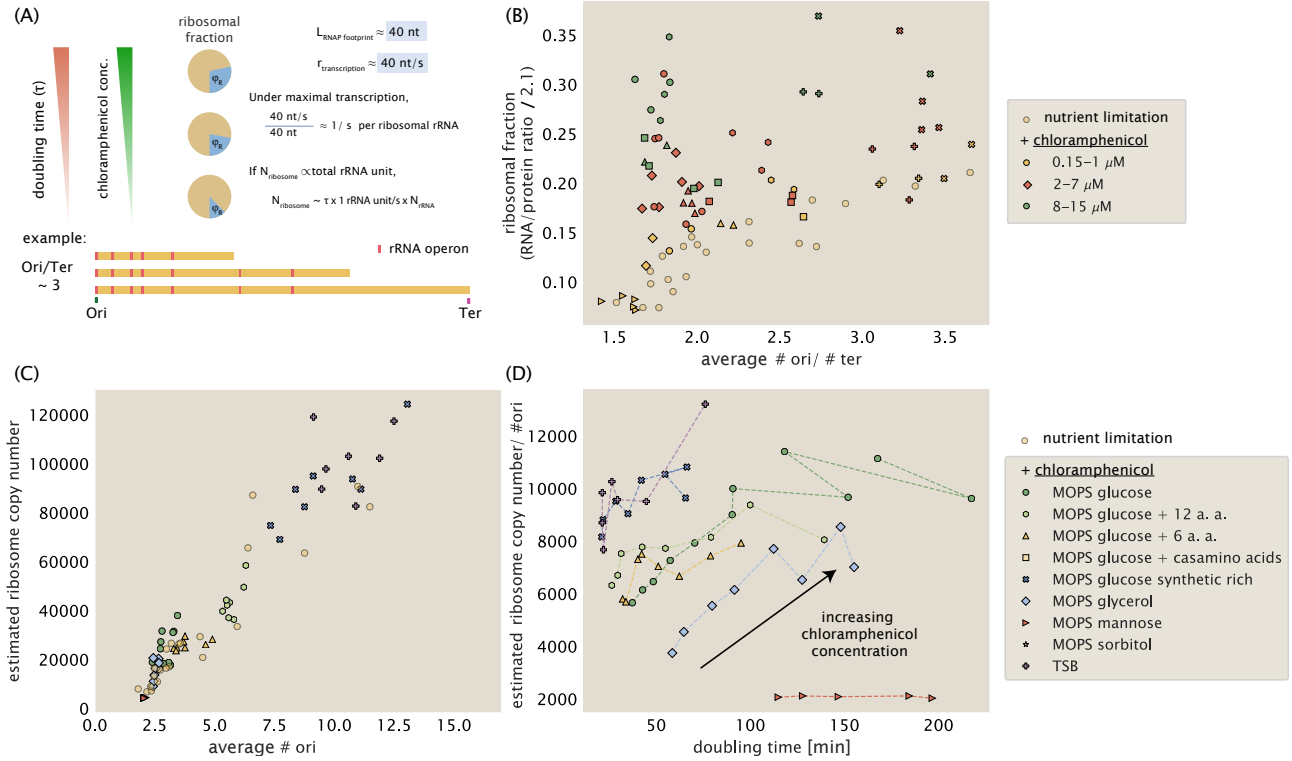

Figure 4. Potential effect of chloramphenicol on ribosomal abundance of if rRNA production is limiting. (A) Schematic of proposed change in ribosomal abundance upon addition of non-lethal does of chloramphenicol. We consider, for example, a ⟨#ori⟩ / ⟨#ter⟩ ratio of 3, which reflects an effective chromosome whose gene dosage is biased moreso to regions near the origin. If ribosome production is limited by rRNA production in particular, then sequestering ribosomes will slow down cell doubling and provide addition time for more rRNA to be made. (B) Estimated ribosomal fraction (≈ RNA/protein ratio x 2.1 ***Dai et al.*** (***2016***)) at as a function of measured ⟨#ori⟩ / ⟨# ter⟩ ratio. Data is split into 'nutrient-limited' growth (pale yellow), low chloramphenicol concentration (yellow, 0.15 - 1 μM), medium chloramphenicol concentration (red, 2 - 7 μM), and high chloramphenicol concentration (red, 8 - 15 μM). Marker type corresponds to the growth media as indicated in part (C). (C) Scaling of estimated ribosomal copy number with ⟨#ori⟩, showing that cells still scale their total protein in accord with apparent growth law (***Si et al., 2017***) irrespective of presence of chloramphenicol. (D) Estimated ribosomal copy number normalized by ⟨#ori⟩. Data shows a media-specific increase in ribosomal abundance per origin with longer doubling times. All data is from (***Si et al., 2017***), and show that average values from each growth condition and chloramphenicol concentration (including data from both strains, MG1655 and NCM3722).