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**[2, 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; [*]Contributed equally

Abstract

RNA Synthesis

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

rRNA

We begin with an estimation of the number of RNA polymerases needed to synthesize the rRNA that serve as catalytic and structural elements of the ribosome. Each ribosome contains three rRNA molecules of lengths 120, 1542, and 2904 nucleotides (BNID: 108093, ***Milo et al.*** (***2010***)), meaning each ribosome contains $\approx$ 4500 nucleotides. As the *E. coli* RNA polymerase transcribes DNA to RNA at a rate of $\approx$ 40 nucleotides per second (BNID: 101904, ***Milo et al.*** (***2010***)), it takes a single RNA polymerase $\approx$ 100 s to synthesize the RNA needed to form a single functional ribosome. Therefore, in a 5000 s division time, a single RNA polymerase transcribing rRNA at a time would result in only $\approx$ 50 functional ribosomal rRNA units – far below the observed number of $\approx 10^4$ ribosomes per cell.

Of course, there can be more than one RNA polymerase transcribing the rRNA genes at any given time. To elucidate the *maximum* number of rRNA units that can be synthesized given a single copy of each rRNA gene, we will consider a hypothesis in which the rRNA operon is completely tiled with RNA polymerase. *In vivo* measurements of the kinetics of rRNA transcription have revealed that RNA polymerases are loaded onto the promoter of an rRNA gene at a rate of $\approx$ 1 per second (BNID: 111997; 102362, ***Milo et al.*** (***2010***)). If RNA polymerases are being constantly loaded on to

the rRNA genes at this rate, then we can make the approximation that ≈ 1 functional rRNA unit is synthesized per second. With a 5000 second division time, this hypothesis leads to a maximal value of 5000 functional rRNA units, still undershooting the observed number of 10^4 ribosomes per cell.

E. coli has evolved a clever mechanism to surpass this kinetic limit for the rate of rRNA production. Rather than having only one copy of each rRNA gene, *E. coli* has seven copies of the operon (BIND: 100352, ***Milo et al.*** (***2010***)) four of which are localized directly adjacent to the origin of replication (***Birnbaum and Kaplan, 1971***). As fast growth requires that multiple copies are being synthesized simultaneously, this means that the total number of rRNA genes can be be on the order of ≈ 10 – 70 at moderate to fast growth rates (***Stevenson and Schmidt, 2004***). Using our standard time scale of a 5000 second division time, we can make the lower-bound estimate that the typical cell will have 7 copies of the rRNA operon. Synthesizing one functional rRNA unit per second per operon, a total of 4×10^4 rRNA units can be synthesized, comfortably above the observed number of ribosomes per cell.

How many RNA polymerases are then needed to constantly transcribe 7 copies of the rRNA genes? We approach this estimate by considering the maximum number of RNA polymerases can be tiling the rRNA genes with a loading rate of 1 per second and a transcription rate of 40 nucleotides per second. Considering that a RNA polymerase has a physical footprint of approximately 40 nucleotides (BNID: 107873, ***Milo et al.*** (***2010***)), we can state that there is ≈ 1 RNA polymerase per 80 nucleotides. With a total length of ≈ 4500 nucleotides per operon and 7 operons per cell, the maximum number of RNA polymerases that can be transcribing rRNA at any given time is ≈ 400. As we will see in the coming sections, the synthesis of rRNA is the dominant requirement of the RNA polymerase pool.

mRNA

To form a functional protein, all protein coding genes must first be transcribed from DNA to form an mRNA molecule. While each protein requires an mRNA blueprint, many copies of the protein can be synthesized from a single mRNA. Factors such as strength of the ribosomal binding site, mRNA stability, and rare codon usage frequency dictate the number of proteins that can be made from a single mRNA, with yields ranging from 10^1 to 10^4 (BNID: 104186; 100196; 106254, ***Milo et al.*** (***2010***)). Computing the geometric mean of this range yields ≈ 1000 proteins synthesized per mRNA, a value that agrees with experimental measurements of the number of proteins per cell (≈ 3×10^6, BNID: 100088, ***Milo et al.*** (***2010***)) and total number of mRNA per cell (≈ 3×10^3, BNID:100064, ***Milo et al.*** (***2010***)).

This estimation captures the *steady state* mRNA copy number, meaning that at any given time, there will exist approximately 3000 unique mRNA molecules. To determine the *total* number of mRNA that need to be synthesized over the cell's lifetime, we must consider degradation of the mRNA. In most bacteria, mRNAs are rather unstable with life times on the order of several minutes (BNID: 104324; 106253; 111927; 111998, ***Milo et al.*** (***2010***)). For convenience, we will assume that the typical mRNA in our cell of interest has a typical lifetime of ≈ 300 seconds. Using this value, we can determine the total mRNA production rate to maintain a steady-state copy number of 3000 mRNA per cell. While we direct the reader to the appendix for a more detailed discussion of mRNA transcriptional dynamics, we state here that the total mRNA production rate must be on the order of ≈ 15 mRNA made every second. In *E. coli*, the average protein is ≈ 300 amino acids in length (BNID: 108986; ***Milo et al.*** (***2010***)), meaning that the corresponding mRNA is ≈ 900 nucleotides which we will further approximate to be ≈ 1000 nucleotides given non-protein coding regions of the mRNA present on the 5′ and 3′ ends. This means that the cell must have enough RNA polymerase molecules about to sustain a transcription rate of ≈ 1.5×10^4 nucleotides per second. Knowing that a single RNA polymerase polymerizes RNA at a clip of 40 nucleotides per second, we arrive at a comfortable estimate of ≈ 250 RNA polymerase complexes needed to satisfy the mRNA demands of the cell. It is worth noting that this number is approximately half of that required to synthesize

enough rRNA, as we saw in the previous section. We find this to be a striking result as these 250 RNA polymerase molecules are responsible for the transcription of the ≈ 4000 protein coding genes which are not ribosome associated.

tRNA

The final class of RNA molecules worthy of quantitative consideration is the the pool of tRNAs used during translation to map codon sequence to amino acid identity. Unlike mRNA or rRNA, each individual tRNA is remarkably short, ranging from 70 to 95 nucleotides each (BNID: 109645; 102340, ***Milo et al.*** (***2010***)). What they lack in length, they make up for in abundance. There are approximately ≈ 3000 tRNA molecules present for each of the 20 amino acids (BNID: 105280, ***Milo et al.*** (***2010***)), although the precise copy number is dependent on the identity of the ligated amino acid. Using these values, we make the estimate that $\approx 5 \times 10^6$ nucleotides are sequestered in tRNA per cell. Unlike mRNA, tRNA is remarkably stable with typical lifetimes *in vivo* on the order of ≈ 48 hours (***Abelson et al., 1974***; ***Svenningsen et al., 2017***) – well beyond the timescale of division. Once again using our rule-of-thumb for the rate of transcription to be 40 nucleotides per second and assuming a division time of ≈ 5000 seconds, we arrive at an estimate of ≈ 20 RNA polymerases to synthesize enough tRNA. This requirement pales in comparison to the number of polymerases needed to generate the rRNA and mRNA pools and can be neglected as a significant transcriptional burden.

RNA Polymerase and σ-factor Abundance

These estimates, summarized in ***Figure 1*** (A), reveal that synthesis of rRNA and mRNA are the dominant forces dictating the number of RNA polymerases needed per cell. For completeness, we can use our estimates of ≈ 400, 250, and 20 RNA polymerases needed to synthesize the required number of rRNAs, mRNAs, and tRNAs, respectively, to state that the typical cell needs to maintain a pool of ≈ 700 RNA polymerases. As is revealed in ***Figure 1*** (B), this estimate is about an order of magnitude below the observed number of RNA polymerase complexes per cell (≈ 5000 - 7000). This disagreement between the estimated number of transcriptionally active RNA polymerases and these observations jibes with recent literature revealing that ≈ 80 % of RNA polymerases in *E. coli* are not transcriptionally active (***Patrick et al., 2015***). This leads us to consider other factors intimately involved in transcription may set the scale of this curious balance.

One such factor we can consider is the influence of σ-factors, namely σ^{70} (RpoD) which is the dominant "general-purpose" σ-factor in *E. coli*. While initially thought of as being solely involved in transcriptional initiation, the past two decades of single-molecule work has revealed a more multipurpose role for σ^{70} including facilitating transcriptional elongation (***Kapanidis et al., 2005***; ***Goldman et al., 2015***; ***Perdue and Roberts, 2011***; ***Mooney and Landick, 2003***; ***Mooney et al., 2005***). ***Figure 1*** (B) is suggestive of such a role as the number of σ^{70} proteins per cell is in close agreement with our estimate of the number of transcriptional complexes needed. In the appendix and supplemental figure XXX [GC: format number later], the slope of the σ^{70} abundance as a factor of the growth rate can be very accurately estimated by factoring in a) the growth-rate dependent size of the proteome and b) the rRNA gene dosage resulting from parallelized replication of the chromosome.

While these estimates and comparison with experimental data reveal an interesting dynamic at play between the transcriptional demand and copy numbers of the corresponding machinery, these findings illustrate that transcription cannot be the rate limiting step in bacterial division. ***Figure 1*** (A) reveals that the availability of RNA polymerase is not a limiting factor for cell division as there is ∼ 10-fold more complexes than needed. Furthermore, if more transcriptional activity was needed to satisfy the cellular requirements, more σ^{70}-factors could be expressed to utilize a larger fraction of the RNA polymerase pool.

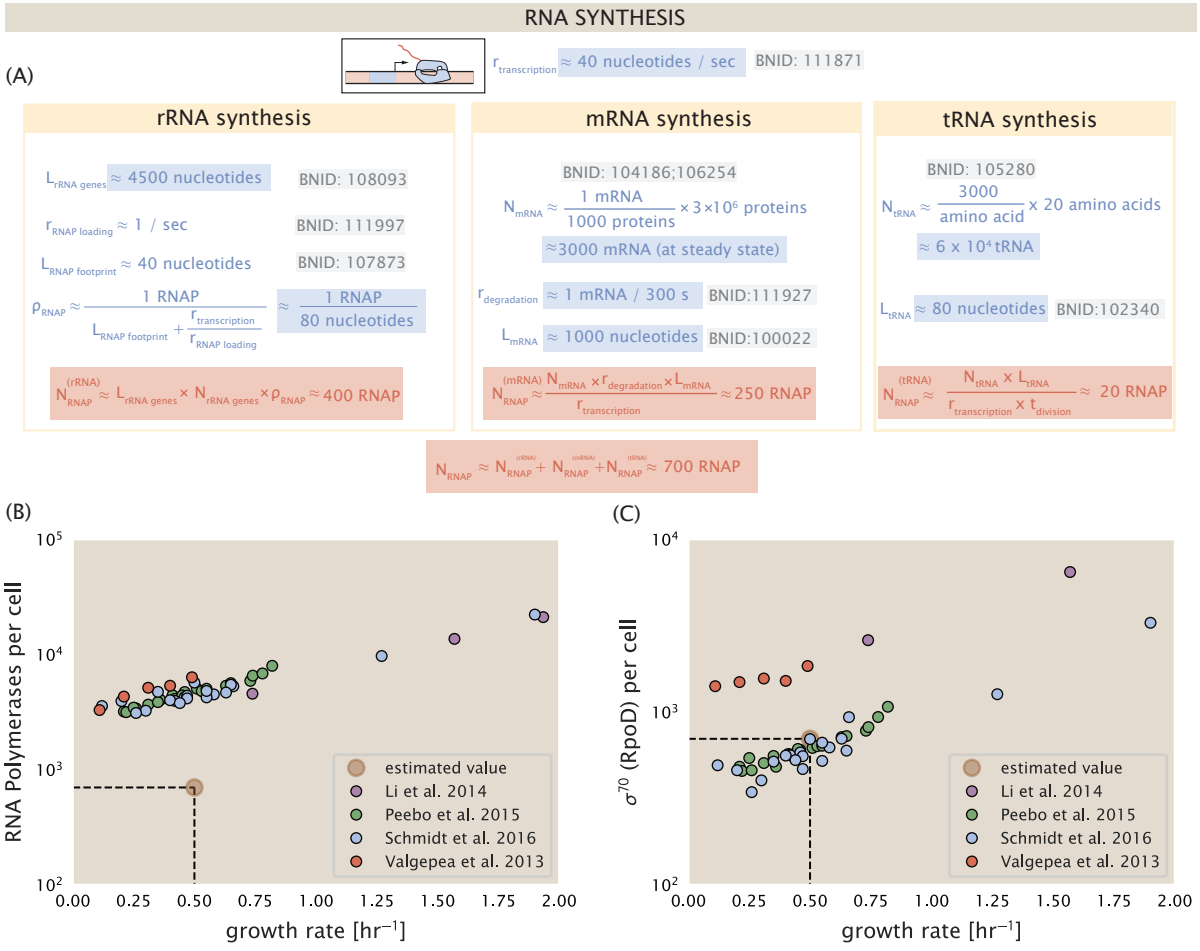

Figure 1. Estimation of the RNA polymerase demand and comparison with experimental data. (A) Estimations for the number of RNA polymerase needed to synthesize sufficient quantities of rRNA, mRNA, and tRNA from left to right, respectively. Bionumber Identifiers (BNIDs) are provided for key quantities used in the estimates. (B) The RNA polymerase core enzyme copy number as a function of growth rate. Colored points correspond to the average number RNA polymerase core enzymes that could be formed given a subunit stoichiometry of $[RpoA]_2[RpoC][RpoB]$. (C) The abundance of σ^{70} as a function of growth rate. Estimated value for the number of RNAP is shown in (B) and (C) as a translucent brown point at a growth rate of 0.5 hr^{-1}.

References

Abelson, H., Johnson, L., Penman, S., and Green, H. (1974). Changes in RNA in relation to growth of the fibroblast: II. The lifetime of mRNA, rRNA, and tRNA in resting and growing cells. *Cell*, 1(4):161–165.

Birnbaum, L. S. and Kaplan, S. (1971). Localization of a Portion of the Ribosomal RNA Genes in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 68(5):925–929.

Goldman, S. R., Nair, N. U., Wells, C. D., Nickels, B. E., and Hochschild, A. (2015). The primary σ factor in *Escherichia coli* can access the transcription elongation complex from solution *in vivo*. *eLife*, 4:e10514.

Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonjia, E., Mekler, V., Ebright, R. H., and Weiss, S. (2005). Retention of Transcription Initiation Factor Σ70 in Transcription Elongation: Single-Molecule Analysis. *Molecular Cell*, 20(3):347–356.

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.

Mooney, R. A., Darst, S. A., and Landick, R. (2005). Sigma and RNA Polymerase: An On-Again, Off-Again Relationship? *Molecular Cell*, 20(3):335–345.

Mooney, R. A. and Landick, R. (2003). Tethering Σ70 to RNA polymerase reveals high in vivo activity of σ factors and Σ70-dependent pausing at promoter-distal locations. *Genes & Development*, 17(22):2839–2851.

Patrick, M., Dennis, P. P., Ehrenberg, M., and Bremer, H. (2015). Free RNA polymerase in *Escherichia coli*. *Biochimie*, 119:80–91.

Perdue, S. A. and Roberts, J. W. (2011). σ^{70}-dependent Transcription Pausing in *Escherichia coli*. *Journal of Molecular Biology*, 412(5):782–792.

Stevenson, B. S. and Schmidt, T. M. (2004). Life History Implications of rRNA Gene Copy Number in *Escherichia coli*. *Applied and Environmental Microbiology*, 70(11):6670–6677.

Svenningsen, S. L., Kongstad, M., Stenum, T. S. n., Muñoz-Gómez, A. J., and Sø rensen, M. A. (2017). Transfer RNA is highly unstable during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Research*, 45(2):793–804.