Phenomenological Growth Laws [1]

Pankaj Mehta

February 1, 2021

In this lecture, we continue our discussion of self-replication. What are the constraints on the ability to self-replicate? How *fast* can bacteria reproduce? Here we will introduce a series of works by the group of Terrence Hwa that derive "phenomenological growth laws" relating the speed of reproduction with constraints produced by the molecular machinery that make a copy of the cell. These are the biophysical constraints placed on Von Neuman's contructor Automaton. Surprisingly, we will find that we can explain experiments without a detailed recourse to molecular mechanism!

[1] Readings: Scot et al . Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science* 2010; Scott and Hwa. Bacterial growth laws and their applications (2012) Basan et al. Overflow metabolism in E. coli results from efficient proteome allocation. *Nature* 2015

Last lecture, we introduced the basic minimum logical components that are required for creating a self-replicating machine. In this lecture, we will ask about how these components are constrained by basic physical principles. Self-replication requires constructing a new physical copy of the organism. Embedding the Von Neumann schema in a real physical world places constraints on how fast and how accurately one can reproduce. In exponential growth phase, this must be done with *limited resources*. The question then becomes are there general ways that thecell must allocate resources in order to be able to self-reproduce. In this lecture, we focus questions of speed or growth rates that are achievable.

Building upon a largely ignored current in microbiology that study bacterial physiology, these studies revitalized the idea that we can find general equations describing this "resource allocation problem" without recourse to any detailed molecular biological detail. The key insights are derived from mixing a few basic biological facts about the biological "constructor" automaton : ribosomes must produce new proteins (including themselves) and that they must also produce the proteins that import the raw ingredients from the bath of parts. Let us see how this works below.

Basic biology of translation

There is some minimal basic biology we have to understand how to proceed. This is the basic biology of translation.

Of course, Figure 1, while informative, is a ridiculous biology cartoon. It shows the ribosome working like a well-oiled, deterministic *"mecahnical machine"*. However, the ribosome is *"thermodynamic machine"* because translation is a thermodynamic/kinetic process that is governed by free-energy differences. Hence, this deterministic picture cannot be true. Furthermore, notice that this is a *nonequilibrium*

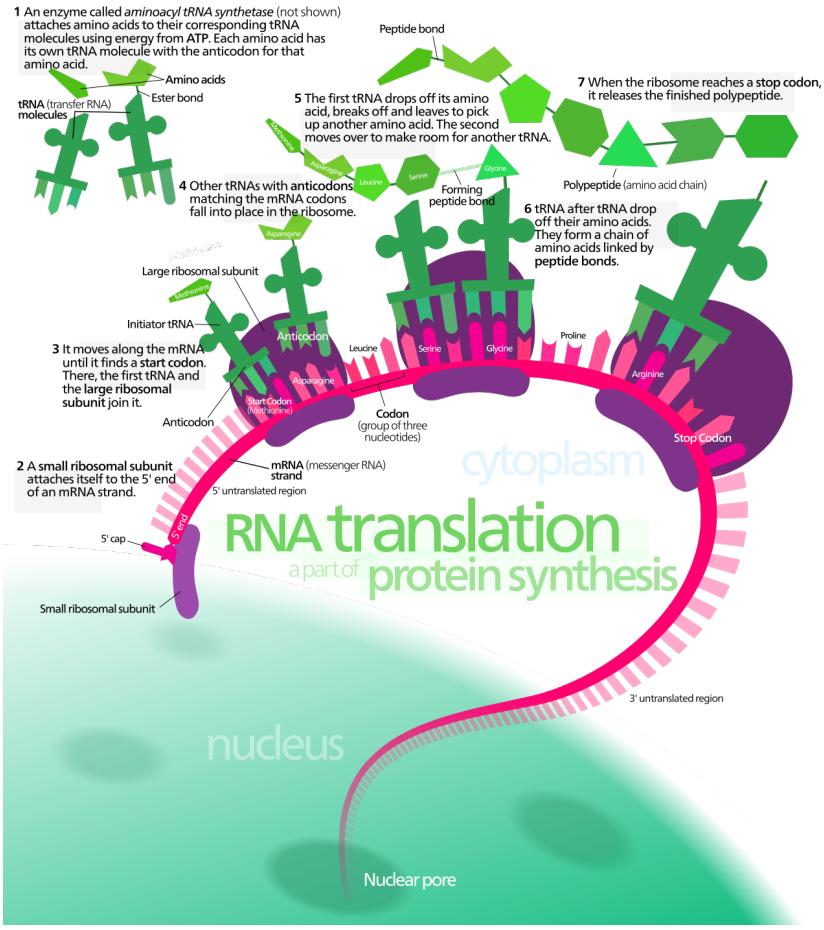

Figure 1: Overview of tranlsation. From Wikipedia under Creative Commons license `https://commons.wikimedia.org/wiki/File:Protein_synthesis.svg`.

process since it consumes energy. We will think about why it must consume energy in great detail in a few lectures. For now, it is useful to think about why would you want to consume some energy. We will return to this when we talk about molecular machines later in the class. For now, just notice the basic steps involved are coordinate by the ribosome. This is really one of the glorious structures in the world.

Basics of translation and a biophysical caveat

Translation is the process through which proteins are made from messenger RNA. This is a complicated process. The figure from Wikipedia shows the basic steps of translation.

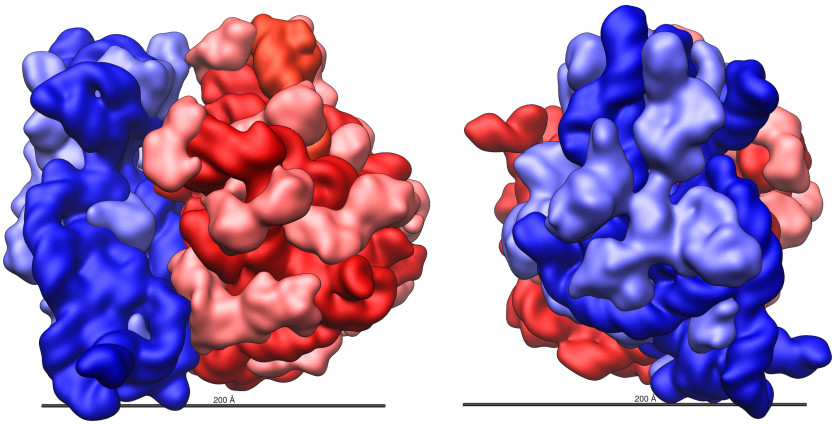

Figure 2: Ribosome has two sub-units. This figure is taken from the Wikipedia entry. By Vossman - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=6865434

The glorious ribosome

The ribosome is an elaborate factory for producing proteins. Ribosomes are about 200A in size and have a highly conserved (i.e. similar in all organisms) structure across all kingdoms of life (e.g. prokaryotes, archaea, and eukaryotes). They are composed of about 65 percent ribosomal RNA and 35 percent ribosomal proteins. These two components are highly-intertwined in the actual structure. The protein stoichometry of the ribosome is also extremely conserved [2]. Finally, we note that nearly all the RNA in an organism is ribosomal RNA (nearly 85 %).

Ribosomes are made of two parts, a small subunit and a large subunit that come together to form the complete ribosome (see Fig 2). This units have different roles

- The small subunit binds to the mRNA and binds to the large subunit.
- The large subunit bind to the small subunit and then binds to the tRNAs and amino acids.

The ribosome stays abound to the nascent mRNA until the whole protein is made. This observation (combined with clever sequencing techniques) has led to interesting new techniques such as riboprofiling [3]

[2] There has been some very interesting work challenging the idea that the protein stoichometry of the ribosomes is always the same. See N. Slavov et al Differential stoichiometry among core ribosomal proteins. *Cell Reports* 13 (2015)

[3] See in particular the work Gene-Wei Li and Jonathan Weissman. The idea is that by looking at where/how many ribosome are bound to an mRNA we can get a proxy for translational rates.

Empirical Growth Laws and their consequences

There are two levels to thinking about the resource allocation problem. The first is extremely empirical. We take some measurements, fit the data, and find relationships between the data. This empirical work requires *no theory or biological interpretation*. The second is try to

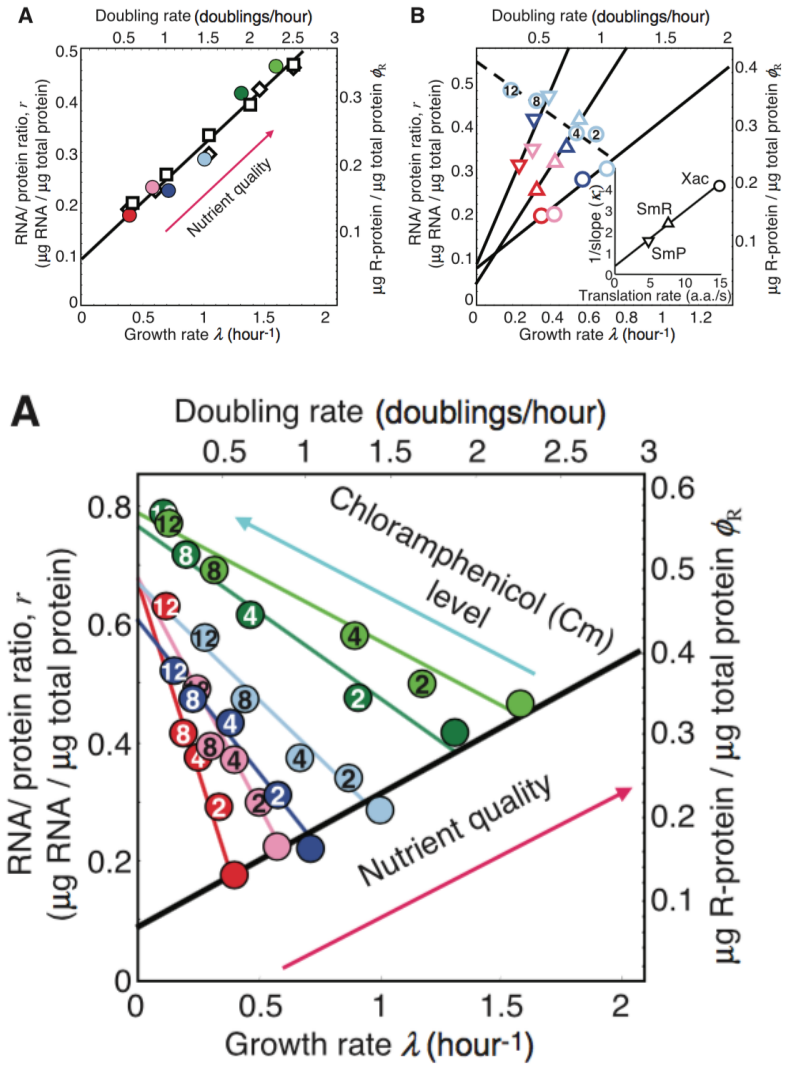

Figure 3: (Top) A. The RNA/protien content (left axis) as a function of growth rate when the media (color) is varied for a single strain EQ2 of *E. coli* (filled circle) as well as data from older experiments (square diamonds). The right vertical axis is obtained by multiplying the left-axis by the conversion factor $\rho = 0.76 \mu g$ protein$/\mu g$ RNA that converts this ratio to a mass fraction of ribosomes. B. This linear relation also holds when you use different strains (upward and downward triangles) but with a different slope that depends on the strain.(Bottom)The same strains as filled circle in the top but now subject to increasing amounts of the translating inhibiting antibiotic (number indicates concentration in μM). Notice in these different linear relations the slope changes substantially but the vertical intercepts do not seem to. Copied from Figure 1+2 from Scott et al. 2015

understand what these laws may arise by thinking about the resource allocation problem.

Growth Law 1

Scott *et al* start by plotting the RNA/protein ratio r versus growth rate when growth rate λ is mediated by changing the *media/nutrient quality*. Figure 3 top show that there is a very simple linear relation between these variables

$$r = r_0 + \frac{\lambda}{\kappa_t}, \tag{1}$$

where we have defined the slope of line by κ_t. It is perhaps more transparent to write this as

$$\lambda = \kappa_t(r - r_0). \tag{2}$$

This shows that κ_t converts the fraction of RNA (a proxy for the fraction of proteins that are in ribosomes) into a growth rate [4]. Read this way, this law states **the growth rate is directionally proportional to the ribosome content**. Furthermore, the proportionality constant depends on the details of the strain that we are using. For this reason, it is reasonable to *associate κ_t with translational efficiency*.

[4] It is actually easy to convert r into the mass fraction of ribosomes $\phi_R = M_R/M$ where M_R is the mass of ribosomes and M is the total mass of proteins. We simply need to fraction of RNA that is ribosomal (0.86 μg rRNA/ 1μ g RNA), the mass of ribosomal protein (r-proteins) per mass of ribosomal RNA (0.53 μg r-protein/1 μg rRNA), and a third factor that accounts for the fact that there are other proteins affiliated with the ribosome besides r-proteins (1.67 μg extended ribosome/ 1 μg r-proteins). Taken together we get that $\rho = 0.86 * 0.53 * 1.67 \mu g\text{protein}/\mu g\text{RNA} = 0.76 \mu g\text{protein}/\mu g\text{RNA}$. This is how the right-hand vertical axis is obtained in Figure 3.

Growth Law 2

In the second growth law (Figure 3 bottom), the growth rate is changed by adding an antibiotic chloramphenicol that effects translation. In this case, we get a different linear growth law

$$r = r_{\max} - \frac{\lambda}{\kappa_n}, \tag{3}$$

where κ_n^{-1} is the slope of this line. Again, it is helpful to re-arrange this expression to get

$$\lambda = \kappa_n(r_{\max} - r) \equiv \kappa_n \Delta r. \tag{4}$$

This law says that **as you decrease translational efficiency, the cell allocates more and more of its protein content to producing ribosomes**. In other words, it seems to be limited by the ability to translate enough proteins.

Predictions and Tests of the Growth Law

We can combine these empirical laws easily. Setting the right hand sides of equations 2 and 4 equal to each other solving for r yields

$$r = \frac{\kappa_n}{\kappa_t + \kappa_n} r_{\max} + \frac{\kappa_t}{\kappa_t + \kappa_n} r_0. \tag{5}$$

This can be substituted into Eq. 4 gives

$$\lambda = \frac{\kappa_\tau \kappa_n}{\kappa_\tau + \kappa_n}(r_{\max} - r_0) \equiv \lambda_c(\kappa_\tau)\frac{\kappa_n}{\kappa_\tau + \kappa_n}, \tag{6}$$

where in the last line we have defined $\lambda_c(\kappa_\tau)$ the maximum growth rate of strain with translational efficiency characterized by κ_τ grown in medium characterized by κ_n. The value of $\lambda_c(\kappa_\tau)$ was found to be correspond to a doubling time of 20 minutes setting an upper bound on the growth rate!! This also makes a simple prediction. If we lower

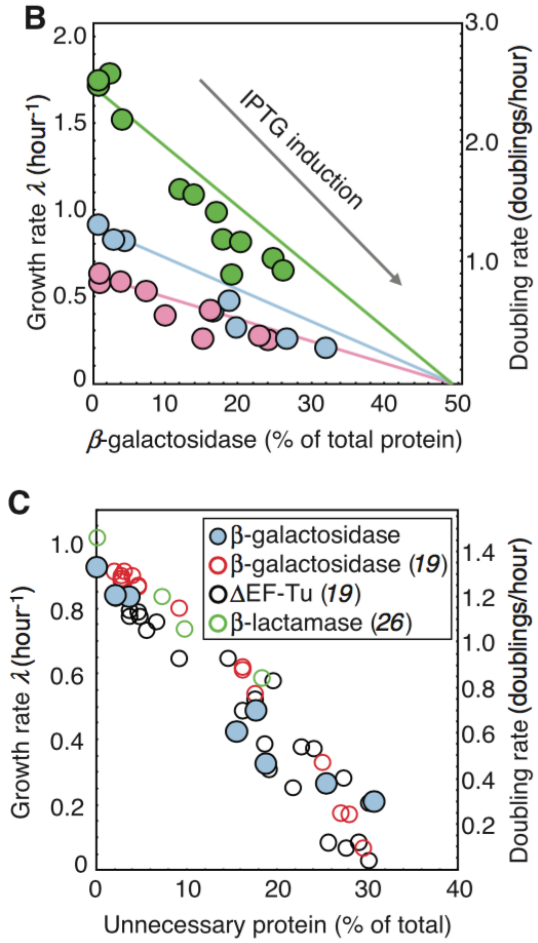

Figure 4: Test of Eq. 7 by over-expression of unnecessary proteins .

$r_{\max}$ than we will change the maximal growth rate that a bacteria can achieve.

So how can we test this. We can reduce $r_{\max}$ by forcing the cell to produce unnecessary protein. If we call this fraction $\phi_U = \rho r_u$, then we know that upon expressing these proteins $r_{\max} \to r_{\max} - r_u$. Then, we have that

$$\begin{aligned} \lambda_c(\phi_U) &= \kappa_\tau(r_{\max} - r_0 - r_u) = \kappa\tau(r_{\max} - r_0)(1 - \frac{r_u}{r_{\max} - r_0}) \\ &= \lambda_c(\phi_U = 0)(1 - \frac{\phi_U}{\phi_c}), \end{aligned} \tag{7}$$

where we have defined $\phi_c = \rho(r_{\max} - r_0) \approx 0.48$ from fits above. **The growth rate should decrease linearly with the fraction of unnecessary protein.** Thus, we have a parameter free prediction for this situation that was tested. The results are show in Figure 4 and agree remarkably!

Notice that growth law 2 says that if you make translations suffi-

ciently inefficient, growth will cease because you cannot increase the ribosomal fraction beyond some maximum $r_{\max}$. This suggests that there is some fixed fraction of proteins ϕ_Q that cannot be allocated to increasing growth with

$$\phi_Q \equiv 1 - \phi_R^{\max} = 1 - \rho r_{\max}, \tag{8}$$

(here ρ is the factor that converts RNA content to ribosomal mass content discusses above). We can also can define a variable protein content that can increase or decrease based on the amount of ribosomes as

$$\phi_P = 1 - \phi_Q - \phi_R. \tag{9}$$

Notice that we have that substituting the definition of ϕ_Q and ϕ_R that

$$\phi_P = 1 - (1 - \rho r_{\max}) - \rho r = \rho(r_{\max} - r) \tag{10}$$

Combining this with Eq. 4 gives

$$\phi_P = \rho \frac{\lambda}{\kappa_n}. \tag{11}$$

This is a clean prediction of our growth laws. **If we vary the growth rate by changing translation inhibition, then the expression of the variable protein fraction should increase linearly with growth rate, with the slope given by κ_n^{-1}.**

This prediction was tested by inserting $\beta - galactosidase$ gene under a constitutive promoter (i.e. promoter that is not regulated by any genetic or feedback mechanism) and adding antibiotic as above. Such a protein should belong to the non-essential variable component ϕ_P. As shown in Figure 5, this is exactly what happens!

Deriving the empirical growth laws from a simple resource allocation model

Having gone through the basic biology of translations and set-up the resource allocation problem, we are ready to start thinking about the resource allocation problem underlying cellular growth (self-reproduction).

Where do this empirical laws come from. Let us think about the basic ingredients

- **Collective state variables** The physiological state of the cell can be summarized in the state variables κ_t, κ_n, and ϕ_R^{X}.
- **Fixed resource fraction that cannot change** Another crucial aspect of the model that there is a fixed fraction of the protein ϕ_Q that must be made *independent* of the growth rate. This is is of course given by $1 - \phi_R^{\max}$.

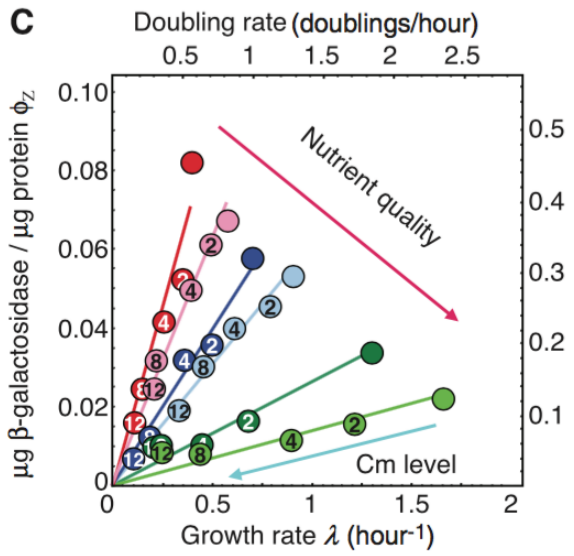

Figure 5: Expression of $\beta - galactosidase$ gene under a constitutive promoter as growth rate is varied by translation-inhibiting antibiotic. A linear response with growth rate is observed as predicted by Eq. 11.

- **Dynamic allocation between ribosomes and variable protein fraction**. Depending on the growth rate λ the cell seems to shift resources between ribosome fraction ϕ_R and the variable protein fraction ϕ_P.

Microscopic expressions for κ_t

So how can we derive expressions for κ_t and κ_n from these observations with some additional assumptions. Let us assume that the fraction of ribosomes allocated to making each kind of protein class $X = R, P, Q$ is f_X with

$$f_P + f_Q + f_R = 1. \tag{12}$$

Furthermore, let us make the key assumption that *the translational efficiency of ribosomes is independent of the protein class.* we have that

$$\frac{d}{dt} M_X = f_X k N_R, \tag{13}$$

where N_R is the number of ribosomes. Since the cells are growing exponentially with rate λ, we know that the above equations give

$$\lambda M_X = f_X k N_R, \tag{14}$$

and defining the mass of the ribosome m_R this becomes

$$\lambda M_R = f_R \lambda M = \frac{f_R k M_R}{m_R}. \tag{15}$$

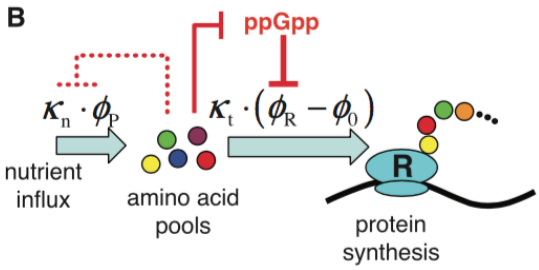

Figure 6: During steady-state exponential growth, efficient resource allocation requires that the nutrient flux $\kappa_n \phi_P$ be flux matched to the amino acid output flux $\kappa_t(\phi_R - \phi_0)$.

Rearranging we get that

$$\lambda = \frac{k}{m_R}\phi_R = \frac{k\rho}{m_R} r \tag{16}$$

Comparing with Eq. 2 (if we set $r_0 = 0$ [5]) then

[5] See HW problems

$$\kappa_t = \frac{k\rho}{m_R} \approx 20.1\,\text{aminoacids/sec} \tag{17}$$

Microscopic expressions for κ_n

To derive this, we will assume that mass must be actively imported into the cell by some enzyme E so that

$$\frac{d}{dt}M = \lambda M = c \cdot J = c k_E M_E \tag{18}$$

where c converts the input flux to mass and we have assumed the flux is proportional to total enzyme activity. $k_E M_E$. We also assume that these enzymes are a fixed fraction of the total protein content

$$\phi_E = \frac{M_E}{M} = \alpha_E \phi_P. \tag{19}$$

Then we have that

$$\lambda = c\alpha_E k_E \phi_P. \tag{20}$$

Comparing with Eq. 11 gives

$$\kappa_n = c\alpha_E k_E \rho. \tag{21}$$

This expression tells us that the fraction of proteins in ϕ_P is dynamically adjusted so that it matches the amino acid output. The putative feedbacks that make this happen are shown in Figure 6.

An integrated theory of growth

We can summarize all this in what Hwa and collaborators call an integrated theory of growth. We are given three state variables κ_t, κ_n,

and $\phi_R^{\max}$ that encodes the translational capacity, the nutritional capacity, and the maximal resource capacity. The equations that become this

$$\begin{aligned}
\phi_P + \phi_R + \phi_U &= \phi_R^{\max} && (22)\\
\lambda(\kappa_t, \kappa_n, \phi_R^{\max}, \phi_U) &= \frac{\phi_R^{\max} - \phi_0 - \phi_U}{\rho} \frac{\kappa_t \kappa_n}{\kappa_t + \kappa_n}, && (23)\\
\phi_R(\kappa_t, \kappa_n, \phi_R^{\max}, \phi_U) &= (\phi_R^{\max} - \phi_0 - \phi_U)\frac{\kappa_n}{\kappa_t + \kappa_n} + \phi_0 = \rho r && (24)\\
\phi_P(\kappa_t, \kappa_n, \phi_R^{\max}, \phi_U) &= (\phi_R^{\max} - \phi_0 - \phi_U)\frac{\kappa_t}{\kappa_t + \kappa_n} && (25)
\end{aligned}$$

which can also be rearranged to give

$$\lambda(\kappa_t, \kappa_n, \phi_R^{\max}, \phi_{UE}) = \lambda(\kappa_t, \kappa_n, \phi_R^{\max}, \phi_U = 0)(1 - \frac{\phi_{UE}}{\phi_R^{\max} - \phi_0}) \quad (26)$$

Somewhat surprisingly, these equations capture all the complexity of gene regulation and resource allocation in exponentially growing E. coli.

What have we learned about cells

- Despite all the complexity cells growth can be described by simple laws because complexity is hidden in a few global physiological variables.
- Cells re-allocate resources between importing raw parts and making ribosomes to make new proteins. This dynamic allocation shows how cells face real physical constraints in reproduction and these often determine global behaviors.
- Even without molecular mechanism, we can make precise predictions about the cells because of these physical resource allocation constraints after putting in a minimum amount of biology.

Homework

1. Explain why this derivation break down when the nutrient media is extremely poor.

2. In deriving Eq. 17, we have set $r_0 = 0$. This results from the fact that we have assumed that all ribosomes are transcribing. Show that if we assume that not all ribosomes are transcribing, that we can get $r_0 \neq 0$. In particular, explain how this changes the equations we have written down and show it does not change the microscopic interpretation of κ_t.

3. *Extending the model to dynamic resource allocation* Add?