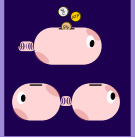


Growth Balance Analysis

Hugo Dourado

Institute for Computer Science and Department of Biology
Heinrich-Heine Universität, Düsseldorf

July 9, 2024

Inria Département MathNum EURIP LEARNING PLANET INSTITUTE

Why yet another "balance analysis"?

Growth Balance Analysis (GBA): simplified framework for nonlinear self-replicating cell models at balanced growth[1].

- **Nonlinear:** includes nonlinear kinetic rate laws.
- **Self-replicating:** metabolism + protein synthesis and dilution of **all** components.
- **Balanced growth:** constant (external and internal) concentrations in time.

A framework, not a model: find common properties to all possible models.

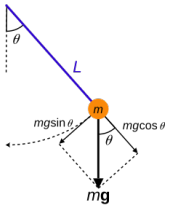
Mathematical simplification: allows analytical study to find fundamental principles.

[1] Dourado & Lercher, An analytical theory of balanced cellular growth, *Nature Communications* 2020.

Mathematical simplification: the least number of variables and equations

Not important for linear problems, but critical for nonlinear problems!

Example: Simple pendulum



Angle θ ("generalized coordinate") completely determines the system state, no need of x,y,z.

Why looking for simplest formulation?

- Easier numerical calculations.
- Independent variables are preferable for analytical methods.
- Deeper understanding of the problem.
- Most "elegant" solution.

Balanced growth (or steady growth)

For a steady-state environment defined by "external" concentrations $\mathbf{x}$:

- Steady-state growth rate μ (1/h), direct measure of fitness.
- Steady-state internal concentrations $\mathbf{c}$ (g/L) of reactants (substrates, products)

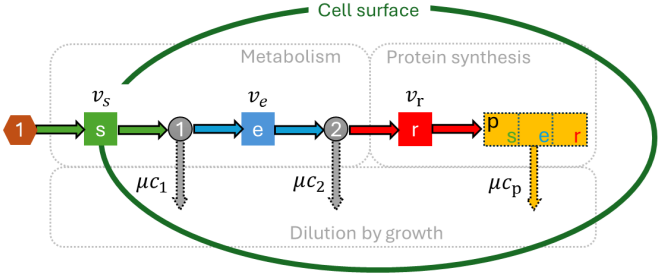
$$c_i = \frac{\text{abundance of "i" (g/cell)}}{\text{volume (L/cell)}} = \text{constant}$$

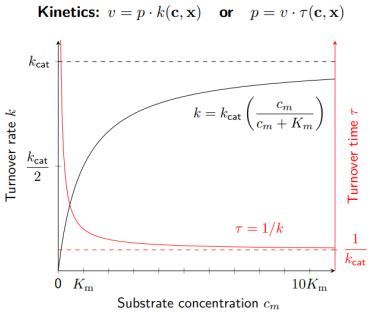
Mass concentrations (not abundances) better describe cell states: i) constant, ii) reaction kinetics depend on concentrations, iii) relate to cell density (g/L).

Matching units for fluxes $\mathbf{v}$: mass per volume per time (g L^{-1} h^{-1}).

"Self-replicator" models: Molenaar et al. *Mol Sys Biol* 2009

"Self-replicating": nonlinear kinetics for transport + metabolism + protein synthesis, with dilution by growth of **all** components (no biomass input, now it is an output).





Optimal state: maximize μ limited by mass conservation, kinetics and total protein.

Self-replicating models must be nonlinear: saturation/dilution trade-off.

Density constraint(s)

Linear models: fixed density of "biomass" (diffuse concept), total protein c_{p}.

Self-replicator models [a.k.a. Molenaar models]: fixed total protein c_{p}.

GBA: fixed cell density ρ (g/L) including **all** components (indicated by experiments[2])

$$\rho = c_{\mathrm{p}} + \sum_m c_m$$

where m are all "non-protein" components, and assumed unique protein composition[3].

GBA units: mass concentration (g/L) is the most convenient unit. To match units, we normalize the stoichiometric matrix $\mathbf{S}$ with the molecular weights $\mathbf{w}$ (g/mol)

$$\mathbf{S}^{\mathrm{total}} \xrightarrow[\text{columns by } \mathbf{w}]{\text{multiply}} \mathrm{diag}(\mathbf{w})\,\mathbf{S}^{\mathrm{total}} \xrightarrow[\text{columns}]{\text{normalize}} \mathbf{M}^{\mathrm{total}} \xrightarrow[\text{external rows}]{\text{exclude}} \mathbf{M}$$

[2] Baldwin et al. *Archives of Microbiology* 1995, Kubitschek et al. *Journal of bacteriology* 1983, Cayley et al. *Journal of Molecular Biology* 1991

[3] Dourado et al. *PLOS Comp Bio* 2023.

The "mass fraction matrix" $\mathbf{M}$

$$\mathbf{M} = \begin{array}{c} \\ i \end{array} \left\{ \begin{array}{c} \overbrace{}^{s} \overbrace{}^{e} \overbrace{\phantom{M_{\mathrm{r}}^m}}^{\mathrm{r}} \\ \left[\begin{array}{ccc} M_s^m & M_e^m & M_{\mathrm{r}}^m \\ \hline 0 & 0 & M_{\mathrm{r}}^{\mathrm{p}} \end{array} \right] \\ \underbrace{}_{j} \end{array} \begin{array}{l} \\ \} m \\ \mathrm{p} \end{array} \right.$$

j = reactions (transport s, enzymatic e, ribosome r)

i = internal reactants ("metabolites" m, protein p)

(upper index = rows, internal reactants)

(lower index = columns, reactions)

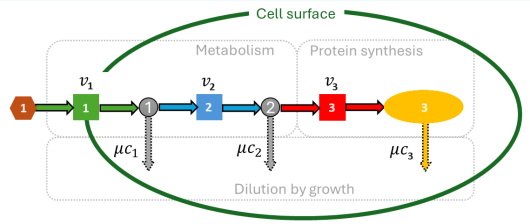
"Mass accretion" vector $\mathbf{a}$: the sum of each column

$$a_j \coloneqq \sum_i M_j^i \quad = \begin{cases} \neq 0 & , j = s \\ = 0 & , j = e \\ = 0 & , j = \mathrm{r} \end{cases} \qquad \Leftrightarrow \qquad \mathbf{w}^\top \mathbf{S}^{\mathrm{total}} = \mathbf{0}^\top$$

Example of a GBA model: "model A"

Three Michaelis-Menten reactions in series

A) Model scheme

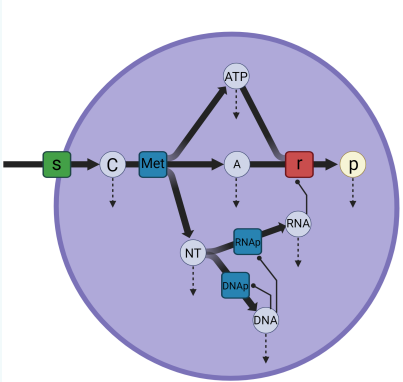


B) Model parameters

$$\mathbf{M} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}, \quad \mathbf{K} = \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 22 & 0 \\ 0 & 0 & 40 \\ 0 & 0 & 0 \end{bmatrix} \end{array} \begin{array}{c} \\ 1 \\ 1 \\ 2 \\ 3 \end{array}, \quad \mathbf{k}_{\text{cat}} = \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 6 & 6 & 5 \end{bmatrix} \end{array}, \quad \rho = 340\ g/L$$

Example of a GBA model: "model B"

Model with replication, transcription, Michaelis-Menten kinetics with activation



$$\mathbf{M} = \begin{array}{c} \begin{matrix} s & Met & DNAp & RNAp & r \end{matrix} \\ \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 & -0.3 \\ 0 & 0.3 & -1 & -1 & 0 \\ 0 & 0.5 & 0 & 0 & -0.7 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} C \\ ATP \\ NT \\ A \\ DNA \\ RNA \\ p \end{matrix}$$

$$\mathbf{K} = \begin{array}{c} \begin{matrix} s & Met & DNAp & RNAp & r \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 11 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 \\ 0 & 0 & 10 & 10 & 0 \\ 0 & 0 & 0 & 0 & 25 \end{bmatrix} \end{array} \begin{matrix} C_{ex} \\ C \\ ATP \\ NT \\ A \end{matrix}$$

$$\mathbf{A} = \begin{array}{c} \begin{matrix} s & Met & DNAp & RNAp & r \end{matrix} \\ \begin{bmatrix} 0 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \end{array} \begin{matrix} DNA \\ RNA \end{matrix}$$

$$\mathbf{k}_{\text{cat}} = \begin{array}{c} \begin{matrix} s & Met & DNAp & RNAp & r \end{matrix} \\ \begin{bmatrix} 13 & 24 & 7 & 17 & 4 \end{bmatrix} \end{array}, \ \rho = 340$$

The general GBA problem

For some given model $(\mathbf{M}, \boldsymbol{\tau}, \rho)$ and environment $\mathbf{x}$:

$$
\begin{aligned}
&\underset{\mathbf{v} \in \mathbb{R}^{\mathrm{r}}, \mathbf{c} \in \mathbb{R}^{\mathrm{p}}}{\text{maximize}} && \mu && \text{(Maximize growth rate)} \\
&\text{subject to:} \\
& && \mathbf{M}\,\mathbf{v} = \mu\,\mathbf{c} && \text{(Flux balance)} \\
& && c_{\mathrm{p}} = \mathbf{v}^{\top} \boldsymbol{\tau}(\mathbf{c}, \mathbf{x}) && \text{(Reaction kinetics and protein sum)} \\
& && \rho = \sum \mathbf{c} && \text{(Constant cell density)} \\
& && \mathbf{v} \odot \boldsymbol{\tau}(\mathbf{c}, \mathbf{x}) \geq \mathbf{0} && \text{(non-negative protein concentrations)} \\
& && \mathbf{c} \geq \mathbf{0} && \text{(non-negative reactant concentrations)}
\end{aligned}
$$

where $\odot$ indicates the element-wise multiplication.

Approximated problem: maximization of μ with global mass conservation

The net mass uptake: enforced by the sum all equations in $\mathbf{M}\,\mathbf{v} = \mu\,\mathbf{c}$

$$v_{\text{uptake}} = v_{\text{in}} - v_{\text{out}} = \sum_{i,j} M_j^i\, v^j = \mu \sum_i c_i = \mu\,\rho \quad ,$$

thus,

$$\mu(\mathbf{v}, \mathbf{c}) = \frac{\sum_{i,j} M_j^i\, v^j}{\rho(\mathbf{c})} \quad .$$

For any given $\mathbf{v}$:

$$\text{maximal } \mu(\mathbf{c}) \Leftrightarrow \text{minimal } \rho(\mathbf{c}) = c_{\text{p}} + \textstyle\sum_m c_m$$

Accounting for kinetics and protein sum:

$$\rho(\mathbf{c}) = \sum_j v_j\, \tau_j(\mathbf{c}) + \sum_m c^m$$

Approximated problem: optimal density ρ

The optimal state must satisfy

$$\boxed{\frac{\partial \rho}{\partial c^m} = \sum_j v_j \frac{\partial \tau_j}{\partial c^m} + 1 = 0 \quad \forall \ m}$$

Economics analogy: marginal cost from kinetic benefit and density cost

marginal cost = marginal kinetic benefit + marginal density cost (= 0 if optimal)

Note: the kinetic benefit is the protein saved due to increased saturation (< 0)

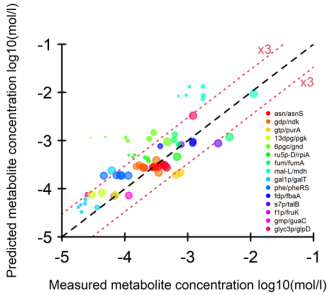
$$\sum_j v_j \frac{\partial \tau_j}{\partial c^m} = \sum_j \left(\frac{\partial p_j}{\partial c^m} \right)_{\mathbf{v}=const.}$$

Optimal density: validation

Simplest case: Michaelis-Menten kinetics and a 1-to-1 reaction-substrate relationship

$$p_j = c^m \left(1 + \frac{c^m}{K_j^m}\right) \tag{1}$$

E. coli enzymes and substrates are close to this optimality[4]



[4] Dourado et al. On the optimality of the enzyme–substrate relationship in bacteria, *PLOS Biology* 2021.

The GBA problem with no alternative pathways: full column rank $\mathbf{M}$

For some given model $(\mathbf{M}, \boldsymbol{\tau}, \rho)$ and environment $\mathbf{x}$:

$$\underset{\mathbf{v} \in \mathbb{R}^{\mathrm{r}}, \mathbf{c} \in \mathbb{R}^{\mathrm{p}}}{\text{maximize}} \quad \mu \qquad \text{(Maximize growth rate)}$$

subject to:

$$\mathbf{M}\,\mathbf{v} = \mu\,\mathbf{c} \qquad \text{(Flux balance)}$$

$$c_{\mathrm{p}} = \mathbf{v}^{\top}\boldsymbol{\tau}(\mathbf{c}, \mathbf{x}) \qquad \text{(Reaction kinetics and protein sum)}$$

$$\rho = \sum \mathbf{c} \qquad \text{(Constant cell density)}$$

~~$\mathbf{v} \odot \boldsymbol{\tau}(\mathbf{c}, \mathbf{x}) \geq \mathbf{0}$~~ ~~(non-negative protein concentrations)~~

$$\mathbf{c} \geq \mathbf{0} \qquad \text{(non-negative reactant concentrations)}$$

Simplest case: there is a inverse $\mathbf{W} = \mathbf{M}^{-1}$, so:

$$\mathbf{v} = \mu\,\mathbf{W}\,\mathbf{c}$$

The GBA problem with no alternative pathways: formulation on $\mathbf{c}$

For invertible $\mathbf{M}$: formulation on $\mathbf{c}$ in few steps

Substituting $\mathbf{v} = \mu\,\mathbf{W}\,\mathbf{c}$ into $c_{\mathrm{p}} = \mathbf{v}^{\top}\boldsymbol{\tau}(\mathbf{c},\mathbf{x})$

$$c_{\mathrm{p}} = \mu\,(\mathbf{W}\,\mathbf{c})^{\top}\,\boldsymbol{\tau}(\mathbf{c},\mathbf{x}) \quad .$$

Solving for μ: we get the objective function $\mu(\mathbf{c},\mathbf{x})$

$$\boxed{\mu(\mathbf{c},\mathbf{x}) = \frac{c_{\mathrm{p}}}{(\mathbf{W}\,\mathbf{c})^{\top}\,\boldsymbol{\tau}(\mathbf{c},\mathbf{x})}} \quad .$$

The only constraint left:

$$\boxed{\rho = \sum \mathbf{c}} \quad .$$

The problem is now completely reformulated on $\mathbf{c}$ as independent* variables

The GBA problem with no alternative pathways: analytical "solution"

Reformulated problem: for some given model $(\mathbf{M}, \boldsymbol{\tau}, \rho)$ and environment $\mathbf{x}$

$$\underset{\mathbf{c} \in \mathbb{R}^{\mathrm{p}}_{+}}{\text{maximize}} \quad \mu(\mathbf{c}, \mathbf{x}) = \frac{c_{\mathrm{p}}}{(\mathbf{W}\,\mathbf{c})^{\top} \boldsymbol{\tau}(\mathbf{c}, \mathbf{x})}$$

subject to:

$$\sum \mathbf{c} = \rho \quad .$$

Analytical conditions for optimal states: using Lagrange multipliers, we find

$$\boxed{\mu\,(\mathbf{W}\,\mathbf{c})^{\top} \frac{\partial \boldsymbol{\tau}}{\partial c_m} + \mu\,\boldsymbol{\tau}^{\top}\,(\mathbf{W}_m - \mathbf{W}_{\mathrm{p}}) + 1 = 0 \quad \forall\ m} \tag{2}$$

With $\sum \mathbf{c} = \rho$: $\mathrm{p} - 1$ algebraic equations on $\mathrm{p} - 1$ variables (solvable).

The GBA problem with no alternative pathways: economics

Substituting $\mathbf{v} = \mu\,\mathbf{W}\,\mathbf{c}$ into the solution (2)

$$\frac{c_\mathrm{p}}{\mu}\frac{\partial \mu}{\partial c^m} = -\sum_j v_j \frac{\partial \tau^j}{\partial c^m} + \mu \sum_j \tau_j \left(W_\mathrm{p}^j - W_m^j\right) - 1 = 0 \quad \forall\, m$$

Economics analogy: new "structural" marginal benefit, the proportional decrease in protein allocation

(marginal) value = kinetic benefit + protein allocation benefit + density cost (= 0 if optimal)

Because of $\mathbf{v} = \mu\,\mathbf{W}\,\mathbf{c}$ and $c_\mathrm{p} = \rho - \sum_m c^m$: increasing c^m also causes a protein allocation decrease

$$-\sum_j \left(\frac{\partial p^j}{\partial c^m}\right)_{\boldsymbol{\tau},\,\mu=const.} = -\sum_j \tau_j \frac{\partial v^j}{\partial c^m} = \underbrace{\mu \sum_j \tau_j\, W_\mathrm{p}^j}_{\text{protein production}} - \underbrace{\mu \sum_j \tau_j\, W_m^j}_{\text{metabolite production}}$$

This contribution is typically very low[a], around 0.03 (explains why predictions of eq.(1) are good).

[a] Dourado, Quantitative principles of optimal cellular resource allocation, *PhD Thesis* 2020.

Equations for balance growth states: model A

1) Original problem: Implicit constraints on μ, involving $v_1, v_2, v_3, c_1, c_2, c_3, x_1$ (6 variables, 5 equations)

$$
\begin{aligned}
v_1 - v_2 &= \mu\, c_1 \\
v_2 - v_3 &= \mu\, c_2 \qquad \text{(mass conservation)} \\
v_3 &= \mu\, c_3 \\
\frac{v_1}{6}\left(1+\frac{1}{x_1}\right) + \frac{v_2}{6}\left(1+\frac{22}{c_1}\right) + \frac{v_3}{5}\left(1+\frac{40}{c_2}\right) &= c_3 \qquad \text{(kinetics and protein sum)} \\
c_1 + c_2 + c_3 &= 340 \qquad \text{(constant cell density)}
\end{aligned}
$$

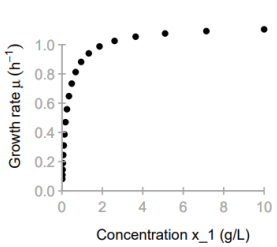
2) GBA: Explicit constraint on $\mu(c_1, c_2, x_1)$ (using $c_3 = 340 - c_1 - c_2$)

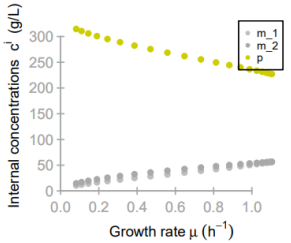
$$
\mu(c_1, c_2, x_1) = \frac{340 - c_1 - c_2}{\frac{1}{6}\left(1+\frac{1}{x_1}\right) + \frac{340-c_1}{6\cdot 340}\left(1+\frac{22}{c_1}\right) + \frac{340-c_1-c_2}{5\cdot 340}\left(1+\frac{40}{c_2}\right)} \qquad \text{(constrained growth rate)}
$$

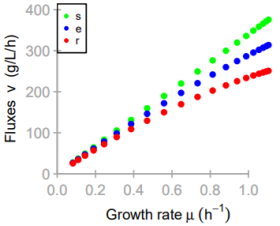
3) Analytical conditions for optimal balanced growth state (system of algebraic equations)

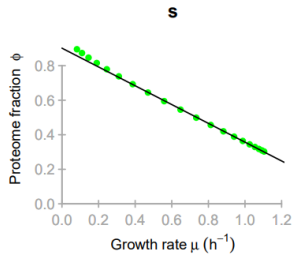
$$
\begin{aligned}
\mu\frac{22}{6}\frac{(340-c_1)}{(c_1)^2} + \mu\left[\frac{1}{6}\left(1+\frac{22}{c_1}\right) + \frac{1}{5}\left(1+\frac{40}{c_2}\right)\right] - 1 = 0 \quad (m=1) \\
\mu\,\frac{40\,(340-c_1-c_2)}{5\,(c_2)^2} + \mu\left[\frac{1}{5}\left(1+\frac{40}{c_2}\right)\right] - 1 = 0 \quad (m=2)
\end{aligned}
$$

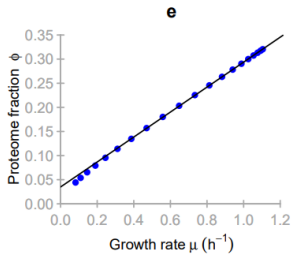
Numerical solutions for different external concentrations x: model A

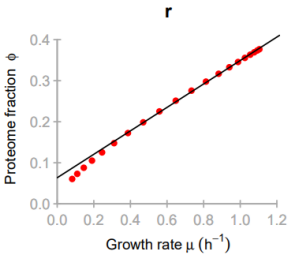






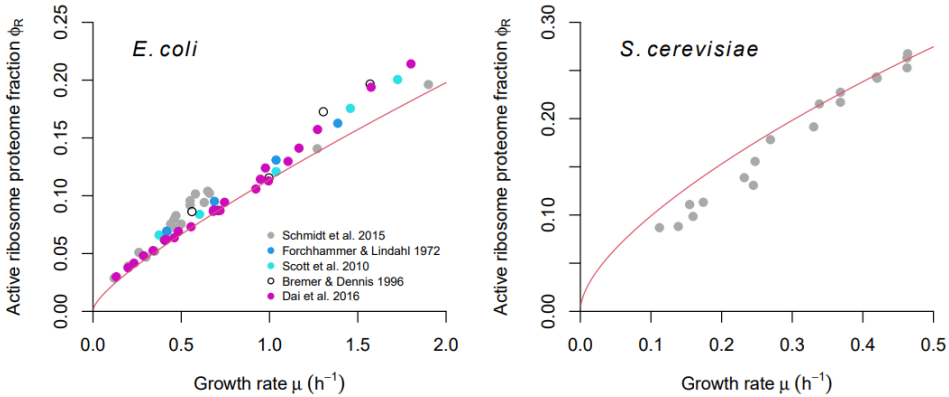






Comparison to data: *E. coli* and yeast ribosomal protein ϕ_{r} vs. μ

in vivo data close to the predicted optimality[5] (red lines, no fitting).



[5] Dourado & Lercher, An analytical theory of balanced cellular growth, *Nature Communications* 2020.

"Growth Control Analysis": holistic view of the growing cell

Metabolic Control Analysis (MCA): perturbations on metabolism (open system).

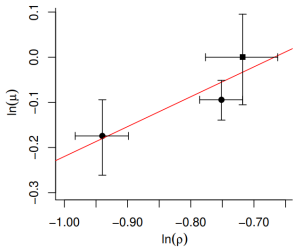
Growth Control Analysis (GCA): perturbations self-replicating system (closed system), all is connected ⇒ analytical expressions[6].

- Growth Control Coefficients Γ: change in μ by perturbing one concentration c_i.
- Growth Adaptation Coefficients A: change in optimal μ^* by changing parameters.

E.g.: changing in the density ρ

$$A_\rho = \frac{\rho}{\mu^*}\frac{\mathrm{d}\mu^*}{\mathrm{d}\rho} = \frac{\rho}{c_\mathrm{p}}\left(1 - \mu\sum_j \tau_j W_\mathrm{p}^j\right)$$

Comparing to *E. coli* data[7] →



[6] Dourado & Lercher, *Nature Communications* 2020, [7] Cayley et. al, *Biophys. J* 2000

The general GBA problem: formulation on $\mathbf{f}$

For some given model $(\mathbf{M}, \boldsymbol{\tau}, \rho)$ and environment $\mathbf{x}$:

$$\underset{\mathbf{v} \in \mathbb{R}^{\mathrm{r}}, \mathbf{c} \in \mathbb{R}^{\mathrm{p}}}{\text{maximize}} \quad \mu \qquad \text{(Maximize growth rate)}$$

subject to:

$$\mathbf{M}\,\mathbf{v} = \mu\,\mathbf{c} \qquad \text{(Flux balance)}$$

$$c_{\mathrm{p}} = \mathbf{v}^{\top}\boldsymbol{\tau}(\mathbf{c}, \mathbf{x}) \qquad \text{(Reaction kinetics and protein sum)}$$

$$\rho = \sum \mathbf{c} \qquad \text{(Constant cell density)}$$

$$\mathbf{v} \odot \boldsymbol{\tau}(\mathbf{c}, \mathbf{x}) \geq \mathbf{0} \qquad \text{(non-negative protein concentrations)}$$

~~$\mathbf{c} \geq \mathbf{0}$~~ ~~(non-negative reactant concentrations)~~

Main trick for analytical "solution": let's define the "flux fractions"

$$\mathbf{f} := \frac{\mathbf{v}}{\mu\,\rho} \quad \left(= \frac{\mathbf{v}}{v_{\text{uptake}}}, \text{kind of "mass yield" w.r.t. net mass uptake, adimensional} \right)$$

The general GBA problem: formulation on $\mathbf{f}$

General formulation on $\mathbf{f}$ in few steps

Substituting $\mathbf{v} = \mu\,\rho\,\mathbf{f}$ into $\mathbf{M}\,\mathbf{v} = \mu\,\mathbf{c}$

$$\rho\,\mathbf{M}\,\mathbf{f} = \mathbf{c} \quad \text{(independent of } \mu\text{)}.$$

Substituting $\mathbf{c} = \rho\,\mathbf{M}\,\mathbf{f}$ into $c_{\mathrm{p}} = \mathbf{v}^{\top}\boldsymbol{\tau}(\mathbf{c}, \mathbf{x})$

$$M_{\mathrm{r}}^{\mathrm{p}} f_{\mathrm{r}} = \mu\,\mathbf{f}^{\top}\boldsymbol{\tau}(\rho\,\mathbf{M}\,\mathbf{f}, \mathbf{x})$$

Solving for μ:

$$\boxed{\mu(\mathbf{f}, \mathbf{x}) = \frac{M_{\mathrm{r}}^{\mathrm{p}} f_{\mathrm{r}}}{\mathbf{f}^{\top}\boldsymbol{\tau}(\rho\,\mathbf{M}\,\mathbf{f}, \mathbf{x})}}$$

The density constraint:

$$\rho = \sum \mathbf{c} \quad \Leftrightarrow \quad \boxed{\mathbf{a}^{\top}\mathbf{f} = 1}$$

The general GBA problem: analytical "solution"

Reformulated problem: for some given model $(\mathbf{M}, \boldsymbol{\tau}, \rho)$ and environment $\mathbf{x}$

$$\underset{\mathbf{f} \in \mathbb{R}^{\mathrm{r}}}{\text{maximize}} \quad \mu(\mathbf{f}, \mathbf{x}) = \frac{M_{\mathrm{r}}^{\mathrm{p}}\, f^{\mathrm{r}}}{\mathbf{f}^{\top} \boldsymbol{\tau}(\rho\, \mathbf{M}\, \mathbf{f}, \mathbf{x})}$$

subject to:

$$\mathbf{a}^{\top} \mathbf{f} = 1$$
$$\mathbf{f} \odot \boldsymbol{\tau}(\rho\, \mathbf{M}\, \mathbf{f}, \mathbf{x}) \geq \mathbf{0} \quad .$$

Analytical conditions for optimal states: using KKT conditions, we find

$$\boxed{\left(M_j^{\mathrm{p}} - \mu\, \tau_j - \mu\, \mathbf{f}^{\top} \frac{\partial \boldsymbol{\tau}}{\partial f^j} + \mu\, \mathbf{f}^{\top} \frac{\partial \boldsymbol{\tau}}{\partial \mathbf{f}}\, \mathbf{f}\, a_j \right) f_j = 0 \quad \forall\, j} \tag{3}$$

Using $\mathbf{a}^{\top} \mathbf{f} = 1$: we have $\mathrm{r} - 1$ algebraic equations on $\mathrm{r} - 1$ variables (solvable).

The general GBA problem: economics

Substituting $\mathbf{v} = \mu\,\rho\,\mathbf{f}$ and $\mathbf{c} = \rho\,\mathbf{M}\,\mathbf{f}$ into the solution (3), we find[7]

$$M_j^{\mathrm{p}} - \mu\,\tau_j - \mathbf{v}^\top \frac{\partial \boldsymbol{\tau}}{\partial \mathbf{c}} \mathbf{M}_j + a_j\,\mathbf{v}^\top \frac{\partial \boldsymbol{\tau}}{\partial \mathbf{c}}\,\mathbf{c}/\rho = 0 \quad \forall\ j$$

Economics analogy: the marginal value of each flux fraction f_j

protein production benefit (increased protein production) + protein cost (protein in j) + kinetic benefit (protein saved) + biomass production benefit (increased biomass production) (= 0 if opt.)

[7] Dourado et al. Mathematical properties of optimal fluxes in cellular reaction networks at balanced growth, *PLOS Comp Biol* 2023.

Grow Control Analysis: Grow Adaptation Coefficient for k_{cat}

We can show from first principles (using the Envelope Theorem)[8] that:

$$A_{k_{\text{cat}}^j} = \frac{k_{\text{cat}}^j}{\mu^*} \frac{\mathrm{d}\mu^*}{\mathrm{d}k_{\text{cat}}^j} = \phi_j$$

Proportional change in μ^* is exactly the same as proportion of protein allocated to j.

[8] Dourado et al. Mathematical properties of optimal fluxes in cellular reaction networks at balanced growth, *PLOS Comp Biol* 2023.

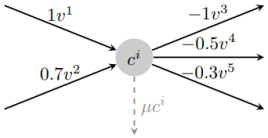
Summary

- **GBA: self-replicating models on independent variables, easier to study.**
- **Analytical conditions for optimal balanced growth (fundamental principles).**
- **Experimental indications that cells do implement near optimal strategies.**
- **Proteins emerge as the "currency" in cell economics from first principles.**

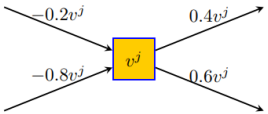
(soon chapter in the EPCB book)

Constraints on GBA

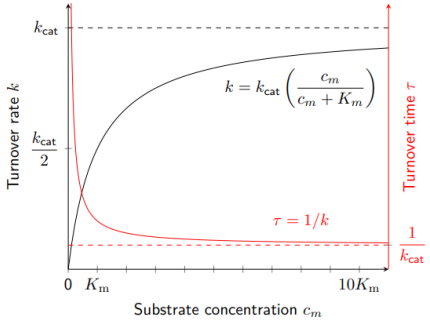
A) Flux balance for each reactant "i"



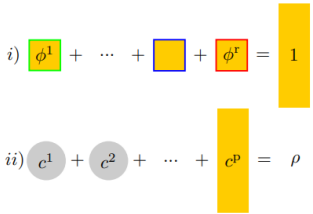
B) Mass balance within each reaction "j"



C) Kinetics: $v = p \cdot k(\mathbf{c}, \mathbf{x})$ **or** $p = v \cdot \tau(\mathbf{c}, \mathbf{x})$



D) Density constraints



Michaelis-Menten kinetics with activation

Based on "Convenience kinetics"[9], we define the Michaelis-Menten kinetics with activation, corresponding "activation constants" $\mathbf{A}$

$$\tau_j = \frac{1}{k_{\text{cat}}^j} \prod_m \left(1 + \frac{A_j^m}{c^m}\right) \left(1 + \frac{K_j^m}{c^m}\right) \prod_n \left(1 + \frac{K_j^n}{x^n}\right)$$

[9] Liebermeister & Klipp, Bringing metabolic networks to life: convenience rate law and thermodynamic constraints, 2006.

Equations for balance growth states: model A

1) Original problem: Implicit constraints on μ, involving $v_1, v_2, v_3, c_1, c_2, c_3, x_1$ (6 variables, 5 equations)

$$
\begin{aligned}
v_1 - v_2 &= \mu\, c_1 \\
v_2 - v_3 &= \mu\, c_2 \qquad \text{(mass conservation)} \\
v_3 &= \mu\, c_3 \\
\frac{v_1}{6}\left(1+\frac{1}{x_1}\right) + \frac{v_2}{6}\left(1+\frac{22}{c_1}\right) + \frac{v_3}{5}\left(1+\frac{40}{c_2}\right) &= c_3 \qquad \text{(kinetics and protein sum)} \\
c_1 + c_2 + c_3 &= 340 \qquad \text{(constant cell density)}
\end{aligned}
$$

2) GBA: Explicit constraint on $\mu(f_2, f_3, x_1)$ (from the density constraint $f_1 = 1$)

$$
\mu(f_2, f_3, x_1) = \frac{f_3}{\frac{1}{6}\left(1+\frac{1}{x_1}\right) + \frac{f_2}{6}\left(1+\frac{22}{340(1-f_2)}\right) + \frac{f_3}{5}\left(1+\frac{40}{340(f_2-f_3)}\right)} \qquad \text{(constrained growth rate)}
$$

3) Analytical conditions for optimal balanced growth state (system of algebraic equations)

$$
\begin{aligned}
\frac{1}{6}\left(1+\frac{22}{340(1-f_2)}\right) + \frac{22 f_2}{6\,[340(1-f_2)]^2} - \frac{40 f_3}{5\,[340(f_2-f_3)]^2} &= 0 \qquad (j=2) \\
1 - \mu\frac{1}{5}\left(1+\frac{40}{340(f_2-f_3)}\right) - \mu\,\frac{40 f_3}{5\,[340(f_2-f_3)]^2} &= 0 \qquad (j=3)
\end{aligned}
$$

Optimal substrate mass concentration = free enzyme mass concentration

The optimal mass concentration balance for minimal ρ :

$$c_m = \frac{p^j \, K_m^j}{K_m^j + c_m} \quad .$$

But this corresponds exactly to the free enzyme mass concentration

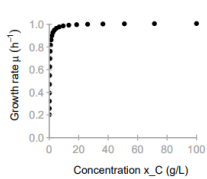
$$p_{\text{free}}^j \coloneqq p^j - p^j \left(\frac{c_m}{c_m + K_m^j} \right) = \frac{p^j \, K_m^j}{K_m^j + c_m} \quad .$$

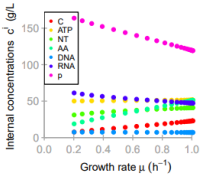
Thus[10],

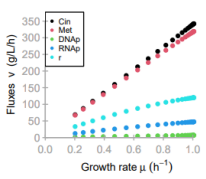
$$c_m = p_{\text{free}}^j \quad .$$

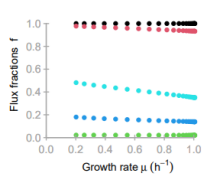
[10] Dourado et al. On the optimality of the enzyme–substrate relationship in bacteria, *PLOS Biology* 2021

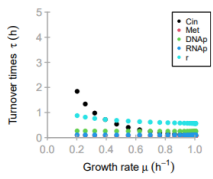
Numerical solutions for different external concentrations x: model B

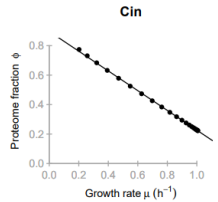


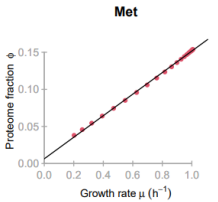


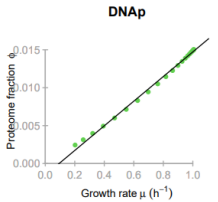


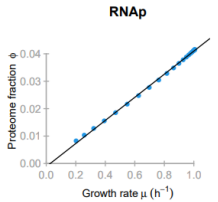


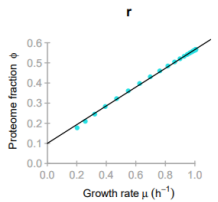












The dynamic generalization: fitness optimization

For some given model $(\mathbf{M}, \boldsymbol{\tau}, \rho)$ and dynamic environment $\mathbf{x}(t)$:

$$\underset{\mathbf{v}(t),\,\mathbf{c}(t)}{\text{maximize}} \quad \int_0^T \mu \, \mathrm{d}t \qquad \text{(Maximize fitness)}$$

subject to:

$$\mathbf{M}\,\mathbf{v} = \mu\,\mathbf{c} + \dot{\mathbf{c}} \qquad \text{(Mass conservation)}$$

$$c_\mathrm{p} = \mathbf{v}^\top \boldsymbol{\tau}(\mathbf{c}, \mathbf{x}) \qquad \text{(Reaction kinetics and protein sum)}$$

$$\rho = \sum \mathbf{c} \qquad \text{(Constant cell density)}$$

Main trick for analytical "solution": define the "generalized fluxes" $\mathbf{q}$ such that

$$\rho\,\mathbf{M}\,\mathbf{q} = \mathbf{c} \quad ,$$

then reformulate the problem on $\dot{\mathbf{q}}, \mathbf{q}, \mathbf{x}$, and solve Euler-Lagrange equations.