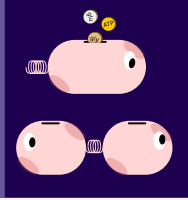
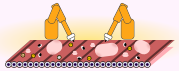
Economic Principles in Cell Physiology

Paris, July 6, 2022



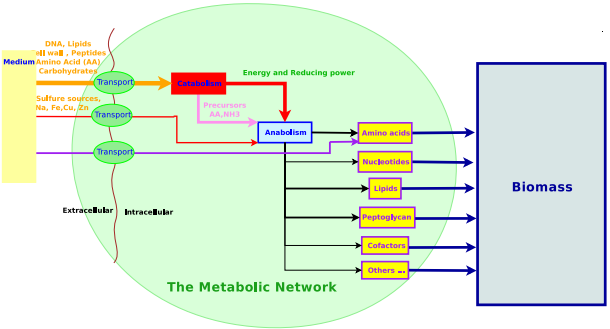
Resource allocation models

Anne Goelzer & Wolfram Liebermeister



LEARNING PLANET INSTITUTE

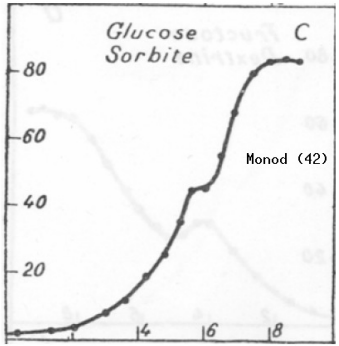
The cell as considered by FBA



$$\begin{aligned} &\text{maximize} && c^T \nu \text{ (growth rate)} \\ &\nu \in \mathbb{R}^m \\ &\text{subject to} \\ &&& \Omega\nu = 0 \\ &&& \alpha_i \leq \nu_i \leq \beta_i \end{aligned}$$

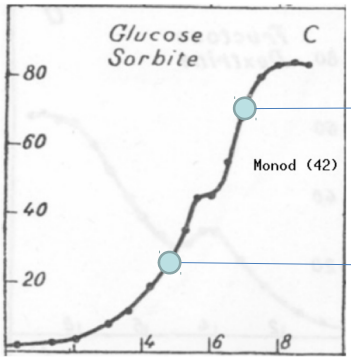
where Ω is the stoichiometry matrix and c is the objective function (as growth rate maximisation).

An effective modeling framework . . .



Can FBA predict some well-established phenomena (as the catabolic repression)?

Monod's experiments is not recovered by FBA



Predicted use of carbon source by other constraint-based modeling methods (FBA)

Sorbitol + Glucose

Sorbitol + Glucose

A relevant modeling framework, but unable to predict some well-established biological phenomena in an autonomous way.

Advantages and drawbacks of Flux Balance Analysis

Advantages

- no parameters;
- include the whole metabolic network through mass conservation: $\Omega.\nu = 0$;
- an optimization problem which can be easily solved.

Drawbacks

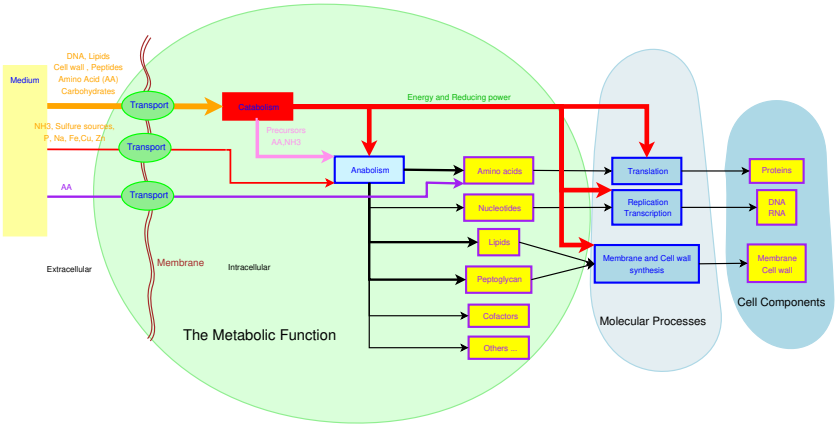
- choice of the objective function (maximization of biomass);
- need to limit superior bounds on uptake fluxes to have a solution;
- genetic regulations are not included.

Conclusion: FBA problem is not enough constrained.

Include other structural constraints active during exponential growth (as mass conservation)

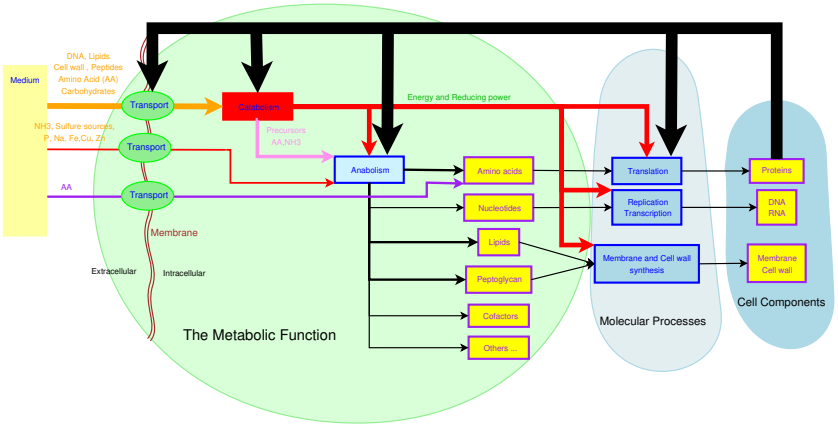
- FBAwMC for *FBA with Molecular Crowding*: Vazquez et al. *Impact of the solvent constraint on Escherichia coli metabolism*. BMC Systems Biology, 2:7. 2008.
- RBA for *Resource Balance Analysis*: Goelzer et al. *Cell design in bacteria as a convex optimization problem*. Automatica. 47(6):1210-1218. 2011.

Description of biomass



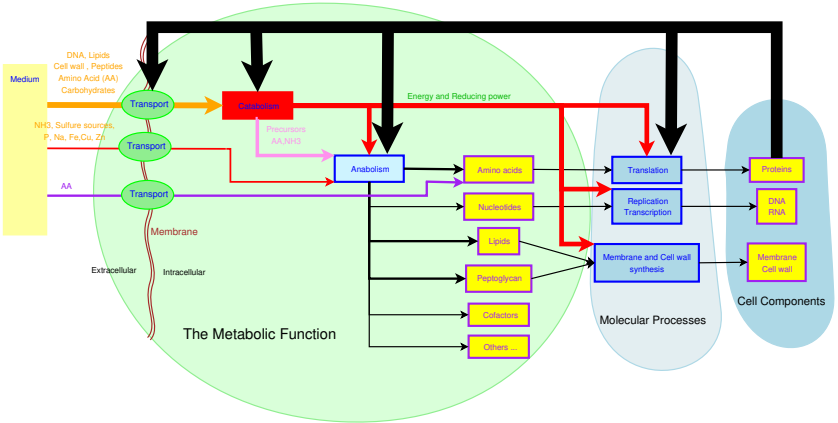
Different molecular processes produce cellular components. . .

An implicit coupling: proteins



. . . by sharing a common resource: proteins.

An implicit coupling: proteins



Need to integrate the protein concentration in the optimization problem.

Three structural constraints during exponential growth

Exponential growth

By definition of the "**balanced**" regimen of exponential growth, we have

- the growth rate μ is constant;

$$\dot{V}(t) = \mu V(t)$$

- the concentrations of the different cellular components are also constant.

A first consequence: for each protein $\mathcal{P}_j$ of concentration P_j,

$$\frac{dP_j(t)}{dt} = \underbrace{p_j(t)}_{\text{Production}} - \underbrace{\mu P_j(t)}_{\text{Dilution}}$$

$\Longrightarrow$ a constant flux of protein production of $p_j = \mu P_j \neq 0$ is required in exponential growth (steady-state).

For all proteins, a total flux equal to $\sum_j p_j = \mu \sum_j P_j$ of protein production is required.

A first structural constraint

Ribosomes are composed of proteins.

Consequence 2: for the ribosomes of concentration R,

- a synthesis flux of $r = \mu R$ is required in steady-state;
- which leads to a constraint on the capability of the protein production

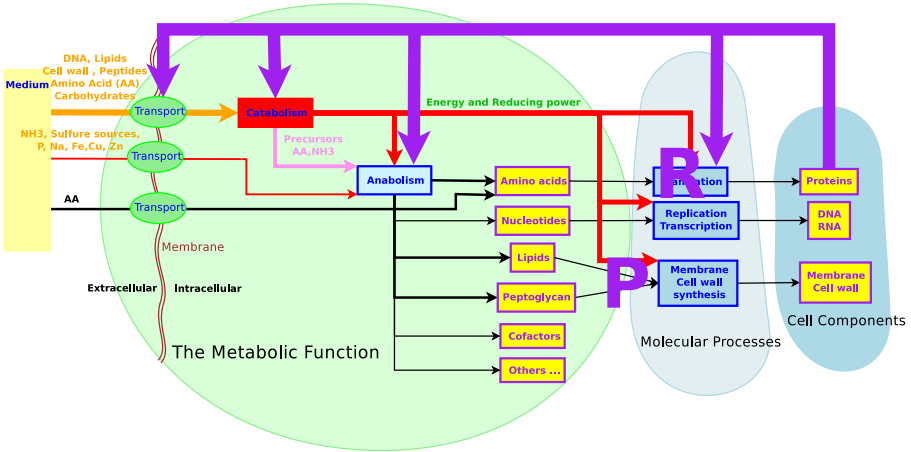
$$\mu R \leq k_T R \quad \Longrightarrow \quad \mu \leq k_T.$$

k_T: is related to the efficiency of the translation of ribosomes.

A first structural constraint on the capability of protein production by the translation apparatus

Let us now distinguish another set of proteins . . .

Another set of proteins P



Impact of nonribosomal proteins

Let us consider another set of proteins $\mathcal{P}$ of concentration P.

k_T, an optimistic superior bound ?

$$\mu(R+P) \leq k_T R \quad \Longrightarrow \quad \mu \leq k_T \frac{R}{R+P}$$

$\Longrightarrow$the superior bound is decreasing.

Impact of variations of P

Objective: Mimic the repression of the protein synthesis.

- Initial state of given concentration P_1

$$\mu(R + P_1) \leq k_T R$$

- Let $\delta P \in [0, P_1)$ such as $P_2 = P_1 - \delta P$.

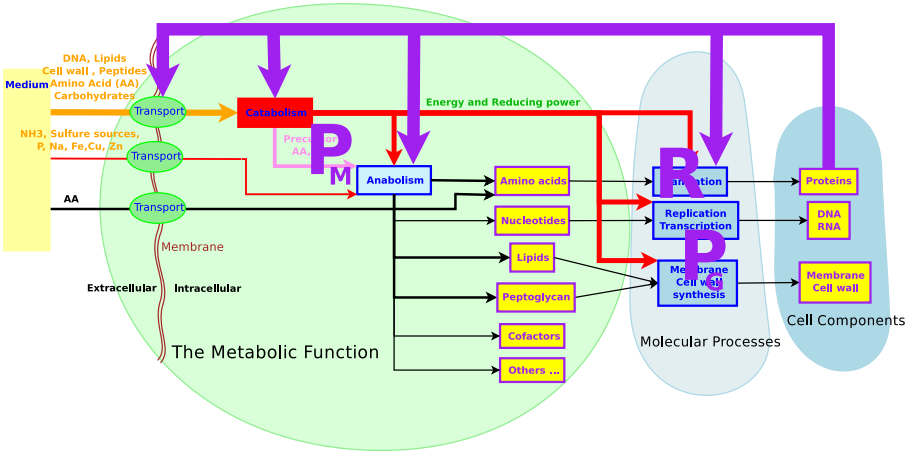
$$\mu(R + P_2 + \delta P) \leq k_T R$$

- by choosing $\delta\mu = \mu \frac{\delta P}{R+P_2}$,

$$(\mu + \delta\mu)(R + P_2) \leq k_T R$$

$\Longrightarrow$ μ can increase.

Two sets of nonribosomal proteins



- $\mathcal{P}_M$ involved in the metabolic network;
- $\mathcal{P}_G$ belonging neither to $\mathcal{P}_M$ nor to $\mathcal{R}$.

A second structural constraint

Let ν the flux of metabolic precursors produced by $\mathcal{P}_M$. So

$$\nu = \mu(\alpha_R R + \alpha_G P_G + \alpha_M P_M)$$

where

- α_R is the amount of precursors required to ribosome synthesis;
- α_M is the amount of precursors required to $\mathcal{P}_M$ synthesis;
- α_G is the amount of precursors required to $\mathcal{P}_G$ synthesis.

A second structural constraint on the capability of production of the metabolic network to provide the metabolic precursors required for the cell growth

An important assumption

Integrate the link between flux and enzyme concentration.
Recall that $\nu = f_M(S,P)P_M$, where $f_M(S,P)$ is the apparent catalytic rate (also called apparent turnover rate) of the enzyme.

Assumption: for a given exponential growth condition, metabolite concentrations are constant. So we simplified the non-linear relation $f_M(S,P)$ by one parameter k_M.

- A mass balance constraint.

$$\nu = \mu(\alpha_R R + \alpha_G P_G + \alpha_M P_M)$$

- A capacity constraint.

$$-k_M^{'} P_M \leq \nu \leq k_M P_M$$

where k_M and $k_M^{'}$ are the apparent catalytic rate in the forward and backward direction of $\mathcal{P}_M$.

A third structural constraint

Let us return to

$$\mu(R+P) \leq k_T R \quad \Longrightarrow \quad \mu \leq k_T \frac{R}{R+P}$$

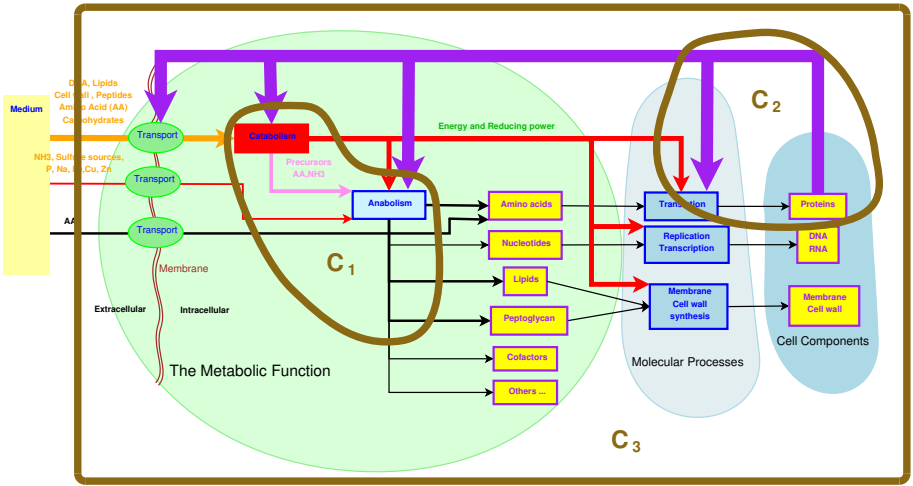
Theoretically, we could have

$$\lim_{R \longrightarrow \infty} k_T \frac{R}{R+P} = k_T.$$

BUT: The intracellular volume is limited.
$\Longrightarrow$ Every cellular component has a maximal concentration.

A third structural constraint on the volume of cellular compartments and membrane occupancy

Three structural constraints

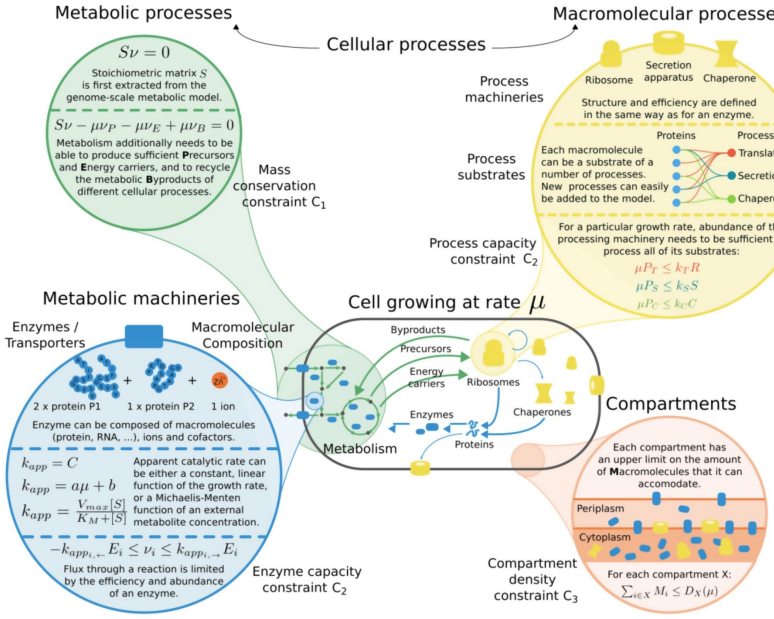


And the mass conservation and the stoichiometry of metabolic network

Generalization to the whole cell

At steady state
(constant growth rate)

Resource Balance Analysis in a nutshell



The RBA optimisation problem

Abundance of housekeeping proteins

Growth rate

Abundance of molecular machines

Metabolic fluxes

Fixed abundances of macro-components (i.e. cell wall precursors, etc.)

$$\text{For fixed } P_G \geq 0,\ \mu \geq 0,$$

$$\begin{aligned}
&\text{find} && Y \in \mathbb{R}^{m+p}_{\geq 0}, \nu \in \mathbb{R}^m, \\
&\text{subject to} && \\
&(C_1) && -\Omega\nu + \mu(C^S_Y Y + C^S_B \bar{B} + C^S_G P_G) = 0 \\
&(C_{2a}) && \mu(C^M_Y Y + C^M_G P_G) - K_T Y \leq 0 \\
&(C_{2b}) && -K'_E Y \leq \nu \leq K_E Y \\
&(C_3) && C^D_Y Y + C^D_G P_G - \bar{D} \leq 0
\end{aligned}$$

Efficiencies of molecular machines (i.e. kapp)

Densities of compartments

The RBA optimisation problem

Abundance of housekeeping proteins

Growth rate

Abundance of molecular machines

Metabolic fluxes

For fixed $P_G \geq 0, \mu \geq 0,$

find $Y \in \mathbb{R}^{m+p}_{\geq 0}, \nu \in \mathbb{R}^m,$

subject to

$$(C_1) \quad -\Omega\nu + \mu(C^S_Y Y + C^S_B \bar{B} + C^S_G P_G) = 0$$

$$(C_{2a}) \quad \mu(C^M_Y Y + C^M_G P_G) - K_T Y \leq 0$$

$$(C_{2b}) \quad -K^{'}_E Y \leq \nu \leq K_E Y$$

$$(C_3) \quad C^D_Y Y + C^D_G P_G - \bar{D} \leq 0$$

Medium setting

(for a transporter uptaking a metabolite S_{ext})

$K_E = \frac{k_{max} S_{ext}}{K_m + S_{ext}}$

Abundance of S_{ext}

Theoretical properties of $P_f(\mu)$

The following theoretical properties are obtained:

1. $P_f(\mu)$ is a Linear Programming problem;
   $\Longrightarrow$ same complexity as the FBA problem, the RBA problem is easily solvable!

2. For a set of given environmental conditions, **there exists a maximum growth rate**
   - **without defining any objective function** (by constrast to FBA);
   - **defined by a trade-off on proteins**;
   - for which a protein distribution (enzyme/ribosome) exists;
   - which can be computed by dichotomy with $P_f(\mu)$;

3. Theoretical predictions of induced/repressed metabolic sub-systems with respect to the concentrations of extracellular nutrients;

4. Every mechanism allowing to save proteins increases the growth rate.

What predicts RBA?

For a set of given extracellular concentration of nutrients, RBA computes

- the maximal growth rate
- the metabolic fluxes including the substrate uptake and by-product secretion rates
- the abundances of molecular machines such as enzymes, transporters, chaperones and proteins involved in the translation apparatus

But if an objective function is added, one can look for the cellular configuration maximizing the production of a compound (metabolite, or protein) of interest at given growth rate.

Prediction of the Monod's curve $\mu = f(\text{Glucose})$

RBA captures the macroscopic/microscopic behavior of bacteria

Monod's experiments

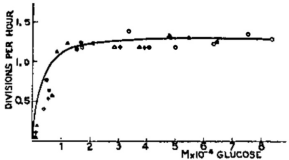
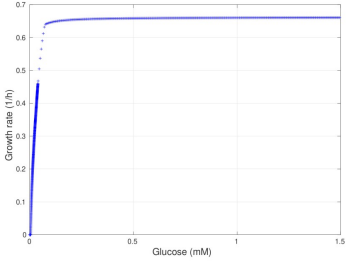


FIG. 4.—Growth rate of *E. coli* in synthetic medium as a function of glucose concentration. Solid line is drawn to equation (2) with $R_K = 1.35$ divisions per hour, and $C_1 = 0.22$ M $\times 10^{-4}$ (11). Temperature 37° C.



Prediction of the relation Ribosome = f(μ)

RBA captures the macroscopic/microscopic behavior of bacteria

Monod's experiments

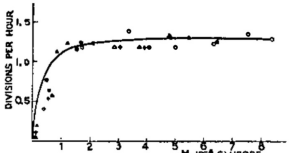
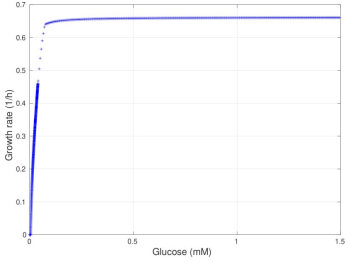
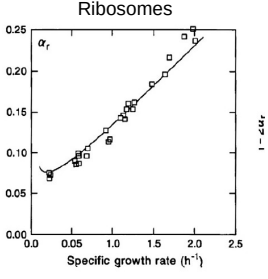


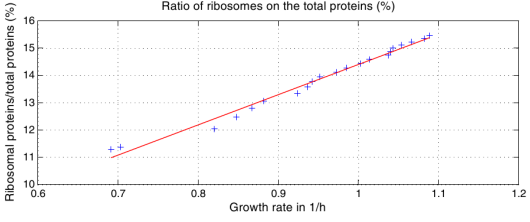
FIG. 4.—Growth rate of *E. coli* in synthetic medium as a function of glucose concentration. Solid line is drawn to equation (2) with $R_K = 1.35$ divisions per hour, and $C_1 = 0.22\ \mathrm{M} \times 10^{-4}$ (11). Temperature 37° C.



Marr's experiments

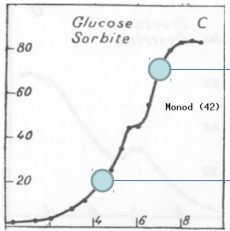


A. Goelzer, V. Fromion / Biochimica et Biophysica Acta 1810 (2011) 978–988



Prediction of a hierarchy in the use of nutrients

Monod's experiments recovered by RBA



	Predicted use of carbon source by RBA	Predicted use of carbon source by other constraint-based modeling methods (FBA)
Sorbitol only	Sorbitol only	Sorbitol + Glucose
Glucose only	Glucose only	Sorbitol + Glucose

RBA predicts a hierarchy in the use of carbon sources

Most induced/repressed sub-systems in the metabolic network coincides with a known regulatory structure

→ Promising to predict the cell phenotype in combined stress conditions

Estimation of model parameters

Parameters to be estimated

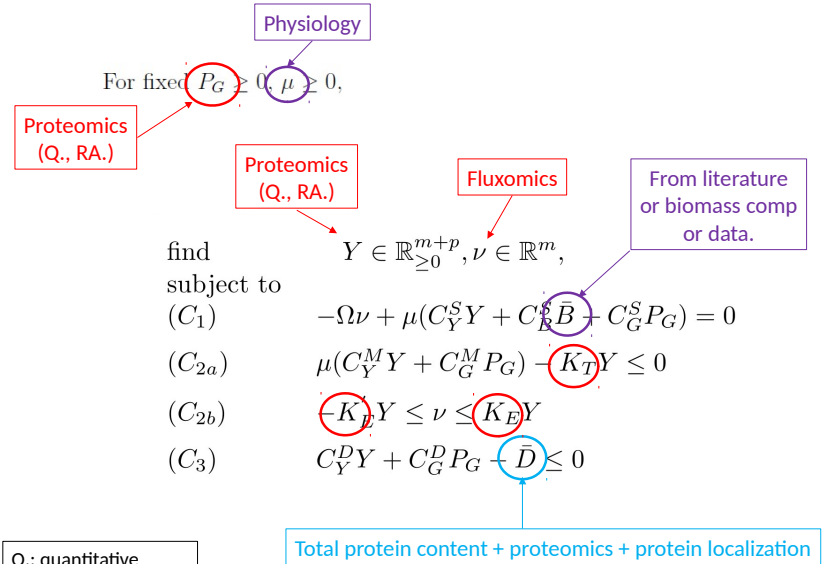
For fixed $P_G \geq 0$, $\mu \geq 0$,

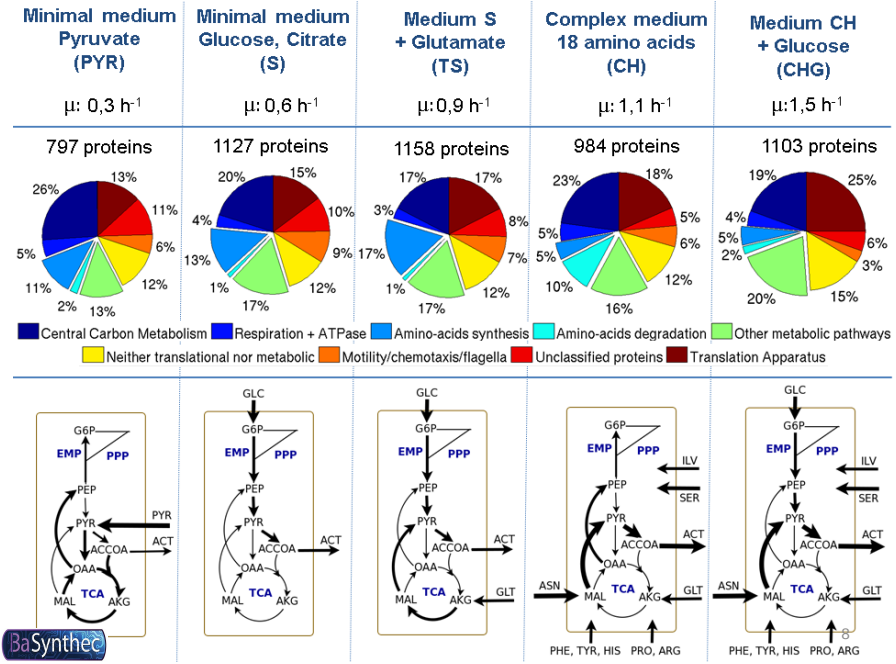
$$
\begin{aligned}
&\text{find} && Y \in \mathbb{R}^{m+p}_{\geq 0}, \nu \in \mathbb{R}^{m}, \\
&\text{subject to} && \\
&(C_1) && -\Omega\nu + \mu(C^S_Y Y + C^S_B \bar{B} + C^S_G P_G) = 0 \\
&(C_{2a}) && \mu(C^M_Y Y + C^M_G P_G) - K_T Y \leq 0 \\
&(C_{2b}) && -K'_E Y \leq \nu \leq K_E Y \\
&(C_3) && C^D_Y Y + C^D_G P_G - \bar{D} \leq 0
\end{aligned}
$$

From the stoichiometry of chemical reactions

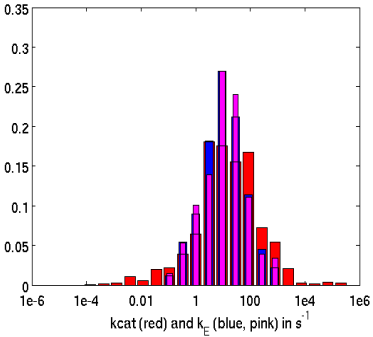
From annotation & bioinformatics

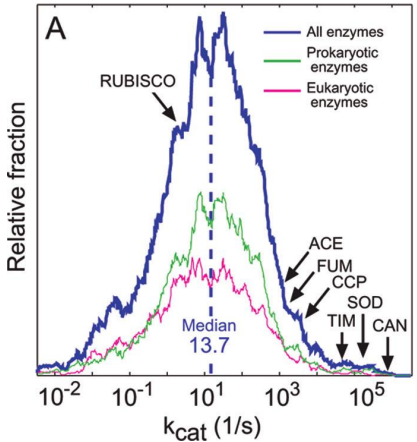
Parameters to be estimated




Minimal medium
Pyruvate
(PYR)
μ: 0,3 h-1
Minimal medium
Glucose, Citrate
(S)
μ: 0,6 h-1
Medium S
+ Glutamate
(TS)
μ:0,9 h-1
Complex medium
18 amino acids
(CH)
μ: 1,1 h-1
Medium CH
+ Glucose
(CHG)
μ:1,5 h-1
797 proteins
1127 proteins
1158 proteins
984 proteins
1103 proteins
Central Carbon Metabolism
Respiration + ATPase
Amino-acids synthesis
Amino-acids degradation
Other metabolic pathways
Neither translational nor metabolic
Motility/chemotaxis/flagella
Unclassified proteins
Translation Apparatus
GLC
G6P
EMP
PPP
PEP
PYR
ACCOA
OAA
MAL
TCA
AKG
ACT
GLT
ILV
SER
ASN
PHE, TYR, HIS
PRO, ARG
BaSynthec

Identification of apparent catalytic rate of 600 enzymes (Consistency with the expected distribution)

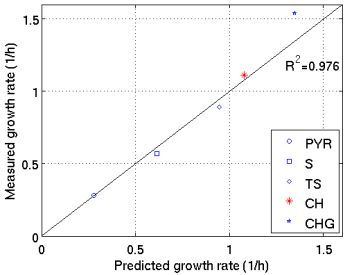


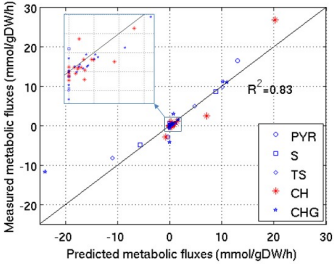


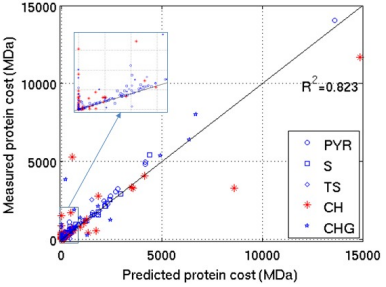
A. Bar-Even, et al. The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters, Biochemistry, 2011, 50 (21), pp. 4402–4410

BaSynthec

Quantitative prediction of the resource allocation







Generation of RBA models

How to build a resource allocation model for another bacterium?

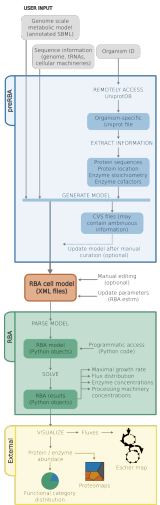
For model building, one needs

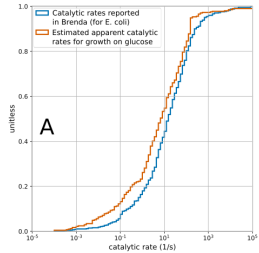
- the annotated genomic sequence
- a genome-scale metabolic model in SBML format
- description of the molecular machines of the macro-molecular processes that you want to include in RBA

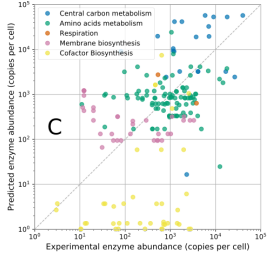
For model calibration, one needs

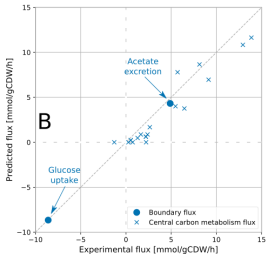
- quantitative proteomic datasets
- uptake and excretion rate of nutrients, fluxomics

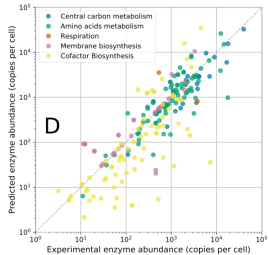
Use RBApy to generate a calibrated RBA model











Conclusion

The RBA framework provides theoretical properties for the cell design in bacteria

- The trade-off on resources structurally limits the growth rate;
- Every mechanism allowing to save resources increases the growth rate.

The RBA framework can efficiently predict for a set of concentrations of extracellular nutrients

- The maximal growth rate;
- The resource repartition of the cell: the abundance of enzymes, transporters and non-metabolic molecular machines as ribosomes and metabolic fluxes;
- The impact of a rational modification of the bacterial strain (e.g. plug-in an entire metabolic pathway).

RBApy makes RBA model generation and simulation accessible for a large diversity of prokaryotes

see https://rba.inrae.fr/ for existing RBA models and code availability.

We are hiring ! (30 months position to develop a RBA model of cancerous and healthy cells)