Modelling and optimization of metabolic pathways in bacteria

Guillaume Jeanne, Sihem Tebbani, Anne Goelzer, Vincent V. Fromion, Didier Dumur

Modelling and Optimization of Metabolic Pathways in Bacteria

Guillaume Jeanne*†, Sihem Tebbani*, Anne Goelzer†, Vincent Fromion† and Didier Dumur*
*Laboratoire des Signaux et Systèmes, CentraleSupélec - CNRS - Univ. Paris-Sud, Univ. Paris-Saclay,
Control Department, Plateau du Moulon, 91190 Gif-sur-Yvette, France
†INRA, UR1404, MaIAGE, Université Paris-Saclay, 78350 Jouy-en-Josas, France

***Abstract*—The rational bacterial strain design is a major challenge in synthetic biology. This paper deals with the optimization of a bacterial strain for specific processes taking place in a bioreactor. Such problems are namely maximizing the growth and the production of a product of interest. First, a model combining the internal behavior of the cells with a bioreactor environment is developed assuming mass balance and biological constraints. This model assumes that the production of proteins can be controlled. The problem is then solved with constant optimization variables and returns an optimal strategy for synthetic strains.**

Keywords: System Biology, Bioprocess Optimization, Bacterial Strain design, Resource Allocation, Cellular Process

I. Introduction

Bioengineering is the application of engineering techniques to biological systems. Biological systems and in particular micro-organisms like bacteria have become a new way to produce interest chemical product. Growing micro-organisms in a bioreactor fed with glucose is common [1], [2] and the challenge is now to design new bacterial strains with better behavior for practical applications [3], [4], [5], [6]. Biologists have usually created synthetic cell strains which are more efficient, more resistant, faster to reproduce, etc., by trial and error approaches. This procedure is time consuming, expensive and does not guarantee any optimality, except the fact that they are proved by experiments.

Among others, the difficulty to anticipate the behavior of the obtained cells and the lack of property in the supposed optimality are reasons why we need to model cells. The goal is to model cellular behavior of a process in a bioreactor and then to use optimization techniques to describe how would perform a better strain. It is fundamental to understand and predict internal behavior of cells in order, then, to design strains optimized according to a given criterion in a given context. Besides, the coupling with a model of bioreactor directly includes the strain design in an application context and allows a point of view different from a pure theoretical or general one, under laboratory conditions.

There are plenty of models for cellular behavior [7], [8], [9], [10]. Steady-state models under constraints [7], [8] are very accurate and can return faithful cellular fluxes values of internal species. Unfortunately, the trade-off between prediction capacity and simplicity turns these models to be very complex to use and to connect to bioreactor models for example. Indeed, they describe well the wild cells, but if we want to design a new strain and compare it to the natural one, the change would be difficult to interpret.

As a matter of fact, to model both the bioreactor and the intracellular state, the idea developed in this paper is to keep the same approach as those faithful steady-state models, i.e. (i) focus on biological main species, (ii) arise biological constraints, and couple it with a dynamical model describing the evolution of internal concentrations. Optimization techniques can then be used on this comprehensible representation to determine optimized strains.

This paper is organized as follows: Section II gives a description of an internal cell model and biological components integrated in this model. Section III presents two optimization problems, namely the improvement of growth and production maximization of the product of interest and the constraints to add in order to represent at best the biological behavior of the cells. Section IV starts by setting the numerical and computational details and then presents the solutions found for two optimization problems. Section V gives some conclusions and perspectives.

II. Cell Description and Process Modeling

This work focuses on processes in batch reactors (i.e. no addition or removal of culture media), in finite time horizon, exclusively. This choice was motivated by the simplicity of the expressions in terms of concentration variation and of the derived optimization problem.

A cell is considered to be an entity with a fixed volume v_{bact} and composed of different kinds of proteins, ribosomes, internal substrate and metabolites. The reactor has a constant volume V_{react} and contains external substrate, the bacteria colony and a product of interest produced by the cells. The total volume of bacteria, denoted $V_{pop}(t)$, equals the number of bacteria $N_{pop}(t)$ multiplied by a unitary volume v_{bact}.

In this paper, the bacteria and culture medium are assumed homogeneous.

A. Mass Balance Equations

$n_\xi(t)$ denotes the amount and $[\xi](t)$ the concentration of species ξ at time t. If ξ is internal to the cell, then

$$[\xi_{int}](t) = \frac{n_{\xi_{int}}(t)}{V_{pop}(t)} \tag{1}$$

Otherwise, if ξ is external, then

$$[\xi_{ext}](t) = \frac{n_{\xi_{ext}}(t)}{V_{react}} \tag{2}$$

Hence, the time derivatives are given by:

$$\frac{d[\xi_{int}]}{dt}(t) = \frac{1}{V_{pop}(t)} \times \frac{dn_{\xi_{int}}}{dt}(t) - \frac{[\xi_{int}](t)}{V_{pop}(t)} \times \frac{dV_{pop}}{dt}(t) \tag{3}$$

and

$$\frac{d[\xi_{ext}]}{dt}(t) = \frac{1}{V_{react}} \frac{dn_{\xi_{ext}}}{dt}(t) \tag{4}$$

The second term of the right hand side of (3) is due to the growth of the colony. In fact, the specific growth rate μ is defined as the logarithmic time derivative of the bacteria population:

$$\mu(t) \triangleq \frac{1}{N_{pop}(t)} \frac{dN_{pop}}{dt}(t) = \frac{1}{V_{pop}(t)} \frac{dV_{pop}}{dt}(t) \tag{5}$$

Macroscopically, this term is often replaced by an empirical Monod or Haldane expression [1]. With such an expression, the growth rate only depends on external parameters, namely the external substrate concentration. In this paper, μ will depend on concentrations of cell internal compounds.

Denoting q the material flows for the whole colony, expressed in number of molecules per hour, the time derivatives of amounts are given by:

$$\frac{dn_{\xi}}{dt}(t) = \pm q_{exchanged}(t) + q_{produced}(t) - q_{consumed}(t) \tag{6}$$

Those fluxes are global to the colony and the same for each bacterium of the colony. Thus, they can be expressed per bacterium. In addition, assuming that the medium is uniform, they can be normalized by the volume of a cell and in this way be defined per unit of volume:

$$\nu(t) = \frac{q(t)}{v_{bact} N_{pop}(t)} = \frac{q(t)}{V_{pop}(t)} \tag{7}$$

with $\nu(t)$ expressed in mole per liter per hour.
Keeping these notations, (6) becomes

$$\frac{1}{V_{pop}(t)} \frac{dn_{\xi}}{dt}(t) = \pm \nu_{exchanged}(t) + \nu_{produced}(t) - \nu_{consumed}(t) \tag{8}$$

To summarize equations (3), (4), (5) and (8),

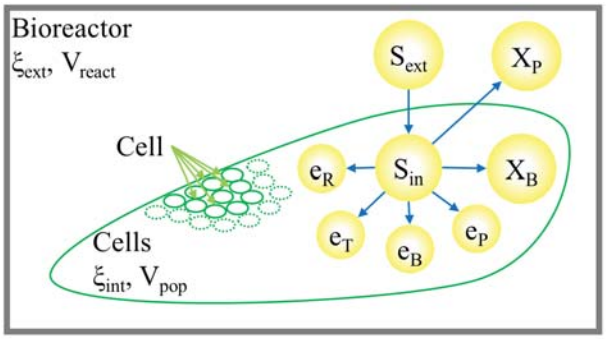
$$\begin{cases} \frac{d[\xi_{int}]}{dt}(t) = \pm\, \nu^{\xi_{int}}_{exchanged}(t) + \nu^{\xi_{int}}_{produced}(t) \\ \qquad\qquad - \nu^{\xi_{int}}_{consumed}(t) - \mu(t)[\xi_{int}](t) \\ \frac{d[\xi_{ext}]}{dt}(t) = \Big(\pm\, \nu^{\xi_{ext}}_{exchanged}(t) + \nu^{\xi_{ext}}_{produced}(t) \\ \qquad\qquad - \nu^{\xi_{ext}}_{consumed}(t) \Big) \times \frac{V_{pop}(t)}{V_{react}} \end{cases} \tag{9}$$

The time dependence is further omitted to simplify notations.

B. Cell internal model and dynamics

1) Resource Balance Analysis Approach: This work is based on the Resource Balance Analysis (RBA) [8] philosophy for the cell internal phenomena modeling and understanding. The RBA approach gives a description of the inner behavior of the cell in steady-state, using a constraints approach easily reducible and adaptable to a bioprocess perspective. From the cells point of view, steady-state means the internal concentrations are constant as well as the growth rate. As a consequence, the overall population growth is exponential.
While respecting all the stoichiometric constraints due to mass balance in the cell, a cell has to (i) produce enough metabolic precursors, macrocomponents, and internal compounds, respecting the mass conservation, (ii) maintain the rate of protein production high enough to ensure the production of necessary protein, (iii) limit its internal mass. Respecting theses constraints, the RBA model finds the optimal repartition of resources between the different cellular processes, especially metabolic pathways and the ribosome concentration in the cell ensuring the highest growth rate in steady-state.
In this paper, we kept the intrinsic constraints while reducing the model dimension. Simplifications come from gathering biological elements into "pools". The concentrations variations are expressed as fluxes and these fluxes are subject to constraints. As said in the introduction, the goal is to obtain a modestly sized model which permits better comprehension.
The pools are presented in Fig. 1 and in the following subsections, and the constraints are defined in paragraph III-B of the Optimization Section. The figure 1 also sums up the exchanges between pools, where the notations will be defined hereafter.

Fig. 1. General representation of the bioreactor and exchanges between species



2) Metabolic Precursors: First of all, there is an internal substrate pool, denoted S_{in}, gathering the metabolic precursors necessary to build bigger elements in the cell (as for example macrocomponents or proteins). Outside the cells, the substrate is denoted S_{ext} and the stoichiometry between S_{in} and S_{ext} is exactly one. It assumes there is no entrance cost for elements. Thus the mass balance is:

$$S_{ext} \xrightarrow{\nu_{import}} S_{in} \tag{10}$$

Remark: the ATP (Adenosine triphosphate), the free amino acids and other tiny elements are all in this pool.

3) Metabolites & Macrocomponents Synthesis: X_P denotes an external chemical species of interest produced by the cells. One mole of S_{in} can proceed α_P mole of X_P, leading to:

$$S_{in} \xrightarrow{\nu_{X_P}} \alpha_P.X_P \tag{11}$$

X_B is the pool of all the internal elements X_{Bi} different from the ribosomes, the proteins, and whose concentration cannot vary much around the nominal value. There are two categories: the heavy macro components (the DNA, the membrane, the cell wall) that are a few but really heavy in equivalent amino acids and all the metabolites (the intermediates and products of metabolism) which are very numerous but much more lightweight in equivalent amino acids. The stoichiometry is given by the α_{B_i} and summed up as follows:

$$S_{in} \xrightarrow{\nu_{X_B}} \sum_i \alpha_{B_i}.X_{B_i} = \alpha_B.X_B \tag{12}$$

4) Protein & Ribosomes Synthesis: The fluxes (10), (11) and (12) are catalyzed by enzymes e_B, e_P and e_T. All these enzymes are produced by ribosomes e_R and consuming a specific amount of S_{in}. We pose that 1 mole of S_{in} produces α_{e_i} mole of e_i.
e_B is the pool of enzymes involved in the whole metabolic network of the cell synthesizing X_B. Their production is modeled by:

$$S_{in} \xrightarrow{\nu_{e_B}} \alpha_{e_B}.e_B \tag{13}$$

e_P are the enzymes producing (from S_{in}) and excreting X_P outside the cell. They are produced from S_{in} as:

$$S_{in} \xrightarrow{\nu_{e_P}} \alpha_{e_P}.e_P \tag{14}$$

e_T are the membranous proteins. Among them, a constant proportion is responsible for the substrate import inside the cell. The reaction scheme is:

$$S_{in} \xrightarrow{\nu_{e_T}} \alpha_{e_T}.e_T \tag{15}$$

e_R are the ribosomes. They follow the same production process as proteins. They are considered as big proteins, 30 times bigger than other proteins in equivalent amino acids [11]. As other proteins, their production is modeled as:

$$S_{in} \xrightarrow{\nu_{e_R}} \alpha_{e_R}.e_R \tag{16}$$

C. Fluxes Expressions & Overall Dynamics

All fluxes expressions are based on a common structure: the concentration of the enzymes which catalyze the reaction multiplied by an efficiency. The formulation of this term differs for the different reactions.

For the import, the efficiency saturates with the external substrate concentration and the internal substrate concentration. The catalysts are a proportion of membranous enzymes e_T. Then, denoting r_{imp} the ratio of proteins causing the import of S among all the membranous proteins, and with v_T, K_T and K_S three constants, the import flux formula is:

$$\nu_{import} = \frac{r_{imp} v_T [S_{ext}]}{[S_{ext}] + K_T(1 + [S_{in}]/K_S)} \times [e_T] \tag{17}$$

The efficiencies for the production of X_B and X_P depend on the concentration of internal substrate. Hence, mass action laws appear:

$$\begin{cases} \nu_{X_B} = k_B[S_{in}] \times [e_B] \\ \nu_{X_P} = k_P[S_{in}] \times [e_P] \end{cases} \tag{18}$$

The enzymes and ribosomes production fluxes are catalyzed by ribosomes. The efficiencies depend on the internal substrate concentration and are controlled by the genes expressions level u_{e_i}. When u_{e_i} is high, the production of e_i is important. These are the inputs of the system. They are positive and bounded by a maximum value, u_{MAX}.
Hence, the production flux of enzymes and ribosomes are:

$$\begin{cases} \nu_{e_B} = u_{e_B}[S_{in}] \times [e_R] \\ \nu_{e_P} = u_{e_P}[S_{in}] \times [e_R] \\ \nu_{e_T} = u_{e_T}[S_{in}] \times [e_R] \\ \nu_{e_R} = u_{e_R}[S_{in}] \times [e_R] \end{cases} \tag{19}$$

Consequently, from (9), fluxes expressions (17), (18) and (19), and growth rate (5), the dynamics of the bioreactor can be expressed with 9 state variables as follows:

$$\begin{cases} \dot{[S_{in}]} = \dfrac{r_{imp} v_T [S_{ext}][e_T]}{[S_{ext}] + K_T(1 + [S_{in}]/K_S)} \\ \qquad - k_P[e_P][S_{in}] - k_B[e_B][S_{in}] \\ \qquad - \displaystyle\sum_{B,P,T,R} u_{e_i}[e_R][S_{in}] - \mu[S_{in}] \\ \dot{[e_B]} = \alpha_{e_B} u_{e_B}[e_R][S_{in}] - \mu[e_B] \\ \dot{[e_P]} = \alpha_{e_P} u_{e_P}[e_R][S_{in}] - \mu[e_P] \\ \dot{[e_T]} = \alpha_{e_T} u_{e_T}[e_R][S_{in}] - \mu[e_T] \\ \dot{[e_R]} = \alpha_{e_R} u_{e_R}[e_R][S_{in}] - \mu[e_R] \\ \dot{[X_B]} = \alpha_B k_B[e_B][S_{in}] - \mu[X_B] \\ \dot{[S_{ext}]} = -\dfrac{r_{imp} v_T [S_{ext}][e_T]}{[S_{ext}] + K_T(1 + [S_{in}]/K_S)} \times \dfrac{V_{pop}}{V_{react}} \\ \dot{[X_P]} = \alpha_P k_P[e_P][S_{in}] \times \dfrac{V_{pop}}{V_{react}} \\ \dot{V_{pop}} = \mu \times V_{pop} \end{cases} \tag{20}$$

As previously said, these dynamics need biological constraints to fully characterize the problem.

Remarks: the problem can easily be generalized to more metabolic pathways, other internal reactions and other internal substrates, by reproducing the previous patterns of equations.

D. Growth Rate depending on Density Constraint

In this model, the bacteria are supposed to duplicate continuously, without perturbation on the enzymes action. The duplication is motivated by the regulation of the internal density. When the internal mass increases, the cell divides and the total volume rises, so that the mass per cell (the density) remains constant.

1) Cell Density: The intracellular mass is expressed in equivalent amino acids and the density D is expressed in equivalent amino acids per unit of volume.
The metabolites and the internal substrate mass is negligible with respect to the proteins and ribosomes one. It simplifies the expression as only the proteins and ribosomes contributions are taken into consideration in the mass. Thus, the density is defined as:

$$D = \frac{\sum\limits_{B,P,T,R} aa_{e_i} n_{e_i}}{V_{pop}} = \sum_{B,P,T,R} aa_{e_i}[e_i] \tag{21}$$

where aa_ξ is the equivalent mass of chemical species ξ in amino acids.
The membranous proteins and the cytosolic ones are all counted for the "internal density" without distinguishing membranous and cytosolic densities.

2) Density regulation: Time derivative of the density in (21) leads to:

$$\dot{D} = \sum_{B,P,T,R} aa_{e_i}\dot{[e_i]} \tag{22}$$

From (20), it comes:

$$\dot{D} = \sum_{B,P,T,R} aa_{e_i}\left(u_{e_i}\alpha_{e_i}[e_R][S_{in}] - \mu[e_i]\right) \tag{23}$$

and thus,

$$\dot{D} = \sum_{B,P,T,R}\left(aa_{e_i}\alpha_{e_i}u_{e_i}\right)[e_R][S_{in}] - \mu D \tag{24}$$

In the literature, the density is observed to be constant in steady-state [12]. Here, the density is supposed to be around a constant value D_d. The evolution of D is supposed to be given by a first-order dynamics with characteristic time τ_D small compared to characteristic times of enzymes dynamics, thus,

$$\dot{D} = -\frac{1}{\tau_D}(D - D_d) \tag{25}$$

Hence, from (24)-(25), the volume growth regulates the density and induce a growth rate equal to:

$$\mu = \frac{[e_R][S_{in}]}{D}\sum_{B,P,T,R}\left(aa_{e_i}\alpha_{e_i}u_{e_i}\right) + \frac{1}{\tau_D}(1 - \frac{D_d}{D}) \tag{26}$$

with D given by (21).
μ depends on the inputs u_{e_i} and on the state variables, i.e. the four $[e_i]$ and $[S_{in}]$.

III. Optimization

A. Optimization Problems

In this paper, the objective is to determine the optimal genes expressions u_{e_i}, that maximize either the final bacteria volume or the X_P produced. The problem is, thus, as follows:

$$\max_{u_{e_B},u_{e_P},u_{e_T},u_{e_R}} J$$
$$\text{subject to} \begin{cases} 0 \leq u_{e_B}, u_{e_P}, u_{e_T}, u_{e_R} \leq u_{MAX} \\ \text{dynamics (20)} \\ \text{growth rate (26)} \\ \text{biological constraints(32)} \end{cases} \tag{27}$$

The criterion J is either $V_{pop}(t = t_f)$, or $[X_P](t = t_f)$, depending on the considered objective. The inputs u_{e_i} are assumed constant along all the culture duration.
Biological constraints that are included are detailed hereafter.

B. Biological Constraints

1) Translation: The translation of messengers RNA into proteins involves ribosomes and thus the global production flux of proteins cannot exceed what the ribosomes can produce. In fact, with k_T the ribosome efficiency, a ribosome efficiency constraint can be given by (28). Note that there is no competition between the genes. This assumption is very simplistic and reflects the absence of messengers RNA in this model.

$$\sum_{B,P,T,R} \nu_{e_i} \leq k_T[e_R] \tag{28}$$

From [13], the ribosome efficiency depends only on the growth rate as follows:

$$k_T(\mu) = v_{k_T}\frac{\mu}{K_{k_T} + \mu} \tag{29}$$

The two constants v_{k_T} and K_{k_T} can be computed using data from [13].
In addition, with (19), the translation constraint is rewritten as:

$$[S_{in}]\sum_{B,P,T,R} u_{e_i} \leq v_{k_T}\frac{\mu}{K_{k_T} + \mu} \tag{30}$$

2) Metabolism Constraints: As mentioned in II-B, X_B gathers all the internal elements (macrocomponents and metabolites) whose concentration is crucial for the cell survival.
A metabolism constraint consists in the control of $[X_B]$ around its initial value X_{B0}. A cell is assumed to be viable as long as the metabolites and the macrocomponents are present in the cell around a given concentration. This given concentration is the initial one and the tolerance is a specific ratio. In this paper, an arbitrary tolerance of 10% is chosen, leading to:

$$\forall t, 0.9 \leq \frac{[X_B](t)}{X_{B0}} \leq 1.1 \tag{31}$$

Indeed, as in [7], the composition of macrocomponents (DNA, cell wall, membrane, etc.) does not change with growth rate.

3) Constraints Summary: Biological constraints impose 2 bounds on state variables and 1 constraint linking command variables and states. These constraints are summarized below:

$$\begin{cases} [X_B] \leq & 1.1 \times X_{B0} \\ [X_B] \geq & 0.9 \times X_{B0} \\ [S_{in}] \sum\limits_{B,P,T,R} u_{e_i} \leq & v_{k_T} \dfrac{\mu}{K_{k_T} + \mu} \end{cases} \tag{32}$$

with μ defined in (26). These constraints are added to problem (27).

IV. Simulation Results

A. Numerical and Computation Details

The initial states for the cell composition are given by the RBA Model [11] in Table I and the reactor initial states are provided in Table II. All parameters values can be found in Table III [11]. g_{CDW} denotes gram of Cell Dry Weight, it is the common way to measure micro-organisms population in constraint-based models and thus cellular volumes of bacteria. aa unit stands for equivalent amino acids. In order to consider

TABLE I
INITIAL STATE VALUES [11]

	$mM.g_{CDW}^{-1}$	($mM\ aa.g_{CDW}^{-1}$)
$[S_{in}]_0$	0.1484	(0.1187)
$[e_B]_0$	$9.0.10^{-3}$	(3.24)
$[e_P]_0$	0	(0)
$[e_T]_0$	$4.8.10^{-3}$	(1.728)
$[e_R]_0$	$7.07.10^{-5}$	(0.7141)
$[X_B]_0$	$2.15.10^{-2}$	(0.1281)

TABLE II
INITIAL REACTOR STATE

$[S_{ext}]_0$	20	mM
$[X_P]_0$	0	mM
V_{reac}	10	L
V_{pop_0}	0.045	g_{CDW}

the same orders of magnitude, the states variables are normalized with the number of equivalent amino acids per molecule of each chemical species (right column of Table I).
In problem (27), equations (20), (26) and (32) are discretized with an explicit Euler method with $0.001h$ sampling time. The four optimal u_{e_i} are then determined using a Sequential Quadratic Programming algorithm [14] (SQP, by means of Matlab *fmincon* function) adding state constraints to ensure concentration positivity. The algorithm is initialized by the wild case (defined hereafter in section IV-B).

TABLE III
PARAMETERS VALUES USED BY THE SOLVER [11]

Notation	Value (Unit)
v_T	$370 \times 3600\ (h^{-1})$
r_{imp}	0.01
K_T	$0.8\ (mmol.L^{-1})$
K_S	$1\ (mmol.L^{-1})$
k_B	$12 \times 3600\ (h^{-1})$
k_P	$12 \times 3600\ (h^{-1})$
aa_{e_B}	$360\ (aa)$
aa_{e_P}	$360\ (aa)$
aa_{e_T}	$360\ (aa)$
aa_{e_R}	$10100\ (aa)$
$aa_{S_{in}}$	$0.8\ (aa)$
aa_{X_B}	$5.97\ (aa)$
aa_{X_P}	$1\ (aa)$

Notation	Value (Unit)
τ_D	$2/60\ (h)$
D_d	$5.682\ (aa.g_{CDW}^{-1})$
v_{k_T}	$27 \times 3600\ (aa)$
K_{k_T}	$0.5\ (aa.h^{-1})$
α_{e_B}	$2.2.10^{-3}$
α_{e_P}	$2.2.10^{-3}$
α_{e_T}	$2.2.10^{-3}$
α_{e_R}	$9.5.10^{-5}$
α_B	$1.67.10^{-4}$
α_P	0.8

TABLE IV
AFFINITIES OF THE WILD CASE, THE OPTIMUM OF THE PROBLEM (27) WITH THE OPTIMUM FOUND

	Wild		max(V) pb.		max($[X_P]$) pb.	
u_{e_B}	186 300	(58%)	108 400	(51%)	81 800	(36%)
u_{e_P}	0	(0%)	0	(0%)	31 600	(14%)
u_{e_T}	99 380	(31%)	74 950	(35%)	86 500	(37%)
u_{e_R}	33 880	(11%)	30 970	(14%)	30 000	(13%)
$\sum u_{e_i}$	319 560	(100%)	214 320	(100%)	229 900	(100%)

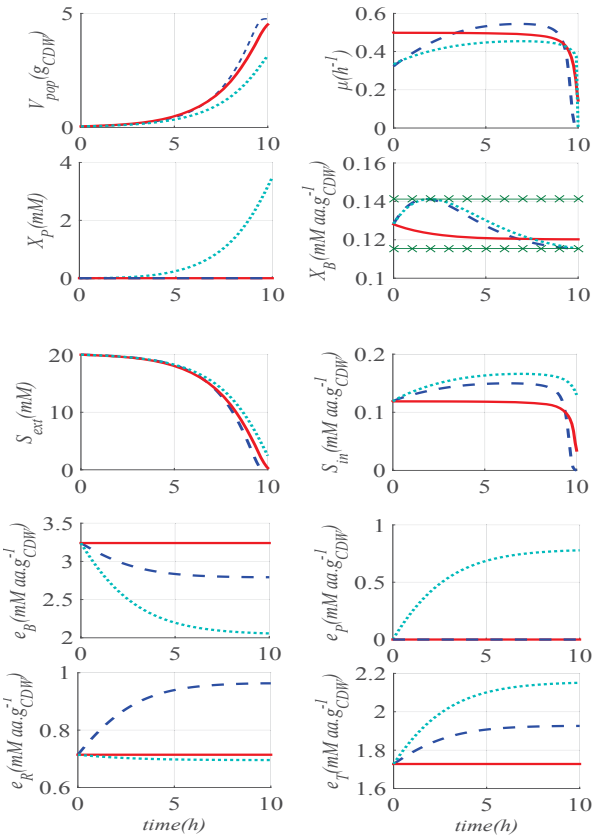


Fig. 2. In dotted turquoise, max Xp. In dashed blue max V. In solid red, wild. Green crossed lines in $X_B(t)$ represent constraint (31).

B. Optimal solutions

Hereinafter, the wild case denotes the cell with inputs computed to keep the enzymes concentration constant. These inputs are supposed to be the natural resource allocation in a genetically non-modified cell. In fact, the concentrations of enzymes are constant in steady-state and the RBA model gives the initial condition as steady-state values for exponential growth. Computing $u_{e_i}^{wild}$ is as simple as setting all the time derivatives to zero in (20).
Optimization results are summarized in the table IV. The ratio over the $\sum\limits_{P,B,R,T} u_{e_i}$ are bracketed.

The evolution of the internal and external concentrations is given in Fig. 2. It compares the evolution of 3 cases: (i) wild strains, (ii) strains optimized for growth, (iii) strains optimized for production.

C. Interpretation

Comparing the strains maximizing the final state volume with the wild cells, we can see that there is not much difference at the final point. The benefit is of 5%, but the strategy is different. Instead of a quasi constant growth rate, decreasing with the external substrate exhaust as in the wild case, the designed strain tries to save resources at the beginning and has a bell-curved growth rate evolution. This curve ends exactly at zero, as the cells manage to consume all the external and internal substrates at final time.

However, the cells maximizing the growth shall end up in bad shape as they do not have any internal resource to go on producing enzymes. This model does not include any cellular process for maintenance. Consequently, since the cell does not spend any cellular resources for maintenance, the optimal cells have no internal resources at final state.

Concerning the strain maximizing the product of interest production, we can see that the growing strategy is similar to the one maximizing the volume. The cells grow and produce X_P simultaneously.

In both optimized cases, the fraction of resources $u_{e_T}/\sum u_{e_i}$ is greater than in the wild case and $u_{e_R}/\sum u_{e_i}$ is also greater, while the total amount of resources is lower. It reflects a storing strategy for the cell. In fact, less ressources are used to produce enzymes. The growth rate is lower at the begnning of the simulation but reaches an higher value than the wild strain at mid-simulation leading to higher final biomass.

An interesting point to notice is that the sum of all inputs is much smaller in designed strains than in wild one. The translation constraint is not saturated in both modified strains compared to the wild one.

Indeed, the ratio $\frac{u_{e_j}}{\sum u_{e_i}}$ is more significant for the concentrations evolutions than the absolute value u_{e_i}. In fact, the inputs are related to the resources allocation in the different proteins production. The sum of the inputs is the total allocation and is, thus, closely linked to the growth rate. The ratios represent the relative repartition of the resources between biomass, product of interest synthesis, import and ribosome synthesis. Hence, same ratios correspond approximately to same cellular strategy, with different growth rates.

Inputs of enzymes devoted to metabolic networks and product of interest synthesis are approximately half of the total inputs in both cases. It looks like the maximizing volume strain is not constrained that much with time and has enough time to consume efficiently its resources even with having supplement resources to produce X_P. The difference is then that the strain optimized for the production does not manage to consume all the substrate as opposed to the maximizing volume strain.

V. Conclusion & Perspectives

In this paper, we propose a simple model coupling a macro-description of bioreactor to the intracellular scale of microorganisms. Using this new model of bioreactor, we investigated the optimal strategies of resource allocation between internal cellular processes maximizing the production of biomass or of a compound of interest. The model leads to an explicit expression of bacterial growth rate with respect to internal biological mechanisms instead of using the standard Monod relation.

Further work will consider model validation on experimental data. Other operating modes for the bioreactor like fed-batch are also still to be explored and tested. Indeed, by reconciling the infracellular and macroscopic scales, we can optimize simultaneously complex strategies of genetic modifications of bacterial strains (u_{e_i}) and of bioreactor control (feed rate).

Biologically, u_{e_i} is related to the cost of messenger RNAs (mRNAs) coding for the protein e_i. It aggregates all the gene transcription process (synthesis of the mRNA from the DNA sequence, the mRNA will further be translated by ribosomes into proteins). The biological realization of the optimal trajectory u_{e_i} remains obviously a biological challenge. However, our model could drive the biological implementation of synthetic u_{e_i}.

References

[1] G. Bastin and D. Dochain. *Adaptive control of bioreactors.* Elsevier, Amsterdam 1990.

[2] B. Srinivasan, D. Bonvin, E. Visser, and S. Palanki. Dynamic optimization of batch processes: II. role of measurements in handling uncertainty. *Computers & chemical engineering*, vol. 27, pp. 27–44,2003.

[3] W. Xie, L. Ye, X. Lv, et al. Sequential control of biosynthetic pathways for balanced utilization of metabolic intermediates in saccharomyces cerevisiae. *Metabolic engineering*, vol. 28, pp. 8–18, 2015.

[4] S. Waldherr, D. A. Oyarzún, and A. Bockmayr. Dynamic optimization of metabolic networks coupled with gene expression. *Journal of theoretical biology*, vol. 365, pp. 469–485, 2015.

[5] W.J. Holtz and J. D. Keasling. Engineering static and dynamic control of synthetic pathways. *Cell*, vol. 140, pp. 19–23, 2010.

[6] K. G. Gadkar, F. J. Doyle III, J. S. Edwards, and R. Mahadevan. Estimating optimal profiles of genetic alterations using constraint-based models. *Biotechnology and bioengineering*, vol. 89, pp. 243–251, 2005.

[7] A. Varma and B.O. Palsson. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Applied and environmental microbiology,* vol. 60, pp. 3724–3731, 1994.

[8] A. Goelzer, V. Fromion, and G. Scorletti. Cell design in bacteria as a convex optimization problem. *Automatica*, vol. 47, pp. 1210–1218, 2011.

[9] P. Unrean and C.J. Franzen. Dynamic flux balancing elucidates NAD(P)H production as limiting response to furfural inhibition in *Saccharomyces cerevisiae* *Biotechnology Journal*, vol. 10, pp. 1248-1258, 2015.

[10] A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences*, vol. 112, pp. 1038–1047, 2015.

[11] A. Goelzer, J. Muntel, V. Chubukov, et al. Quantitative prediction of genome-wide resource allocation in bacteria. *Metabolic engineering*, vol. 32, pp. 232–243, 2015.

[12] H.E. Kubitschek, W.W. Baldwin, S.J. Schroeter, and R. Graetzer. Independence of buoyant cell density and growth rate in *Escherichia coli*. *Journal of Bacteriology*, vol. 158, pp. 296–299, 1984.

[13] H. Bremer and P.P. Dennis. Modulation of chemical composition and other parameters of the cell by growth rate. In F.C. Neidhart, editor, *Escherichia coli and salmonella: cellular and molecular biology*, pp. 1553–1569. American Society of Microbiology Press, Washington D.C., USA, 2nd edition, 1996.

[14] R. Fletcher *Practical methods of optimization.* John Wiley & Sons, 2013.