Optimization of metabolic networks in biotechnology applications

Carlos Pozo, Gonzalo Guillén-Gosálbez*, Albert Sorribas and Laureano Jiménez

Departament d'Enginyeria Química (EQ), Escola Tècnica Superior d'Enginyeria Química (ETSEQ), Universitat Rovira i Virgili (URV), Campus Sescelades, Avinguda Països Catalans, 26, 43007 Tarragona, Spain

Problem statement

In recent years, the use of genetic manipulation techniques has led to significant improvements in the production of certain biochemical products. However, the intrinsic complexity of metabolic networks (see for instance *Figure 1*) makes the visual inference of the most promising genetic changes not a straight-forward task. Because of this, so far mutation and selection of new processes have been made in a trial-and-error basis (Polisetty et al., 2008), which has led to non-optimal solutions. Hence, one expects that actual biological processes could be further improved if quantitative design principles for the modification of the genes (i.e., which genes should be modified and how) were provided by a more rational approach like optimization (*metabolic engineering*).

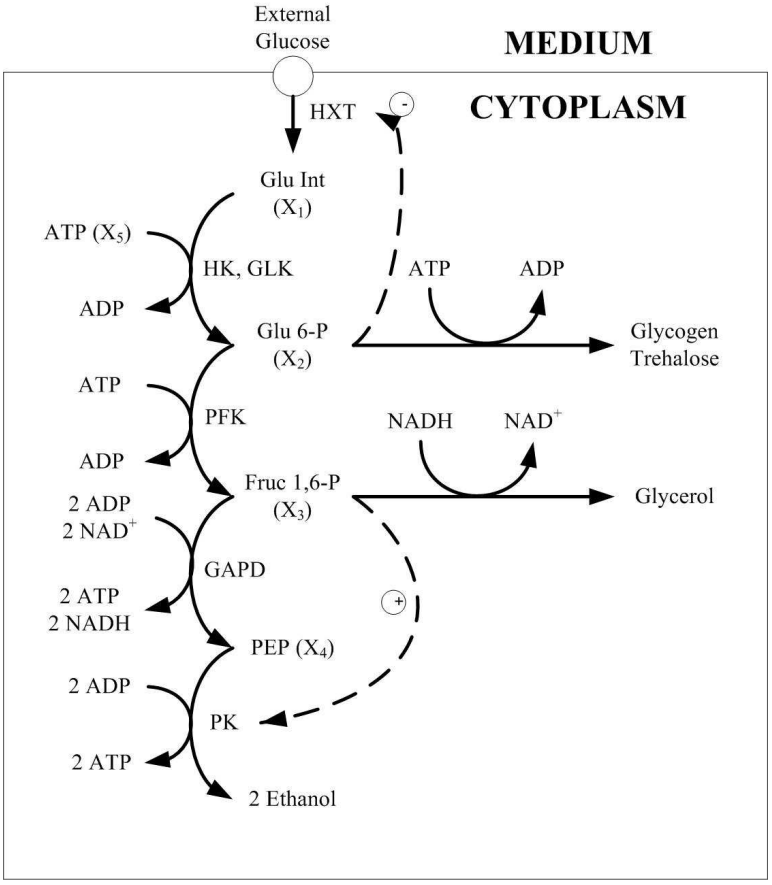

Figure 1. Fermentation pathway in Saccharomyces cerevisiae.

One of the key steps in this approach is the selection of an appropriate mathematical model. In this context, the use of the general mass action (GMA) formalism shows a good compromise between accuracy and simplicity. For a metabolic network containing n internal metabolites whose concentration X_i can vary with the time t due to the action of p flows, the GMA representations is as follows:

$$\frac{dX_i}{dt} = \sum_{r=1}^{p} \mu_{ir} v_r \qquad i = 1,\ldots,n \tag{1}$$

where μ_{ir} is the stoichiometric coefficient of the metabolite i in the process r and indicates the number of molecules of metabolite i involved in such a process. The velocity v_r can be described using different representations, but, the so-called power-law formalism (Voit, 1992) is an appropriate one:

$$v_r = \gamma_r \prod_{j=1}^{n+m} X_j^{f_{rj}} \qquad r = 1,\ldots,p \tag{2}$$

In this representation, γ_r is an apparent rate constant for flow r, and f_{rj} is the kinetic order of metabolite j in process r, which allows quantifying its effect on the considered rate. Note that contributions of the m (independent) external metabolites are also accounted for in this representation.

By introducing equation (2) into equation (1) and assuming steady state conditions for the metabolic network, one obtains a GMA model as follows:

$$\frac{dX_i}{dt} = \sum_{r=1}^{p} \left(\mu_{ir} \gamma_r \prod_{j=1}^{n+m} X_j^{f_{rj}} \right) = 0 \qquad i = 1,\ldots,n \tag{3}$$

The optimization of these systems aims at finding the appropriate changes in the enzyme activities (continuous variables) that optimize (maximize/minimize) a certain objective function (typically, the synthesis rate of the desired product). The enzyme activities obtained in the optimization solution can be implemented in the real system by tuning the expressions of the corresponding genes. The original formulation is an NLP model; however, binary variables can be added in order to limit the number of enzymes that can be modified simultaneously. This gives rise to an MINLP formulation.

References

- Polisetty, P., Gatzke, E., Voit, E., 2008. Yield optimization of regulated metabolic systems using deterministic branch-and-reduce methods. Biotechnology and Bioengineering 99(5), 1154-1169.
- Voit, E., 1992. Optimization in integrated biochemical systems. Biotechnology and Bioengineering 40(5), 572-582.