UDC 681.511.2:577.124.24:547.477.1:582.288
ISSN 1330-9862

original scientific paper

A Novel Approach to Design of Overexpression Strategy for Metabolic Engineering. Application to the Carbohydrate Metabolism in the Citric Acid Producing Mould *Aspergillus niger*

Néstor V. Torres[1,*], *Eberhard O. Voit*[2], *Carlos Glez-Alcón*[3] *and Felipe Rodríguez*[1]

[1] Grupo Tecnología Bioquímica y Control Metabólico, Departamento de Bioquímica y Biología Molecular, Facultad de Biología, Universidad de La Laguna, 38206 La Laguna, Tenerife, Islas Canarias, España.

[2] Department of Biometry and Epidemiology, Medical University of South Carolina, Charleston, South Carolina 29425-2503, USA.

[3] Departamento de Estadística, Investigación Operativa y Computación. Facultad de Matemáticas, Universidad de La Laguna, 38206 La Laguna, Tenerife, Islas Canarias, España.

Received: August 24, 1997
Accepted: June 17, 1998

Summary

A metabolic model for the production of citric acid in Aspergillus niger is optimized with respect to the production rate of citrate. The model makes use of the S-system representation of biochemical systems, which renders it possible to use linear programming to optimize the process. While originally nonlinear, the optimization problem is reduced in such a way that methods of linear programming can be used. The optimizations lead to profiles of enzyme activities that are compatible with the physiology of the cells, which guarantees their viability and fitness, and yield higher rates of the desired final end products than the original systems. It was found that maintaining the metabolite pools within narrow physiological limits (20% around the basal steady state level) and allowing the enzyme concentrations to vary within the range of 0.1 to 50 times their basal values it is possible to triple the glycolytic flux while maintaining 100% yield of substrate transformation. To achieve these improvements it is necessary to modulate seven or more enzymes simultaneously.

Keywords: optimization, linear programming, metabolic engineering, citric acid, *A. niger*

Introduction

A goal in any biotechnological design for obtaining of a given product is its optimization. Until now the most important method used for this purpose has been the improvement of strains by the classical techniques of random mutation and screening. However, this approach has reached its limits in many cases *(1)*. This scenario is currently changing due to the developments in molecular genetics, which give us the opportunity to introduce heterologous genes and regulatory elements. This situation configures the new paradigm of metabolic engineering *(2)*.

However the combination of the information on biosynthetic pathways and the experimental tools provided by molecular genetics are not sufficient by themselves to design an efficient strategy of genetic manipulation aimed at the overproduction of a certain product. Traditionally, models have been represented within the Michaelis-Menten formalism, which has proved useful for the characterization of isolated reaction mechanisms *in vitro*. These models can, in principle, address questions about which enzymes should be altered in order to improve a desired product or flux. However, the ra-

* Author to whom correspondence should be addressed: tel./fax: +34-922-318334, E-mail: ntorres@ull.es

tional functions underlying these approaches have turned out to be mathematically inconvenient *(3–5)*. In particular, the complex nonlinear structure of these models has limited the attempts to optimize metabolic systems with mathematical methods to a small number of cases with poor results *(4-6)*. What it is needed thus is a theoretical framework able to integrate the available information and to direct the exploration of the best changes to be made. In this respect some theoretical approaches involving the quantitative understanding of the cellular metabolism have been applied recently to overcome these limitations. In one of them, the Metabolic Control Analysis, Kacser and Acerenza *(7)* presented a method for achieving increases in metabolite production, and more recently Stephanopoulos and Vallino *(8)*, and Small and Kacser, *(9)* have developed general concepts and methodologies to improve cellular properties of microorganisms in a directed fashion. On the other hand, now within the framework of the Biochemical System Theory (BST) *(10-12)*, optimization methods were developed for biochemical processes represented as S-systems *(13-15)*. The great advantage of optimizations with S-systems is the fact that the optimization problem is strictly linear, even though S-system models themselves are nonlinear and rich enough to model virtually any set of differentiable functions or differential equations *(16)*. In this paper an alternative modeling and optimization approach based on S-system models directly derived from Biochemical System Theory will be presented and the application of this method to the citric acid production by *Aspergillus niger* will be illustrated.

The Optimization Approach: Application to the Maximization of Citric Acid Production by *Aspergillus niger*

The implementation of the method can be divided into two main steps.

1. Design of the mathematical model, translation into the corresponding S-system form and quality assessment

The first step in this process is the formulation of a kinetic model of the pathway under consideration. Biochemical pathways are traditionally formulated as differential equations containing Michaelis-Menten rate laws and their generalization. The resulting models integrate kinetic data and other available information about metabolites and effector concentrations, as well as fluxes obtained from experimental observations *in vivo* or *in vitro*.

An alternative to this approach is the formulation of a pathway in the form of an S-system *(12,17,18)*, which is possible in a straightforward analytical manner *(12,18)*. Once the model is in S-system form it is possible to carry out both analytical *(10-12)* and numerical quality assessment, which involves the investigation of: *i)* the stability of the steady-state; *ii)* the robustness of the model, defined as whether the model is able to tolerate small structural changes; and *iii)* the dynamic features characterizing transient responses of the system to perturbations.

In our present case, we develop an S-system model of the carbohydrate metabolism in *A. niger* in conditions of citric acid production *(19,20)*. The model addresses the citric acid production stage, when the system attains a well-characterized steady-state in which citrate production is the only activity of quantitative importance *(21)*. Fig. 1 shows a mechanistic representation of the system that includes most features of documented relevance *(21,22)*. Most of the data used in the model were obtained from the *Aspergillus niger* strain B60, which has been extensively described in the literature *(21)*.

In the S-system model, each net rate law for synthesis and degradation is represented by a product of power-law functions of all variables that have an influence upon the net rate law in question. The rate of a given process, V_i is written as:

$$V_i = \alpha \prod_{j=1}^{n+m} X_j^{g_{i,j}} \qquad /1/$$

where X_j are variables modeling enzymes, metabolites, or effectors, that affects the rate in question. The indices $j = 1,...,n$ refer to dependent variables, while $j = n+1,...,$ $...,n+m$ refer to independent variables. Those processes that characterize reactions forming a metabolite X_i are aggregated to give a single rate law for net synthesis, V_i. Similarly, those rate laws that characterize reactions removing a metabolite X_i are aggregated to give a single rate law for net degradation, V_{-i}. The descriptive equations for a biochemical system can then be written in terms of power-law functions as follows:

$$\frac{dX_i}{dt} = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{i,j}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{i,j}}, \quad i = 1,\ldots,n \qquad /2/$$

The parameters α_i and $g_{i,j}$ are the rate constants and kinetic orders associated with the rate law for net synthesis of X_i. Similarly β_i and $h_{i,j}$ are associated with the rate law for net degradation of X_i. The kinetic order parameters $g_{i,j}$ and $h_{i,j}$ are defined as:

$$g_{i,j} = \left(\frac{\partial V_i}{\partial X_j} \frac{\partial X_j}{\partial V_i} \right)_0 \qquad /3/$$

$$h_{i,j} = \left(\frac{\partial V_{-i}}{\partial X_j} \frac{\partial X_j}{\partial V_{-i}} \right)_0 \qquad /4/$$

where the additional index 0 indicates that a quantity is measured at the steady-state level of metabolite concentration.

The steady-state solution of equation /2/ can be obtained by analytical methods *(10,11)* and the steady-state flux through any pool X_i is directly obtained as a product of power-law functions in the known concentrations at steady-state. The S-system representation that derives from the pathway characteristics and from the available kinetic information is *(19,20)*:

$$dX_1/dt = 0.459\ X_8^{0.23}\ X_9^{0.0009}\ X_{10} - 1.833\ 10^3\ X_1^{2.194}\ X_2^{-1.333}\ X_{11}^{0.2}\ X_{12}^{0.8} \qquad /5/$$

$$dX_2/dt = 3.25\ 10^{-4}\ X_1^{2.665}\ X_2^{-1.667}\ X_{12} - 1.861\ X_2^{0.9}\ X_8^{0.056}\ X_{13}\ X_{21}^{0.65}\ X_{22}^{-0.3}$$

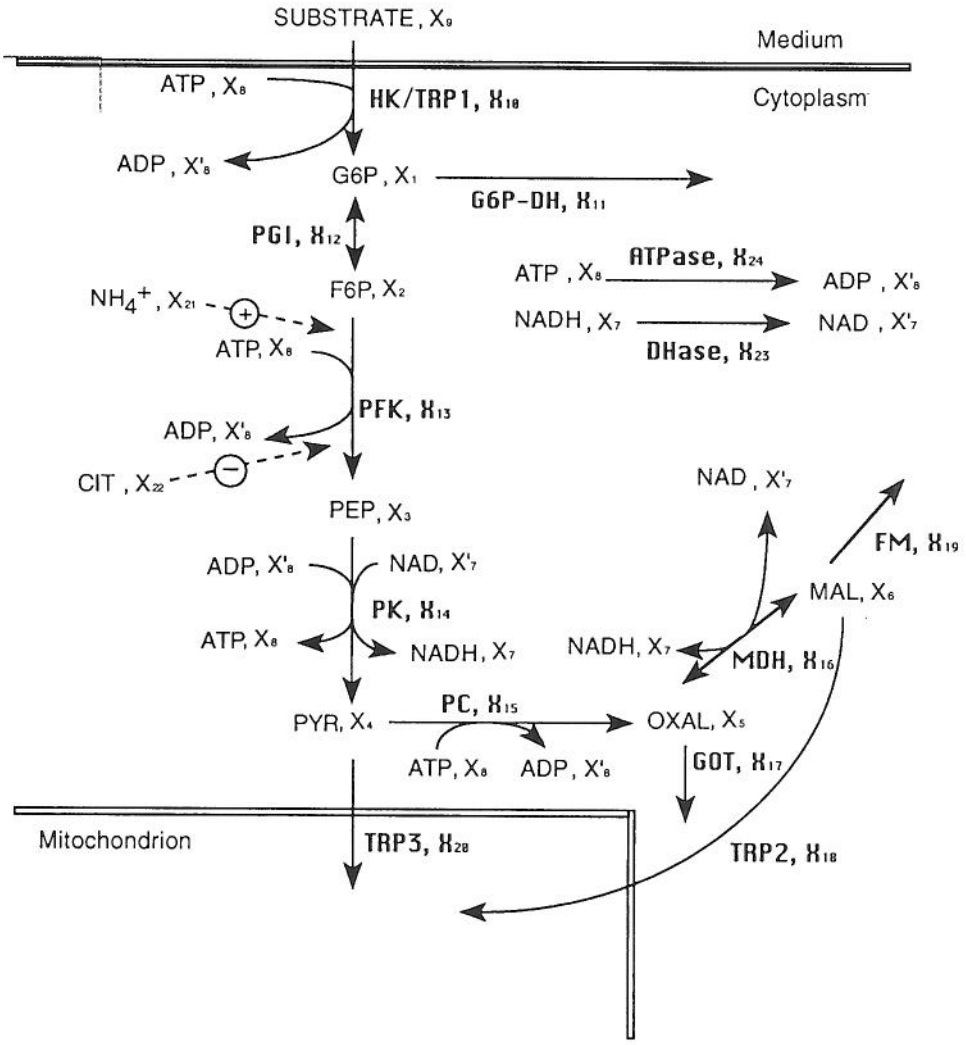

Fig. 1. Mechanistic model of the carbohydrate metabolism in *Aspergillus niger* under conditions of citric acid production. Arrows indicate reactions. Dependent variables (thin typeface) are numbered from 1-8: Substrate,X_9; G6P,X_1, glucose-6-phosphate; F6P,X_2, fructose-6-phosphate; PEP,X_3, phosphoenol pyruvate; PYR,X_4, pyruvate; OXAL,X_5, oxalacetate; MAL,X_6, malate; NADH,X_7, NADH; NAD,X'_7, NAD; ATP,X_8, ATP and ADP,X'_8, ADP. Independent variables (bold typeface) are numbered from 9-24: HK/TRP1,X_{10}, hexokinase-substrate transport step; G6P--DH,X_{11}, glucose-6-phosphate dehydrogenase; PGI,X_{12}, phosphoglucose isomerase; PFK,X_{13}, phosphofructokinase; PK,X_{14}, pyruvate kinase; PC,X_{15}, pyruvate carboxylase; MAL,X_{16}, malate dehydrogenase; GOT,X_{17}, aspartate aminotransferase; TRP2,X_{18}, malate transport; FM,X_{19}, fumarase; TRP3,X_{20}, pyruvate transport; NH_4^+,X_{21}, ammonium; CIT,X_{22}, citrate; DHase,X_{23}, dehydrogenases; ATPase,X_{24}, ATPases. For further information see *(21,22)*.

$$dX_3/dt = 1.861\ X_2^{0.9}\ X_8^{0.056}\ X_{13}\ X_{21}^{0.65}\ X_{22}^{-0.3} - 0.21\ X_3^{0.9}\ X_7^{-0.028}\ X_8^{-3.196}\ X_{12}^{0.8}\ X_{14}$$

$$dX_4/dt = 0.42\ X_3^{0.9}\ X_7^{-0.028}\ X_8^{-3.196}\ X_{14} - 2.538\ X_4^{0.836}\ X_8^{0.161}\ X_{15}^{0.509}\ X_{20}^{0.486}$$

$$dX_5/dt = 1.895\ X_4^{0.7}\ X_8^{0.315}\ X_{15} - 4.23\ 10^{14}\ X_5^{3.2}\ X_6^{-2.2}\ X_7^{3.788}\ X_{16}^{0.945}\ X_{17}^{0.054}$$

$$dX_6/dt = 2.04\ 10^{15}\ X_5^{3.333}\ X_6^{-2.33}\ X_7^{4.005}\ X_{16} - 0.005\ X_6^{0.813}\ X_{18}^{0.979}\ X_{19}^{0.02}$$

$$dX_7/dt = 0.204\ X_3^{0.873}\ X_7^{-0.028}\ X_8^{-4.06}\ X_{14} - 8.68\ 10^6\ X_5^{1.618}\ X_6^{-1.136}\ X_7^{2.203}\ X_{16}^{0.486}\ X_{23}^{0.513}$$

$$dX_8/dt = 0.841\ X_3^{0.9}\ X_7^{-0.028}\ X_8^{-3.196}\ X_{14} - 0.308\ X_2^{0.225}\ X_4^{0.179}\ X_8^{0.257}\ X_9^{0.294}\ X_{10}^{0.312}\ X_{13}^{0.25}\ X_{15}^{0.256}\ X_{21}^{0.162}\ X_{22}^{-0.075}\ X_{24}^{0.18}$$

The numerical values of the kinetic orders $g_{i,j}$ and $h_{i,j}$ and of the rate constants α_i and β_i were obtained from published experimental data as shown elsewhere *(19)*. The steady-state solution of this model is consistent with experimental observations *(20)*, and eigenvalue analysis *(10,11)* confirms its local stability.

2. *Linearization, linear programming and optimization*

The key advantage of formulating the biochemical pathway as an S-system model is the fact that the steady-state in this representation is characterized by a system of linear algebraic equations *(11)*. S-system models are characterized by a system of nonlinear equations which become linear, under a logarithmic transformation, in the steady-state *(11)*. In addition to the steady-state equations, typical objective functions and constraints on metabolites, activities and stoichiometric relationships can be formulated as linear equations and inequalities, such that the problem reduces to one of straightforward linear optimization. Therefore optimization of yields and fluxes can be carried out with readily available linear optimization programs *(15)*. Upon formulating a biochemical pathway as an S-system, the problem of optimizing a particular flux under typical constraints reads as follows:

Linear Program

(1) maximize ln(flux)

subject to

(2) steady-state equations, expressed in logarithms of variables

(3) ln(dependent or independent variable) ≤ constant

(4) ln(dependent or independent variable) ≥ constant

(5) ln(dependent or independent variable) = constant

(6) ln(variable) unrestricted

(7) ln(flux) ≤ constant

(8) ln(flux) ≥ constant

(9) ln(flux) unrestricted

(10) ln(flux1/flux2) ≤ constant

In this formulation, (1) is a typical objective function that is linear in the logarithms of the involved dependent and independent variables. Constraint (2) assures that the optimized system is in a steady-state, no matter what the altered enzyme concentrations are, while (3) and (4) constrains variables to stay within certain limits. Constraint (5) forces the variable to be at a given value, whereas (6) is an option that permits any real value for the logarithm of a variable and thus any positive real value for the variable itself. Constraints (7) – (9) are the corresponding ones on fluxes, and (10) states that the logarithm of the flux ratio flux1/flux2 should remain below a certain limit. The optimization is executed with any of the available linear optimization packages, such as LINDO PC 5.3 (LINDO Systems Inc.) or KLP (Kinetics Software, 1390 Fell Street, 103, San Francisco, CA 94117).

The implementation of the above in our present case is as follows.

(1) Objective function. – Even though there is a direct correlation between the glycolytic flux and citrate synthesis *(23)* a more accurate criterion of productivity is maximization of the reactions that synthetize the precursors of citrate, namely, the synthesis of mitochondrial pyruvate ($V_{4,0}$) and malate ($V_{6,0}$). To implement this objective we first define the corresponding rate equations. Expressed in terms of the power-law formalism, they take the form:

$$V_{4,0} = \alpha_4 \cdot X_4^{g_{4,4}} X_{20}^{g_{4,20}}$$
$$V_{6,0} = \alpha_6 \cdot X_6^{g_{6,6}} X_{18}^{g_{6,18}} \qquad /6/$$

From the fluxes and their kinetics characteristics at the basal steady state it is a straightforward process to calculate the rate constants α_i' and the kinetic orders $g_{i,j}$ (i = 4,6; j = 4,6,18,20) *(10,19,20,24)*. As a result, we obtain the following expressions:

$$V_{4,0} = 0.8697\ X_4^{0.98} X_{20}^{1}$$
$$V_{6,0} = 0.0047\ X_6^{0.83} X_{18}^{1} \qquad /7/$$

At the same time, we need to take into account the stoichiometry of the citrate synthase reaction, which in our model consumes a molecule of mitochondrial malate and a molecule of mitochondrial pyruvate to produce a molecule of citrate, and impose the additional constraint that these two rates should be equal. This requires in our notation:

$$V_{4,0} = V_{6,0} \qquad /8/$$

After taking logarithms in Eqs. /6/ and rearranging terms we obtain

$$0.98\ y_4 + y_{20} - 0.83\ y_6 - y_{18} = -5.2031 \qquad /9/$$

as the new stoichiometric constraint.

To optimize the system under these conditions, we thus maximize

$$\ln(0.8697) + 0.98\ y_4 + y_{20} \qquad /10/$$

and require that the equal rate constraint Eq. /9/ be satisfied.

(2) Steady-state constraints. – The second step is the formulation of the steady-state equations expressed in terms of the logarithms of the variables. In both cases above considered these constraints take the form:

$$-2.194\ y_1 + 1.333\ y_2 + 0.23\ y_8 + 0.0009\ y_9 + y_{10} - 0.2\ y_{11} - 0.8\ y_{12} = -5.523$$

$$2.665\ y_1 - 2.567\ y_2 - 0.056\ y_8 + y_{12} - y_{13} - 0.65\ y_{21} + 0.33\ y_{22} = 8.651$$

$$0.9\ y_2 - 0.9\ y_3 + 0.028\ y_7 + 3.252\ y_8 + y_{13} - y_{14} + 0.65\ y_{21} - 0.3\ y_{22} = -2.179$$

$$0.9\ y_3 - 0.836\ y_4 - 0.028\ y_7 - 3.357\ y_8 + y_{14} - 0.509\ y_{15} - 0.486\ y_{20} = 1.797$$

$$0.7\ y_4 - 3.2\ y_5 + 2.2\ y_6 - 3.788\ y_7 + 0.315\ y_8 + y_{15} - 0.945\ y_{16} - 0.054\ y_{17} = 33.039$$

$$3.33\ y_5 - 3.143\ y_6 + 4.005\ y_7 + y_{16} - 0.979\ y_{18} - 0.0204\ y_{19} = -40.447 \qquad /11/$$

$$0.873\ y_3 - 1.618\ y_5 + 1.132\ y_6 - 2.231\ y_7 - 4.06\ y_8 + y_{14} - 0.486\ y_{16} - 0.513\ y_{23} = 17.563$$

$$-0.225\ y_2 + 0.9\ y_3 - 0.179\ y_4 - 0.028\ y_7 - 3.454\ y_8 - 0.294\ y_9 - 0.312\ y_{10} - 0.25\ y_{13} + y_{14} - 0.256\ y_{15} - 0.162\ y_{21} + 0.075\ y_{22} - 0.180\ y_{24} = -1.003$$

where $y_i = \ln(X_i)$ for i = 1,2,3,...,24.

(3) Constraints on enzyme concentrations. – Next we determine which enzyme concentrations are allowed to vary. In the first step, we needed to know the best possible optimized solution, and allowe all enzymes to change between 10% its basal value and 50 times that value. This range of variation is consistent with current DNA recombinant techniques. The only exceptions were X_{11}, X_{17} and X_{19}. These enzymes divert flux from the pathway leading to citrate and were maintained constant at the basal values (we assumed that their activities are necessary to guarantee cell viability). By the same token, the phosphofructokinase effectors citrate (X_{22}) and ammonium (X_{21}) were kept constant at their basal values. The mathematical formulation of these constraints in the logarithmic domain is:

$$\begin{aligned} y_9 &= 5.669; & y_{17} &= -3.0365 \\ -4.135 \le y_{10} \le -2.525; & & -2.919 \le y_{18} \le 1.686 & \\ y_{11} &= -4.074; & y_{19} &= -3.912 \\ 0.425 \le y_{12} \le 5.030; & & -5.404 \le y_{20} \le -0.798 & \\ -4.893 \le y_{13} \le -0.288; & & y_{21} &= 0.854 \\ -4.120 \le y_{14} \le -1.897; & & y_{22} &= 1.726 \\ -6.502 \le y_{15} \le -1.897; & & -2.302 \le y_{23} \le 2.302 & \\ -0.580 \le y_{16} \le 4.025; & & -2.302 \le y_{24} \le 2.302 & \end{aligned} \qquad /12/$$

(4–6) Constraints on metabolite pools. – It is also necessary to limit the range of variation of the metabolite concentrations. The overexpression of enzymes normally associated with the overproduction of a given product provokes changes in the intermediate metabolite pools that have kinetics effects on the rest of the metabolism. This usually leads to poor yield in product synthesis. For this reason we set limits for the ranges of these metabolite pools; they were allowed to vary up to 20% about the steady-state level, a threshold small enough to ensure no significant changes in the other processes affecting the overall yield. The mathematical formulation of these constraints in the logarithmic domain is:

$$\begin{aligned} -3.912 \le y_1 \le -0.693; & & -9.755 \le y_5 \le -6.536 & \\ -5.298 \le y_2 \le -2.079; & & -0.862 \le y_6 \le 2.356 & \\ -6.214 \le y_3 \le -2.995; & & -4.961 \le y_7 \le -1.742 & \\ -3.729 \le y_4 \le -0.510; & & -2.302 \le y_8 \le 0.916 & \end{aligned} \qquad /13/$$

Since the optimization itself does not address questions of stability, the steady-state solution of the S-system must be checked with respect to stability and possibly with respect to robustness. Both types of analyses can be executed analytically *(11,12)* or with computational means. Unstable steady-state solutions are normally discarded. Since the actual fermentation system is stable, emerging instabilities in the altered system must be due to constraints on some variables or fluxes that are too slack.

(7–10) Constraints on fluxes. – In the optimization process the system deviates rather far from the original steady-state operating point and as a consequence of the particular approximation that underlies the S-system model, some of the stoichiometric relationships that hold in the basal solution are no longer satisfied. In order to avoid significant stoichiometric violations we add a restriction that reflects the required dependency between fluxes. We demand that the flux $V_{3,4}$ be less than twice the input flux $V_{9,1}$:

$$\frac{V_{3,4}}{2V_{9,1}} < 1 \qquad /14/$$

which is equivalent to:

$$\frac{V_{9,1}}{V_{3,4}} > 0.5 \qquad /15/$$

From Eq. /5/ the power law expressions for $V_{9,1}$ and $V_{3,4}$ are obtained. After taking logarithms and rearranging we obtain:

$$-0.9\ y_3 + 0.0281\ y_7 + 3.426\ y_8 + 0.00093\ y_9 + y_{10} - y_{14} > -0.780 \qquad /16/$$

as the flux stoichiometric constraint. To optimize the system under these conditions, we thus maximize Eq. /6/ and require that the equal rate constraint Eq. /9/ and the flux stoichiometric constraint Eq. /16/ be satisfied.

Results

The optimum solution

Fig. 2 shows the optimized solution. In this solution the glucose uptake rate ($V_{9,1}$; 0.016 μmol/min mg dry weight) increases (2.54 times) with respect to the basal solution, and at the same time the citrate rate ($V_{6,0}$; 0.016 μmol/min mg dry weight) production is 3.33 times higher than the basal solution. Moreover, the final balance shows that all substrate is converted into citrate. Thus, at the steady-state of the optimum solution the organism consumes more glucose and produces citrate at a higher rate and with a better yield than under the original, basal conditions.

In the optimum solution the system is more efficient than in the basal steady-state; the highest amplification factor is 5.57 (X_{21}; ATPase), whereas the others range between 1 and 5. That means that no drastic changes are needed to increase citrate production by more than three times. It is also noted that the fluxes through G6P-DH (X_{11}), GOT (X_{17}), and fumarase (X_{19}), remain almost unchanged (around 20% of the basal steady-state). This is due to the fact that the activities are kept unchanged for all enzymes that divert flux from the main glycolytic stream. This quasi-constancy is crucial for cellular viability since it ensures that the remaining metabolic processes proceed in the same way at the basal steady-state level.

In summary, one can conclude that the optimized solution produces citrate at a rate more than three times greater than the basal system. Furthermore, it consumes more substrate but does not require significantly higher protein synthesis than the basal solution. The yield of substrate transformation is 100% and thus higher than in the basal solution.

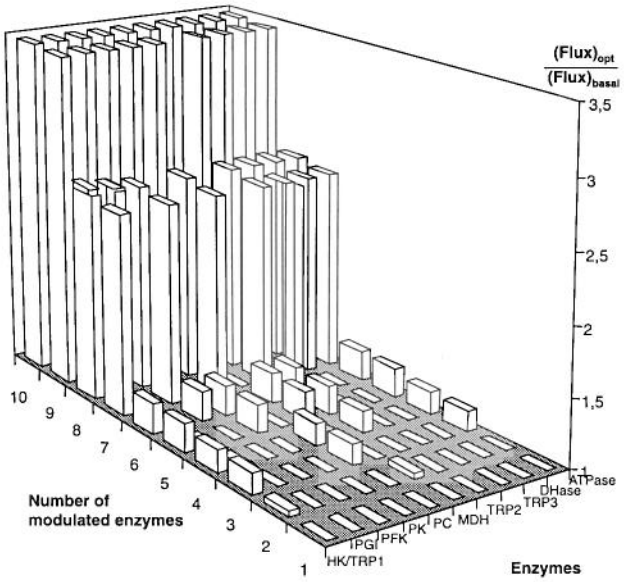

Fig. 2. Optimized enzyme profile in the carbohydrate metabolism of *A. niger*. The objective function maximizes the cytosolic pyruvate synthesis $V_{3,4}$, and the equality of the flux rates $V_{4,0}$ and $V_{6,0}$. Also a flux stoichiometric constraint, $V_{9,1}/V_{3,4} > 0.5$ was imposed. The metabolite pools were always within 20% about the basal steady-state level and the enzymes activities were allowed to vary between 0.1 and 50 times the basal levels. Only variables allowed to vary are shown. Enzyme activities and precursors fluxes are given in μmol/min mg dry weight and concentrations in mM. See text for discussion.

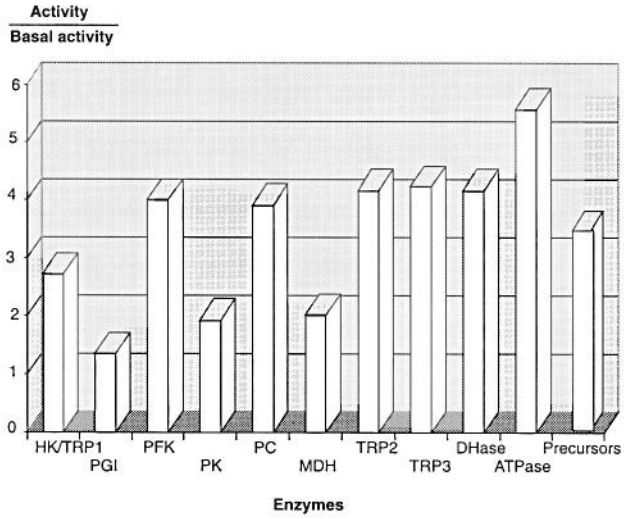

Fig. 3. Optimum solutions obtained with combinations of one to ten enzymes involved in the carbohydrate metabolism of *A. niger* under conditions of citric acid accumulation. Metabolite pools were allowed to vary up to 20% about the basal steady-state, while the enzyme activities were allowed to vary from 0.1 to 50 times the basal steady-state activity. See text for discussion.

Optimization of subsets of enzymes

The previous section assumed that we had the capabilities of changing any or all enzymes to prescribed degrees. Arguments of practicability suggest that this is rarely the case. In reality, some enzyme activities can be altered relatively easy, whereas others are very difficult to manipulate.

We systematically searched the best solution involving up to ten enzymes or transport steps. Fig. 3 shows the best solution obtained in each case. It is evident that no significant increase in the rate of citric acid production can be obtained unless at least seven enzymes are modulated (2.48 times increase over the basal rate). Most of these improved solutions involve alterations in X_{15}, X_{18} and X_{24}, and all except one involve X_{10}. This is an important finding that will be discussed below.

Discussion

While some of the results on citric acid production are not surprising, others are interesting by themselves, and they also indicate the potential of the proposed method. First, the results demonstrate that the method is working and produces reasonable results. Second, they offer some insight in the design of the pathway, because the identification of enzymes or transport steps that are absolutely necessary for improvements in yield, in a sense, is analogous to an analysis in which enzymes of a pathway are rate limiting. Finally, the results provide some indication of how difficult it might be to implement an optimized solution *in vivo*. For instance, our analysis suggests that significant improvements in citric acid yield require the simultaneous modification of at least seven enzymes. Furthermore, alterations in four particular activities (hexokinase-substrate transport, pyruvate carboxylase, malate transport and ATPases) seem almost unavoidable. Such information is very valuable since it lets the researcher estimate a projected benefit against the time and cost effort necessary to improve the performance of the process of interest. The analysis also reveals which enzyme and transport activities are rather insignificant with respect to improvements in yield. Even through technologies for manipulating these steps may be available and cheap, such effort would be wasted.

Application of this approach for the design of an overexpression strategy to the production of citric acid by *A. niger* implies that the metabolic pathway leading to the accumulation of citric acid is already very close to optimal. This is evident in the rate of citric acid production and in yield of glucose transformation into citric acid, where the wild strain is superior to any of the optimized solutions. Given, however, that in the case of biotechnological production the initial substrate consists of cheap, low quality sugars, it is more profitable to allow for a high rate of citrate production than a bigger yield, particularly when the yield transformation in the optimized solution is also significant. This situation is not very surprising if we take into account that the whole process and the strain have been the subject of many studies aimed at improving efficiency. It should be stressed, though, that other processes such as the synthesis of antibiotics by filamentous fungi, is less optimized, both in terms of rate of production and yield. The constrained optimization in the case of citric acid production essentially confirmed experimental expectations. It did not produce surprising insights, but it is encoura-

ging that the method produced reasonable results. It is hoped that similar analyses will suggest quantitative means for improving yield in other microbial processes that reach beyond experimental intuition.

Of course, there is no absolute guarantee that the suggested enzyme modifications will qualitatively and quantitatively produce the predicted results. Assumptions in setting up the model, uncertainties in the kinetic data used for parameter estimation, and effectors that for simplicity are considered constant could affect the altered system. It should be noted here that the S-system formulation introduces an error in the flux stoichiometry, as soon as we deviate from steady-state. In this study the optimized solutions are away from the operating point and some of the stoichiometric constraints are violated. In our study we have imposed all the stoichiometric constraints but in this case the system was unable to move away from the basal solution. The crucial point is to make sure that important stoichiometric constraints are not violated too badly. The rationale is that the pathway is not isolated within *A. niger* and that any excess material will be easily absorbed and that the substrate demand satisfied from outside, especially in case of compounds like ATP or NAD.

Which constraints are active in a particular case depends on the numerical specification of the model. In order to investigate the question in our case, it is useful to distinguish between constraints imposed on metabolites and on enzymes on the one hand and the stoichiometric constraints on the other. In all optimizations considered here, increasing the boundaries for metabolites caused significant increases in the optimal production rates (results not shown). By contrast, increase in the upper limits of enzyme concentration did not affect the optimal solutions. Thus, in this particular system, the metabolite constraints turned out to be limiting. Leaving the system »stoichiometrically« unconstrained caused the optimum solutions to attain larger values of the production rates (results not shown). However in a recent study we applied a non-linear direct optimization approach, based on a stochastic multi-start search algorithm, to a related system, the production rates of ethanol, glycerol and carbohydrates by *Saccharomyces cerevisiae (25)*. When this optimization technique was applied to the Michaelis-Menten model representations of the system without imposing any constraint, we found a close agreement between the results obtained using the IOM method also without stochiometric constraint imposed. In fact, qualitatively, both optimization approaches rendered the same profile of enzymes to be modulated, while quantitatively, agreement was almost total.

Furthermore, it is important to be aware that the results are approximations, and if we predict an increase by a factor of 3, this only means that the increase is around this figure. Finally, it is also imaginable that secondary pathways that are irrelevant under natural conditions could assume a dominant role in the altered system and that this would decrease or annihilate the predicted improvements.

In spite of these caveats, we believe that the proposed method has its merit. It is a quantitative tool that integrates all available information in a logical fashion and can serve as a screening tool that has a good chance of providing a reasonable ranking of the importance of enzymes for optimizing processes of biotechnological relevance.

Acknowledgements

This work has been supported by a research grant from CICYT, ref. BIO96-1458 and by a grant from the Gobierno de Canarias, ref. COF 1997/08.

List of symbols

Metabolites

G6P: glucose-6-phosphate
F6P: fructose-6-phosphate
PEP: phosphoenol pyruvate
PYR: pyruvate
OXAL: oxalacetate
MAL: malate
CIT: citrate

Enzymes and transport steps

HK/TRP1: hexokinase-substrate transport system
G6P-DH: glucose-6-phosphate dehydrogenase
PGI: phosphoglucose isomerase
PFK: phosphofructokinase
PK: pyruvate kinase
PC: pyruvate carboxylase
MDH: malate dehydrogenase
GOT: aspartate aminotransferase
TRP2: malate transport
FM: fumarase
TRP3: pyruvate transport
Dhase: dehydrogenases

References

1. N. H. Barton, M. Turelli, *Annu. Rev. Genet. 23* (1989) 337-370.
2. J. Bailey, *Science, 252* (1991) 1668-1674.
3. M. Savageau: A critique of the enzymologist's test tube. In: *Fundamental of Medical Cell Biology*, E. E. Bittar (Ed.), *Vol. 3A*, JAI Press Inc., Greenwich, Connecticut (1992) pp. 45-108.
4. S. Schuster, R. Heinrich, *J. Math. Biol. 29* (1991) 425-442.
5. S. Schuster, R. Schuster, R. Heinrich, *J. Math. Biol. 29* (1991) 443-455.
6. G. Pettersson, *Eur. J. Biochem. 206* (1992) 289-295.
7. H. Kacser, L. Acerenza, *Eur. J. Biochem. 216* (1993) 361-367.
8. G. Stephanoupoulos, J. J. Vallino, *Science, 252* (1991) 1675-1681.
9. J. R. Small, H. Kacser, *Eur. J. Biochem. 226* (1994) 649-656.
10. M. A. Savageau, *J. Theor. Biol. 25* (1969) 365-369.
11. M. A. Savageau, *J. Theor. Biol. 25* (1969) 370-379.
12. M. Savageau: *Function & Design in Molecular Biology*, Addison-Wesley, Reading, Mass (1976).
13. C. Hatzimanikatis, C. Floudas, J. E. Bailey, *AIChE J., 42* (1996) 1277-1292.
14. L. Regan, I. D. L. Bogle, P. Dunhill, *Comp. Chem. Engin. 175-6* (1993) 627-637.

15. E. O. Voit, *Biotechnol. Bioeng. 40* (1992) 572-582.
16. M. A. Savageau, E. O. Voit, *Math. Biosci. 87* (1987) 83-115.
17. F. Shiraishi, M. A. Savageau, *J. Biol. Chem. 267* (1992) 22934-22943.
18. E. O. Voit: *Canonical Nonlinear Modeling. S-system Approach to Understanding Complexity*, Van Nostrand Reinhold, New York (1991).
19. N. V. Torres, *Biotechnol. Bioeng. 44* (1994) 104-111.
20. N. V. Torres, *Biotechnol. Bioeng. 44* (1994) 112-118.
21. C. P. Kubicek, M. Röhr, *CRC Crit. Rev. in Biotechnol. 3* (1986) 331-373.
22. M. Mattey, *CRC Crit. Rev. Biotechnol. 12* (1992) 87-132.
23. G. Schreferl-Kunar, M. Grotz, M. Rohr, C. P. Kubicek, *FEMS Microbiol. Lett. 59* (1989) 297-300.
24. A. Sorribas, M. A. Savageau, *Math. Biosci. 86* (1989) 127-145.
25. F. Rodríguez-Acosta, C. Regalado, N. V. Torres, *J. Biotechnol.* (1998) In press.

Modeliranje prekomjerne ekspresije u metaboličkom inženjeringu s primjenom na metabolizam ugljikohidrata u proizvodnji limunske kiseline s pomoću plijesni *Aspergillus niger*

Sažetak

Optimiran je metabolički model brzine proizvodnje limunske kiseline u plijesni Aspergillus niger. Biokemijski su sustavi u modelu prikazani u S-sustavu, što omogućava primjenu linearnog programiranja u optimiranju procesa. Iako je izvorni model nelinearan, pretvorbom u S-sustav problem je pojednostavljen tako da omogućava primjenu metoda linearnog programiranja. Ono daje takve profile enzimskih aktivnosti koji su kompatibilni s fiziologijom stanica, što jamči njihov visok stupanj preživljavanja i izdržljivosti, a i veće prinose željenih gotovih proizvoda od izvornih sustava. Ustanovljeno je održavanje metaboličkih puteva unutar uskih fizioloških granica (20%-tno odstupanje od osnovne ravnotežne razine) dopuštajući da koncentracije enzima variraju u rasponu 0,1–50 puta od njihovih osnovnih vrijednosti, što omogućuje utrostručenje glikolitičkog toka uz zadržavanje 100%-tne pretvorbe supstrata. Da bi se postigla ta poboljšanja, potrebno je modulirati sedam ili više enzima.