An Indirect Optimization Method for Biochemical Systems: Description of Method and Application to the Maximization of the Rate of Ethanol, Glycerol, and Carbohydrate Production in *Saccharomyces cerevisiae*

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

[1]Grupo de 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; telephone and fax: +34-22-603704; e-mail: ntorres@ull.es
[2]Department of Biometry and Epidemiology, Medical University of South Carolina, Charleston, South Carolina 29425-2503
[3]Departamento de Estadística, Investigación Operativa y Computación. Facultad de Matemáticas, Universidad de La Laguna, 38206 La Laguna, Tenerife, Islas Canarias, España

Received 23 July 1996; accepted 18 January 1997

Abstract: Three metabolic models for the production of ethanol, glycerol, and carbohydrates in yeast are optimized with respect to different production rates. While originally nonlinear, all three optimization problems are 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. In order to increase ethanol rate production at least three times, six enzymes must be modulated. By contrast, when the production of glycerol or carbohydrates is optimized, modulation of just one enzyme (in the case of glycerol) or two enzymes (in the case of carbohydrates) is necessary to yield significant increases in product flux rate. Comparisons of our results with those obtained from other methods show great similarities and demonstrate that both are valid methods. The choice of one or the other method depends on the question of interest. © 1997 John Wiley & Sons, Inc. *Biotechnol Bioeng* **55**: 758–772, 1997.

Keywords: Optimization; metabolic systems; linear programming; S-system representation; ethanol; glycerol; carbohydrates; *Saccharomyces cerevisiae*

INTRODUCTION

The development of strains with increased performance in the production of a desired product is one of the main tasks in biotechnology. The procedure for this task has traditionally been a series of iterations of mutagenesis followed by selection. Although this procedure has been enormously successful in the past—with some of the new strains producing over 100 times more than the original parent strain—the rate of progress in increasing yield has slowed down considerably in many cases. Confounding this trend is the fact that the experimenter has a priori no control over the mutations produced or their effects. It is becoming increasingly evident that new approaches to this problem are necessary.

Development in two fields of investigation are promising. The first one is a significantly better understanding of the structure of metabolic systems and of the kinetics and thermodynamics of the chemical reactions that take place in living cells. In many cases this understanding is not merely qualitative but quantitative; that is, it can be expressed in terms of kinetics equations. The second development is occurring in molecular biology. Current techniques and the development of numerous useful vectors enable the microbiologist not only to change the protein content of a given organism but also to alter its enzymatic profile specifically in such a way that the synthesis of a given end product or intermediate is enhanced. The combination of these two developments permits modifications of the metabolic structure of an organism and improvements in biotechnological yields.

Kinetic information has been traditionally expressed within the mathematical framework of Michaelis–Menten equations and their generalizations, and these equations are sometimes integrated in models of comprehensive biochemical systems (e.g. Galazzo and Bailey, 1990; Heinrich et al., 1977; Reich and Sel'kov, 1981). These models can, in principle, address questions about which enzymes should be

Correspondence to: N. V. Torres
Contract grant sponsor: Ministerio de Educación y Cultura, CICYT
Contract grant number: B1096-1458

 CCC 0006-3592/97/050758-15

altered in order to improve a desired product or flux. However, the rational functions underlying these approaches have turned out to be mathematically inconvenient (e.g., Savageau, 1992; Schuster et al., 1991). In particular, the complex nonlinear structure of these models has limited attempts to optimize metabolic systems with mathematical methods to a small number of cases with scarse results (Pettersson, 1992; Schuster et al, 1991; Schuster and Heinrich, 1991).

To alleviate this situation, new approaches to capturing the behavior of biochemical systems have been devised. Of particular relevance among these are biochemical systems theory (BST; Savageau, 1969a, 1971, 1976) and metabolic control analysis (MCA; Heinrich and Rapoport, 1974; Kacser et al., 1973). Within the framework of MCA, Westerhoff and Kell (1987) proposed an approach for designing optimal enzyme activity profiles leading to a maximal yields. This approach requires knowledge of the kinetic properties of enzymes and is not suitable for large changes in activities, as they are possible with genetic manipulations. More recently Kacser and Acerenza (1993) presented a strategy (henceforth called the Kacser and Acerenza method) for increasing the production of a metabolite of interest, while preserving conditions necessary for cell viability. Their method enables the researcher to specify the minimum number of enzymes to be manipulated by using fairly basic information about the pathway structure of the system.

Within the framework of BST, optimization methods were developed for biochemical processes represented as S-systems (Hatzimanikatis et al., 1996; Regan et al., 1993; Torres et al., 1996; Voit, 1992). We recently applied one of these methods to the citric acid production in *Aspergillus niger* (Torres et al., 1996) and predicted alternate enzyme profiles that promised higher yields. Our results also indicated that to achieve such yields at least seven enzymes had to be made available in prescribed quantities.

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 (Savageau et al., 1987a). The relative simplicity of the optimization problem derives from the fact that the steady-state equations of S-systems are linear when represented in logarithmic coordinates and that relevant constraints on variables and fluxes also become linear upon logarithmic transformation. As a consequence, the entire theory of linear programming applies, and optimizations can be executed very efficiently with widely available software.

If a biochemical system is not represented as an S-system, such an optimization procedure is usually not possible. For instance, there are no general methods even to compute steady states analytically if the systems are modeled with Michaelis–Menten kinetics or in the otherwise closely related form of a generalized mass action (GMA) system. Since steady-state conditions are important in many types of optimization, this lack of analytical computability is a real hindrance. The researcher is thus faced with the decision either to forego the advantages of linear optimization or to make the pathway model adhere to the S-system form. In the former case, problems of nonlinear programming are involved which may or may not be solveable. In the latter case, the researcher has to ask whether the S-system form is an appropriate representation of a biochemical or metabolic pathway. As with any model, this question cannot be answered in general, outside the answer that the quality depends on the given situation.

Very often, Michaelis–Menten (MM) models are considered the gold standard for biochemical analysis, and other approaches are measured against this standard. In particular, discrepancies between a MM model and an alternative representation are almost automatically considered a problem with the latter. This strong reliance on MM models may in some instances be unjustified, since these models, like any other models, are based on assumptions that may not always be satisfied. MM models have had a long history of successful application to pathways in vitro, but there are also well-documented cases in which their applicability and analytical features have been questioned (e.g., Hill et al., 1977; Shiraishi and Savageau, 1992).

Power-law models are relatively younger, but they do have a 25-year track record of successful application to biochemical phenomena. Comparative analyses have shown that the dynamic and steady-state responses of power-law models are rather similar to those of MM models, if variations about a common operating point are not too extreme. For large deviations, power-law and MM models may predict significantly different responses, but since no good experimental data are available to assess their accuracy, it would be premature automatically to identify the MM models as closest to the truth. There has been extensive discussion in the literature about the advantages and disadvantages of alternative models for biochemical systems. Between the MM and power-law models (GMA and S-systems), the question of underlying mechanisms has been an issue. The former were often considered more appropriate because they were based on intuitively appealing concepts of the formation and splitting of enzyme complexes, whereas the latter were argued to yield a valid and often more general reflection of the complexity encountered in biochemical systems. Between the two variants within BST, GMA systems were sometimes considered to be more realistic since they model each flux with an individual power-law term, whereas S-systems aggregate all synthesizing fluxes of a pool into one power-law term and all degradating fluxes into a second power-law term. For small variations about an operating point of choice, typically a steady state, the differences between the three models are negligible, but for larger variations the differences become more pronounced. In particular, the S-system formulation is less accurate with respect to flux stoichiometry, whereas the GMA formulation is less accurate with respect to mass conservation, at least when measured against MM models. Overall, it can only be stated with certainty that all models are approxima-

tions, and as such simplifications, of the true metabolic pathway, whose characteristics are not known in sufficient detail to declare one of the models superior to the other two in all situations.

Whichever model may be most appropriate for analyzing a biochemical system, it seems indisputable that optimization questions are most easily approached with S-system models for reasons outlined above. This suggests the question: If we use the S-system strategy to optimize a pathway, but the more appropriate model in fact is of MM or GMA type, how significant an error do we have to expect? This question is the object of the present analysis. It is addressed in the following manner. We perform optimizations based on S-system models and subsequently compare the results with results of corresponding MM and GMA models. The article thus serves two purposes: first, to extend the previous methods to the indirect optimization of other nonlinear system description and, second, to evaluate the accuracy of optimizations that are solely based on S-systems. We apply this *indirect optimization method* (IOM) to ethanol, glycerol, and carbohydrate production in *Saccharomyces cerevisiae* and show with different optimization criteria how predictions based on S-system optimization compare with simulation results of the corresponding MM and GMA models.

INDIRECT OPTIMIZATION METHOD

The implementation of the method can be subdivided into three main steps, which are diagramatically represented in Figure 1.

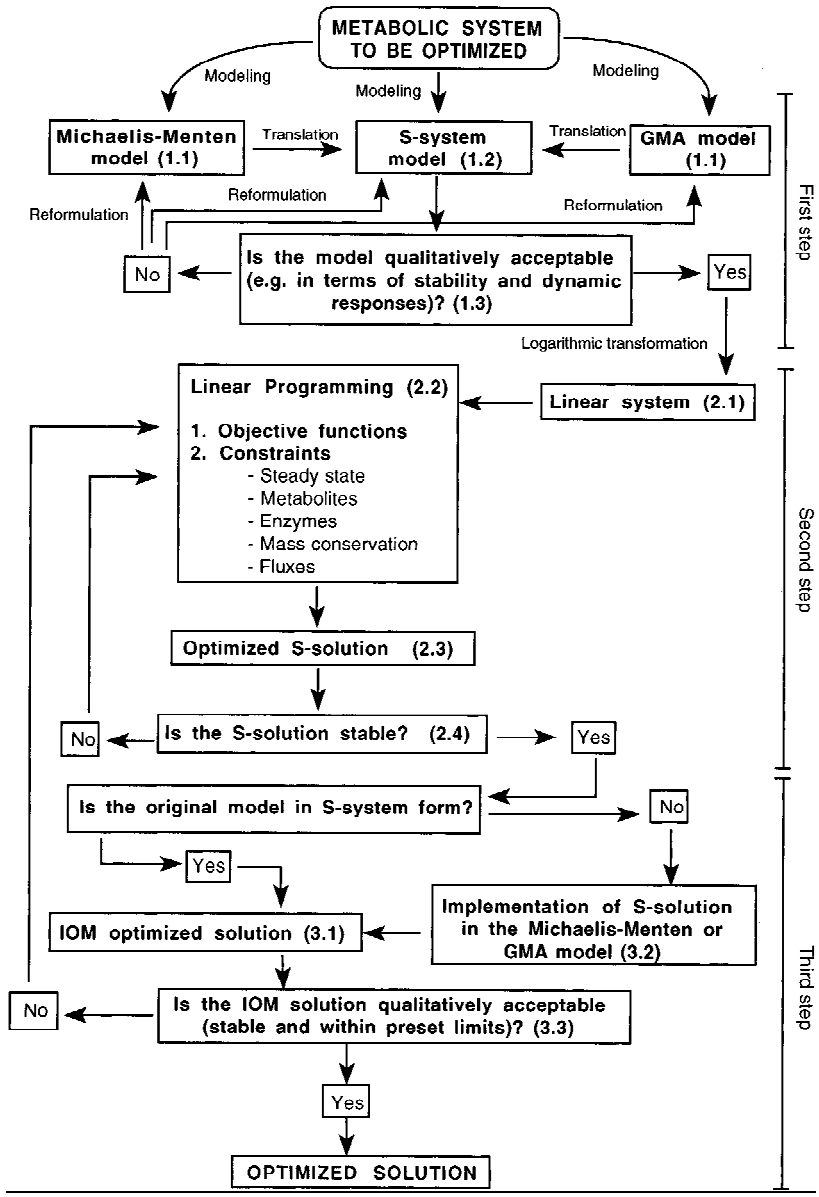

Figure 1. Flow diagram of the indirect optimization method (IOM).

1. Design of the mathematical model, translation into the corresponding S-system form, and quality assessment. Biochemical pathways are traditionally formulated as differential equations containing MM rate laws and their generalizations. 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 (step 1.1). A first alternative to this approach is the formulation of a pathway in the form of a GMA system. This representation can be developed directly from experimental data or it can be obtained from a corresponding MM model through approximation. In the GMA formulation, the representation of fluxes is simplified from sometimes unwieldy rational functions to products of power-law functions that are better suited for mathematical analysis.

An important shortcoming of both, the MM and the GMA approaches, is the fact that it is apparently not possible to compute steady states with analytical methods. This limitation can be overcome when we design models directly, or reformulate previous MM or GMA models, as S-systems (Savageau, 1976; Shiraishi and Savageau, 1992; Voit, 1991), which is possible in a straightforward, analytical manner (step 1.2; Savageau, 1976; Voit, 1991). The S-system representation is particularly advantageous in that it facilitates analytical and numerical quality assessments, which allow us to check (i) the stability of the steady state, which is a fundamental prerequisite for any model of an actual experimental system; (ii) the robustness of the model, indicating whether the model is able to tolerate small structural changes; and (iii) dynamic features that characterize the transient responses to temporary perturbations or permanent alterations. These analyses often pinpoint problems of consistency and reliability of the mathematical representation (step 1.3; see Ni and Sabvageau, 1996a,b; Okamoto and Savageau, 1984, 1986).

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 (Savageau, 1969b). The formulation of these linear equations constitutes step 2.1. In addition to the steady-state equations, typical objective functions and constraints on fluxes and metabolites can be formulated as linear equations or linear inequalities (step 2.2), so that the entire problem becomes one of straightforward linear optimization (see Torres et al., 1996; Voit, 1992). 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) $\leqslant$ constant
(4) ln(dependent or independent variable) $\geqslant$ constant
(5) ln(dependent or independent variable) $=$ constant
(6) ln(variable) unrestricted
(7) ln(flux) $\leqslant$ constant
(8) ln(flux) $\geqslant$ constant
(9) ln(flux) unrestricted
(10) ln(flux1/flux2) $\leqslant$ constant

In this formulation, (1) is a typical objective function that is linear in the logarithms of the involved dependent and independent variables; (2) assures that the optimized system is in a steady state, no matter what the altered enzyme concentrations are; (3) and (4) constrain variables to stay within certain limits; (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; (7)–(9) are the corresponding constraints on fluxes; and (10) represents that the logarithm of the flux ratio flux1/flux2 should remain below a certain limit. Numerical examples for these constraints are discussed in the following section where we design linear programs for the optimization of ethanol, glycerol, and carbohydrate production in *S. cerevisiae* under conditions of a suspended cell culture at pH 4.5.

The optimization (step 2.3) 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 optimized enzyme concentration profiles and their corresponding steady states are ''S-solutions'' in the sense that they constitute the optimized steady state of the S-system model, which may differ from the steady state of the original MM or GMA model. These differences typically become apparent as inaccuracies in the flux stoichiometry at branchpoints, which is due to the aggregation of fluxes in the S-system formalism (cf. Savageau et al., 1987a,b). It has been documented, however, that the S-system approximation in many cases is sufficiently accurate for predicting the dynamic and steady-state properties of the system over a reasonably wide range of variation in conditions (Sorribas and Savageau, 1989; Voit and Savageau, 1987).

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 (Savageau, 1969b, 1976) or with computational means (e.g., Torres et al., 1996). Unstable steady-state solutions are normally discarded (step 2.4). 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. The system is to be revisited (step 2.2) with more stringent constraints.

It is noted that if we allow changes in the constraints up to several multiples of the original, it is necessary to check for differences between the S-system representation and the GMA or MM model. If significant discrepancies in numeri-

cal predictions or even in stability are detected, it might be advisable to execute the optimization in several smaller steps.

3. Transfer of results to the original model. In cases where the pathway is modeled originally as an S-system, this step is unnecessary (step 3.1). In other cases, the S-system may have been derived as an approximation of a different pathway model. In this latter situation, the enzyme concentration profile of the S-solution is now used as input to the original model (step 3.2). In the most common case of a Michaelis–Menten representation, this is accomplished by substituting the optimized enzyme concentrations in a simulation program that allows dynamical evaluations of the MM model. If the original model is expressed in the GMA representation, this can be done straightforwardly by using the simulation package ESSYNS (Voit et al., 1990).

The steady state of the GMA or MM model, implemented with the new enzyme concentrations, cannot typically be computed analytically, but it can be solved numerically, for instance with a Newton method. Alternately, the differential equations can be solved numerically until the system has approached the steady state sufficiently closely (step 3.1). Either way, the result is an optimized steady-state solution that is consistent with the GMA or MM model, respectively. This solution is not necessarily the true optimum, since it is computed via S-system approximation, which is accompanied by some unknown approximation error. This is the cost to be paid for reducing the arbitrarily complex nonlinearities of a traditional biochemical system model to a linear program. The following examples will demonstrate the differences between the steady states of the original and the S-system models are often small in comparison to the experimental accuracy and in light of other uncertainties involved in any modeling effort (see discussion in Torres et al., 1996).

A possible outcome of the present step is that some of the metabolite concentrations exceed the imposed limits or that the traditional model is unstable (step 3.3). If the discrepancies in metabolites are physiologically significant, or if the system is unstable, some of the constraints in step 2.2 must be changed accordingly. Once a satisfactory solution is obtained, the optimized profile of enzyme concentrations serves as a target in the development of modified strains of the microorganism.

In summary, the previously developed optimization method (Regan et al., 1993; Voit, 1992; Torres et al., 1996), which assumed that an S-system represented the true pathway with sufficient accuracy, was extended here to situations in which the S-system representation is an approximation of another model, such as an MM or GMA model. By performing the optimization with the approximating S-system model, the resulting solution is only an ''approximately optimal solution,'' but it is obtained in a straightforward fashion that circumvents the often considerable difficulties posed by the nonlinearity of other models.

THE EXPERIMENTAL SYSTEM

The proposed optimization method is now applied to ethanol, glycerol, and carbohydrate production in the yeast *S. cerevisiae.* The specific system under consideration (see Fig. 2) produces the target metabolites through the anaerobic fermentation of glucose in a suspended cell culture at pH 4.5 (Galazzo and Bailey, 1990, 1991).

The metabolic system and the experimental conditions were chosen for a number of reasons. First, the kinetics and fluxes involved in the formation of ethanol, glycerol, and carbohydrates in vivo as well as the intracellular concentrations of substrates, intermediate metabolites, and effectors at the considered steady-state conditions are well known (Galazzo and Bailey, 1990, 1991). Moreover, the steady state of this system was recently characterized in some detail (Cascante et al., 1995; Curto et al., 1995), and the quality of different alternate representations was analyzed (Sorribas et al., 1995). These studies showed that the steady state is locally stable, as indicated by eigenvalues with all the real parts negative, thus meaning that the system variables will return to their reference values after a small perturbation in

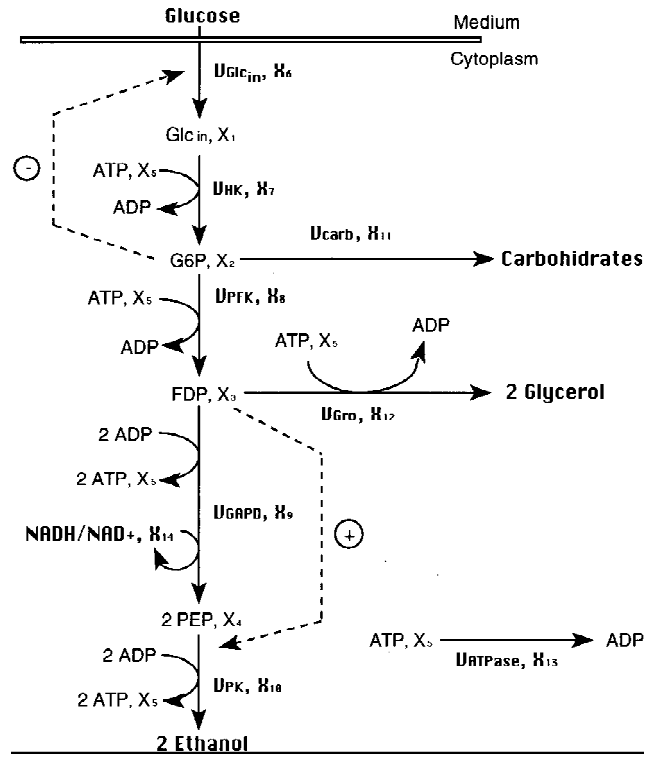

Figure 2. Anaerobic fermentation pathway of the yeast *S. cerevisiae*, according to Galazzo and Bailey (1990,1991). The suspended cell culture transforms glucose to ethanol, glycerol, and polysaccharides at pH 4.5. Metabolites (regular typeface) are numbered 1–5: X_1 (Glc_{in}, intracellular glucose); X_2 (G6P, glucose-6-phosphate); X_3 (FDP, fructose diphosphate); X_4 (PEP, phosphoenol pyruvate); X_5 (ATP). Enzymes and other effectors (bold typeface) are numbered from 6 to 14: X_6 (V_{in}, sugar transport system); X_7 (V_{HK}, hexokinase); X_8 (V_{PFK}, phosphofructokinase); X_9 (V_{GADP}, glyceraldehyde 3-phosphate dehydrogenase); X_{10} (V_{PK}, pyruvate kinase); X_{11} (V_{carb}, glycogen synthetase); X_{12} (V_{gro}, glycerol 3-phosphate dehydrogenase, proportional to PK); X_{13} (V_{ATPase}, ATPase); X_{14} (NADH/NAD^+).

the dependent concentrations. Furthermore, the system is characterized by low parameter sensitivities: only 4 of the 55 sensitivities have absolute values between 1 and 5, whereas all the other sensitivities have values lower than 1 (Sorribas et al., 1995). These sensitivities quantify the effects of changes in the parameters of the system, thereby providing a measure of the robustness of the model. Finally, the system quickly returns to the predisturbance steady state after an increase or decrease in a dependent or independent variable, a response that is to be expected from a well-functioning system. In summary, regarding the system response after an increase or decrease of a dependent or independent variable, the present system showed quick returns to the predisturbance steady state, the response being an appropriate behavior. In summary, the present system is well determined under the chosen conditions, and the previously developed mathematical models seem to correspond well to the pathway in vivo.

Second, the metabolic system is complex enough for exploring nontrivial aspects of reliability, practical implementation, and accuracy of the method. The metabolic pathway has two branchpoints, one leading to glycogen and trehalose (V_{carb}) and the other leading to glycerol (V_{gro}). Also, there is an inhibitory feedback loop from glucose-6-phosphate to V_1^+, and an activating feedforward loop from fructose diphosphate to V_4^- (see Fig. 2). Finally, the pathway contains some splitting reactions, as well as some steps, associated with coupled cofactors, in which the molecularity is 2 and 3, respectively. These features are very interesting because they create a situation that allows us to test various aspects of the IOM strategy.

Third, a considerable body of academic and technical knowledge is available, which renders feasible the actual implementation of DNA vectors in modified strains of *S. cerevisiae* with optimized profiles of enzyme activities (Davies et al., 1992; Heinish, 1986; Schaaff et al., 1989).

APPLICATION TO ETHANOL, GLYCEROL AND CARBOHYDRATE PRODUCTION IN *SACCHAROMYCES CEREVISIAE*

Mathematical Description

The Michaelis–Menten Representation

The mathematical model of the pathway, expressed in terms of MM rate laws, was recently published by Galazzo and Bailey (1990, 1991) and is not reproduced here. For our comparison of methods, this model is taken as the ''true'' representation of the pathway, and we will not discuss whether this representation or a power-law model is more adequate. The rate equations and the parameter values used here are also taken from Galazzo and Bailey (1990, 1991).

The S-System Model

The corresponding S-system and GMA models for suspended cells at pH 4.5 were developed by Curto and co-workers (1995). The S-system is given as a set of five differential equations with two power-law terms each:

$$\begin{aligned}
dX_1/dt &= 0.8122\, X_2^{-0.2344}\, X_6 \\
&\quad -2.866999105\, X_1^{0.7464}\, X_5^{0.0243}\, X_7 = V_1^+ - V_1^- \\
dX_2/dt &= 2.866999105 X_1^{0.7464}\, X_5^{0.0243}\, X_7 \\
&\quad -0.524046743\, X_2^{0.739}\, X_5^{-0.394}\, X_8^{0.999}\, X_{11}^{0.001} \\
&\quad = V_2^+ - V_2^- \\
dX_3/dt &= 0.522758878\, X_2^{0.7318}\, X_5^{-0.3941}\, X_8 \\
&\quad -0.0148443752\, X_3^{0.584}\, X_4^{0.03}\, X_5^{0.119}\, X_9^{0.944}\, X_{12}^{0.056} \\
&\quad X_{14}^{-0.575} = V_3^+ - V_3^- \\
dX_4/dt &= 0.022074642\, X_3^{0.6159}\, X_5^{0.1308}\, X_9\, X_{14}^{-0.6088} \\
&\quad -0.094712323\, X_3^{0.05}\, X_4^{0.533}\, X_5^{-0.0822}\, X_{10} = V_4^+ - V_4^- \\
dX_5/dt &= 0.09133492\, X_3^{0.333}\, X_4^{0.266}\, X_5^{0.024}\, X_9^{0.5}\, X_{10}^{0.5}\, X_{14}^{-0.304} \\
&\quad -3.211579932\, X_1^{0.198}\, X_2^{0.196}\, X_5^{0.372}\, X_7^{0.265}\, X_8^{0.265} \\
&\quad X_{11}^{0.0002}\, X_{13}^{0.47} = V_5^+ - V_5^-
\end{aligned} \tag{1}$$

The kinetic orders in this representation were calculated by Curto and co-workers (1995) from published experimental data, and the values of α_i and β_i were determined from the kinetic order values and steady-state concentration values of metabolites, as is commonly done in power-law approximations (e.g., Voit, 1991, Ch. 2). The steady state of the S-system model under the given conditions is consistent with experimental observations (Galazzo and Bailey, 1990, 1991), and eigenvalue analysis confirmed its stability (Sorribas et al., 1995). The quality of the representation was subsequently analyzed with standard tools of biochemical system theory, and it was shown that the model captured a realistic picture of the pathway in vivo (Sorribas et al., 1995).

Maximization of fluxes

Objective Functions

The first step in setting up a linear program is the definition of the objective function. We considered three optimization tasks:

Maximization of the rate of ethanol production: The rate of ethanol production is given directly by the flux through the pyruvate kinase reaction, V_4^- [cf. Eq. (1)]:

$$V_4^- = 0.094712323\, X_3^{0.05}\, X_4^{0.533}\, X_5^{-0.0822}\, X_{10} \tag{2}$$

The objective function for this linear program in logarithmic coordinates reads

$$0.05\, y_3 + 0.533\, y_4 - 0.0822\, y_5 + y_{10} \tag{3}$$

Maximization of the rate of glycerol and carbohydrate production: In these cases we need to use the corresponding, power-law representation of the individual fluxes (Curto et al., 1995):

$$V_{gro} = 0.092717816\, X_3^{0.05}\, X_4^{0.533}\, X_5^{-0.0822}\, X_{12} \quad (4)$$

$$V_{carb} = 8.903845725 \times 10^{-4}\, X_2^{8.6107}\, X_{11} \quad (5)$$

After taking logarithms in the above expressions, we obtain, as the linearized objective functions,

$$0.05\, y_3 + 0.533\, y_4 - 0.0822\, y_5 + y_{12} \quad (6)$$

$$8.6107\, y_2 + y_{11} \quad (7)$$

for the glycerol and carbohydrate production, respectively.

Steady-State Constraints

The second step is the formulation of the steady-state constraints, again expressed in terms of the logarithms of the variables. In the three cases considered these constraints take the following form:

$$\begin{aligned}
&-0.7464\, y_1 - 0.2344\, y_2 - 0.024\, y_5 + y_6 - y_7 = 1.2612\\
&0.7464\, y_1 - 0.739\, y_2 + 0.418\, y_5 + y_7\\
&\quad - 0.999\, y_8 - 0.001\, y_{11} = -1.699\\
&0.7318\, y_2 - 0.584\, y_3 + 0.03\, y_4 - 0.513\, y_5 + y_8 - 0.944\\
&\quad y_9 - 0.056\, y_{12} + 0.575\, y_{14} = -3.5615\\
&0.5659\, y_3 - 0.533\, y_4 + 0.213\, y_5 + y_9 - y_{10} - 0.608\, y_{14}\\
&\quad = 1.4564\\
&\quad - 0.198\, y_1 - 0.196\, y_2 + 0.333\, y_3\\
&\quad + 0.266\, y_4 - 0.348\, y_5 - 0.265\, y_7 - 0.265\, y_8\\
&\quad + 0.5\, y_9 + 0.5\, y_{10} - 0.002\, y_{11} - 0.47\, y_{13}\\
&\quad - 0.304\, y_{14} = 33.039 \quad (8)
\end{aligned}$$

Constraints on Enzyme Concentrations

Next we define which enzyme concentrations may vary and by how much. We generally allow the enzymes to vary between 1 and 50 times the base values. These limits are consistent with the range of variations made possible by recombinant DNA technologies (Guarante et al., 1980). In all cases, the values of those enzymes diverting flux from the target product were kept constant at the base steady-state values. This was so in order to ensure that the rest of metabolism remains unperturbed. With these settings, we obtain the following constraints on the control variables for each of the flux maximization tasks.

Ethanol maximization:

$$2.980 \leq y_6 \leq 6.892$$

$$4.226 \leq y_7 \leq 8.138$$

$$3.456 \leq y_8 \leq 7.368$$

$$3.910 \leq y_9 \leq 7.822$$

$$8.143 \leq y_{10} \leq 12.055$$

$$y_{11} = 2.66$$

$$y_{12} = 5.313$$

$$3.222 \leq y_{13} \leq 7.134$$

Glycerol maximization: In this particular case we initially allowed the enzymes to vary in concentration between 1 and 50 times their base value. However, analysis of the optimized solution revealed significant differences between the S-solution and the solution to the MM model (step 2.4), and we subsequently changed the maximum to 10 times the base value, thereby avoiding the former problems:

$$2.980 \leq y_6 \leq 5.283$$

$$4.226 \leq y_7 \leq 6.529$$

$$3.456 \leq y_8 \leq 5.758$$

$$y_9 = 3.91$$

$$y_{10} = 8.143$$

$$y_{11} = 2.66$$

$$5.313 \leq y_{12} \leq 7.615$$

$$3.222 \leq y_{13} \leq 5.525$$

Carbohydrate maximization:

$$2.980 \leq y_6 \leq 6.892$$

$$4.226 \leq y_7 \leq 8.138$$

$$3.456 \leq y_8 \leq 7.368$$

$$y_9 = 3.91$$

$$y_{10} = 8.143$$

$$2.66 \leq y_{11} \leq 6.57$$

$$y_{12} = 5.313$$

$$3.222 \leq y_{13} \leq 7.134$$

Constraints on Metabolite Concentration

Analogous to the constraints on enzymes, we limit the range of variation of the dependent variables, which represent the metabolite concentrations. For all three optimization tasks, the lower and upper limits of the dependent variables, X_1, . . ., X_5, were set to 0.8 and 1.2 the base values, which corresponds to 20% variation about the steady-state levels. There is no hard evidence on how wide a range in a metabolite can be tolerated without causing losses in growth rate or fitness and subsequently leading to poor yield or rate in product synthesis. We arbitrarily allow changes of up to 20% around the steady-state levels. This threshold presumably can be considered small enough to avoid significant changes in the overall metabolism in vivo.

The ratio NADH/NAD, X_{14}, was set constant and treated as an independent control variable. This ratio was considered constant since fluorescence measurements performed by Galazzo and Bailey (1990) indicated small variations in this ratio. Accordingly, the mathematical formulation, in logarithmic coordinates, leads to the following constraints on the metabolite concentrations:

$$\begin{aligned}
-3.604 &\leq y_1 \leq -3.199\\
-0.213 &\leq y_2 \leq 0.192\\
1.989 &\leq y_3 \leq 2.395\\
-4.879 &\leq y_4 \leq -4.474\\
-0.109 &\leq y_5 \leq 0.304
\end{aligned}$$

Stoichiometric Conservation Constraint

In the optimization process, the system deviates 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 base solution are no longer satisfied. In order to avoid significant stoichiometric deviations, we add a restriction that reflects the required dependency between fluxes. To this end, we impose a stoichiometric conservation constraint in the following form:

$$(V_1^+)/(V_4^-) > 0.5$$

This constraint ensures that the rate V_4^- (ethanol production) cannot be greater than twice the input flux, which accounts for the splitting of each fructose diphosphate molecule into two molecules of phosphoenol pyruvate (X_4).

After substituting from Equations 1 and taking logarithms, we obtain as the flux constraint

$$-0.2344\ y_2 - 0.05\ y_3 - 0.533\ y_4 + 0.0822\ y_5 + y_6 - y_{10} > -2.84204967 \quad (9)$$

To optimize the system under these conditions, we thus maximize Equation (3), (6), and (7) and require that the stoichiometric conservation constraint [Equation (9)] be satisfied.

It should be noted that different stoichiometric conservation constraints could be imposed instead, or in addition to, Equation (9). For instance, we could require that the flux through the branchpoint fructose diphosphate, (FDP; X_3) cannot be greater than the glucose input flux [$(V_3^-)/V_1^+) < 1$] or that the flux toward glycerol should be less than twice the glucose uptake [$(V_{gro})/(2 \cdot V_1^+) < 1$]. Again, these constraints become linear upon logarithmic transformation. Only Equation (9) was considered here; the other options are considered in the Discussion.

RESULTS

The results of our analysis can be evaluated in different ways. The main purpose of this study was to show that different types of biochemical system models can indirectly, yet validly, be optimized by optimizing the corresponding S-system model. This is important, because the optimization of S-system models constitutes a straightforward linear problem, whereas the optimization of the other models is nonlinear and thus more complicated. A secondary, though related, result of our analysis is the observation that Michaelis–Menten, GMA, and S-system models are dynamically rather similar. Of course, all three coincide at the operating point at which the models are defined to be equivalent. But it is interesting to note that they also are similar qualitatively and quantitatively when the systems move away from the operating point. Since the three models have different mathematical structures, there is no guarantee that this is always the case, but the present observation of similarity, along with a large body of previous experience with these models, suggests that all three models can be employed with some reliability. The comparison of the three models will not be discussed further, and the following sections focus on the primary goal of optimizing production rates.

Ethanol Maximization

Table I shows the results of the maximization of ethanol. At step 2.2 the lower and upper limits of the metabolite concentrations (X_1 to X_5) were set to 0.8 and 1.2 the base values (20% variation about the steady-state levels), while the NADH/NAD ratio, X_{14}, was kept constant. The enzyme concentrations were allowed to vary within a range of 1 and 50 times the base values. The flux stoichiometric conservation constraint was the expressed as shown in Equation (9).

Inspection of eigenvalues confirms that the optimized S-solution is stable. All rates of synthesis increase, and the rate of ethanol production is enhanced by a factor of 3.2. Changing the lower or upper limits of the enzyme concentrations (with ranges varying from 0.01 to 50) does not alter these enhancements (details not shown), and the control variables remain unaltered. When we raised the limits for the metabolite concentrations, allowing up to 90% variation about the base levels, the ethanol production increased to 5.23 the original value (details not shown). In all cases, the yield of ethanol production (expressed as $100 \cdot 0.5 \cdot V_4^-/V_{1^+}$) was 100, and the deviation of the total output flux from the total input flux was 2.1%. The latter results indicate the relatively small magnitude of error introduced by aggregating fluxes in the S-system representation.

The optimum profile of enzyme concentrations is also shown in Table I. It can be seen that the optimized solution requires only moderate levels of overexpression of the enzymes involved (factors ranging from 3.15 to 4.25) which can be attained without major difficulties in yeast (see Davies and Brindle, 1992; Guarante et al., 1980; Heinish, 1986; Schaaff et al., 1989). This profile was then implemented in both the GMA version of the system (Curto et al., 1995) and the original Michaelis–Menten model (Galazzo and Bailey, 1990). In both cases we obtained a stable steady state with metabolite and flux profiles shown in Table I.

In the GMA model, the optimized solution corresponds to a steady state with an ethanol rate 3.14 times greater than the base steady state. In this case, only one metabolite deviates slightly from the imposed boundaries (X_5, ATP). The solution of the Michaelis–Menten model corresponds to a steady state that leads to an increased rate of ethanol production (3.48 times the base steady-state rate). This solution is close to the limits of physiological acceptability for the pools. Three intermediates exceed the imposed limit of 20% variation around the steady state: X_2, glucose-6-phosphate showing the biggest deviation observed (73%); X_3, fructose-diphosphate, with 42% deviation; and X_4, phosphoenol–pyruvate, with 57% deviation. It is interesting to note that most fluxes are only slightly elevated in the Michaelis–Menten model, while V_{carb} is significantly higher (83.84

Table I. Optimization of the carbohydrate metabolism of *S. cerevisiae* under conditions of ethanol production in suspended cells at pH 4.5.

Metabolite	Base, $X_{i,\text{Base}}$	S-solution, $X_i/X_{i,\text{Base}}$	GMA solution, $X_i/X_{i,\text{Base}}$	MM solution, $X_i/X_{i,\text{Base}}$	K&A solution, $X_i/X_{i,\text{Base}}$
X_1	3.4×10^{-2}	0.8	0.8	0.96	1
X_2	1.01	1.2	1.18	1.73	1
X_3	9.14	1.2	1.17	1.42	1
X_4	9.5×10^{-3}	1.2	1.15	1.57	1
X_5	1.13	0.8	0.77	0.85	1
Enzyme	$X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$
X_6	19.7	3.15	3.15	3.15	3.39
X_7	68.5	3.58	3.58	3.58	3.39
X_8	31.7	2.42	2.42	2.42	3.40
X_9	49.9	2.94	2.94	2.94	3.54
X_{10}	3440	2.82	2.82	2.82	3.54
X_{13}	25.1	4.25	4.25	4.24	—
Flux	Base	$V_{ij}/(V_{ij})_{\text{Base}}$	$V_{ij}/(V_{ij})_{\text{Base}}$	$V_{ij}/(V_{ij})_b$	$V_{ij}/(V_{ij})_b$
V_1^+	15.96	3.02	3.03	3.48	3.39
V_{gro}	1.77	1.11	1.08	1.25	1
V_{carb}	$1.4 \cdot 10^{-2}$	4.80	4.22	83.84	1
$\boldsymbol{V_4^-}$	**30.11**	**3.20**	**3.14**	**3.48**	**3.54**

Note: The baseline, reference steady state, the S-solution, the GMA solution, the Michaelis–Menten (MM) solution, and the solution according to the method of Kacser and Acerenza (1993) (K&A) are given. The objective here was to maximize ethanol synthesis V_4^- [cf. Equations (2) and (3)]. The metabolite pools were allowed to vary up to 20% about their baseline steady-state levels and the enzyme activities between 1 and 50 times the base levels. Only variables allowed to vary are shown. Enzyme activities and fluxes are given in m*M*/min and concentrations in m*M*. The ethanol yield of each optimized solution was 94.32% (Base); 100% (S-system); 97.76% (GMA); 95.82% (MM); and 94.32% (K&A).

times the base value) than predicted by both the S-solution (4.8) and the GMA solution (4.22). The overall ethanol yield is 95.82%, which is comparable to the yield of the base steady state, but with a rate of ethanol production that is more than three times higher. In any case we can reduce the observed deviation in the metabolite concentrations at the MM model either by imposing more narrow limits to the metabolite concentrations constraints or eventually by imposing more constraints.

Kacser and Acerenza (1993) acknowledge limitations in the applicability of their optimization method to pathways involving bimolecularities, as they occur with coupled cofactors such as ATP/ADP and NAD/NADH. The limitation is particularly relevant in the present case since these couplings occur at various points of the pathway. In order to overcome this problem, the authors suggest that in many instances the concentrations of the cofactors could be treated as control variables, which is a dangerous assumption (see Shiraishi and Savageau, 1992). To maintain quasi-constancy, without being too restrictive, we limited the variation of the ATP/ADP pool to 20% about the base steady state, but kept the NAD/NADH pool concentration constant at the optimized solution, as explained above (see the section on Constraints on Metabolite Concentration). The resulting optimal solution was generally found to be in good agreement with the solution of the Kacser and Acerenza (1993) method as judged by the amplification factors in enzyme activities and fluxes, $X_i/X_{i,\text{Base}}$ and $V_{ij}/(V_{ij})_b$. An exception is V_{carb}, which remains unaffected in Kacser and Acerenza's approach, whereas it is increased by a factor of about 83.84 in our solution. However, the magnitude of this flux in relation to the total final output flux is not significant (0.019%), which makes this deviation irrelevant. Furthermore, in the present optimization procedure, as well as in the following procedures, the amplification factor $X_i/X_{i,\text{Base}}$ does not decrease for each flux from output to input as predicted by the Kacser and Acerenza method. This is presumably due to the fact that the splitting reactions (X_9 and X_{12}) change molecularity. This aspect will be considered in more detail in the Discussion.

Glycerol Maximization

In addition to the obvious task of increasing ethanol production rate, we also employed the IOM approach to produce larger quantities of glycerol and carbohydrates. We are aware that glycerol and carbohydrates are secondary outputs and consider the system primarily as a model system that allows us to study the versatility of the method in optimizing and redesigning a given metabolic pathway with respect to a chosen aim.

Again we set the lower and upper limits of the dependent

variables (X_1 to X_5) to 0.8 and 1.2 the base values (20% variation about the steady-state levels), while the NADH/NAD ratio, X_{14}, is set constant. Also, the limits of variation in enzyme concentrations were allowed to vary between 1 and 50 times the base values, except for the previous branch point (X_{11}) and the enzymes downstream from fructose diphosphate (X_3), namely, X_9, X_{10}, and X_{13}, which were kept constant. This design facilitates the flux toward glycerol without decreasing other fluxes below the base steady-state level. The flux stoichiometric conservation constraint was again expressed as shown in Equation (9).

The main result was quite surprising: The S-solution was stable, but the corresponding GMA and MM solutions were not. This prompted us to change the upper limits for variations in enzymes, with the result that the GMA and MM solutions were unstable for upper limits above 10 and 15 times the base enzyme activity, respectively. This fact could be explained in terms of the dynamics of the system. We realized that intracellular glucose accumulates steadily while the ATP concentration diminishes to almost zero: Any increase in the activity of X_{12} of more than 10 times causes most of the fructose diphosphate to be converted into glycerol. This leaves no chance for ATP to be synthetized, and consequently no glucose phosphorylation is possible at the hexokinase step, which in turn causes instability of the system.

A second explanation for the instability could be given: X_{12} is an enzyme whose activity is already quite high at the base steady state (203 m*M*/min). A 50 times increased activity may constitute an unreasonable overexpression that negatively affects the entire organism. Even if the enzyme activity is increased 15-fold, the MM solution is unacceptable, since phosphoenol pyruvate (X_4) and ATP (X_5) exceed by far (131% and 91%, respectively) the allegedly reasonable maximal increase by 20%. Whatever the reason for the discrepancy the results suggested a revision of the numerical definition of the constraints.

When we let both enzymes vary up to a maximum of 10, the system was stable, but the metabolite constraints were violated in some instances. The most significant deviation was observed in the case of X_5 where deviations of 59% and 62% were detected in the GMA and MM models, respectively. Also, X_2 attained a steady-state value above the 20% allowed, although the magnitude in this case was less (43% and 30%, for GMA and MM, respectively; see Table II). Of course, when the upper limit for variation in enzyme activity was reduced, the deviations also diminished (results not shown). In the S-solution, the only enzyme to be overexpressed is glycerol-3-phosphate dehydrogenase (X_{12}) whose concentration should be raised to 10 times the steady-state base value. Again, as was discussed before, this value is well within the limits of actual feasibility. It can thus be

Table II. Optimization of the carbohydrate metabolism of *Saccharomyces cerevisiae* under conditions of ethanol production in suspended cells at pH 4.5.

Metabolite	Base, $X_{i,\text{Base}}$	S-solution, $X_i/X_{i,\text{Base}}$	GMA solution, $X_i/X_{i,\text{Base}}$	MM solution, $X_i/X_{i,\text{Base}}$	K&A solution, $X_i/X_{i,\text{Base}}$
X_1	3.4×10^{-2}	0.8	0.92	0.89	1
X_2	1.01	0.8	0.57	0.70	1
X_3	9.14	1.2	1.06	0.93	1
X_4	9.5×10^{-3}	1.2	0.74	1.01	1
X_5	1.13	0.96	0.41	0.38	1
Enzyme	$X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$
X_6	19.7	1.20	1.20	1.20	1.43
X_7	68.5	1.49	1.49	1.49	1.43
X_8	31.7	1.47	1.47	1.47	1.43
X_{12}	203.0	10	10	10	8.81
X_{13}	25.1	1.0	1.0	1.0	—
Flux	Base	$V_{ij}/(V_{ij})_{\text{Base}}$	$V_{ij}/(V_{ij})_b$	$V_{ij}/(V_{ij})_b$	$V_{ij}/(V_{ij})_b$
V_1^+	15.96	1.26	1.37	1.31	1.43
$\boldsymbol{V_{\text{gro}}}$	**1.77**	**10.95**	**9.05**	**8.81**	**8.81**
V_{carb}	0.014	0.15	0.007	0.049	1
V_4^-	30.11	1.11	0.92	0.88	1

Note: The baseline, reference steady state, the S-solution, the GMA solution, the Michaelis-Menten (MM) solution, and the solution according to the method of Kacser and Acerenza (1993) (K&A) are given. The objective here was to maximize glycerol synthesis V_{gro} [cf. Equations (4) and (6)]. The metabolite pools were allowed to vary up to 20% about the base steady-state level and the enzyme activities between 1 and 50 times the base levels, except for X_{12}, which was allowed to vary between 1 and 10. Only variables allowed to vary are shown. Enzyme activities and fluxes are given in m*M*/min and concentrations in m*M*. The glycerol yield of each optimized solution was 5.54% (Base); 48.18% (S-system); 36.63% (GMA); 37.29% (MM); and 37.29% (K&A).

concluded that in the optimized solution the rate of glycerol production increases by a factor of 10.95, while ethanol production (V_4^-) remains almost unaltered. The rate of carbohydrate synthesis decreases to 15% of the base value and glucose uptake rate, V_1^+, increases by 26%. The GMA solution exhibits a similar pattern, but with smaller increases in V_{gro} and larger decreases in V_4^- and V_{carb}. Overall, all solutions produce more glycerol with some loss in ethanol synthesis in some cases and, thus, constitute a more sufficient system if the target product is glycerol.

It is worthwhile noting that in order to obtain the above increases in fluxes, the main enzyme to be modulated is X_{12}, which is the step responsible for glycerol synthesis from fructose diphosphate. The other enzymes remain almost unaltered, which greatly simplifies the experimental implementation of the improved solution. In the optimized solution of the MM model, the glycerol production is more than 8 times greater than at the base steady state. Also remarkable is the fact that the overall yield of the conversion of glucose into glycerol is 37.29%, which is considerably higher than the 5.54% of the base steady state. These results suggest that the present system is suitable for glycerol production and that it involves lesser changes in enzyme activities than for the optimization of ethanol.

The solution of the S-system, GMA, and MM model were found to be strongly correlated with that obtained from application of Kacser and Acerenza's method.

Carbohydrate Maximization

Again we define the lower and upper limits of the dependent variables (X_1 to X_5) as 0.8 and 1.2 of the base values and keep the NADH/NAD ratio (X_{14}) constant. The limits for the control variables are set at 1 and 50 times the base values, except for the glycerol branchpoint (X_{12}), and the enzymes downstream from glucose diphosphate, X_2, namely, X_8, X_9, X_{10}, and X_{13}, are kept constant. As in the glycerol maximization, this design facilitates the direct conversion of glucose into carbohydrate at the expense of the glycerol and ethanol synthesis.

Table III shows some of these optimizations. In the S-solution, the only enzyme that has to be overexpressed is glycogen synthetase (X_{11}), the enzyme that diverts the flux; its concentration should be raised to the maximum of 50 times over the original steady-state value. While such an increase approaches the limits of technical and physiological feasibility (see Guarante et al., 1980), it is important to note that any significant increase in this enzyme can be expected to show improvements in the rate of carbohydrate synthesis. In our particular solution, the production of carbohydrates increases by a factor of 234.33, while ethanol and glycerol production and glucose uptake remain almost unaltered. Overall, this solution produces more carbohydrates without a concomitant loss in ethanol synthesis. Thus, if the end products of interest are ethanol and carbo-

Table III. Optimization of the carbohydrate metabolism of *S. cerevisiae* under conditions of ethanol production in suspended cells at pH 4.5.

Metabolite	Base, $X_{i,\text{Base}}$	S-solution, $X_i/X_{i,\text{Base}}$	GMA solution, $X_i/X_{i,\text{Base}}$	MM solution, $X_i/X_{i,\text{Base}}$	K&A solution, $X_i/X_{i,\text{Base}}$
X_1	3.4×10^{-2}	0.8	0.84	0.86	1
X_2	1.01	1.19	1.02	1.09	1
X_3	9.14	1.14	1.10	1.13	1
X_4	9.5×10^{-3}	1.20	1.06	1.13	1
X_5	1.13	1.09	0.92	0.97	1
Enzyme	$X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$	$X_i/X_{i,\text{Base}}$
X_6	19.7	1.15	1.15	1.15	1.18
X_7	68.5	1.30	1.30	1.3	1.18
X_{11}	14.31	50.0	50	50	106.8
X_{13}	25.1	1.0	1.0	1.0	1
Fluxes	Base	$V_{ij}/(V_{ij})_{\text{Base}}$	$V_{ij}/(V_{ij})_{\text{Base}}$	$V_{ij}/(V_{ij})_b$	$V_{ij}/(V_{ij})_b$
V_1^+	15.96	1.10	1.14	1.17	1.18
V_{gro}	1.77	1.07	1.02	1.07	1
$\boldsymbol{V_{carb}}$	**0.014**	**234.33**	**59.38**	**106.77**	**106.77**
V_4^-	30.11	1.10	1.04	1.07	1

Note: The baseline, reference steady state, the S-solution, the GMA solution, the Michaelis-Menten (MM) solution, and the solution according to the method of Kacer and Acerenza (1993) (K&A) are given. The objective here was to maximize carbohydrate synthesis V_{carb} [cf. Equations (5) and (7)]. The metabolite pools were allowed to vary up to 20% about the base steady-state level and the enzyme activities between 1 and 50 times the base levels. Only variables allowed to vary are shown. Enzyme activities and fluxes are given in m*M*/min and concentrations in m*M*. The ethanol yield of each optimized solution was 0.087% (Base); 18.74% (S-system); 4.58% (GMA); 8.09% (MM); and 7.97% (K&A).

hydrates, the S-solution characterizes a system that is more efficient than the baseline model. The same holds true for the GMA and the MM solutions, although to a lesser degree. The optimized enzyme profile in these systems leads to an increased carbohydrate production of 59.38 times (GMA) and 106.77 times (MM) the original, which constitutes 4.58% and 8.09% of the total output flux, respectively. The other fluxes are rather similar in the s-solution, the GMA and the MM solution, except for V_{carb}, whose GMA value is less than in the S-solution. The MM solution of the present optimization is stable, and the concentrations of all intermediates (X_1 to X_5) satisfy the predefined 20% constraint on variation (see Table III). Overall, the optimized system shows a higher rate of carbohydrate synthesis and better yields than the base system. However, the system is rather inefficient in transforming glucose into carbohydrates.

A comparison of increases in enzyme activities predicted by the Kacser and Acerenza (1993) method and the MM solution shows fairly good agreement. In both cases, the enzyme to be modulated is X_{11}, while the other enzymes remain practically unaltered in the optimized solutions. It is noted that, in order to obtain this increase in carbohydrate production, the only enzyme to be modulated is the one responsible for carbohydrate synthesis from glucose diphosphate (X_{11}), while the other enzymes remain almost unaltered. This greatly simplifies the actual implementation of the improved solution.

Optimization of Subsets of Enzymes

The solutions obtained in the previous section require quantitatively prescribed changes in all enzyme activities. Such an implementation necessitates considerable experimental work and requires either the introduction of strong promoters or the use of additional copies of the relevant genes, for instance, by integration or by using multicopy vectors. A pertinent question, therefore, is whether similar results could be achieved with the modulation of fewer enzymes. Accordingly, we systematically searched for the minimum subset of enzymes necessary to produce the previously calculated optimum solution or a solution that would only be slightly inferior to the optimum. The results for the maximization of ethanol, glycerol, and carbohydrates are shown in Table IV. For each number of modulated enzymes, the combination of steps shown in Table IV yields a stable, feasible solution with the highest target flux.

Table IV.A shows that in order to obtain significant increases in the ethanol flux, the entire set of six enzymes has to be modulated. This result follows the same pattern that emerged when we optimized citric acid production in *A. niger* using the same mathematical procedures (Torres et al., 1996). Table IV.A suggests in which order the enzymes should be modulated to obtain progressively better solutions. Foremost is the substrate uptake rate (X_6), followed closely by ATPase activity (X_{13}). These two steps increase

Table IV. Optimized solutions obtained for combinations of enzymes involved in ethanol (A), glycerol (B), and carbohydrate (C) production in *S. cerevisiae.*

	No. of enzymes	Modulated enzymes	$V_4^+/(V_4^+)_{base}$
A. Ethanol maximization	1	X_6,	1.12
	2	X_6, X_{13}	1.13
	3	X_6, X_8, X_{10}	1.13
	4	X_6, X_8, X_9, X_{10}	1.16
	5	X_6, X_7, X_9, X_{10}, X_{13}	1.29
	6	X_6, X_7, X_8, X_9, X_{10}, X_{13}	3.54
	No. of enzymes	**Modulated enzymes**	$V_{gro}/(V_{gro})_{base}$
B. Glycerol maximization	1	X_{12},	5.91
	2	X_6, X_{12}	7.93
	3	X_6, X_7, X_{12}	8.34
	4	X_6, X_7, X_8, X_{12}	8.81
	5	X_6, X_7, X_8, X_{12}, X_{13}	8.81
	No. of enzymes	**Modulated enzymes**	$V_{carb}/(V_{carb})_{base}$
C. Maximization of carbohydrates	1	X_{11}	33.17
	2	X_6, X_{11}	106.77
	3	X_6, X_7, X_{11}	106.77
	4	X_6, X_7, X_{11}, X_{13}	106.77

Note: Metabolite pools were allowed to vary up to 20% about the base steady state, while the enzyme activities were allowed to vary from 1 to 50 times the baseline steady-state activity (except for glycerol production, for which the upper limit for X_{12} was set to 10 times the base value).

flux by about 10%. The next enzymes to be modulated should be phosphofructokinase (X_8) and the pyruvate kinase (X_{10}). While least important when modulated alone, additional modulation of GADP (X_9) and the hexokinase (X_7) leads to significant flux increases.

An entirely different situation occurs when we maximize glycerol (Table IV.B) or carbohydrate (Table IV.C) production. In the first case we found that by modulating just one enzyme we can already amplify glycerol production over five times. Subsequent modulation of X_6 and X_7 further increases glycerol synthesis to about eight times the base value. In fact, modulation of three enzymes produces essentially the same flux as modulation of five enzymes.

When we optimize carbohydrate production, modulation of X_{11} alone produces a 33-fold increase in flux, while the simultaneous modulation of X_6 and X_{11} further elevates the flux to its maximum (106.77 times the base flux). These results are of practical interest, because they save a significant amount of experimental work when the task is to optimize glycerol or carbohydrate production. They are also interesting from an academic point of view. It appears that all available experimental means have been explored to optimize the system with respect to ethanol production and that further improvements will require rather complicated alterations of several enzymes simultaneously. By contrast, the system has not been optimized as thoroughly with respect to glycerol and carbohydrates, and even ''simple'' means of improvement may not have been fully exploited yet.

It is interesting to note that all effective optimizations of the rate of glycerol production (Table IV.B) include modulation of X_{12}, the enzymatic step involved in the synthesis of glycerol from fructose-1,6-diphosphate, while all optimizations of carbohydrate production (Table IV.C) involve X_{11}, the enzyme responsible for the synthesis of carbohydrates from glucose-6-phosphate. This observation suggests that manipulation of the last step is a strong candidate for modification, if not other information is available. By the same token, the uptake activity (X_6) is the second most influential enzyme in both cases.

DISCUSSION

The question of which mathematical model is the best representation of a metabolic pathway in vivo is yet unanswered. Traditionally, Michaelis–Menten models and their generalizations have often been taken as ''true'' descriptions, but quantitatively more sophisticated analyses in recent times have cast some doubt on the universal adequacy of these established models, (e.g., Hill et al., 1977; Savageau, 1992). As alternatives, various types of power-law models have been proposed, and again, it is not yet decided which of these might be the best choice for a particular purpose. If one agrees that it is too early to identify the best modeling structure, one has to ask how methods developed for one model might translate into other models. One such comparison has been discussed in this article.

Expanding on a previous method for optimizing biochemical and metabolic network models in S-system form, we have applied this approach to the optimization of other types of biochemical network models. The key idea of this method is to approximate the original model with an S-system model, to optimize this approximate model, and to use the resulting profiles of independent variables in the original model.

The result of this indirect method is not guaranteed to be the true optimum, but our findings demonstrate that the approximate enzyme profiles significantly improve the performance of the original model. The quality of the present results is supported by a considerable body of evidence indicating that the approximating S-system is very often a valid system representation. The advantage of using the S-system representation as an intermediate structure is a great reduction in complexity from nonlinear to linear optimization. The latter is straightforward, whereas the former is often plagued by mathematical and computational problems.

In order to test the quality of our method, we performed several comparisons. Using different optimization tasks, we computed the constrained optimum for the approximating S-system model, and subsequently specified the original model with the optimized enzyme profiles. The resulting steady state of the original model then indicated how closely the S-solution matched the results of the original model. We also compared the results with a method recently proposed by Kacser and Acerenza (1993).

In almost all cases, the agreement between the optimized S-solution and the optimized solution for the other models was good. The only case of a significant discrepancy occurred in the optimization of glycerol production, for which the S-solution was stable but the corresponding Michaelis–Menten solution was not. This discrepancy prompted a review of the optimal enzyme profile and a subsequent reoptimization with altered constraints. In Figure 1, this corresponded to an additional loop (going back to step 2.2 and changing constraints on enzyme concentrations) until a stable steady-state was encountered. While no specific guidelines were developed for how to alter the constraints, such redefinitions are no unsurmountable task. Since the approximation is by construction exact at the baseline steady state, a stable, feasible solution always exists, and the redefinition of constraints is a closely confined task.

A well-known feature of the S-system representation is the aggregation of fluxes at branchpoints, which may lead to inaccuracies in the stoichiometry of fluxes. This aggregation is an absolute requirement for the linear optimization of S-system models. Without aggregation, the model has the form of a GMA model, which may or may not be more accurate than the S-system model, but which does not permit the explicit computation of steady states. Therefore, a straightforward linear optimization is impossible.

The magnitude of the difference between the original and the S-system model depends on the pathway under consideration and the distance between the baseline and the opti-

mized steady states. Because of the nonlinear nature of the models involved, general quantitative predictions about optimal states are difficult. However, in the case of S-systems, the optimization problem is linear, and it is known that all solutions are bounded by linear hyperplanes, in logarithmic space. These hyperplanes are defined by the steady-state and flux equations, as well as constraints imposed on the dependent and independent variables. The underlying theory ensures that optimal solutions are found at intersections of these hyperplanes, this being the reason why and how the simplex algorithm of linear optimization works.

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. By contrast, increases in the upper limits of enzyme concentrations did not affect the optimal solutions. Thus, in this particular system, the metabolite constraints turned out to be limiting. Constraints on the stoichiometry of fluxes had similar effects to those observed in a previous study (Torres et al., 1996). Leaving the system ''stoichiometrically'' unconstrained caused the optimum solutions to attain larger values of the production rates (results not shown). However, when other types of flux constraints were imposed (see below), the solution remained unaffected.

It is easy to measure the degree to which the flux stoichiometry in the approximating S-system deviates at a given branchpoint and under specified conditions. If the deviation is too severe, additional constraints can be employed to control the error. In the present study, such a constraint was given as Equation (9). Other constraints were imposed and their effects on the optimal enzyme profile evaluated. For instance, we limited the flux through the branchpoint FDP, X_3, to be less than the glucose input: $(V_3^-)/(V_{1^+}) < 1$. As a second option, we required that the flux toward glycerol be less than twice the glucose uptake: $(V_{gro})/(2 \cdot V_1^+) < 1$. Third, we forced the flux toward the target carbohydrates to be less than the glucose uptake flux: $(2 \cdot V_{carb})/(V_1^+) < 1$. These additional constraints did not affect the optimal enzyme profiles much when ethanol or carbohydrates were the target compounds. By contrast, the fluxes in the S-solution were decreased in the optimization of glycerol when additional constraints were imposed. In this particular case, the MM solution showed other problems as well. In particular, the solution was stable only when the maximal amplification of enzyme activities was limited to 15, as opposed to 50 or more in other optimizations. These results point to discrepancies between the original and the S-system model that are of a numerical nature as long as the system is close to the base steady state but become structural if the system is moved too far away from the original state.

Kacser and Acerenza (1993) discuss limitations of their optimization method that derive from the occurrence of some rather common kinetic features of metabolic systems. These include (i) biomolecularity associated with coupled cofactors, (ii) feedbacks, (iii) allosteric enzymes, (iv) substrate cycles, and (v) nonlinear versus linear kinetics of enzymes, as they occur in enzyme–enzyme interactions. In the procedure proposed here, the occurrence of such features is explicitly taken into account through the model design of the pathway and does not constitute a limitation. For instance, the experimental system of the present study involves bimolecularity associated with coupled cofactors; there are a feedback and a feedforward modulation; and the kinetics of some of the enzymes are allosteric. These features are automatically implemented in the S-system representation and thus are an integral part of the optimized linear steady-state system. The experimental pathway does not contain enzyme–enzyme interactions, but these would not constitute a limitation to the generality of the method either, since it has been shown that any type of kinetics or enzyme–enzyme interaction can be validly represented with an S-system model (Sorribas and Savageau, 1989). This generality directly translates into the proposed optimization method.

At any rate, it is interesting to note that the differences between the results from the method of Kacser and Acerenza (1993) and from the IOM approach outlined here are minor. Given that both methods are based on different types of assumptions and that the true mathematical nature of biochemical systems is unknown, it is as yet impossible to decide which method produces results that are closer to the true optimum.

Preliminary results in our laboratory indicate that the optimization strategy applied here can directly be applied to other types of systems, such as bioreactors, for which the underlying processes and pathway structures are different. This could be of considerable relevance, since bioreactors permit extensive control over the system, thereby providing a better chance of successfully implementing predicted solutions. The linear treatment of nonlinear systems, via the S-system representation, furthermore suggests generalizations of the method toward streamlined multiobjective or multilevel optimization (e.g., Candler and Norton, 1977, Candler and Townsley, 1982; Clark, 1990; Clark and Westerberg, 1990).

References

Candler, W., Norton, R. 1977. *Multilevel programming and development policy,* World Bank Staff Working Paper No. 258. IBRD, Washington, DC.

Candler, W., Townsley, R. 1982. A linear two-level programming problem. Comput. Ops. Res. **9**: 59–76.

Cascante, M., Curto, R., Sorribas, A. 1995. Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using Biochemical Systems Theory and Metabolic Control Analysis: Steady-state analysis. Math. Biosc. **130**: 51–69.

Clark, P. A. 1990. Bilevel programming for steady-state chemical process design. II. Performance study for nondegenerate problems. Comput. Chem. Eng. **14**: 99–109.

Clark, P. A., Westerberg, A. W. 1990. Bilevel programming for steady-

state chemical process design. I. Fundamentals and algorithms. Comput. Chem. Eng. **14**: 87–97.

Curto, R., Sorribas, A., Cascante, M. 1995. Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using Biochemical Systems Theory and Metabolic Control Analysis: Model definition and nomenclature. Math. Biosc. **130**: 25–50.

Davies, S. E. C., Brindle, K. M. 1992. Effects of overexpression of phosphofructokinase on glycolysis in the yeast *Saccharomyces cerevisiae*. Biochemistry **331**: 4729–4735.

Galazzo, J. L., Bailey, J. E. 1990. Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae.* Enz. Microbiol. Technol. **12**: 162–172.

Galazzo, J. L., Bailey, J. E. 1991. Errata. Enz. Microbiol. Technol. **13**: 363–371.

Guarante, G., Gail, L., Roberts, T. M., Ptashne, M. 1980. Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β-globin. Cell **20**: 543–553.

Hatzimanikatis, C., Floudas, C. A., Bailey, J. E. 1996. Analysis and design of metabolic reaction networks via mixed-integer linear optimization. AIChE J. **42** (5): 1277–1292.

Heinish, J. 1986. Isolation and characterization of the two structural genes encoding for phosphofructokinase in yeast. Mol. Gen. Genet. **202**: 75–80.

Heinrich, R., Rapoport, T. A. 1974. A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. Eur. J. Biochem. **42**: 89–95.

Heinrich, R., Rapoport, S., Rapoport, T. A. 1977. Metabolic regulation and mathematical models. Prog. Biophys. Molec. Biol. **32**: 1–82.

Hill, C. M., Waight, R. D., Bardsley, W. G. 1977. Does any enzyme follow the Michaelis–Menten equation? Mol. Cell. Biochem. **15**: 173–178.

Kacser, H., Acerenza, L. 1993. A universal method for achieving increases in metabolite production. Eur. J. Biochem. **216**: 361–367.

Kacser, H., Burns, J. 1973. The control of flux. Symp. Soc. Exp. Biol. **27**: 65–104.

Ni, T.-C., Savageau, M. A. 1996a. Model assessment and refinement using strategies from biochemical system theory: Application to metabolism in human red blood cells. J. Theoret. Biol. **179**: 329–368.

Ni, T.-C., Savageau, M. A. 1996b. Application of biochemical system theory to metabolism in human red blood cell: Signal propagation and accuracy of representation. J. Biol. Chem. **271**(14): 7927–7941.

Okamoto, M., Savageau, M. 1984. Integrated function of a kinetic proofreading mechanism: Steady-stage analysis testing internal consistency of data obtained *in vivo* and *in vitro* and predicting parameter values. Biochemistry **23**: 1701–1709.

Okamoto, M., Savageau, M. 1986. Integrated function of a kinetic proofreading mechanism: Double-state proofreading by isoleucyl-tRNA synthetase. Biochemistry **25**: 1969–1975.

Pettersson, G. 1992. Evolutionary optimization of the catalytic efficiency of enzymes. Eur. J. Biochem. **206**: 289–295.

Reich, J. G., Sel'kov E. E. 1981. Energy Metabolism of the Cell. A Theoretical Treatise. Academic, London.

Regan, L., Bogle, I. D. L., Dunnill, P. 1993. Simulation and optimization of metabolic pathways. Comp. Chem. Eng. **175-6**: 627–637.

Savageau, M. 1969a. Biochemical system analysis, I. Some mathematical properties of the rate law for the component enzymatic reactions. J. Theor. Biol. **25**: 365–369.

Savageau, M. 1969b. Biochemical system analysis, II. The steady-state solutions for an *n*-pool system using a power law approximation. J. Theor. Biol. **25**: 370–379.

Savageau, M. 1971. Concepts relating the behavior of biochemical systems to their underlying molecular properties. Arch. Biochem. Biophys. **145**: 612–621.

Savageau, M. 1976. Biochemical system analysis: A study of function and design in molecular biology. Addison-Wesley, Reading, MA.

Savageau, M. 1992. A critique of the enzymologist's test tube, pp. 45–1087. In: E. E. Bittan (ed.), Fundamentals of medical cell biology, vol. 3A. JAI Press, Greenwich, CT.

Savageau, M., Voit, E. O. 1987. Recasting nonlinear differential equations as S-systems: A canonical nonlinear form. Math. Biosci. **87**: 83–115.

Savageau, M., Voit, E. O., Irvine, D. H. 1987a. Biochemical system theory and metabolic control theory, 1. Fundamental similarities and differences. Math. Biosci. **86**: 127–145.

Savageau, M., Voit, E. O., Irvine, D. H. 1987b. Biochemical system theory and metabolic control theory, 2. The role of summation and connectivity relationships. Math. Biosci. **86**: 147–169.

Schaaff, I., Heinish, J., Zimmermann, F. K. 1989. Overproduction of glycolytic enzymes in yeast. Yeast **5**: 285–290.

Schuster, S., Heinrich, R. 1991. Minimization of intermediates concentrations as a suggested optimality principle for biochemical networks. I. Theoretical analysis. J. Math. Biol. **29**: 425–442.

Schuster, S., Schuster, R., Heinrich, R. 1991. Minimization of intermediates concentrations as a suggested optimality principle for biochemical networks. II. Time hierarchy, enzymatic rate laws, and erythrocyte metabolism. J. Math. Biol. **29**: 443–455.

Shiraishi, F., Savageau, M. 1992. The tricarboxylic acid cycle in *Dictyostelium discoideum.* I. Formulation of alternative kinetic representations. J. Biol. Chem. **267**: 22912–22918.

Sorribas, A., Savageau, M. 1989. Strategies for representing metabolic pathways within biochemical system theory: Reversible pathways. Math. Biosci. **94**: 161–193.

Sorribas, A., Curto, R., Cascante, M. 1995. Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using Biochemical Systems Theory and Metabolic Control Analysis: Model validation and dynamic behaviour. Math. Biosci. **130**: 71–84.

Torres, N. V., Voit, E. O., González-Alcón, C. 1996. Optimization of nonlinear biotechnological processes with linear programming: Application to citric acid production by *Aspergillus niger.* Biotechnol. Bioeng. **49**: 247–258.

Voit, E. O. 1991. Canonical Nonlinear Modeling: S-System Approach to Understanding Complexity. Van Nostrand Reinhold, New York.

Voit, E. O. 1992. Optimization in integrated biochemical systems. Biotechnol. Bioeng. **40**: 572–582.

Voit, E. O., Savageau, M. 1987. Accuracy of alternative representations for integrated biochemical systems. Biochemistry **26**: 6869–6880.

Voit, E. O., Irvine, D. H., Savageau, M. 1990. User's Guide to ESSYNS. 2nd edition. Medical University of South Carolina Press, Charleston, SC.

Westerhoff, H. V., Kell, D. B. 1987. Matrix method for determining steps most rate-limiting to metabolic fluxes in biotechnological processes. Biotechnol. Bioeng. **30**: 101–107.