Analysis and Prediction of the Effect of Uncertain Boundary Values in Modeling a Metabolic Pathway

Pedro de Atauri,1 Albert Sorribas,2 Marta Cascante1

1Departament de Bioquímica i Biologia Molecular, Facultat de Ciències Químiques, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain; telephone: +34-93-4021593; e-mail: marta@sun.bq.ub.es
2Departament de Ciències Mèdiques Bàsiques, Universitat de Lleida, Avinguda Rovira Roure 44, 25198-Lleida, Spain

Received 10 March 1999; accepted 19 September 1999

Abstract: The integration of large quantities of biological information into mathematical models of cell metabolism provides a way for quantitatively evaluating the effect of parameter changes on simultaneous, coupled, and, often, counteracting processes. From a practical point of view, the validity of the model’s predictions would critically depend on its quality. Among others, one of the critical steps that may compromise this quality is to decide which are the boundaries of the model. That is, we must decide which metabolites are assumed to be constants, and which fluxes are considered to be the inputs and outputs of the system. In this article, we analyze the effect of the experimental uncertainty on these variables on the system’s characterization. Using a previously defined model of glucose fermentation in Saccharomyces cerevisiae, we characterize the effect of the uncertainty on some key variables commonly considered to be constants in many models of glucose metabolism, i.e., the intracellular pH and the pool of nucleotides. Without considering if this variability corresponds to a possible true physiological phenomenon, the goal of this article is to illustrate how this uncertainty may result in an important variability in the systemic responses predicted by the model. To characterize this variability, we analyze the utility and limitations of computing the sensitivities of logarithmic-gains (control coefficients) to the boundary parameters. With the exception of some special cases, our analysis shows that these sensitivities are good indicators of the dependence of the model systemic behavior on the parameters of interest. © 2000 John Wiley & Sons, Inc. Biotechnol Bioeng 68: 18–30, 2000.

Keywords: model validation; sensitivity analysis; mathematical modeling; glucose fermentation

INTRODUCTION

A highly regulated and coupled network of enzyme-catalyzed reactions and selective transport systems accomplishes the metabolic activities of living cells. In each case, natural selection has lead to optimal systems that allows the cell to be competitive in its natural environment. However, an optimal design in a natural setting is not necessarily the optimum design from a technological point of view. The possibility of modifying a natural process to achieve a new capability, from the optimization of batch cultures to the design of new therapeutic drugs, is an important goal in biotechnology (Bailey, 1991).

A rational approach to metabolic engineering must analyze the characteristics of the reaction network before attempting specific genetic manipulations. Mathematical models provide a support tool for critically exploring the available experimental data. One of the crucial points for developing useful mathematical models in metabolic engineering is to develop strategies for appropriately validating such models (Savageau, 1971; Shiraishi and Savageau, 1992a,b,c,d, 1993; Curto et al., 1995, 1997, 1998a,b; Cascante et al., 1995; Sorribas et al., 1995; Ni and Savageau, 1996a,b). The implications of the impossibility of developing a complete, detailed model of a metabolic system including all the cellular processes that have an impact in a given pathway have been considered in a systematic manner by Schlosser and Bailey (1990). Apart from that, there are different potential sources of inaccuracies when building a mathematical model. Among others, the more critical are: 1) the selected mathematical representation is not adequate for characterizing the biological system, 2) there are inaccuracies on the selected kinetic parameter values, and 3) the model contains simplifications that compromise its validity.

The inadequacy of the mathematical formalism chosen to achieve the goals desired in the characterization of the system can be another source of inaccuracies. The choice of a particular mathematical formalism is critical and determines the kind of analysis that can be performed on the model. Among other requirements, the mathematical representation must facilitate computing the profile of parameter sensitivities, since this profile can be used as a diagnostic tool for suggesting portions of the model that need further improvement (Ni and Savageau, 1996a,b; Cascante et al., 1995; Curto et al., 1997). The inaccuracies due to uncertainties in
the kinetic parameter values are also critical for evaluating the model predictions (Small and Fell, 1990; Thomas and Fell, 1994; Pettersson, 1996a,b; Petkov and Maranas, 1997). In addition, the conceptual model can incorporate different simplifications that can compromise the model predictions. These simplifications lead us to consider some variables as constant parameters, disconnecting the metabolic pathway from the rest of the metabolic processes. While necessary for making feasible the construction of a model, these simplifications can be a potential source of biased results. The use of auxiliary processes that allows a buffering effect on key variables can overcome some of these problems (Voit and Ferreira, 1998). However, when the involved variables participate in different kinetic processes, a detailed analysis of the effects of uncertainties in these variables is necessary as a validation step for the model.

In this article we will not focus on a detailed analysis of a particular physiological system, but on the implications resulting from considering arbitrary boundaries in models of metabolic pathways that incorporate complex enzyme kinetics equations. This can be the case, for instance, of considering constant the amount of nucleotides in a given model. This may be necessary for simplifying the model, but the uncertainties related to the experimental determination of this parameter can have a dramatic effect on the system’s properties. A critical step in the process of model definition is checking the system’s behavior for different values of these parameters. Without considering if uncertainties correspond to physiological phenomena, when it comes to considering the model’s utility, parameter uncertainty can be accepted if this does not involve dramatic changes in qualitative behavior. Otherwise, further refinements would be needed to clarify the appropriate values. From a practical point of view, the most common situations leading to these simplifications include considering constraints between dependent intermediates and fixing parameter values, e.g., by considering a particular fast reaction to be in equilibrium or by considering the variability of a concentration metabolite to be negligible. Implicitly or explicitly, through a mathematical model representing the metabolic system all these fixed parameter uncertainties spread through the calculations of the properties of the system and may compromise its validity. At this point, it is critical to check the effects of these simplifications.

MATERIAL AND METHODS

Model Description

As a case example, we shall use a model of the anaerobic fermentation pathway from glucose to ethanol, glycerol, and polysaccharides in Saccharomyces cerevisiae published by Galazzo and Bailey (1990, 1991). This model is a combination of experimental measurements and literature research, and contains some simplifications resulting from the experimental conditions used, e.g., the absence of a nitrogen source. Although many of the involved parameters were specifically determined in this study, others were assumed to be equal to those known in other cells. The experimental conditions, and the resulting models, were chosen to explore the difference between immobilized yeast cells entrapped in calcium alginate and suspended cells, both in two different cell environments defined by external pH 4.5 and 5.5. As a reference system, we will only consider the immobilized cell at pH 5.5 condition. The scheme for the fermentation pathway is shown in Figure 1, and a description of the kinetic model is shown in the Appendix section. In this model, we indicate the law-rules used to represent the involved kinetics and the metabolite constraints and equilibrium relationships considered. Corresponding to the experimental conditions indicated above, the total amount of adenine nucleotides (AN) and the intracellular pH (pH\textsuperscript{in}) are considered constants, introducing a necessary simplification that defines a boundary of the studied system.

![Figure 1. Anaerobic fermentation pathway of yeast Saccharomyces cerevisiae from glucose to ethanol, glycerol, and polysaccharides. This metabolic pathway involves five dependent concentrations and eight fluxes. Chemical species: Glc, glucose; G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; FDP, fructose-1,6-diphosphate; PEP, phosphoenolpyruvate; poolpol, polysaccharide pool (glycogen + trehalose). Enzymes/pathway steps: \( V_{\text{in}} \), glucose uptake; \( V_{\text{hk}} \), hexokinase; \( V_{\text{pfk}} \), phosphofructokinase; \( V_{\text{gapd}} \), glyceraldehyde 3-phosphate dehydrogenase; \( V_{\text{pk}} \), pyruvate kinase; \( V_{\text{pol}} \), polysaccharide storage (glycogen + trehalose); \( V_{\text{gol}} \), glycerol production; \( V_{\text{ATPase}} \), ATP consumption.](image-url)
Although the model considers fixed quantities for these parameters, their interval of uncertainty can be obtained from the original article. This makes this model appropriate as an example for analyzing the effect of selecting boundaries in a metabolic model. In the model of Galazzo and Bailey (1990, and references therein) the concentrations of sugar phosphate intermediates and the summation of unbound ATP and ADP were estimated from data obtained by in vivo $^{31}$P nuclear magnetic resonance spectroscopy (NMR) (Galazzo and Bailey, 1989). In vivo pH$^\text{in}$ was determined from the chemical shift of intracellular P$_i$ resonance using an NMR titration curve, where the error in the determination of pH is expected to be less than 0.2 pH units (Shanks and Bailey, 1988). In the original model, the considered value of pH$^\text{in}$ is 6.80 for immobilized cells at external pH 5.5. If we take an error lower than 0.2 pH units in the determination of pH, the uncertainty interval around the reference value experimentally determined may be estimated to be between 6.7 and 6.9. For AN, these authors estimate a mean value of 3 mM, with an uncertainty interval between 2.7 and 3.3 mM. Since no experimental data exist on the statistical distribution, we will assume a uniform distribution of the possible values of these parameters within their uncertainty interval. Assumption of a different distribution would lead, essentially, to similar qualitative conclusions.

**Model Equations and Steady-State Adjustments**

The model is defined as a set of ordinary differential equations that incorporates the different kinetic expressions for each rate-law (see Appendix). To evaluate model performance, we will consider changes with respect to two different categories of parameters: the boundary parameters (AN and pH$^\text{in}$) and the step activities ($V_{\text{max}}$). At this point, it is important to introduce a difference between parameters that determine the system’s design and parameters that are used to explore system behavior (Ni and Savageau, 1996a,b). Design parameters are those parameters that are considered to change between two different situations of interest. For instance, a pH$^\text{in}$ equal to 6.9 determines a different environment than a pH$^\text{in}$ equal to 7.1. Then, pH$^\text{in}$ is considered a design parameter in our case. In each condition defined by a different value of pH$^\text{in}$, we are interested in checking the system’s performance. Behavior parameters are those parameters used to check the system response at different values of the design parameters. In our case, the response to a change in a $V_{\text{max}}$ is used as a measurement of performance for comparing different designs. Then, $V_{\text{max}}$ is a behavior parameter in this case. The distinction between design and behavior parameters is important.

Each time we change the value of a design parameter we must assure that an adjustment is introduced so that the different designs are comparable (Savageau, 1974, 1977; Irvine and Savageau, 1985a,b; Ni and Savageau, 1996a,b; Hlavacek and Savageau, 1995, 1996, 1997). Since in the original model of Galazzo and Bailey the steady-state values were obtained experimentally, some adjustments were required so that the model could match these values for a given set of design parameters. In that case, the ratio between reduced and oxidized NAD (NADH/NAD$^+$) and the $V_{\text{max}}$ of several enzymes were considered as model adjustable parameters. In the Appendix, an asterisk identifies each adjustable parameter. For instance, each time a different design parameter AN or pH$^\text{in}$ is considered, we must adjust the model to maintain the same reference steady-state changing the value of the model adjustable parameters. Otherwise, the resulting comparisons are not valid.

**Steady-State Behavior Characterization**

Steady-state characterization involves quantification of the system response to changes in external variables through the use of sensitivity analysis. This characterization on the original model is difficult because of the complexity of the involved equations. As an alternative, the tools provided by biochemical systems theory (BST) and metabolic control analysis (MCA) are more convenient (for a complete account of methods, see Cascante et al., 1989a,b; Fell, 1997; Curto et al., 1995, 1997, 1998a,b; Sorribas et al., 1995; Cascante et al., 1995; and references therein). In many aspects, BST and MCA use quasi-equivalent representations and yield similar results. Operationally, BST and MCA differ in including an explicit representation for the rate-equations. In that sense, BST provide a differential equation model in power-law form that allows model simulation and a straightforward analysis of the system’s properties. MCA considers an implicit representation of the underlying processes leading to a compact set of matrix equations to derive the steady-state characterization.

In both approaches, the steady-state characterization is provided by computing systemic sensitivities, defined as a normalized, or eventually nonnormalized, partial derivative of any systemic property; for example, an intermediate metabolite pool or a steady-state flux, with respect to any independent variable. As a general definition, a systemic sensitivity is computed as:

$$S(F, p) = \left( \frac{d\ln(F)}{d\ln(p)} \right)_0 = \left( \frac{dF}{dp} \right)_0 \cdot \frac{p_0}{F_0}$$

where $F$ can be any systemic property, and $p$ can be any behavior parameter, for example, an external effector or an enzyme activity (Savageau and Sorribas, 1989). The subscript $0$ refers to the nominal operating point. Because BST uses an explicit power-law for representing the underlying processes, parameter sensitivities include the case in which $p$ is a rate constant or a kinetic order (see below).

Systemic sensitivities are called *logarithmic gains* and *parameter sensitivities* in BST, and *control coefficients* or *combined response coefficients* in MCA. The reader is referred to the literature for an account of equivalencies and for comparing the operating similitude and differences between both approaches (Sorribas and Savageau, 1989a,b,c;
Curto et al., 1995; Savageau, 1991, 1992). In the following, we use the BST nomenclature, keeping in mind that logarithmic-gains in BST correspond to control coefficients and response coefficients in MCA.

Systematic sensitivities depend on the local sensitivities defined as normalized partial derivative of any individual reaction rate \( v_i \) with respect to any dependent or independent variable that affects this rate. In MCA, these sensitivities are known as elasticities, and in BST are known as kinetic orders. If we consider that \( v_i \) is a function of a set of variables and parameters, \( v_i(X_1, X_2, \ldots) \), a kinetic order is defined as:

\[
    f_{ij} = \left( \frac{d \ln(v_i)}{d \ln(X_j)} \right)_{0}
\]

The methods for computing systematic sensitivities from the local sensitivities can be found in the literature (see, for instance, Cascante et al., 1989a,b; Savageau and Sorribas, 1989). These methods have been implemented in a function package, SYMMET (under development in our laboratory) using Mathematica 3.0 (Wolfram, 1996).

Steady-state characterization can be complemented by analyzing nonnormalized second-order sensitivities, i.e., the sensitivities of systemic sensitivities with respect to fixed parameters considered as boundaries of the metabolic system analyzed (Savageau, 1971, 1976):

\[
    \left( \frac{dS(F, p)}{d(\text{boundary parameter})} \right)_{0}
\]

These sensitivities will be computed numerically by infinitesimally perturbing the boundary parameter around the operational point. Because boundary parameters determine the system’s design, steady-state adjustment is required after any change in their values. These changes assure comparability between the different scenarios considered. In the Results section, all the required computations were performed using Mathematica 3.0.

**Analysis of the Effect of Changes in Boundary Parameters**

The second-order sensitivities of logarithmic-gains predict a linear dependency of the logarithmic gains with regard to changes in boundary parameters around the operational steady-state (Salvador, 1996). The actual logarithmic gains corresponding to each value of a boundary parameter within the uncertainty interval were computed directly from the kinetic model by using SYMMET, according to the methods described above. These logarithmic gains were used as reference for comparison with the values predicted by the sensitivity coefficients.

Stability of a given steady-state is a fundamental condition for the existence of a system and for ensuring that the parameters used in the model are compatible with a realistic situation (Sorribas and Savageau, 1989a; Savageau and Sorribas, 1989; Savageau, 1972). When a value in boundary parameters changes, the model is adjusted to preserve the reference steady state but the stability of the resulting model may be compromised. Then, in addition to the analysis of systemic sensitivities, a stability analysis through the same parameters’ intervals is necessary to identify critical parameter regions. The local stability of the nominal steady state can be evaluated by examination of the eigenvalues of the characteristic equation after a linearization of the local representation of the system (Savageau, 1976). If the real parts of all eigenvalues are negative, the nominal steady-state is locally stable. In this case, the system will return to its steady-state following small perturbations.

**RESULTS AND DISCUSSION**

The results presented in this article are organized in two blocks. First, we compute the steady-state characterization at the nominal operational point of the model through logarithmic gains and through the second-order sensitivities of logarithmic gains with regard the boundary parameters. Second, using the kinetic model, we systematically recomputed the values of the logarithmic gains within the interval of uncertainty of the boundary parameters. These values are compared with the values predicted from second-order sensitivities of logarithmic gains. Since stability of the reference steady state can be compromised when the value of boundary parameters is changed, the local stability of the model in the range of uncertainty of these boundary parameters is also checked.

**Steady-State and Steady-State Behavior Characterizations**

The steady-state characterization of the reference model of the anaerobic fermentation pathway described in the above section was published elsewhere (Curto et al., 1995; Cascante et al., 1995; Sorribas et al., 1995). We use these results as a reference. Logarithmic gains with respect to each step activity (\( V_{\text{max}}^{\text{step}} \)) are shown in Table I. All of them have absolute values between 0.1 and 2.5, with the exception of the high values observed in the gains of phosphoenolpyruvate (PEP). Regarding ethanol production, it should be noted that the steps displaying a higher logarithmic gain were glucose uptake (\( V_{\text{in}} \)), phosphofructokinase (\( V_{\text{pfk}} \)) and ATP consumption (\( V_{\text{ATPase}} \)).

The logarithmic-gain values can be used to tentatively predict the steady-state system response to small perturbations in any of the independent variables. From a theoretical point of view, these predictions are strictly true only if the changes in the independent variables are infinitesimal. In practice, although they may be inaccurate for large changes, in many cases they suffice for predicting the system response within a physiological range of variation of the independent variables (see for instance Sorribas and Savageau, 1989a; Curto et al., 1998b). In our case, we concentrated our attention in the step activities as behavior characterizations.
parameters to check for system behavior. Logarithmic gains are a function of the underlying kinetic orders and they can be easily computed once the system scheme is defined (see, for instance, Cascante et al., 1995). In general, the strategy for analyzing a given system consists in fixing an operating point in which the kinetic orders are computed from the existing information. Once computed, the logarithmic gains characterize the response of the system at this point.

According to their definition, kinetic orders depend on the boundary parameters. Consequently, logarithmic gains will change as we change any of these conditions. Despite this, in many cases their values may be relatively similar for a given range of boundary conditions. If this is the case, it is an indication that the system response does not change qualitatively within the considered range.

As a validation procedure, it may be a good strategy to introduce an intermediary step in which kinetic orders are computed as a function of critical variables. This may be the case of expressing the kinetic orders as a function of AN and pH. If these dependencies are made explicit, the corresponding logarithmic gains will be a function of AN and pH. Accordingly, we can compute the sensitivity of a given logarithmic gain to changes in those design parameters. These sensitivities will be computed as nonnormalized second-order sensitivities with the general form indicated in Eq. (3).

In the considered model, the analysis of second-order sensitivities shows high values with respect to AN and pH (see Tables II and III), with the exception of these of hexokinase (V_{hk}) and V_{gapd} that are almost all zero and are not shown in these tables. In particular, the second-order sensitivities of the different logarithmic gains in the case of PEP are extremely high. Regarding ethanol production, second-order sensitivities respect to infinitesimal changes in AN or in pH show, in both cases, the highest values for V_{pk}.

If these dependencies are made explicit, the corresponding logarithmic gains will be a function of AN and pH. In the considered model, the analysis of second-order sensitivities shows high values with respect to AN and pH (see Tables II and III), with the exception of these of hexokinase (V_{hk}) and V_{gapd} that are almost all zero and are not shown in these tables. In particular, the second-order sensitivities of the different logarithmic gains in the case of PEP are extremely high. Regarding ethanol production, second-order sensitivities respect to infinitesimal changes in AN or in pH show, in both cases, the highest values for V_{pk}.

Table I. Flux and concentration logarithmic gains computed at the nominal steady state described in the Appendix.

<table>
<thead>
<tr>
<th>Activities</th>
<th>( V_{\text{max}}^{\text{in}} )</th>
<th>( V_{\text{max}}^{\text{hk}} )</th>
<th>( V_{\text{max}}^{\text{pk}} )</th>
<th>( V_{\text{max}}^{\text{gapd}} )</th>
<th>( V_{\text{max}}^{\text{pk}} )</th>
<th>( V_{\text{max}}^{\text{pol}} )</th>
<th>( V_{\text{max}}^{\text{gol}} )</th>
<th>( V_{\text{max}}^{\text{ATPase}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{\text{in}} )</td>
<td>0.43</td>
<td>0.00</td>
<td>0.25</td>
<td>0.00</td>
<td>−0.03</td>
<td>0.21</td>
<td>0.03</td>
<td>0.11</td>
</tr>
<tr>
<td>( V_{\text{hk}} )</td>
<td>0.43</td>
<td>0.00</td>
<td>0.25</td>
<td>0.00</td>
<td>−0.03</td>
<td>0.21</td>
<td>0.03</td>
<td>0.11</td>
</tr>
<tr>
<td>( V_{\text{pk}} )</td>
<td>0.37</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
<td>−0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.18</td>
</tr>
<tr>
<td>( V_{\text{gapd}} )</td>
<td>0.37</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
<td>0.03</td>
<td>0.02</td>
<td>−0.03</td>
<td>0.18</td>
</tr>
<tr>
<td>( V_{\text{pol}} )</td>
<td>0.37</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
<td>−0.07</td>
<td>0.07</td>
<td>−0.03</td>
<td>−0.12</td>
</tr>
<tr>
<td>( V_{\text{gol}} )</td>
<td>0.63</td>
<td>0.00</td>
<td>−0.28</td>
<td>0.00</td>
<td>0.03</td>
<td>0.07</td>
<td>−0.03</td>
<td>0.18</td>
</tr>
<tr>
<td>( V_{\text{gol}} )</td>
<td>0.37</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
<td>−0.97</td>
<td>0.02</td>
<td>0.97</td>
<td>0.18</td>
</tr>
<tr>
<td>( V_{\text{ATPase}} )</td>
<td>0.20</td>
<td>0.00</td>
<td>0.86</td>
<td>0.00</td>
<td>0.15</td>
<td>−0.44</td>
<td>−0.15</td>
<td>0.37</td>
</tr>
<tr>
<td>Glc</td>
<td>0.90</td>
<td>−2.12</td>
<td>0.48</td>
<td>0.00</td>
<td>−0.07</td>
<td>0.06</td>
<td>0.07</td>
<td>0.27</td>
</tr>
<tr>
<td>FDP</td>
<td>0.83</td>
<td>0.00</td>
<td>0.74</td>
<td>−2.46</td>
<td>0.01</td>
<td>0.21</td>
<td>−0.01</td>
<td>0.68</td>
</tr>
<tr>
<td>PEP</td>
<td>25.7</td>
<td>0.00</td>
<td>60.6</td>
<td>0.01</td>
<td>−34.3</td>
<td>−20.5</td>
<td>−8.65</td>
<td>−22.9</td>
</tr>
<tr>
<td>ATP</td>
<td>0.20</td>
<td>0.00</td>
<td>0.86</td>
<td>0.00</td>
<td>0.15</td>
<td>−0.44</td>
<td>−0.15</td>
<td>−0.63</td>
</tr>
</tbody>
</table>

High values of PEP and ethanol production (V_{pk}) logarithmic gains are indicated in bold face.

Table II. Second-order sensitivities of flux and concentration logarithmic gains respect to infinitesimal changes in pH in at the nominal steady-state described in the Appendix.

<table>
<thead>
<tr>
<th>Activities</th>
<th>( V_{\text{max}}^{\text{in}} )</th>
<th>( V_{\text{max}}^{\text{pk}} )</th>
<th>( V_{\text{max}}^{\text{pol}} )</th>
<th>( V_{\text{max}}^{\text{gol}} )</th>
<th>( V_{\text{max}}^{\text{ATPase}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{\text{in}} )</td>
<td>0.45</td>
<td>−1.18</td>
<td>−0.13</td>
<td>0.21</td>
<td>0.13</td>
</tr>
<tr>
<td>( V_{\text{hk}} )</td>
<td>0.45</td>
<td>−1.18</td>
<td>−0.13</td>
<td>0.21</td>
<td>0.13</td>
</tr>
<tr>
<td>( V_{\text{pk}} )</td>
<td>0.76</td>
<td>−2.00</td>
<td>−0.22</td>
<td>0.36</td>
<td>0.22</td>
</tr>
<tr>
<td>( V_{\text{gapd}} )</td>
<td>0.76</td>
<td>−2.00</td>
<td>−0.22</td>
<td>0.36</td>
<td>0.22</td>
</tr>
<tr>
<td>( V_{\text{pol}} )</td>
<td>0.50</td>
<td>1.32</td>
<td>0.14</td>
<td>−0.23</td>
<td>−0.14</td>
</tr>
<tr>
<td>( V_{\text{gol}} )</td>
<td>0.76</td>
<td>−2.00</td>
<td>−0.22</td>
<td>0.36</td>
<td>0.22</td>
</tr>
<tr>
<td>( V_{\text{ATPase}} )</td>
<td>1.54</td>
<td>−4.05</td>
<td>−0.44</td>
<td>0.72</td>
<td>0.44</td>
</tr>
<tr>
<td>Glc</td>
<td>0.86</td>
<td>−2.26</td>
<td>−0.25</td>
<td>0.40</td>
<td>0.25</td>
</tr>
<tr>
<td>FDP</td>
<td>−1.37</td>
<td>3.59</td>
<td>0.39</td>
<td>−0.64</td>
<td>−0.39</td>
</tr>
<tr>
<td>PEP</td>
<td>155.</td>
<td>−175.</td>
<td>−92.5</td>
<td>13.4</td>
<td>15.1</td>
</tr>
<tr>
<td>ATP</td>
<td>1.54</td>
<td>−4.05</td>
<td>−0.44</td>
<td>0.72</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Each second-order sensitivity is in correspondence with the logarithmic gains in Table I. High values of second-order sensitivities of PEP and ethanol production (V_{pk}) logarithmic gains are indicated in bold face.
Second-order sensitivities of $V_{\text{in}}$ and $V_{\text{ATPase}}$ logarithmic gains with regard to $\text{pH}^m$ are also relatively high. These high second-order sensitivities predict important changes in the relative qualitative importance of the different steps in the pathway associated with the uncertainty interval for the boundary parameters.

### Analysis of Second-Order Sensitivity Predictions

Second-order sensitivities predict the change in systemic properties as a result of changing the design parameters. From a theoretical point of view, these sensitivities are an approximation to the actual behavior of the system at the operating point. For illustrative purposes, it would be interesting to compare the result of this analysis with the system’s behavior obtained by simulation of the original model. This comparison will help in understanding the utility of these sensitivities as a validation tool in the considered model.

First, the logarithmic-gains for the reference model are computed systematically at different values of the boundary parameters $\text{pH}^m$ and AN. This is made by selecting a value for each parameter and by computing the corresponding kinetic orders. In each case, the appropriate adjustments are made so that the reference steady state is preserved. After these computations, the logarithmic gains are obtained as indicated in the Methods section.

Once the logarithmic gains are computed, the analysis is first made through one-dimensional graphical plots to show the effect of each design parameter. Then, through two-dimensional graphical plots, we show the joint effect of simultaneous changes in both parameters. Figure 2 shows the step logarithmic gains of PEP and ethanol production ($V_{\text{pk}}$). This analysis clearly shows a qualitative change in those systemic properties, including some positive-negative changes. This is especially remarkable with respect to glucose uptake ($V_{\text{in}}$) and phosphofructokinase ($V_{\text{pfk}}$) perturbations. In Figure 2, the predictions of the sensitivities of logarithmic gains at the nominal operational point appear as tangent lines to the series of logarithmic gains. In general, these results clearly show the predictive validity of the second-order sensitivity analysis. In some special cases, it may occur that the change in logarithmic gains is not monotonous. For instance, the dependency of logarithmic gains of PEP with respect to changes in $V_{\text{pfk}}$ by ranging AN (see Fig. 2c). In those cases, the prediction from the second-order sensitivities is questionable.

Figures 3 and 4 show the analysis of the logarithmic gains of PEP and ethanol production ($V_{\text{pk}}$) on the two-dimensional region described by simultaneously considering the $\text{pH}^m$ and AN uncertainty intervals. Contour lines identify equivalent logarithmic gain values. From these results it is immediately seen that $\text{pH}^m$ has a higher effect than AN on the final values. Moreover, the curvature of the lines also shows that the dependency is not necessarily monotonous. Before attempting an interpretation of these results, a stability analysis is done in parallel with the computation of logarithmic gains. Analysis of stability through examination of eigenvalues shows areas where the real parts of the eigenvalues of the characteristic equation are positive. These two areas of instability are represented in black in Figures 3 and 4. One of these areas shows positive eigenvalues and corresponds to the maximum $\text{pH}^m$ and minimum AN interval values. For the considered interval of parameter values, this area is a small region. However, it is important to note that it may become relevant for a simultaneous $\text{pH}^m$ increase and AN decrease out of the considered limits. The second area, with a positive eigenvalue, corresponds to a region of low $\text{pH}^m$ and AN values. Both unstable areas correspond to extreme values of AN and $\text{pH}^m$, that lie outside the considered uncertainty ranges of both parameters. Irregularities in the contour lines in this limit region in Figures 3 and 4 correspond to limitations in the graphical representation.

All these results show that the second-order sensitivities corresponding to PEP and some fluxes present a pattern of behavior that is strongly dependent on $\text{pH}^m$ or AN (Tables II and III). The computation of the logarithmic gains within the range of values inside the interval of uncertainty of the two boundary parameters confirm the strong dependency predicted by the second-order sensitivities of logarithmic gains (Figs. 2, 3, 4). These results also show that these sensitivities are a good prediction of the system’s response. This result agrees with the underlying theory that relates these sensitivities with an approximation of the actual changes in the system.

In principle, computation of second-order sensitivities suffices for checking this situation in this particular case. A complementary stability analysis helps in identifying parameter combinations leading to unstable steady-states. The combination of both analyses leads to a diagnostic of the
model in the considered uncertainty region for the design parameters.

**Guidelines for Model Improvement**

The results presented in this article show that the uncertainties associated with boundary parameters may have a significant effect on the characterization of systemic properties. Even when the range of uncertainty is narrow, these effects can lead to a variety of different qualitative behaviors. This can be quantified by changes in the logarithmic gains computed within the considered range. In those cases, the predictive value of the results is questionable. Second-order sensitivities are an appropriate tool for characterizing the dependency of systemic responses from system parameters. We have shown that changes in logarithmic gain values are correctly predicted by the second-order sensitivities and that these are an alternative to an exhaustive exploration.

![Figure 2. Two-dimension plots of systemic coefficients with respect to $V_{\text{max}}^{\text{p}k}$, $V_{\text{in}}^{\text{max}}$, $V_{\text{ATP}ase}^{\text{max}}$, $V_{\text{pol}}^{\text{max}}$ and $V_{\text{pk}}^{\text{max}}$, through a range of pH in or AN. Logarithmic gains appears as succession of points, diamonds, squares, or triangles. Lines tangents to these successions in the middle value of pH in or AN correspond to the second-order sensitivities from Tables II and III.](image-url)

- (a): Logarithmic gains of PEP through a range of pH in from 6.55 to 7.05.
- (b): Logarithmic gains of ethanol production ($V_{\text{pk}}$) through a range of pH in from 6.55 to 7.05.
- (c): Logarithmic gains of PEP through a range of AN from 2.4 to 3.6.
- (d): Logarithmic gains of $V_{\text{pk}}$ through a range of AN from 2.4 to 3.6.
of the parameter region. In some extreme cases, however, it has been observed that the system behavior predicted from the second-order sensitivities deviates significantly from the actual system behavior. This is observed in our model for values of the boundary parameters that correspond to the limit region between stability and instability of the system.

Stability analysis offers a complementary vision of the influence of design parameters of the system response. We have shown that unstable steady-states may appear for particular combinations of the considered parameters within the considered region of variability. These instability regions represent a sort of forbidden combination for the considered parameters, since an unstable steady-state is unexpected to occur in this system under physiological conditions.

The combined results of sensitivity and stability analyses allow identifying a high variability of system behavior associated with small changes in the values of the boundary parameters. If we consider that this observation is not acceptable as reproducing a true physiological phenomenon,
then it is necessary to consider the possible explanations for this inconsistency so that a change in the model can correct this problem.

First, we should consider if the variables that have been defined as boundary parameters in the model were in fact dependent on other system variables. If this is the case, assuming fixed values for them in our model we may be altering the actual regulatory pattern of the system, which may result in the anomalous behavior observed. A solution for this situation is to incorporate additional processes that refine the system description. For instance, Bailey and co-workers refined their original model maintaining AN as a constant, but expanding it by making pH\textsuperscript{in} depend on ATP (Schlosser et al., 1994). Although this is an improvement of the original model, further refinements may consider the processes related to the AN pool. Alternatively, the strategy suggested by Voit and Ferreira (1998) could be applied. However, in our case buffering the pH\textsuperscript{in} and AN within a narrow range would not correct some of the problems detected.

Figure 4. Logarithmic gains and stability through two-dimension contour plots confronting a range of pH\textsuperscript{in} from 6.55 to 7.05 and a range of AN from 2.4 to 3.6. Contour lines identify the logarithmic gains. Black areas correspond to unstable regions; a big area to low pH\textsuperscript{in} and AN and an extreme area to the maximum pH\textsuperscript{in} and minimum AN values. a: Logarithmic gains of PEP with respect to V\textsubscript{max}, where contour lines separate five units. b: Logarithmic gains of ethanol production (V\textsubscript{pk}) with respect to V\textsubscript{max}, where contour lines separate 0.05 units. c: Logarithmic gains of PEP with respect to V\textsubscript{max}\textsubscript{ATPase}, where contour lines separate five units. d: Logarithmic gains of V\textsubscript{pk} with respect to V\textsubscript{max}\textsubscript{ATPase}, where contour lines separate 0.05 units.
CONCLUSIONS

Mathematical modeling is a powerful tool for studying complex systems. Model development requires a sound knowledge of the biological problem and great experience. This experience is required to capture the essential components of the problem so that the model accurately represents the target system without unnecessary complications. These complications should be avoided mainly because mathematical analysis is painful if the model of the system incorporates complicated equations.

A typical exercise in modeling is simplifying several processes by considering a constant amount of some pool. This is particularly common if this pool represents a sort of boundary, source, or sink for the biological system. This strategy is also used with internal metabolites that apparently do not change during the measured response. This is the case of considering rapid equilibrium in many instances. Without discussing the appropriateness of these simplifications in the different models, we have addressed the possible implications of considering a given value for this constant variable. The analysis of a model of ethanol fermentation pathway in *Saccharomyces cerevisiae* is an example of a nongeneral but possible case where, within the experimentally measured uncertainty range, the system can display an important variety of behaviors. In the most critical cases, the system can even enter a parameter region in which no stable steady-state exists.

Our results point to the problem of giving too much credit to a parameter value without critically checking the implications of this selection. In the example examined in this article, it seemed obvious that a fixed AN value of 3 mM was a good choice. However, a critical examination of the alternative values, in combination with the feasible pH\textsuperscript{in} values, reveals a potential problem in the model. Whether the observed variability in system behavior is admissible or not depends on the biology of the studied system. However, in many cases, we expect a relatively robust behavior, at least in qualitative terms, within a given parameter region. In those cases, results similar to those reported in this article lead to serious doubts about the utility of the model.

APPENDIX

In this appendix we show a version of the model from Galazzo and Bailey corresponding to the immobilized pH 5.5 conditions (1990, 1991). The corresponding kinetics, metabolite levels, and fluxes are taken from the original article and personal communication with the authors. Steady-state values agree with previously published analysis of the same model (Curto et al., 1995).

Mass Balance Equations

\[
\frac{d\text{Glc}}{dt} = V_{\text{in}} - V_{\text{hk}} \\
\frac{d\text{G6P}}{dt} = V_{\text{hk}} - V_{\text{pfk}} - V_{\text{pol}} \\
\frac{d\text{FDP}}{dt} = V_{\text{pfk}} - V_{\text{gapd}} - \frac{1}{2} V_{\text{gol}} \\
\frac{d\text{PEP}}{dt} = 2V_{\text{gapd}} - V_{\text{pk}} \\
\frac{d\text{ATP}}{dt} = -V_{\text{hk}} - V_{\text{pfk}} + 2V_{\text{gapd}} + V_{\text{pk}} - V_{\text{pol}} - V_{\text{ATPase}}
\]

Steady state:

<table>
<thead>
<tr>
<th>Intermediates (mM)</th>
<th>fluxes (mM/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glc = 0.17</td>
<td></td>
</tr>
<tr>
<td>G6P = 3.05</td>
<td></td>
</tr>
<tr>
<td>FDP = 5.38</td>
<td></td>
</tr>
<tr>
<td>PEP = 0.499</td>
<td></td>
</tr>
<tr>
<td>ATP = 1.92</td>
<td></td>
</tr>
<tr>
<td>V_{\text{in}} = 34.33</td>
<td></td>
</tr>
<tr>
<td>V_{\text{hl}} = 34.33</td>
<td></td>
</tr>
<tr>
<td>V_{\text{pol}} = 8.48</td>
<td></td>
</tr>
<tr>
<td>V_{\text{pfk}} = 25.85</td>
<td></td>
</tr>
<tr>
<td>V_{\text{gol}} = 3.63</td>
<td></td>
</tr>
<tr>
<td>V_{\text{gapd}} = 24.03</td>
<td></td>
</tr>
<tr>
<td>V_{\text{ATPase}} = 27.48</td>
<td></td>
</tr>
<tr>
<td>V_{\text{pk}} = 48.06</td>
<td></td>
</tr>
</tbody>
</table>
Dependency on Sum of Adenosine Nucleotides and Intracellular pH

The original model considers equilibrium between ATP, ADP, and AMP. Furthermore, a constant pool of adenosine nucleotides (AN) is assumed. The value of the equilibrium constant for adenylate kinase (K_eq) and AN determine the relationship between ATP, ADP and AMP:

- Adenilate kinase: \( K_{eq} = \frac{ADP^2}{ATP \cdot AMP} \),
  
  \( K_{eq} = 1 \), \( K_{eq} \) is the equilibrium constant

- Sum of nucleotides: \( AN = ATP + ADP + AMP \)

These relationships allow eliminating ADP and AMP in model. They are substituted for expressions depending on ATP, \( K_{eq} \) and AN:

\[
ADP = \frac{K_{eq} \cdot ATP}{2 - \sqrt{4 \cdot AN \cdot K_{eq} \cdot ATP - 4 \cdot K_{eq} \cdot ATP^2 + K_{eq}^2 \cdot ATP^2}}
\]

and

\[
AMP = AN - ADP - ATP
\]

where \( K_{eq} = 1 \).

After these substitutions, in the final model the activity of pyruvate kinase (\( V_{pk} \)), phosphofructokinase (\( V_{pfk} \)), and glyceraldehyde 3-phosphate-dehydrogenase (\( V_{gapd} \)) depends on AN (see below). On the other hand, \( V_{pk} \) is also affected by the intracellular pH (\( pH^m \)) that appears in the expression of the allosteric constant (\( L_{pk} \)) and in the binding affinity constant of PEP (\( K_{R,PEP} \)). \( V_{pfk} \) is also affected by \( pH^m \) as it appears in the expression of the allosteric constant (\( L_{pfk} \)) (see below).

Model Processes Associated with a Particular Kinetic Rate-Law

**Hexokinase:** \( V_{hk} = \frac{V_{pk}^\max}{0.00062 + 0.11 + 0.1} \cdot \frac{1}{\text{Glc ATP} + \text{Glc} + \text{ATP} + 1} \),

\( V_{pk}^\max = 68.50 \text{ mM min}^{-1} \)

**Phosphofructokinase:**

\( V_{pfk} = \frac{50 \text{ ATP G6P} \cdot R_{pfk}}{R_{pfk}^2 + L_{pfk} \cdot T_{pfk}^2} \)

\( R_{pfk} = 1 + 0.3 \text{ G6P} + 16.67 \text{ ATP} + 50 \text{ G6P ATP} \),

\( T_{pfk} = 1 + 0.00015 \text{ G6P} + 16.67 \text{ ATP} + 0.0025 \text{ G6P ATP} \),

\( L_{pfk} = -\left( 1658.22 - E^{-0.42 + 1.17 \text{ pH}^m} \right) \left( 1 + 0.76 \text{ AMP} + 1 + 40 \text{ AMP} \right)^2 \)

\( V_{pfk} \) is an allosteric enzyme described by the concerned transition model of Hess and Plesser (1978), based on the allosteric theory of Monod et al., (1965). In this model, \( V_{pfk} \) is the fractional saturation, \( R_{pfk} \) is the equilibrium constant of the R conformation, \( T_{pfk} \) is the equilibrium constant of the T conformation, and \( L_{pfk} \) is the equilibrium constant of R to T conversion.

**Glyceraldehyde 3-phosphate dehydrogenase:**

\( V_{gapd} = \frac{V_{gapd}^\max \cdot V_{gapd}^t}{\text{G6P}} \)

\( V_{gapd}^\max = 49.90 \text{ mM min}^{-1} \)

\[ v_{gapd} = \frac{1}{1 + \frac{0.25}{FDP} + \frac{1 + \frac{0.25 \cdot 1 + r_{NN} \cdot 0.09}{0.0003}}{1 + \frac{0.05 \cdot 1 + r_{NN} \cdot 0.09}{0.0003}}}} \]

where \( r_{NN} \) is the ratio between reduced and oxidized NAD (NADH/NAD\(^+\)).

**Pyruvate kinase (ethanol production):**

\[ V_{pk} = \frac{\text{PEP}}{0.02 \cdot K_{R,PEP} \cdot \text{ADP} \cdot R_{pk} + 0.004 \cdot L_{pk} \cdot \text{PEP} \cdot \text{ADP} \cdot T_{pk}} + \frac{1}{1 + 10^{-8.02} \cdot 10^{0.04} \cdot \text{FDP} + 0.02 \cdot \text{PEP} \cdot \text{ADP}} \]

\[ R_{pk} = 1 + \frac{\text{PEP}}{K_{R,PEP}} + 0.2 \cdot \text{ADP} \]

\[ T_{pk} = 1 + 0.02 \cdot \text{PEP} + 0.2 \cdot \text{ADP} + 0.004 \cdot \text{PEP} \cdot \text{ADP} \]

\[ L_{pk} = L_0 \left( \frac{1 + 0.05 \cdot \text{FDP}}{1 + 5 \cdot \text{FDP}} \right)^2 \]

\[ L_0 = 0.33 \cdot R_i \left( 0.2 \cdot L_i - R_i \cdot T_i \right) - 0.004 + 3.006 \cdot T_i \]

\[ T_i = 11.83 - 8.72 \cdot \text{pH}^m + 1.61 \cdot \text{pH}^m \]

\[ L_i = 1.79 \left( 1 + 10^{0.04 \cdot \text{pH}^m - 6.16} \right) \cdot K_{R,PEP} = 0.05581 + \frac{1}{1 + 10^{0.04 \cdot \text{pH}^m - 6.16}} \]

\( V_{pk} \) is, like \( V_{pfk} \), an allosteric enzyme described by the concerned transition model, where, \( L_0 \) is the constant of the R and T conformations in a substrate and modifier-free solution, and \( K_{R,PEP} \) is the binding affinity constant to PEP.

Model Processes Associated with Simplifications That Do Not Correspond to an Individual Enzyme Reaction

**Glucose uptake:** \( V_{in} = \frac{V_{max} \left( 1 - 3.7 \cdot \text{G6P} \cdot 45.6 \right)}{45.6} \)
Polysaccharide storage (trehalose + glycogen):

\[ V_{pol} = V_{\text{gly}}^\text{v} (V_{\text{tre}} + V_{\text{gly}}), \quad V_{\text{tre}} = 0.1 \, \text{gly}, \]

\[ V_{\text{gly}} = \frac{G6P^{0.25}}{304.44 + G6P^{0.25}} \left( 0.56 \left( 1 + 1.1 \frac{G6P}{G6P} \right) + 1 \right)^{-1} \]

Glycerol production: \[ V_{\text{gol}} = V_{\text{gly}}^\text{v} \]

It was assumed that \( V_{\text{gol}} / V_{\text{pk}} \) is constant.

Consumption of ATP: \[ V_{\text{ATPase}} = V_{\text{ATPase}}^\text{v} \cdot \text{ATP} \]

References


