Computational Flux Balance Analysis (FBA) of new representative objective functions using a multiple compartmental objective approach and its application to *Saccharomyces cerevisiae* biological behavior

Carlos E. García Sánchez\(^1\)*, César A. Vargas García\(^2\), Henry Argüello Fuentes\(^3\) and Rodrigo Torres Sáez\(^3\)

*Corresponding author. E-mail: carlos.garcia6@correo.uis.edu.co*

\(^1\)Escuela de Ingeniería Química, \(^2\)Escuela de Ingeniería de Sistemas e Informática, \(^3\)Escuela de Química, Universidad Industrial de Santander, Carrera 27 con Calle 9, Bucaramanga, Santander, Colombia

Introduction

In metabolic engineering searching for new techniques that predict quantitatively metabolic behavior of microorganisms under different environmental conditions has great importance. For this purpose, the most used technique has been flux balance analysis (FBA) 1. FBA assumes that cells have evolved to achieve an optimal behavior owing to evolutionary pressure (i.e. cells regulate its fluxes toward optimal metabolic flux distribution that maximizes some cell objective) 2. For this reason, FBA depends essentially on the selection of the objective function. In general, simulations using FBA methodologies have been consistent with experimental data 3. Until now, most analyses have assumed maximization of cell growth (i.e. biomass production), being this function the most appropriated (4-9). However, not always optimization based on cell growth can be performed 7, in some cases other objective functions have to be carried out (i.e. 10, 11). The problem for creating different types of objective functions from experimental data has been previously analyzed, for example, for searching coefficients of importance (ColS). In this case, the hypothesis consider that a determined flux is maximized by the organism as part of its cellular objective 12, or with BOSS method, where an objective function is created from a stoichiometric network, constraints and a set of experimental data 13. However, in these approaches objective functions are very dependent on the set of particular data, and are not easily interpretable from a physiological point of view. In this work, a predictive model of objective function using flux balance analyses is presented. This model considers combinations of cellular objectives defined according to cellular behavior. Therefore, we modeled mathematically cellular behavior of *S.cerevisiae* by using a cellular objective function composed by combination of different cellular objectives and/or fundamental metabolic routes.

Methods

Experimental data used in this work were obtained from experimental studies with *S. cerevisiae* (14-25). Subsequently, a carbon flux analysis was carried out to the data sets, in order to avoid using of experimental error or uncertainty in the computational measurements. Sets of data with error higher than 10% in the mass balance were discarded. In total, 46 set of experimental data were used. These data represented several experimental conditions, including diverse oxygen availability, substrate uptake, and measured fluxes experimentally determined. As metabolic model of *S. cerevisiae*, a compartmentalized genomic model iMM904 was used, which includes 1577 enzyme reactions and 1228 metabolites, distributed in 8 compartments: extracellular space, cytosol, mitochondrion, peroxisome, nucleus, endoplasmic reticulum, Golgi apparatus and vacuole 26. In order to adjust our model to global cellular objective, we took 5 compartments (cytosol, mitochondrion, nucleus, Golgi apparatus and peroxisome) with high number of reactions that contributes to this aim. Subsequently, some preliminary tests were carried out in order to explore influence of objective functions from compartments in the model. After preliminary tests we decided, based on the effect of the following objective functions: biomass production, secondary metabolite excretion, and ATP and reducing power (NADH+NADPH) production/consumption working with three compartments: cytosol, mitochondria and peroxisome. Each one of the possible objective functions was taken as a combination of objective functions in each selected compartment, and represented as: \(Z = F_1 + F_2 + F_3\), where possible functions are: \(F_1\), \(F_2\) and \(F_3\), respectively from cytosol, mitochondria and peroxisome. In our model, fourteen, four and six possible functions were selected for cytosol, mitochondria and peroxisome. In all cases, an empty set was included. For evaluating performance of every possible objective function, substrate entries and oxygen exchange (if this was available) were initially settled. Subsequently, a flux balance analysis was applied using every objective function, and biomass production and experimental metabolic fluxes were compared. The proposed algorithm carried out on Matlab (The Mathworks, Inc.), and it is based on Cobra toolbox. As solver of linear programming was used the free solver package glpk, with glpkmes as link with Matlab.

Results and Discussions

For evaluating precision of obtained estimations with different explored objective functions, performance of each possible objective function in the flux balance analysis was measured. For this aim we considered: firstly, error percentage of prediction on biomass production, and secondly, error in the prediction of the rest of known metabolic fluxes for the set of data, and comparing the vectors by means of a Euclidian distance between them. As an example, in Table 1 are presented...
10 combinations of objective functions from compartments with better predictions of cellular growth and metabolic fluxes in experiments with \textit{S. cerevisiae} under anaerobic conditions.

Table 1 Combinations of functions from compartments with lower mean error in the prediction of cellular behavior in experiments under anaerobic conditions.

<table>
<thead>
<tr>
<th>Cytosol</th>
<th>Mitochondria</th>
<th>Peroxisome</th>
<th>(e_{\text{biomass}})</th>
<th>(d_{\text{others}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>max (biomass production)</td>
<td>None</td>
<td>max (NADH+NADPH production)</td>
<td>19.16</td>
<td>14.02</td>
</tr>
<tr>
<td>max (biomass production)</td>
<td>None</td>
<td>max (ATP consumption)</td>
<td>19.17</td>
<td>14.01</td>
</tr>
<tr>
<td>max (biomass production)</td>
<td>None</td>
<td>max (fatty acid production)</td>
<td>19.30</td>
<td>13.99</td>
</tr>
<tr>
<td>max (biomass production)</td>
<td>None</td>
<td>None</td>
<td>19.39</td>
<td>14.00</td>
</tr>
<tr>
<td>max (biomass production)</td>
<td>None</td>
<td>min (ATP consumption)</td>
<td>19.39</td>
<td>13.99</td>
</tr>
<tr>
<td>max (biomass production)</td>
<td>None</td>
<td>min (NADH+NADPH consumption)</td>
<td>27.32</td>
<td>13.93</td>
</tr>
<tr>
<td>max (ethanol production)</td>
<td>None</td>
<td>min (NADH+NADPH consumption)</td>
<td>70.13</td>
<td>11520.98</td>
</tr>
<tr>
<td>min (NADH+NADPH production)</td>
<td>None</td>
<td>min (ATP consumption)</td>
<td>74.02</td>
<td>21.66</td>
</tr>
<tr>
<td>max (acetate production)</td>
<td>None</td>
<td>None</td>
<td>79.59</td>
<td>10932.34</td>
</tr>
<tr>
<td>max (succinate production)</td>
<td>None</td>
<td>min (NADH+NADPH consumption)</td>
<td>83.32</td>
<td>10942.25</td>
</tr>
</tbody>
</table>

* Means of percentage error on the estimation of biomass.  
" Mean Euclidian distance from difference between known and FBA-estimated exchange metabolic fluxes.

In Table 1 can be observed that maximization of cell growth is present in the best combinations. This caused that the prediction error of cellular growth were lower than 20%. Moreover, mitochondria do not contribute with maximization or minimization on objective functions under anaerobic conditions, which it agrees with cell behavior under these environmental conditions. On the other hand, prediction error of the rest of metabolic fluxes was very high. Most of experimental values from obtained fluxes did not overcome 15-20 (mmol/gDW*h), and differences on mean Euclidian differences between predictions and values higher than 11 were obtained. Nevertheless, it is possible to distinguish different magnitude orders among values of these errors. In the anaerobic system analyzed, the 6 best options in growth prediction, containing maximization of cell growth combined linearly with a function from peroxisome as objective function, presented also error values in the prediction of the rest fluxes with the lowest magnitude order.

Bibliography