Journal of Process Control 21 (2011) 1483–1492

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

# Identification of minimal metabolic pathway models consistent with phenotypic data

Zita I.T.A. Soons*, Eugénio C. Ferreira, Isabel Rocha

*IBB – Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal*

ARTICLE INFO

*Article history:*
Received 1 December 2010
Received in revised form 29 May 2011
Accepted 30 May 2011
Available online 28 June 2011

*Keywords:*
Elementary modes
Generating vectors
Controlled random search
Model reduction
Metabolism
*Escherichia coli*

ABSTRACT

In metabolic systems, the cellular network of metabolic reactions together with constraints of (ir)reversibility of enzymes determines the space of all possible steady-state phenotypes. Analysis of large metabolic models, however, is not feasible in real-time and identification of a smaller model without loss of accuracy is desirable for model-based bioprocess optimization and control. To this end, we propose two search algorithms for systematic identification of a subset of pathways that match the observed cellular phenotype relevant for a particular process condition. Central carbon metabolism of *Escherichia coli* was used as a case-study together with three phenotypic datasets obtained from the literature. The first search method is based on ranking pathways and the second is a controlled random search (CRS) algorithm. Since we wish to obtain a biologically realistic subset of pathways, the objective function to be minimized is a trade-off between the error and investment costs. We found that the CRS outperforms the ranking algorithm, as it is less likely to fall into local minima. In addition, we compared two pathway analysis methods (elementary modes versus generating vectors) in terms of modelling accuracy and computational intensity. We conclude that generating vectors have preference over elementary modes to describe a particular phenotype. Overall, the original model containing 433 generating vectors or 2706 elementary modes could be reduced to a system of one to three pathways giving a good correlation with the measured datasets. We consider this work as a first step towards the use of detailed metabolic models to improve real-time optimization, monitoring, and control of biological processes.



## 1. Introduction

Most mathematical models used for optimization and control of biotechnological processes are relatively simple and generally ignore the complex interactions between the extracellular environment and the thousands of intracellular enzymes and metabolites. The lack of this information in bioreactor monitoring and control can have a profound impact on biological systems and lead to poor bioreactor control performance. Nevertheless, the use of methods based on large models in process monitoring and control is nowadays limited due to their complexity and the lack of appropriate methodologies. The challenge of the development of a large-scale modelling strategy that predicts cellular phenotypes is not yet solved and is addressed here in the view of bioprocess control.

Genome-scale stoichiometric models are currently the best approximation to a representation of the metabolic capabilities of the cell. However, stoichiometric models represent an infinite number of possible phenotypes and systems biology tools need to be applied such that the simulation matches the phenotypes in given conditions. Also, most tools in systems biology are designed for steady-state applications, whereas the aim of process control requires a dynamic approach. Although dynamics are not addressed explicitly in this work, the model is formulated such that it can be easily extended as such. Moreover, as a consequence of the complexity of the models, the computational intensity is high. The model simulations are therefore too slow for some applications, such as online monitoring and control. Several model reduction approaches can be used to simplify models for use in process control, like the use of lumped reactions, sensitivity analysis tools [1], singular perturbation theory [2], and elimination of the dynamics of some processes based on their time scales [3].

Tools that have the potential to solve some of the above problems may come from metabolic pathway analysis. Metabolic pathway analysis is the discovery and analysis of meaningful routes in metabolic networks. It is becoming increasingly important for assessing network properties and linking the cellular phenotype to the corresponding genotype. Amongst several concepts elementary mode (EM) analysis [4], extreme pathway analysis (EP) [5], and the concept of generating vectors (GVs) [6] are promising tools. The first two tools have been evaluated by several authors, amongst

* Corresponding author. Tel.: +351 253 604 422; fax: +351 253 604 429.
*E-mail addresses:* zita@deb.uminho.pt, zita.soons@gmail.com (Z.I.T.A. Soons), ecferreira@deb.uminho.pt (E.C. Ferreira), irocha@deb.uminho.pt (I. Rocha).

0959-1524/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2011.05.012