Mathematical Biosciences 218 (2009) 50–58

Contents lists available at ScienceDirect

Mathematical Biosciences

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

Metabolic Engineering with power-law and linear-logarithmic systems

Alberto Marin-Sanguino [a,*], Nestor V. Torres [b], Eduardo R. Mendoza [c,d], Dieter Oesterhelt [a]

[a] Max Planck Institute of Biochemistry, Department of Membrane Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Bayern, Germany
[b] Grupo de Tecnologia Bioquimica, Departamento de Bioquimica y Biologia Molecular, Facultad de Biologia, Universidad de La Laguna, 38206 La Laguna, Tenerife, Islas Canarias, Spain
[c] Department of Computer Science, University of the Philippines Diliman, Diliman, Quezon City 1101, Philippines
[d] Physics Department and Center for Nano-Science, Ludwig Maximillians University, Geschwister-Scholl-Platz 1, 80539 Munich, Germany

ARTICLE INFO

Article history:
Received 30 June 2008
Received in revised form 9 December 2008
Accepted 16 December 2008
Available online 11 January 2009

Keywords:
Generalized mass action
S-system
Linlog
(Log)linear
Metabolic Engineering
Biochemical Systems Theory

ABSTRACT

Metabolic Engineering aims to improve the performance of biotechnological processes through rational manipulation rather than random mutagenesis of the organisms involved. Such a strategy can only succeed when a mathematical model of the target process is available. Simplifying assumptions are often needed to cope with the complexity of such models in an efficient way, and the choice of such assumptions often leads to models that fall within a certain structural template or formalism. The most popular formalisms can be grouped in two categories: power-law and linear-logarithmic. As optimization and analysis of a model strongly depends on its structure, most methods in Metabolic Engineering have been defined within a given formalism and never used in any other.

In this work, the four most commonly used formalisms (two power-law and two linear-logarithmic) are placed in a common framework defined within Biochemical Systems Theory. This framework defines every model as matrix equations in terms of the same parameters, enabling the formulation of a common steady state analysis and providing means for translating models and methods from one formalism to another. Several Metabolic Engineering methods are analysed here and shown to be variants of a single equation. Particularly, two problem solving philosophies are compared: the application of the design equation and the solution of constrained optimization problems. Generalizing the design equation to all the formalisms shows it to be interchangeable with the direct solution of the rate law in matrix form. Furthermore, optimization approaches are concluded to be preferable since they speed the exploration of the feasible space, implement a better specification of the problem and exclude unrealistic results.

Beyond consolidating existing knowledge and enabling comparison, the systematic approach adopted here can fill the gaps between the different methods and combine their strengths.

1. Introduction

Metabolic Engineering aims to improve the performance of biotechnological systems through rational manipulation, as opposed to random mutagenesis or trial and error methods. In order to achieve such a goal for complex systems, a mathematical model is needed. Simplifying assumptions are often needed to cope with the complexity of such models in an efficient way, and the choice of such assumptions often leads to models that fall within a certain structural template or formalism. In addition to the advantages of simplification, developing systematic formalisms results in structurally homogeneous models for which standardized algorithms can be developed. This has encouraged the search for fast approximate methods that take advantage of structural regularities. These methods usually provide results that are close to the accurate solutions found by slower general purpose methods [27], and furthermore enable an interactive exploration of the problem. Several methods have been proposed in which an equation can be used to calculate the manipulations needed to obtain a certain steady state [9,3,12]. To achieve the same goal, optimization approaches have also been proposed [30,24,13,15].

The most popular formalisms can be grouped in two categories: power-law [31] and linear-logarithmic [25]. In power-law models, rates and variables are linearized in logarithmic axes, in other words, they become linear in a log–log plot. The power-law formalism includes several variants, of which Generalized Mass Action (GMA) and S-systems are the most common. Linear-logarithmic models are based on a mixed linearization in which the reaction rates, and optionally some variables, stay in a Cartesian axis while the rest of the variables are transformed into a logarithmic scale. There are two variants of the linear-logarithmic formalism, namely (log)linear and linlog. Although all these model types can be obtained through a wide variety of methods [7,14,34,31], they are all based in the same information and their ability to portray different kinetic laws has been explained in terms of approximation the-

* Corresponding author.
E-mail addresses: amarin@biochem.mpg.de, amsanguino@gmail.com (A. Marin-Sanguino), ntorres@ull.es (N.V. Torres), mendoza@lmu.de (E.R. Mendoza), oesterhe@biochem.mpg.de (D. Oesterhelt).

0025-5564/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.mbs.2008.12.010

ory by deriving their fundamental equations through Taylor series [22]. In fact, even such a popular kinetic representation as the Hill equation has been shown to be a particular case of a Taylor series [22].

The first obstacle to overcome when comparing or combining different formalisms is the existence of different notations. Power-law models are normally presented within the notation specific to Biochemical Systems Theory (BST) while linear-logarithmic follow that of Metabolic Control Analysis (MCA). These two frameworks have parallel histories which have been converging during the last 15 years and have been shown to be deeply related [1]. MCA, originally intended to be used without a dynamic model [10], is mainly based on matrix relations between local sensitivities (elasticities) and global sensitivities (response and control coefficients) [16]. Furthermore, the classification of the variables within its scope favors their separation according to biological properties (enzymes, metabolites, effectors, etc.). BST, on the other hand, favors the explicit formulation of dynamic models and classifies variables according to their mathematical role, depending on whether they are constant (independent variables) or time dependent (dependent variables). An enzyme concentration can thus be an independent variable, if it is considered to be constant during the considered time scale, or a dependent variable, if its synthesis and degradation are featured in the model. The MCA approach yields equations with a more immediate biological interpretation, but which have to be modified whenever the biological assumptions change (interactions among enzymes, moiety conservations, etc.), this has led to many slightly different interpretations of the basic formalism [28]. BST keeps a unified formulation that remains consistent independently of the choices made by the modeler [31,32]. Despite their differences, these two frameworks have led to developments within the field of Metabolic Engineering that are sometimes equivalent and sometimes complementary. Given the potential benefit of combining such developments, and since a unified notation for MCA and BST is not to be expected in a near future, this work will use BST as a unifying framework. The reason for such a choice is that the less intuitive variable grouping is largely compensated by the flexibility and generality gained in return. Furthermore, BST goes beyond approximate rates as it can deal with exact representations through detailed mechanistic modelling [20] or recasting arbitrary non-linear functions [19,6,20].

In the next section, the four considered formalisms will be presented in a consistent manner such that all the rate laws of every formalism will depend on the same parameters. This will enable us to present some basic results from BST and MCA in a form that holds for all cases. Finally, a brief overview of the similarities and differences between the power-law and linear-logarithmic formalisms will be presented. This will establish a single framework in which different methods can be not only compared, but also combined.

2. Theoretical framework

Mathematical models in Metabolic Engineering are often given as systems of differential equations according to the form:

$$\dot{\mathbf{x}}_\mathbf{d} = \mathbf{N} \cdot \mathbf{v}. \tag{1}$$

The reaction rates, transport fluxes, etc. collected in vector $\mathbf{v}$ do normally depend on many different factors, some involved in the dynamics of the system (dependent variables, $\mathbf{x_d}$) while others remain constant (independent variables $\mathbf{x_i}$). The subindices $\mathbf{d}$ and $\mathbf{i}$ will be used for dependent and independent variables in subsequent formulas.

It is important to note that there can be dependencies among the rows in $\mathbf{N}$ when there are conservation relations. In such cases, the derivatives of some dependent variables can be written as a linear function of the rest [16] and eliminated from the system. From now on, we will simplify the notation by assuming that the stoichiometric matrix has been thus reduced. This will greatly simplify the notation without any loss on generality.

The complexity of $\mathbf{v} = \mathbf{f}(\mathbf{x_d}, \mathbf{x_i})$ can imply an important risk to the mathematical tractability of the problem, for this reason, it is often simplified to a non-mechanistic, approximated function with a standard structure such that the resulting equations comply with the specifications of a given formalism.

2.1. Derivation of the formalisms

In the following sections, the different rate laws will be consistently derived making use of Taylor series. As a result, they will all be defined around a chosen reference state (usually a steady state) as a function of the variables of the system and two kinds of parameters: kinetic orders and rate constants. Kinetic orders characterize the response of rates to changes in the variables and are defined as:

$$f_{i,j} = \left|\frac{\partial \ln v_i}{\partial \ln x_j}\right|_0 = \left|\frac{\partial v_i}{\partial x_j}\frac{x_j}{v_i}\right|_0, \tag{2}$$

where subindex 0 indicate the value of a magnitude in the steady state. All the kinetic orders of a model can be grouped in a matrix F.

The rate constants can be grouped as a vector $\boldsymbol{\gamma}$, and will have a different definition for every formalism but will always be determined by the rate values at the chosen reference state.

To achieve a compact notation, vectors are often collected in diagonal matrices. These are composed of zeros save for their main diagonal, which contains the elements of the corresponding vector. Each of these matrices will be represented with the same letter as the vector but capitalized, $\mathbf{V} = \mathrm{diag}(\mathbf{v})$.

2.1.1. GMA

The GMA formalism is based in a log–log Taylor series:

$$\ln\frac{v_i}{v_i^0} \simeq \left|\frac{\partial \ln v_i}{\partial \ln x_1}\right|_0 \ln\frac{x_1}{|x_1|_0} + \cdots + \left|\frac{\partial \ln v_i}{\partial \ln x_n}\right|_0 \ln\frac{x_n}{|x_n|_0}, \tag{3}$$

where super or subindex 0 indicate the value of a magnitude in the steady state. The rate law is therefore linear in the log–log space:

$$\mathbf{w} = \mathbf{F_d}\mathbf{y_d} + \mathbf{F_i}\mathbf{y_i} + \mathbf{g}, \tag{4}$$

where $\mathbf{w} = \ln\mathbf{v}$, $\mathbf{y} = \ln\mathbf{x}$ and $\mathbf{g} = \ln\gamma$. Here, ln of a vector $\mathbf{u}$ denotes the vector with components $\ln u_i$ The rate becomes a power-law by undoing the logarithmic transformation:

$$v_i = \gamma_i x_1^{f_1} \ldots x_n^{f_n}, \tag{5}$$

where $\gamma_i = \frac{|v_i|_0}{|x_1|_0^{f_1}\cdots|x_n|_0^{f_n}}$. This expression can be simplified to $\gamma_i = |v_i|_0$ when the variables are normalized by their steady state values. This normalization does not alter any of the properties of the system that will be discussed in this paper, but prevent some forms of robustness analysis. Particularly, the information about robustness to changes in the kinetic orders are lost for the reference state.

The rate law can be substituted in Eq. (1) to obtain a GMA model.

2.1.2. S-systems

S-systems are a particular case of GMA in which there are only two fluxes per equation:

$$\dot{\mathbf{x}}_\mathbf{d} = \mathbf{v}^+ - \mathbf{v}^-. \tag{6}$$