Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Process Control 14 (2004) 717–728
JOURNAL OF PROCESS CONTROL
www.elsevier.com/locate/jprocont

Dynamic metabolic modelling under the balanced growth condition

A. Provost *, G. Bastin

Cesame, Université catholique de Louvain, av. G. Lemaitre 4, B1348 Louvain-La-Neuve, Belgium

Abstract

The issue of bioprocess dynamical modelling is addressed in the situation where measurements of extracellular species are the only available data besides measurements of the biomass itself. This question is investigated under the assumption that a metabolic network connecting the extracellular species is available. A metabolic flux analysis is first performed in order to test the consistency of the metabolic network. The elementary flux modes are computed and translated into a set of macro-reactions connecting the extracellular substrates and products. Then a dynamical model, compatible with the underlying metabolic network, is build on the basis of these macro-reactions. The approach is illustrated with the example of CHO cells cultivated in stirred flasks on a serum-free medium.

Keywords: CHO cells; Elementary flux modes; Metabolic modelling; Metabolic flux analysis

1. Introduction

We are concerned in this paper with the identification of mathematical models of biological processes in the very common situation where measurements of extracellular species in the culture medium are the only available data besides measurements of the biomass itself. The aim of this paper is to investigate this question under a metabolic viewpoint. Therefore, as a starting point for our analysis, we assume that a metabolic network connecting the measured species is available. As a matter of illustration to our discussion we shall consider the example of Chinese Hamster Ovary (CHO) (see e.g. [1]) cells cultivated in batch mode in stirred flasks. The measured extracellular species are the two main substrates (glucose and glutamine) and the three most significantly released metabolites (lactate, ammonia, alanine).

The issue of bioprocess modelling from extracellular measurements has been considered for a long time in the literature. In classical macroscopic models (see Fig. 1) the cells are just viewed as a catalyst for the conversion of substrates into products which is represented by a set of chemical "macro-reactions" that directly connect extracellular substrates and products without paying much attention to the intracellular behaviour. Dynamical mass balance models are then established on the basis of these macro-reactions by identifying appropriate kinetic models from the experimental data. Such macroscopic models rely on the category of so-called "unstructured models" in the standard terminology reported in the important discussion paper of Bailey [2]. The goal of macroscopic modelling is clearly to derive simple models that have been proved of paramount importance in bioengineering for the design of on-line algorithms for process monitoring, control and optimization. However, as noted above, a big drawback is that the macroscopic models are often derived without really taking care of what happens inside the cells. In this context, during the last decade, a new trend in mathematical modelling has emerged by focusing on the so-called "Metabolic Flux Analysis" [3, Chapter 8], where intracellular fluxes are computed from the measured extracellular fluxes by using the stoichiometry of a metabolic network supposed to govern the system. In some sense, Metabolic Flux Analysis rather relies on structured modelling though it essentially is a steady state analysis.

In this paper, our purpose will be to throw a bridge between macroscopic dynamical modelling and metabolic flux analysis.

A "full modelling" approach could be considered. This means that we could try to find a global dynamical

* Corresponding author.
E-mail address: provost@auto.ucl.ac.be (A. Provost).

0959-1524/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2003.12.004

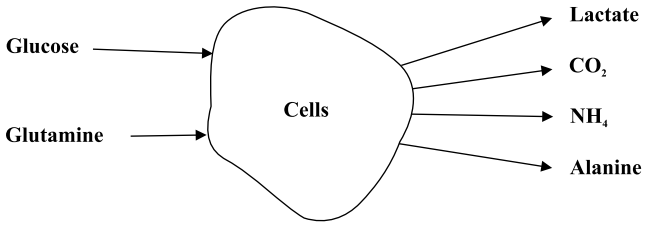

Fig. 1. Macroscopic representation of biochemical reactions (inspired from Fig. 5.28 in [4]).

model which describes the full metabolic network, involving a separate state variable for each intracellular species and a separate kinetic model for each intracellular reaction. Such an approach, although conceivable, is in some sense "ill-posed" because the intracellular kinetics are in general not structurally identifiable without intracellular measurements. We shall rather follow a "reduced modelling" approach where the model is based on a set of macro-reactions which are compatible with the underlying metabolic network and supported by a preliminary metabolic flux analysis.

The aim of this paper is to present the methodology of the model development in a systematic and reproducible way. In Sections 2 and 3, we give the necessary basic concepts and definitions, namely the general dynamic mass balance model of metabolic systems and the balanced growth condition. Then the model development proceeds in three main steps which are successively presented in Sections 4–6. A metabolic flux analysis is first performed in order to check the consistency of the assumed metabolic network with the experimental data. In a second step, the elementary metabolic flux modes are computed and translated into a set of elementary macro-reactions connecting the extracellular substrates and products. Finally, in a third step, a classical dynamical model is established on the basis of the macro-reactions. An interesting consequence of this modelling approach is to allow the prediction of the time evolution of end-products which are not measured. The procedure will be presented through the case study of CHO-320 cells cultivated in serum free medium.

2. The general dynamic metabolic model

For the clarity of the presentation, we shall consider a simple fictional example of metabolic network represented in Fig. 2. The metabolic network involves two groups of nodes: boundary nodes and internal nodes. Boundary nodes can be further separated into initial and terminal nodes. Initial nodes correspond to the external substrates that are consumed but not produced. Terminal nodes correspond either to extracellular products released in the culture medium or intracellular products

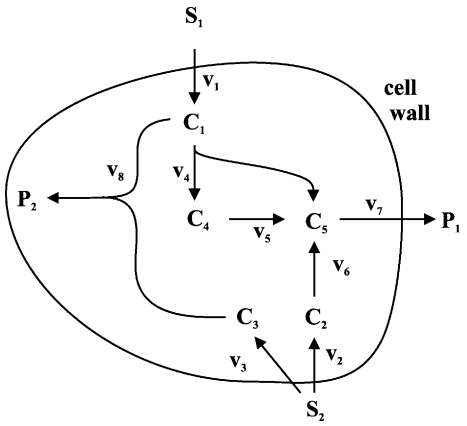

Fig. 2. A simple fictional metabolic network.

which form the cellular material during the growth. In the example of Fig. 2 there are

- two extracellular substrates: S_1 and S_2;
- five intracellular metabolites: C_1, C_2, C_3, C_4, C_5;
- one terminal extracellular product P_1;
- one terminal intracellular product P_2.

The set of the cellular reactions involved in this network can equivalently be written as

$$S_1 \rightarrow C_1$$

$$S_2 \rightarrow C_2$$

$$S_2 \rightarrow C_3$$

$$C_1 \rightarrow C_4 + C_5$$

$$C_4 \rightarrow C_5$$

$$C_2 \rightarrow C_5$$

$$C_5 \rightarrow P_1$$

$$C_1 + C_3 \rightarrow P_2$$

The cells are cultivated in a stirred tank reactor. Without loss of generality, we assume an unitary volume of the culture medium (or equivalently that all the balance equations are written "per unit of liquid reactor volume").

As usual, the cell density or concentration is denoted X. The substrate concentrations in the reactor liquid volume are denoted s_i and collected into a vector $s = (s_1, s_2, \ldots, s_{n_s})$. The mass fractions of the intracellular metabolites inside the cells are denoted c_i and collected into the vector $c = (c_1, c_2, \ldots, c_{n_c})$. These fractions are the ratio between the mass of each species and the mass of the cells. Similarly we define the vector $p = (p_1, p_2, \ldots, p_{n_p})$ of end-product concentrations where p_i denotes a concentration in the reactor liquid volume for extracellular terminal products or a mass fraction for intracellular terminal products.

Assuming an exponential growth with a constant specific growth rate μ in batch mode, we have the following mass balance equations:

$$\begin{aligned}\frac{\mathrm{d}X}{\mathrm{d}t} &= \mu X \\ \frac{\mathrm{d}s}{\mathrm{d}t} &= -v_s X\end{aligned} \tag{1}$$

$$\begin{aligned}\frac{\mathrm{d}(cX)}{\mathrm{d}t} &= NvX \\ \frac{\mathrm{d}p}{\mathrm{d}t} &= v_p X\end{aligned} \tag{2}$$

The first equation $\frac{\mathrm{d}X}{\mathrm{d}t} = \mu X$ is the classical growth equation.

The second equation $\frac{\mathrm{d}s}{\mathrm{d}t} = -v_s X$ represents the uptake of the substrates by the cells with v_s the vector of the specific uptake rates.

The third equation $\frac{\mathrm{d}(cX)}{\mathrm{d}t} = NvX$ represents the mass balances around the intracellular metabolites with $v = (v_1, v_2, \ldots, v_{n_v})$ the vector of the intracellular specific reaction rates (called metabolic fluxes) and N the corresponding $n_c \times n_v$ stoichiometric matrix.

The fourth equation $\frac{\mathrm{d}p}{\mathrm{d}t} = v_p X$ represents the formation of end products with v_p the vector of the specific accumulation rates for intracellular products or excretion rates for extracellular products.

In order to make these definitions and notations fully clear, the model equations corresponding to the network of Fig. 2 is given as follows:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu X$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = -\begin{pmatrix} v_1 \\ v_2 + v_3 \end{pmatrix} X, \quad v_s = \begin{pmatrix} v_1 \\ v_2 + v_3 \end{pmatrix}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} c_1X \\ c_2X \\ c_3X \\ c_4X \\ c_5X \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & 0 \end{pmatrix}}_{N} \underbrace{\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{pmatrix}}_{v} X$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} v_7 \\ v_8 \end{pmatrix} X, \quad v_p = \begin{pmatrix} v_7 \\ v_8 \end{pmatrix}$$

From this example, it appears clearly that v_s and v_p can be written as $v_s = N_s v$ and $v_p = N_p v$ with appropriate matrices N_s and N_p of dimensions $n_s \times n_v$ and $n_p \times n_v$ respectively:

$$N_s = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$N_p = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Moreover, the mass balance equation for the intracellular metabolites can be rewritten as

$$\frac{\mathrm{d}(cX)}{\mathrm{d}t} = \frac{\mathrm{d}c}{\mathrm{d}t}X + c\frac{\mathrm{d}X}{\mathrm{d}t} = NvX \tag{3}$$

or upon division by X and using $\frac{\mathrm{d}X}{\mathrm{d}t} = \mu X$:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = Nv - \frac{c}{X}\frac{\mathrm{d}X}{\mathrm{d}t} = Nv - \mu c \tag{4}$$

The general dynamic metabolic model is written as

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu X$$

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -N_s v X$$

$$\frac{\mathrm{d}c}{\mathrm{d}t} = Nv - \mu c$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = N_p v X$$

This model is valid for a batch reactor. Obviously, it could be easily extended to fed-batch or continuous reactors by inserting appropriate inflow, outflow and dilution terms.

3. The balance growth paradigm

The so-called *balanced growth* condition is a fundamental assumption of metabolic engineering which is widely discussed and motivated in the literature (see e.g. [3]). The condition states that, during the growth of the cells, the internal metabolites are supposed to be at quasi-steady state. This is mathematically expressed by a quasi-static approximation of the mass balance equation of the internal metabolites:

$$\frac{\mathrm{d}c}{\mathrm{d}t} \simeq 0 \quad \Rightarrow \quad Nv - \mu c = 0$$

In terms of singular perturbations, the fast dynamics of the model are the expression of a fast turnover of the intracellular metabolites (compared with the substrate uptakes). While $Nv - \mu c = 0$ is the slow manifold attracting the trajectories of the model in the state-space. The term "balanced growth" is used to explicit the fact that the intracellular cell composition (represented by the vector c) remains almost invariant during the growth phase.

It is furthermore documented in the literature (see e.g. [4]) that the dilution term μc is small (by at least one order of magnitude) with respect to the other fluxes affecting the same metabolite. Therefore, the term μc can be neglected and the equation reduces to

$$Nv = 0$$

This simple linear steady state relation between the intracellular fluxes around the intracellular metabolites is the fundamental equation that underlies the metabolic flux analysis.

4. Metabolic flux analysis

Metabolic flux analysis (MFA) is the name given to the exercise of calculating unknown intracellular fluxes from measured extracellular fluxes by applying the steady state mass balance equation $Nv = 0$. Basically, this equation is under-determined with the number n_v of unknowns v_i larger than the number n_c of equations. The flux vector v can be determined only if additional constraints are introduced. As mentioned above, our concern in this paper is modelling when measurements of extracellular species in the culture medium are the only available data. As we shall illustrate with an experimental application hereafter, we can clearly assume that the specific uptake and excretion rates v_{si} and v_{pi} of those measured extracellular species can be estimated from the data and collected in a vector v_m such that

$$Pv = v_m$$

for some appropriate $\dim(v_m) \times n_v$ matrix P which is just a sub-matrix of

$$\begin{pmatrix} N_s \\ N_p \end{pmatrix}$$

For example, let us assume that the three extracellular species S_1, S_2, P_1 of the hypothetical network of Fig. 2 are measurable. Then:

$$v_m = \begin{pmatrix} v_{s1} \\ v_{s2} \\ v_{p1} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 + v_3 \\ v_7 \end{pmatrix}$$

and

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Let us now consider the linear system:

$$\begin{pmatrix} N \\ P \end{pmatrix} v = \begin{pmatrix} 0 \\ v_m \end{pmatrix}$$

and assume that

$$\operatorname{rank}\begin{pmatrix} N \\ P \end{pmatrix} = n_v$$

In such a case, the system becomes determined and a solution

$$v = \left[\begin{pmatrix} N \\ P \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} N \\ P \end{pmatrix} \right]^{-1} \begin{pmatrix} N \\ P \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} 0 \\ v_m \end{pmatrix}$$

can be computed. In our example of Fig. 2, the solution is

$$v_1 = v_{s_1}$$

$$v_2 = v_6 = \tfrac{1}{3}(v_{p_1} - 2v_{s_1} + 2v_{s_2})$$

$$v_3 = v_8 = \tfrac{1}{3}(2v_{s_1} - v_{p_1} + v_{s_2})$$

$$v_4 = v_5 = \tfrac{1}{3}(v_{s_1} + v_{p_1} - v_{s_2})$$

$$v_7 = v_{p_1}$$

If $\operatorname{rank}\begin{pmatrix} N \\ P \end{pmatrix}$ is smaller than the number of unknowns, then it is necessary to use additional constraints based on the available prior knowledge on the metabolism. This is illustrated in the following experimental application.

4.1. Experimental application to CHO cells

In order to explain the MFA methodology with a real life application, we consider the example of CHO cells cultivated in batch mode in stirred flasks. In accordance with the literature on mammalian cells, the central metabolism of CHO cells (e.g. [10–12]) is represented by the metabolic network depicted in Fig. 3. This metabolic network describes only the part of the metabolism concerned with the utilisation of the two main energetic nutrients (glucose and glutamine). The metabolism of the amino-acids provided by the culture medium is not considered. This network is essentially made of four fundamental pathways: the glycolysis pathway, the glutaminolysis pathway, the TCA cycle and the nucleotides synthesis pathway (e.g. [13]).

Since the goal is a steady state flux analysis, all the intermediate species which belong to the network but are not located at branch points are omitted without loss of generality.

In this network, there are

- two initial substrates: glucose and glutamine;
- four terminal extracellular products: lactate, alanine, NH_4 and CO_2;
- two terminal intracellular metabolites: purine and pyrimidine nucleotides;
- twelve internal metabolites: glucose 6-phosphate, dihydroxy-acetone phosphate, glyceraldehyde 3-phosphate, ribose 5-phosphate, pyruvate, acetyl-coenzyme A, citrate, α-ketoglutarate, malate, glutamate, oxaloacetate, aspartate.

The measured extra-cellular species are the two substrates (glucose and glutamine) and the three most significantly released metabolites (lactate, ammonia, alanine). The experimental data of these substrates and products as well as the cell density, collected during the *exponential growth phase*, are available for three cultures and are shown in Fig. 4.

In accordance with the quasi-steady state assumption, constant specific uptake and excretion rates are assumed during the exponential growth. By integrating

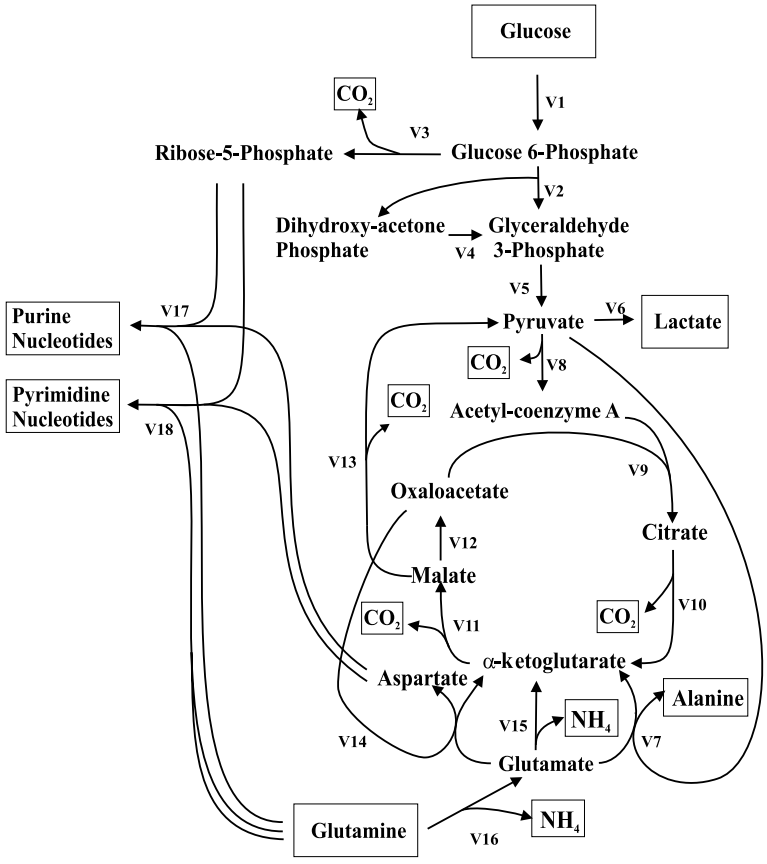

Fig. 3. Metabolic network of CHO cells.

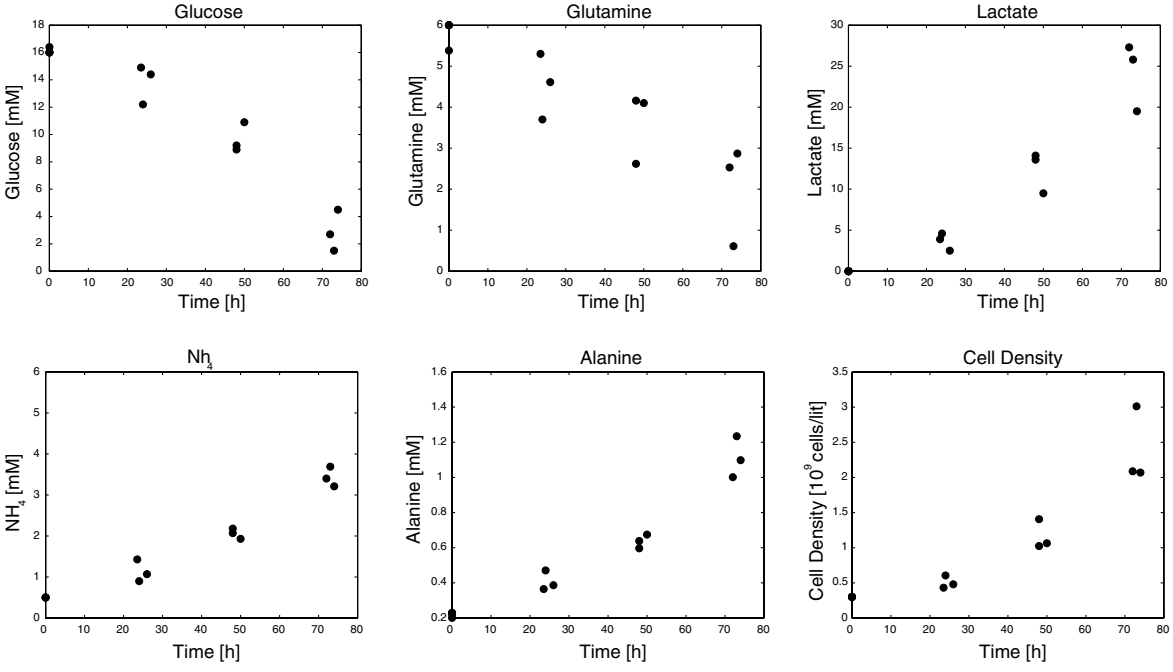

Fig. 4. Experimental data.

Eqs. (1) and (2), the following linear regression equations are obtained:

$$s(t) = -a_s X(t) + b_s \quad \text{with } a_s = \frac{v_s}{\mu} \text{ and}$$
$$b_s = \frac{X(0)v_s}{\mu} + s(0)$$

$$p(t) = a_p X(t) + b_p \quad \text{with } a_p = \frac{v_p}{\mu} \text{ and}$$
$$b_p = -\frac{X(0)v_p}{\mu} + p(0)$$

From the experimental data of Fig. 4, the specific uptake and excretion rates (v_s and v_p) are computed by linear regression during the growth phase (see Fig. 5) and given in Table 1.

The cell density computed with an exponential growth model is also shown in Fig. 5:

$$\frac{dX}{dt} = \mu X \quad \text{with } \mu = 0.6912 d^{-1}$$

The steady state flux balance equations $Nv = 0$ at the internal nodes of the network are expressed as follows:

$$\begin{aligned}
&\text{Glucose 6-Phosphate} : v_1 - v_2 - v_3 = 0\\
&\text{Dihydroxy-acetone Phosphate} : v_2 - v_4 = 0\\
&\text{Glyceraldehyde 3-Phosphate} : v_2 + v_4 - v_5 = 0\\
&\text{Ribose 5-Phosphate} : v_3 - v_{17} - v_{18} = 0\\
&\text{Pyruvate} : v_5 + v_{13} - v_6 - v_7 - v_8 = 0\\
&\text{Acetyl-coenzyme A} : v_8 - v_9 = 0\\
&\text{Citrate} : v_9 - v_{10} = 0\\
&\alpha\text{-Ketoglutarate} : v_{10} + v_7 + v_{15} + v_{14} - v_{11} = 0\\
&\text{Malate} : v_{11} - v_{12} - v_{13} = 0\\
&\text{Glutamate} : v_{16} - v_{15} - v_{14} - v_7 = 0\\
&\text{Oxaloacetate} : v_{12} - v_9 - v_{14} = 0\\
&\text{Aspartate} : v_{14} - v_{17} - v_{18} = 0
\end{aligned} \tag{5}$$

Glucose, lactate and alanine specific rates are directly given in Table 1:

$$v_1 = 4.0546 \text{ mM}/(\text{d} \times 10^9 \text{ cells}),$$
$$v_6 = 7.3949 \text{ mM}/(\text{d} \times 10^9 \text{ cells}) \tag{6}$$

$$v_7 = 0.2686 \text{ mM}/(\text{d} \times 10^9 \text{ cells}) \tag{7}$$

The NH_4 specific excretion rate in Table 1 is the sum of v_{15} and v_{16}:

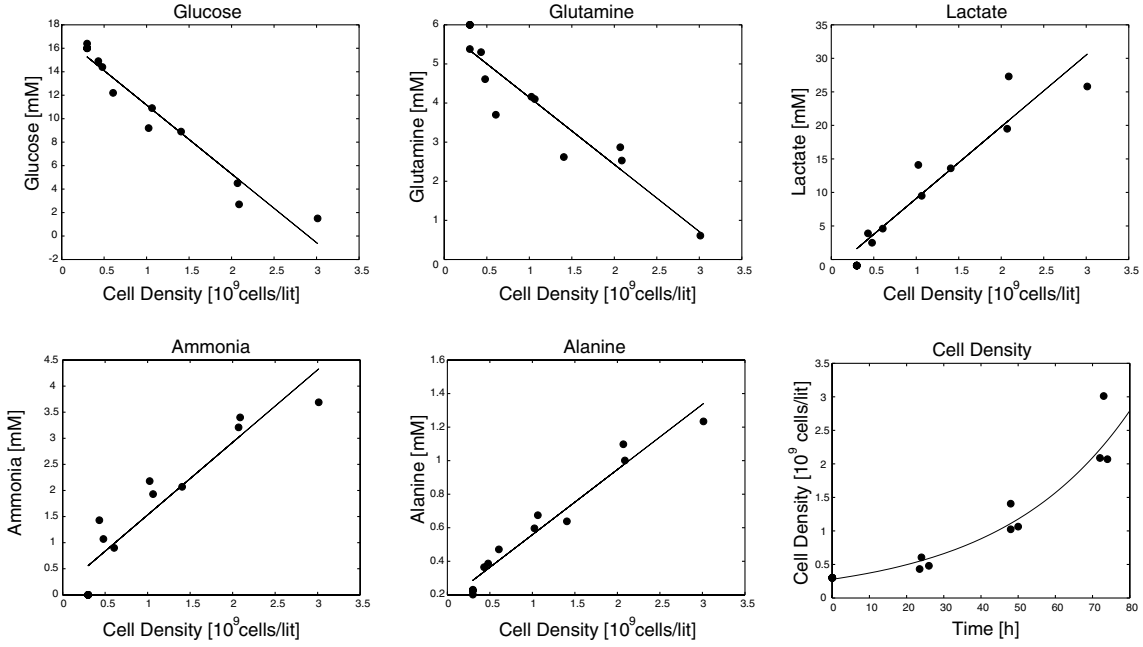

Fig. 5. Regression curves.

Table 1
Specific uptake and excretion rates computed by linear regression

Glucose (mM/(d×10^9 cells))	Glutamine (mM/(d×10^9 cells))	Lactate (mM/(d×10^9 cells))	NH_4 (mM/(d×10^9 cells))	Alanine (mM/(d×10^9 cells))
4.0546	1.1860	7.3949	0.9617	0.2686

Table 2
Metabolic fluxes of CHO cells during the growth phase (mM/(d×10^9 cells))

v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{17}	v_{18}
4.0546	3.7692	0.2854	3.7692	7.5383	7.3949	0.2686	0.3472	0.3472	0.3472	1.1051	0.6327	0.4724	0.2854	0.2038	0.7579	0.1427	0.1427

$$v_{15} + v_{16} = 0.9617 \text{ mM}/(\text{d} \times 10^9 \text{ cells}) \quad (8)$$

and the specific uptake of glutamine is

$$v_{16} + v_{17} + 2v_{18} = 1.1860 \text{ mM}/(\text{d} \times 10^9 \text{ cells}) \quad (9)$$

Since DNA and RNA are made up with equal shares of purine nucleotides and pyrimidine nucleotides, it is a natural assumption to consider that

$$v_{17} = v_{18} \quad (10)$$

Remark that, as above, we are in the form of a well-posed linear system:

$$\begin{pmatrix} N \\ P \end{pmatrix} v = \begin{pmatrix} 0 \\ v_m \end{pmatrix}$$

with system (5) for $Nv = 0$ and Eqs. (6)–(10) for $Pv = v_m$. By solving this system, we get a single well defined positive solution which is given in Table 2.

We see that we get plausible positive values for the fluxes which are comparable to the values obtained in metabolic flux analysis of other mammalian cells [11] and therefore enhance the validity of the proposed network.

Furthermore we are able to predict the specific production rate of the end-metabolites which are not measured, namely CO_2 and nucleotides:

$$v_{CO_2} = v_3 + v_8 + v_{10} + v_{11} + v_{13} = 2.5574 \text{ mM}/(\text{d} \times 10^9 \text{ cells})$$

$$v_{\text{nucl}} = v_{17} + v_{18} = 0.2854 \text{ mM}/(\text{d} \times 10^9 \text{ cells})$$

This value of v_{nucl} allows for a further quantitative validation of the model. Indeed, assuming that the average nucleotide molar mass is about 340 g/mol, the nucleotide mass production rate is approximately $340 \times 0.2854 \times 10^{-3}$ g/(d×10^9 cells×lit) $\simeq 0.1$ g/(d×10^9 cells×lit). We get that the amount of nucleotides in one cell is about 100 pg/cell, which is totally comparable with the values available in the literature (for instance in [12], the mass of RNA + DNA in Hybridoma cells is estimated at 30–40 pg/cell).

5. Elementary flux modes

The flux distributions v which satisfy the fundamental steady state equation $Nv = 0$ are necessarily non-negative vectors v that belong to the kernel of the matrix N. Therefore the space of admissible solutions is the polyhedral cone which is the intersection of the kernel of N and the positive orthant. This means that each solution v can be expressed as a non-negative linear combination of a set of vectors e_i, called generating vectors, which form the *unique* convex basis [6] of the polyhedral cone:

$$v = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n \quad \text{with } \lambda_k \geqslant 0 \text{ for } k = 1, \ldots, n$$

In our prototype example of Fig. 2, the stoichiometric matrix N is

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & 0 \end{pmatrix}$$

The dimension of the kernel of N is

$$\dim(\ker N) = n_v - n_c = 8 - 5 = 3 \quad (11)$$

The convex basis involves three vectors which are easily computed and are given in Table 3.

The vectors of the convex basis define the so-called *elementary flux modes* (also called spanning pathways) which are the simplest metabolic paths that are able to connect the substrates with the end-products. The determination of the convex basis vectors of metabolic networks has been studied extensively in the literature. Typical references are [5,7,8] among others. The three elementary flux modes corresponding to Table 3 are represented in Fig. 6.

Under steady state conditions, these elementary flux modes can be simplified by eliminating the internal metabolites, giving a set of macro-reactions that connect the extra-cellular substrates and the end-products:

$e_1 : S_1 + S_2 \rightarrow P_2$

$e_2 : S_2 \rightarrow P_1$

$e_3 : S_1 \rightarrow 2P_1$

In the case of complex metabolic networks, the computation of the convex basis is less trivial and requires the use of efficient automatic algorithms. Furthermore,

Table 3
The three convex basis vectors

	e_1	e_2	e_3
v_1	1	0	1
v_2	0	1	0
v_3	1	0	0
v_4	0	0	1
v_5	0	0	1
v_6	0	1	0
v_7	0	1	2
v_8	1	0	0

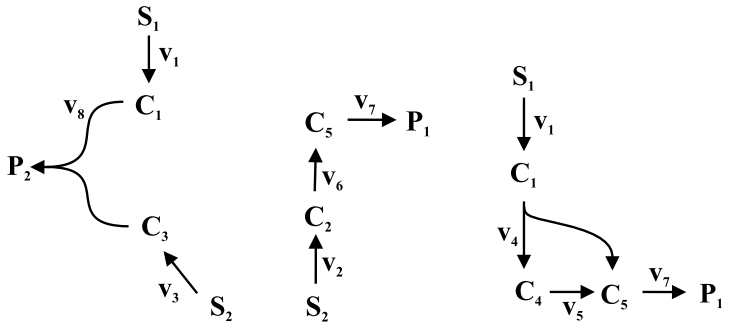

Fig. 6. The three elementary flux modes e_1, e_2, and e_3.

it commonly arises that the dimension of the convex basis is larger than the kernel dimension. These issues are illustrated with our application to CHO cells.

5.1. Application to CHO cells

For the metabolic network of CHO cells (Fig. 3), the stoichiometric matrix N is as follows:

$$N = \begin{pmatrix}
1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\
0 & 0 & 0 & 0 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & -1
\end{pmatrix}$$

Here the elementary flux modes are computed with the C-program Metatool available at the following URL: http://www.bioinf.mdc-berlin.de/projects/metabolic/metatool which description is available on the same URL.

Although the dimension of $\ker(N)$ is 6, the convex basis involves seven vectors listed in Table 4.

The seven corresponding elementary flux modes of CHO cells are expressed as follows:

e_1. Glucose → Glucose6P
Glucose6P → DihydroxyacetoneP + Glyceraldehyde3P
DihydroxyacetoneP → Glyceraldehyde3P
2 Glyceraldehyde3P → 2 Pyruvate
2 Pyruvate → 2 Lactate

e_2. Glucose → Glucose6P
Glucose6P → DihydroxyacetoneP + Glyceraldehyde3P
DihydroxyacetoneP → Glyceraldehyde3P
2 Glyceraldehyde3P → 2 Pyruvate
2 Pyruvate → 2 AcetylcoenzymeA + 2 CO_2
2 Oxaloacetate + 2 AcetylcoenzymeA → 2 Citrate
2 Citrate → 2 α-Ketoglutarate + 2 CO_2
2 α-Ketoglutarate → 2 Malate + 2 CO_2
2 Malate → 2 Oxaloacetate

e_3. Pyruvate + Glutamate → Alanine + α-Ketoglutarate
α-Ketoglutarate → Malate + CO_2
Malate → Pyruvate + CO_2
Glutamine → Glutamate + NH_4

e_4. Pyruvate → Lactate
α-Ketoglutarate → Malate + CO_2
Malate → Pyruvate + CO_2
Glutamate → α-Ketoglutarate + NH_4
Q → Glutamate + NH_4

e_5. Pyruvate → AcetylcoenzymeA + CO_2
AcetylcoenzymeA + Oxaloacetate → Citrate
Citrate → α-Ketoglutarate + CO_2
2 α-Ketoglutarate → 2 Malate + 2 CO_2
Malate → Oxaloacetate
Malate → Pyruvate + CO_2
Glutamate → α-Ketoglutarate + NH_4
Glutamine → Glutamate + NH_4

Table 4
The seven convex basis vectors

	e_1	e_2	e_3	e_4	e_5	e_6	e_7
v_1	1	1	0	0	0	1	1
v_2	1	1	0	0	0	0	0
v_3	0	0	0	0	0	1	1
v_4	1	1	0	0	0	0	0
v_5	2	2	0	0	0	0	0
v_6	2	0	0	1	0	0	0
v_7	0	0	1	0	0	0	0
v_8	0	2	0	0	1	0	0
v_9	0	2	0	0	1	0	0
v_{10}	0	2	0	0	1	0	0
v_{11}	0	2	1	1	2	1	1
v_{12}	0	2	0	0	1	1	1
v_{13}	0	0	1	1	1	0	0
v_{14}	0	0	0	0	0	1	1
v_{15}	0	0	0	1	1	0	0
v_{16}	0	0	1	1	1	1	1
v_{17}	0	0	0	0	0	1	0
v_{18}	0	0	0	0	0	0	1

e_6. Glucose → Glucose6P
Glucose6P → CO_2 + Ribose5P
α-Ketoglutarate → Malate + CO_2
Malate → Oxaloacetate
Oxaloacetate + Glutamate → Aspartate + α-Ketoglutarate
Glutamine → Glutamate + NH_4
Ribose5P + 2 Glutamine + Aspartate → Purine

e_7. Glucose → Glucose6P
Glucose6P → CO_2 + Ribose5P
α-Ketoglutarate → Malate + CO_2
Malate → Oxaloacetate
Oxaloacetate + Glutamate → Aspartate + α-Ketoglutarate
Glutamine → Glutamate + NH_4
Ribose5P + Glutamine + Aspartate → Pyrimidine

By eliminating the internal metabolites between the reactions, the following set of fundamental *macro-reactions* that connect the extracellular substrates and the end-products is obtained:

e_1: Glucose → 2 Lactate
e_2: Glucose → 6 CO_2
e_3: Glutamine → Alanine + 2 CO_2 + NH_4
e_4: Glutamine → Lactate + 2 CO_2 + 2 NH_4
e_5: Glutamine → 5 CO_2 + 2 NH_4
e_6: Glucose + 3 Glutamine → Purine + 2 CO_2 + NH_4
e_7: Glucose + 2 Glutamine → Pyrimidine + 2 CO_2 + NH_4

6. Dynamical macroscopic model

On the basis of this set of macro-reactions, a classical dynamical model of biological reactors [9] can be established as

$$\frac{\mathrm{d}\xi(t)}{\mathrm{d}t} = Kr(t) + u(t) \tag{12}$$

This model expresses the mass balance of the extracellular species inside the reactor with $\xi^{\mathrm{T}} = (\xi_1, \ldots, \xi_n)$ the vector of the species concentrations in the reactor liquid volume; $r^{\mathrm{T}}(t) = (r_1(t), \ldots, r_m(t))$ the vector of the macro-reaction rates; K the stoichiometric matrix of the macro-reaction network; $u(t)$ the net exchange of the species with the outside.

The stoichiometric matrix K is easily computed as

$$K = \begin{pmatrix} -N_s \\ N_p \end{pmatrix} E$$

where E denotes the matrix made up of the vectors of the convex basis of the underlying metabolic network.

For our example of batch cultures of CHO cells in flasks, we have

$$\xi = \begin{pmatrix} \text{Glucose (G)} \\ \text{Glutamine (Q)} \\ \text{Lactate (L)} \\ \text{Ammonia } (NH_4) \\ \text{Alanine (A)} \\ CO_2 \\ \text{Nucleotides (Nucl)} \end{pmatrix},$$

$$K = \begin{pmatrix} -1 & -1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 & -1 & -3 & -2 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 6 & 2 & 2 & 5 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \tag{13}$$

while $u_i(t) = 0$ for all species except for CO_2 where it represents the CO_2 gaseous outflow rate. We assume that the reaction rates $r_i(t)$ of the seven reactions satisfy simple Michaëlis–Menten kinetics. This means that the rates are expressed as

$$r_i = a_i \frac{GX}{k_{Gi} + G} \quad \text{for reactions } e_1\text{–}e_2$$

$$r_i = a_i \frac{QX}{k_{Qi} + Q} \quad \text{for reactions } e_3\text{–}e_5$$

$$r_i = a_i \frac{GQX}{(k_{Gi} + G)(k_{Qi} + Q)} \quad \text{for reactions } e_6\text{–}e_7$$

where a_i are the maximum specific reaction rates, k_{Gi}, k_{Qi} are the Michaëlis (or half-saturation) constants. The mass balance state space model for the measured extracellular species is thus written as

$$\begin{aligned}
\frac{dG(t)}{dt} = & -a_1 \frac{GX}{k_{G1} + G} - a_2 \frac{GX}{k_{G2} + G} \\
& - a_6 \frac{GQX}{(k_{G6} + G)(k_{Q6} + Q)} \\
& - a_7 \frac{GQX}{(k_{G7} + G)(k_{Q7} + Q)} \\
\frac{dQ(t)}{dt} = & -a_3 \frac{QX}{k_{Q3} + Q} - a_4 \frac{QX}{k_{Q4} + Q} - a_5 \frac{QX}{k_{Q5} + Q} \\
& - 3a_6 \frac{GQX}{(k_{G6} + G)(k_{Q6} + Q)} \\
& - 2a_7 \frac{GQX}{(k_{G7} + G)(k_{Q7} + Q)} \\
\frac{dL(t)}{dt} = & \; 2a_1 \frac{GX}{k_{G1} + G} + a_4 \frac{QX}{k_{Q4} + Q} \\
\frac{dN(t)}{dt} = & \; a_3 \frac{QX}{k_{Q3} + Q} + 2a_4 \frac{QX}{k_{Q4} + Q} + 2a_5 \frac{QX}{k_{Q5} + Q} \\
& + a_6 \frac{GQX}{(k_{G6} + G)(k_{Q6} + Q)} \\
& + a_7 \frac{GQX}{(k_{G7} + G)(k_{Q7} + Q)} \\
\frac{dA(t)}{dt} = & \; a_3 \frac{QX}{k_{Q3} + Q}
\end{aligned} \tag{14}$$

In order to complete the model, it remains to select numerical values for the 16 parameters a_i, k_{Gi} and k_{Qi} from the experimental data. In view of the small amount of available data, it is obvious that the model is widely over-parameterized. But under the balanced growth condition, it clearly makes sense to assume that the macro-reactions proceed almost at their maximal rate during the exponential growth phase. The half-saturation constants k_{Gi} and k_{Qi} are therefore selected small enough to be ineffective during the growth phase but large enough to avoid stiffness difficulties in the numerical simulation of the model. Here we have set the half-saturation constants at

$$k_{Gi} = k_{Qi} = 0.1 \text{ mM} \quad \forall i$$

Thus, as long as $G \gg 0.1$ mM and $Q \gg 0.1$ mM, we have the following linear relation between the *measured* specific uptake/excretion rates and the macro-reaction rates a_i:

$$\begin{pmatrix} 4.0546 \\ 1.1860 \\ 7.3949 \\ 0.9617 \\ 0.2686 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 & -1 & -3 & -2 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{pmatrix} \tag{15}$$

In addition, in accordance with equality (10), we have the natural constraint

$$a_6 = a_7 \tag{16}$$

The parameters a_i must be computed as a non-negative solution of the following system:

$$\begin{pmatrix} 4.0546 \\ 1.1860 \\ 7.3949 \\ 0.9617 \\ 0.2686 \\ 0 \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & -1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 & -1 & -3 & -2 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}}_{K'} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{pmatrix} \tag{17}$$

System (17) is underdetermined. The particular non-negative and minimal norm solution of (17) is

a_1^0	a_2^0	a_3^0	a_4^0	a_5^0	a_6^0	a_7^0
3.5956	0.1736	0.2686	0.2038	0	0.1427	0.1427

The simulation results with these parameter values are shown in Fig. 7. Remark that the fifth reaction Glutamine $\rightarrow$ 5 CO_2 + 2 NH_4 is inactive.

The general solution of (17) is written as

$$a_i = a_i^0 + \alpha k_i, \quad i = 1, \ldots, 7, \;\; \alpha \in \mathbb{R} \tag{18}$$

where the vector k is an arbitrary basis vector of $\ker(K')$ which is one-dimensional. A possible choice of k is

$$k = \begin{pmatrix} -0.5 \\ 0.5 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \tag{19}$$

The general solution of (17) is thus written:

$$a_1 = 3.5956 - 0.5\alpha$$

$$a_2 = 0.1736 + 0.5\alpha$$

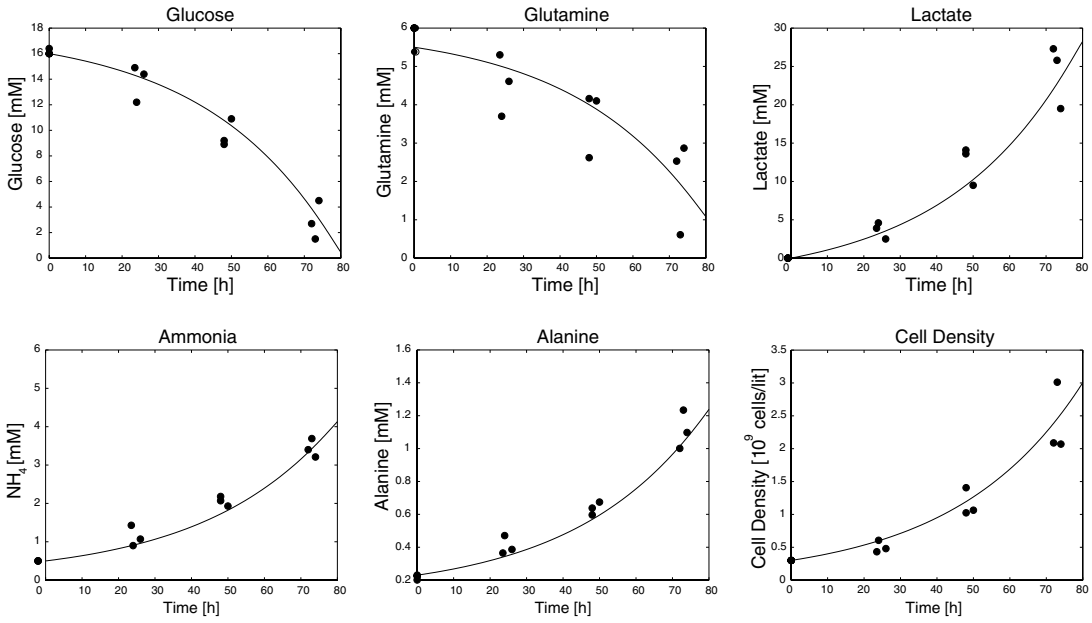

Fig. 7. Experimental data and simulation results.

$$a_3 = 0.2686$$

$$a_4 = 0.2038 + \alpha$$

$$a_5 = -\alpha$$

$$a_6 = 0.1427$$

$$a_7 = 0.1427$$

In order to keep a non-negative solution, we clearly have to constrain the parameter α in the interval:

$$0 \geqslant \alpha \geqslant -0.2038$$

7. Conclusion

The identification of mathematical models for bioprocesses from macro-reactions has been investigated under a metabolic viewpoint. The proposed approach has been illustrated with the example of CHO cell metabolism.

Such models provide a firm basis for the design of online monitoring and optimization of cell culture processes. In particular, the underlying metabolic basis of the model should be helpful in determining how to alter the culture environment so as to achieve robust control and maintain optimal conditions.

Nevertheless, it must be mentioned that the model developed in this paper is clearly valid only during the growth phase and should be carefully extrapolated to other culture phases. This will be the subject of further investigation.

Acknowledgements

This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors.

The authors thank Prof. Y.-J. Schneider, Prof. S. Agathos and Dr. F. Verhoeye (Groupe de Bioingnierie cellulaire, UCL) for numerous highly stimulating discussions and advices on this work and for providing the experimental data.

References

[1] M.J. Keen, N.T. Rapson, Development of a serum-free medium for large scale production of recombinant protein from Chinese Hamster Ovary cell line, Cytotechnology 17 (1995) 153–163.
[2] J. Bailey, Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities, Biotechnology Progress 14 (1998) 8–20.
[3] G. Stephanopoulos, J. Nielsen, A. Aristidou, Metabolic Engineering: Principles and Methodologies, Academic Press, San Diego, 1998.
[4] J.E. Bailey, D.F. Ollis, Biochemical Engineering Fundamentals, second ed., McGraw-Hill, New York, 1986.
[5] S. Schuster, C. Hilgetag, On elementary flux modes in biochemical reaction systems in steady state, Journal of Biological Systems 2 (1994) 165–182.
[6] T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.
[7] C.H. Schilling, S. Schuster, B.O. Palsson, R. Heinrich, Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era, Biotechnology Progress 15 (1999) 296–303.
[8] S. Schuster, T. Dandekar, D.A. Fell, Detection of elementary flux modes in biochemical networks: a promising tool for pathway

analysis and metabolic engineering, Trends in Biotechnology 17 (1999) 53–60.
[9] G. Bastin, D. Dochain, On-line estimation and adaptive control of bioreactors, Elsevier, Amsterdam, 1990.
[10] F. Gòdia, J.J. Cairó, Metabolic engineering of animal cells, Bioprocess and Biosystems Engineering 24 (2002) 289–298.
[11] C. Zupke, G. Stephanopoulos, Intracellular flux analysis in hybridomas using mass balances and in vitro C13 NMR, Biotechnology and Bioengineering 45 (1995) 292–303.
[12] H.P.J. Bonarius, V. Hatzimanikatis, K.P.H. Meesters, C.D. de Gooijer, G. Schmid, J. Tramper, Metabolic flux analysis of hybridoma cells in different culture media using mass balances, Biotechnology and Bioengineering 50 (1996) 299–318.
[13] A.A. Namjoshi, W.S. Hu, D. Ramkrishna, Unveiling steady-state multiplicity in hybridoma cultures: the cybernetic approach, Biotechnology and Bioengineering 81 (1) (2003) 80–91.