Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

DETERMINATION OF MACROSCOPIC REACTION SCHEMES: TOWARDS A UNIFYING VIEW

Ph. Bogaerts[a], M. Rooman[a], V. Vastemans[a] and A. Vande Wouwer[b]
[a] *Service de Chimie générale et Biosystèmes, Université Libre de Bruxelles, CP 165/55, 50, Av. F. D. Roosevelt, B-1050 Brussels, Belgium (E-mail: Philippe.Bogaerts@ulb.ac.be)*
[b] *Service d'Automatique, Faculté Polytechnique de Mons, Boulevard Dolez, 31, 7000 Mons, Belgium (E-mail: Alain.VandeWouwer@fpms.ac.be)*

Abstract: The systematic determination of macroscopic biological reaction networks from experimental data records of the evolution of a set of external substrates and cell products has received increasing attention in recent years. The purpose of this paper is to review existing methods, highlighting the potential connection between them using the concept of equivalence of reaction schemes and discussing potential extensions. *Copyright © 2008 IFAC*

Keywords: reaction networks, mathematical modelling, parameter estimation, biological systems

1. INTRODUCTION

Two approaches on the determination of macroscopic reaction schemes can be distinguished: one in which the available knowledge about the metabolic network is reduced to a 'minimal' set of macroscopic reactions linking external substrates and cellular products (Haag *et al.*, 2005a and 2005b; Provost and Bastin, 2004; Provost *et al.*, 2006), and the other which attempts to directly establish (i.e. without using detailed knowledge about the metabolic network) a macroscopic reaction scheme from experimental measurements of the time evolution of several substrates and external cellular products (Bogaerts and Vande Wouwer, 2001; Hulhoven *et al.* 2005; Bernard and Bastin; 2005a and 2005b). Each approach has its own advantages and drawbacks. The first approach requires the availability of detailed information about the metabolic network of the microorganisms under consideration, but allows, with very little additional effort, the derivation of minimal macroscopic reaction schemes which are consistent with this prior information. The second approach is mostly of interest in situations where prior information about the metabolism is incomplete or unavailable (this situation is frequent in the bioproduction of pharmaceuticals, where genetically manipulated strains are used for expressing products of interest, such as recombinant proteins, antibodies, etc.) and in environmental or agro-food processes where a large range of populations of microorganisms (bacteria, moulds, etc) are involved.

In this study, we focus attention on this second approach and try to connect, in a unifying way, the results presented in (Bogaerts and Vande Wouwer, 2001; Hulhoven *et al.* 2005; Bernard and Bastin; 2005a and 2005b). In particular, we show how the results of these studies could be advantageously combined and extended and, for this purpose, we introduce the concept of equivalence of reaction schemes. As an end result, we propose a systematic procedure for the determination of a macroscopic reaction scheme (and associated kinetics) which could be used for the subsequent design of monitoring and control tools (including simulation, state estimation, optimization, control and supervision).

The text is organised as follows. Section 2 reviews existing results for determining C-identifiable reaction schemes (Bogaerts and Vande Wouwer, 2001; Hulhoven *et al.* 2005) and a method for determining the number of reactions and partial information on the stoichiometry (Bernard and Bastin; 2005a and 2005b). Pros and cons of these methods are also discussed in this section. Section 3 introduces the concept of equivalent reaction schemes. Section 4 proposes a way towards a unifying approach, taking advantages of the above mentioned methods and trying to circumvent their problems. Section 5 introduces an example before drawing some conclusions in Section 6.

2. PRELIMINARIES

Consider the macroscopic reaction scheme (Bastin and Dochain, 1990) given by

$$\sum_{i \in R_k} \left(-\nu_{i,k} \right) \xi_i \xrightarrow{\varphi_k} \sum_{j \in P_k} \nu_{j,k} \xi_j \quad k \in [1,M] \quad (1)$$

where M is the number of reactions;
ξ_i the i-th component;

φ_k the k-th reaction rate;

$\nu_{i,k}$ and $\nu_{j,k}$ the pseudo-stoichiometric (or yield) coefficients (positive when associated to a component which is produced, negative when it is consumed).

The system of mass balances for each of the components ξ_i can be written in the following matrix form

$$\frac{d\,\xi(t)}{d\,t} = K\,\varphi(\xi,t) - D(t)\,\xi(t) + u(t) \qquad (2)$$

where $\xi \in \Re^N$ is the vector of concentrations;

$K \in \Re^{NxM}$ is the pseudo-stoichiometric coefficients matrix ($N \geq M$);

$\varphi \in \Re^M$ is the vector of reaction rates;

$D \in \Re$ is the dilution rate;

$u(t) = F(t) - Q(t)$ with $F \in \Re^N$, the vector of external feed rates and $Q \in \Re^N$, the vector of gaseous outflow rates.

The following developments will assume that:

- $D(t)$ and $u(t)$ are known;
- $\xi(t)$ is measured at discrete times;
- $\varphi(\xi,t)$ is unknown;
- the rank of K is equal to M, i.e. the reactions are linearly independent;
- K is partially or completely unknown and has to be identified independently of the kinetics.

In the following, we first present the main results that are currently available and their pros and cons.

2.1. Determination of the number of reactions and of partial information on the stoichiometry (Bernard and Bastin; 2005a and 2005b)

The first step in the procedure consists in the determination of the number of columns of K, i.e., the number of independent reactions that are distinguishable from the available data using principal component analysis (PCA). To this end, equation (2) can be integrated between two time instants t and $t+T$

$$\xi(t+T) - \xi(t) + \int_t^{t+T} \left(D(\tau)\xi(\tau) - u(\tau)\right) d\tau = K \int_t^{t+T} \varphi(\xi(\tau))\, d\tau \qquad (3)$$

or in compact form $x(t) = Kw(t)$. This latter equation can be expressed at several measurement times $t_k, k = 1,...,L$ (with L>N), by defining the NxL matrix $X = \left[x(t_1) .... x(t_L)\right]$ and the (unknown) MxL matrix $W = \left[w(t_1) .... w(t_L)\right]$, i.e.

$$X = KW \qquad (4)$$

Assuming that W has full rank (i.e., the reactions are independent), the number of reactions is given by the number of nonzero eigenvalues of XX^T.
In practice, several influencing factors have to be taken into account, such as the measurement noise, filtering and interpolation errors, etc., so that the number of reactions is selected according to the first M eigenvalues representing a total variance larger than a fixed confidence threshold.

The second step of the procedure is the determination of the K matrix itself, or at least part of it. The eigenvectors ρ_i associated to the eigenvalues σ_i of the matrix XX^T forms an orthonormal basis that spans K, i.e. there exists a MxM matrix G such that

$$K = \rho G \qquad (5)$$

where the columns of the matrix ρ are the eigenvectors ρ_i. In order to make the matrix G uniquely identifiable, it is necessary to introduce additional structural constraints based on *a priori* biological knowledge of the system.

2.2. Determination of C-identifiable reaction schemes (Bogaerts and Vande Wouwer, 2001; Hulhoven *et al.* 2005)

When $rank\ K = M$, there always exists a partition $K^T = \left[\, K_a^T\ K_b^T \,\right]$ where $K_a \in \Re^{MxM}$ is of full row rank. This partition implies corresponding partitions of ξ and u, i.e., $\xi^T = \left[\, \xi_a^T\ \xi_b^T \,\right]$ and $u^T = \left[\, u_a^T\ u_b^T \,\right]$.
Given such a partition of K, the following matrix equation

$$C\,K_a + K_b = O \qquad (6)$$

(where $O \in \Re^{(N-M)xN}$ is a null matrix) has always a unique solution $C \in \Re^{(N-M)xM}$, which can be used to define an auxiliary vector $z \in \Re^{N-M}$:

$$z(t) = C\,\xi_a(t) + \xi_b(t) \qquad (7)$$

whose dynamics is independent of the kinetics:

$$\frac{d\,z(t)}{dt} = -\,D\,z(t) + C\,u_a(t) + u_b(t) \qquad (8)$$

On the basis of equations (7-8), measurements of the component concentrations and the knowledge of the inputs, an estimation $\hat{C}$ of the matrix C can be computed. Estimates $\hat{K}_a$ and $\hat{K}_b$ are then deduced from equation

(6). The uniqueness of these estimates is called C-identifiability.

A procedure for determining a subset of C-identifiable reaction schemes is based on the use of a sufficient condition deduced from (Chen and Bastin, 1996), which can be stated as follows:

A reaction scheme is C-identifiable if there exists a partition $K^T = \left[K_a^T \; K_b^T \right]$ where $K_a \in \Re^{MxM}$ is full rank and does not contain any unknown coefficient of *K*.

Under this sufficient condition, excluding all prior information on *K* compels K_a to take the following form, after appropriate permutation of its rows:

$$K_a = diag\,\{\pm 1, \pm 1, ..., \pm 1\} \tag{9}$$

where a coefficient ± 1 corresponds to either a product (+1) or a substrate (-1) with respect to which the reaction is normalized.
The determination of C-identifiable reaction schemes goes through a systematic screening of all the possible $\xi_a \in \Re^M$ candidates from the set of *N* available components. This screening has to be repeated for all the possible number of reactions (as *M* is *a priori* unknown as well), leading to an overall number of combinations

$$\sum_{M=1}^{N-N_0-1} \binom{N-N_0}{M} = \sum_{M=1}^{N-N_0-1} \frac{(N-N_0)!}{M!(N-N_0-M)!} \tag{10}$$

where N_0 represents the number of components that are consumed in some reactions and produced in others so that they cannot belong to ξ_a.
For each scenario, a maximum-likelihood estimation of the pseudo-stoichiometry is achieved (Bogaerts *et al.*, 2003), and for a given *M*, the candidate can be selected by comparing the residual values of the cost function at the optimum.
This procedure can be refined in several ways including the consideration of sign constraints on the elements of K_b.

2.3. Pros and cons

Usual criticisms about method 2.2 are the following:

- the test of all possible numbers of reactions is lengthy;
- for a given number of reactions, the method only generates a subset of the C-identifiable schemes as a sufficient (but not necessary) condition of C-identifiability is used;
- the most meaningful (i.e. biologically interpretable) reaction scheme could be non C-identifiable.

The main drawback associated with method 2.1 is that, except for the determination of the number of reactions, it requires some *a priori* knowledge about the reaction scheme that is sought, and the method is therefore not as systematic as 2.2.

3. EQUIVALENT REACTION SCHEMES

In order to generalize the methodology explained in subsection 2.2, the concept of equivalent reaction schemes is introduced:

Definition 3.1
K and *K'* represent two equivalent reaction schemes if for any $\varphi(\xi)$ there exists a $\varphi'(\xi)$ (and conversely) such that

$$K\,\varphi(\xi) = K'\,\varphi'(\xi) \tag{11}$$

Note that, from a general point of view, *K* and *K'* must of course have the same number of rows but may exhibit different numbers of columns.

Property 3.1
If there exists a regular square matrix *P* such that

$$K' = K\,P \tag{12}$$

then *K* and *K'* are equivalent.

Indeed, for any $\varphi(\xi)$, one has to consider

$$\varphi'(\xi) = P^{-1}\,\varphi(\xi) \tag{13}$$

such that (11) holds.

On the basis of this definition and this property, three important consequences can be highlighted.

1) For any reaction scheme characterized by a stoichiometry matrix *K* (even non C-identifiable) such that rank(*K*) = *M*, there necessarily exists an equivalent reaction scheme $K' = K\,K_a^{-1}$ (where $K_a \in \Re^{MxM}$ is a full rank submatrix of *K*) which contains an identity submatrix $I_M \in \Re^{MxM}$. Indeed, those schemes would be equivalent using the property 3.1 with $P = K_a^{-1}$ and, based on the decomposition $K^T = \left[K_a^T \quad K_b^T \right]$ introduced in subsection 2.2, the equivalent reaction scheme would be described by the stoichiometry matrix

$$K' = \begin{bmatrix} I_M \\ K_b\,K_a^{-1} \end{bmatrix} \tag{14}$$

2) Some of the reaction schemes generated by the systematic methodology recalled in subsection 2.2 could be equivalent. In order to limit the search to non equivalent

reaction schemes, the following procedure can be followed:

- Determine any reaction scheme containing an identity submatrix of the form (14).
- Find any singular submatrix living in $\Re^{MxM}$ of this stoichiometry matrix. As this submatrix could not be transformed into an identity stoichiometry submatrix through a regular mapping *P*, the corresponding reaction scheme is not equivalent to the previously determined reaction scheme.
- Repeat the previous step until there is no more singular submatrix corresponding to a partition which has not been already analyzed.

3) For a given number *M* of reactions, the "true" reaction scheme (or the most interpretable one from a biological point of view) will, in most of the cases, not be recovered but only its equivalent exhibiting a submatrix of the form (9). Nevertheless, this solution only differs from the "true" (or most meaningful) one by the mapping based on an unknown matrix *P*, according to property 3.1.

4. TOWARDS A UNIFYING PROCEDURE

The discussion of Section 2 shows that the number of macroscopic reactions can be determined using PCA of the experimental data at hand. Based on this preliminary analysis, it is possible to proceed with one of the two procedures for determining a macroscopic reaction scheme:

(a) The first procedure (subsection 2.1) yields a biologically consistent reaction scheme (especially its number of reactions), but requires additional constraints (based on prior knowledge) which might not be available.

(b) The second procedure (subsection 2.2) yields a most likely C-identifiable reaction scheme. This scheme will generally allow a good representation of the experimental data, but can lack a biological interpretation, if biologically consistent reaction schemes do not satisfy the C-identifiability condition. Even if they would be C-identifiable (which is sometimes the case), they can be in a form which does not contain any identity submatrix (as in 14 or 9) and we don't know at this stage how do determine the mapping *P*, which would allow us to recover them.

As a by-product of the procedure exposed in subsection 2.1, the eigenvectors associated with the eigenvalues of the XX^T matrix form an orthonormal basis which spans the possible set of biological reactions. As a matter of fact, the C-identifiable reaction scheme belongs to this reaction space, and could play the role of an alternative basis. In (Helias and Bernard, 2007), a set of known reactions (with stoichiometric coefficients available from expert knowledge) is confronted to this orthonormal basis so as to select the reactions that belong to the image of *K*.

In the sequel, we will only consider the case when no additional expertise is known and no *a priori* assumptions are made. Using what is developed in section 3, we then propose the following general procedure:

1) Determine the lowest number *M* of reactions allowing to reproduce the experimental data at hand, based on the PCA analysis of subsection 2.1.

2) Given the number of reactions identified in the preceding step, determine the most likely reaction scheme of the form (14), i.e., containing an identity submatrix.

At this point of the procedure, two options arise.

3a) The first one corresponds to the case when the determination of a biologically meaningful reaction network is not required. Indeed, if the main motivation of the modeller is to determine a macroscopic model for control (in a broad sense), it can be sufficient to determine $K\varphi$ as a whole, for instance using a C-identifiable scheme for *K* of the form (14) and, subsequently, identifying a kinetic structure φ. In order to have enough degrees of freedom, such a structure has to be in the form of a combination of sufficiently flexible basis functions. For instance, one could use radial basis artificial neural networks or general kinetic functions like those proposed in (Grosfils *et al.*, 2007) but without any positivity constraint on the kinetic constants (thus allowing negative kinetic model structures). In this case, one could wonder why using a two-step procedure (first determining *K* containing an identity submatrix, then identifying general combinations of basis functions for φ) instead of directly identifying a global model $K\varphi$. The big advantage of the two-step procedure is the following. When turning to the identification of such "black-box" models, whatever the one- or two-step procedure, the results of the parameter estimation procedure becomes very dependent of the quality of the parameter initial guess (due to the necessity to numerically solve a nonlinear optimization problem with the risks of finding local minima). Therefore, combinations of basis functions linearizable w.r.t. their parameters are of primary interest. This is the case with radial basis artificial neural networks (after a non supervised training of the centers and widths of the Gaussian activity functions) or with general kinetic model structures like the ones proposed in (Grosfils *et al.*, 2007). When applying such linearizable models to the whole $K\varphi$, one has to estimate the time derivatives of the whole state ξ, whereas in the case of identifying such linearizable models only for φ, it is sufficient to use the time derivatives of *M* components of ξ (the ones corresponding to the rows of the identity submatrix of (14)). Moreover, if there are several reaction schemes equivalent to the best one identified in step 2), then the one which should be used is the one whose identity submatrix rows correspond to the

variables whose time derivatives are the easiest to estimate (depending on the availability and quality of experimental data).

Finally, coming out of step 2), the other option is the following.

3b) The targeted reaction scheme is the most meaningful one from a biological point of view. It can be C-identifiable or not but, in any case, it must be equivalent to the one identified in step 2). Therefore, this last step consists in determining both the mapping matrix *P* and the (biologically meaningful) kinetic model structures φ' such that the equivalence relations (11) and (12) hold. The way to estimate efficiently such results will be detailed in a future publication.

5. AN EXAMPLE

Let us consider the model of Sonnleitner and Käppeli (1986) which describes the behaviour of *Saccharomyces cerevisiae* yeast cultures. The three reactions of the scheme are the following:

$$\begin{aligned} &G + k_5 O_2 \xrightarrow{\varphi_1} k_1 X + k_7 P \\ &G \xrightarrow{\varphi_2} k_2 X + k_4 E + k_8 P \\ &E + k_6 O_2 \xrightarrow{\varphi_3} k_3 X + k_9 P \end{aligned} \tag{15}$$

where *G*, O_2, *X*, *E* and *P* stand for glucose, oxygen, biomass (yeast cells), ethanol and carbon dioxide respectively. Considering the state vector

$$\xi^T = \begin{bmatrix} G & O_2 & X & E & P \end{bmatrix} \tag{16}$$

the stoichiometry matrix is given by

$$K = \begin{bmatrix} -1 & -1 & 0 \\ -k_5 & 0 & -k_6 \\ k_1 & k_2 & k_3 \\ 0 & k_4 & -1 \\ k_7 & k_8 & k_9 \end{bmatrix} \tag{17}$$

Note that, given the necessary and sufficient condition of C-identifiability given in (Chen and Bastin, 1996), this reaction scheme can easily be shown not to be C-identifiable considering neither column of *K* contains *M*=3 known elements. Using the numerical values given by the same authors, *K* corresponds to

$$K = \begin{bmatrix} -1 & -1 & 0 \\ -0.3968 & 0 & -1.1040 \\ 0.49 & 0.05 & 0.72 \\ 0 & 0.48 & -1 \\ 0.5897 & 0.4621 & 0.6249 \end{bmatrix} \tag{18}$$

Considering the first three rows as submatrix K_a and the last one as submatrix K_b, an equivalent scheme with identity submatrix of the form (14) is given by

$$K' = K\,K_a^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.5133 & 0.4718 & -0.6654 \\ -0.4889 & -0.9151 & -0.5353 \end{bmatrix} \tag{19}$$

Assuming that a PCA of the experimental data at hand has led to the conclusion that 3 reactions are sufficient to describe the data, the second step of section 4 could lead to the equivalent C-identifiable scheme (19), thanks to the method described in subsection 2.2. However, this scheme has no biological meaning:

$$\begin{aligned} &0.5133E + 0.4889P \xrightarrow{\varphi'_1} G \\ &0.9151P \xrightarrow{\varphi'_2} O_2 + 0.4718E \\ &0.6654E + 0.5353P \xrightarrow{\varphi'_3} X \end{aligned} \tag{20}$$

According to steps 3a) and 3b) presented in Section 4, two solutions arise:

- The first one consists in determining "black-box" combinations of basis functions describing $\varphi'_1, \varphi'_2, \varphi'_3$ using the equivalent stoichiometry matrix given in (19). Note that any other equivalent reaction scheme containing an identity submatrix could be used, corresponding to three variables whose time derivatives should be estimated. If, for instance, the user owns a glucose-ethanol analyzer and a pO_2 probe, but no biomass measurement device, then another equivalent reaction scheme should be considered, whose identity submatrix would correspond to the first, second and fourth rows.

- The second one consists in identifying both the mapping $P = K_a^{-1}$ and biologically meaningful kinetic model structures $\varphi_1, \varphi_2, \varphi_3$.

6. CONCLUSIONS

In this paper, several methods allowing the derivation of a macroscopic reaction scheme from experimental data are first reviewed. Attention is focused on methods aiming at deducing a macroscopic reaction scheme on the basis of extracellular measurements, rather than on methods based on some *a priori* detailed knowledge of the metabolic network and trying to map this knowledge into a macroscopic framework. Pros and cons of such existing

methods are discussed, leading to the proposition of a new unifying procedure. The latter consists of three main steps. The first one determines the minimal number of reactions, so as to reproduce the experimental data at hand, on the basis of a PCA, as proposed in (Bernard and Bastin; 2005a and 2005b). The second step generalizes the results presented in (Bogaerts and Vande Wouwer, 2001; Hulhoven *et al.* 2005) and proposes, for the number of reactions determined in the preceding step, the best reaction scheme containing an identity submatrix. A discussion of the concept of equivalent reaction schemes shows that this result is representative of the "best" solution, whatever it is C-identifiable or not. The third and last step can be tackled in two different ways. The first one aims at deriving a global reaction term $K\varphi$ and consists in determining "black-box" combinations of basis functions (without requiring any biological meaning) φ, based on the particular stoichiometry identified in the preceding step. This step-by-step procedure is shown to be more efficient from a parameter estimation point of view than a "one shot" identification of a global model $K\varphi$. Alternatively, a biologically meaningful reaction network can be derived. This involves the estimation of, on the one hand, the mapping P from the stoichiometry determined in the second step to the "true" or "more interpretable" one and, on the other hand, the biologically meaningful kinetic model structures.

Future work will focus on the proposition of efficient methods to tackle this last case, allowing, despite the lack of any C-identifiability property, to get a first estimate of biologically meaningful kinetic parameters independently of the stoichiometry mapping.

7. ACKNOWLEDGEMENT

This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.V. Vastemans benefits from a FRIA grant from the Fund for Scientific Research (FRS). M. Rooman is Research Director of the FNRS.

8. REFERENCES

Bastin, G. and D. Dochain (1990). *On-line estimation and adaptive control of bioreactors*. Elsevier, Amsterdam.

Bernard, O., and G. Bastin (2005a). Identification of reaction networks for bioprocesses: determination of a partially unknown pseudo-stoichiometric matrix. *Bioprocess Biosyst. Eng.*, **27(5)**, 293-301.

Bernard, O., and G. Bastin (2005b). On the estimation of the pseudo-stoichiometric matrix for macroscopic mass balance modelling of biotechnological processes. *Math.Biosci.*, **193(1)**, 51-77.

Bogaerts, Ph., and A. Vande Wouwer (2001). Systematic generation of identifiable macroscopic reaction schemes. *Proceedings of the 8th IFAC Conference on Computer Applications in Biotechnology (CAB8), Montréal (Canada)*, 24-27/6/2001, pp. 13-18.

Bogaerts, Ph., Delcoux J.-L. and Hanus R. (2003). Maximum likelihood estimation of pseudo-stoichiometry in macroscopic biological reaction schemes. *Chem. Eng. Sc.*, **58(8)**, 1545-1563.

Chen, L. and G. Bastin (1996). Structural identifiability of the yield coefficients in bioprocess models when the reaction rates are unknown. *Math. Biosci.*, **132**, 35-67.

Delcoux, J.-L., R. Hanus and Ph. Bogaerts (2001). Equivalence of reaction schemes in bioprocesses modelling. *Proceedings of the Fourth International Symposium on Mathematical Modelling and Simulation in Agricultural and Bio-Industries (M^2SABI'01), Haifa (Israel),* 12-14/6/2001.

Grosfils, A., A. Vande Wouwer and Ph. Bogaerts (2007). On a general kinetic model structure for macroscopic biological reaction rates. *Journal of Biotechnology*, **130**, 253-264.

Haag, J., A. Vande Wouwer and Ph. Bogaerts (2005a). Dynamic modeling of complex biological systems: A link between metabolic and macroscopic description. *Math. Biosci.*, **193(1)**, 25-49.

Haag, J., A. Vande Wouwer and Ph. Bogaerts (2005b). Systematic procedure for the reduction of complex biological reaction pathways and the generation of macroscopic equivalents. *Chem. Eng. Sci.*, **60(2)**, 459-465.

Helias, A. and O. Bernard (2007). Assessing the main reactions in a bioprocess: Application to cheese ripening. *Proceedings of the 10th IFAC International Symposium on Computer Applications in Biotechnology (CAB2007)*, *Cancun (Mexico),* 4-6/6/2007.

Hulhoven, X., A. Vande Wouwer and Ph. Bogaerts (2005). On a systematic procedure for the predetermination of macroscopic reaction schemes. Bioprocess *Biosyst. Eng.*, **27**, 283-291.

Provost, A. and G. Bastin (2004). Dynamic metabolic modelling under the balanced growth condition. *Journal of Process Control*, **14(7)**, 717-728.

Provost, A., G. Bastin, S.N. Agathos and Y.J. Schneider (2006). Metabolic design of macroscopic bioreaction models: application to Chinese hamster ovary cells. *Bioprocess Biosyst. Eng.*, **29(5-6)**, 349-66.

Sonnleitner, B. and O. Käppeli (1986). Growth of *Saccharomyces cerevisiae* is limited by its respiratory capacity: formulation and verification of a hypothesis. *Bioprocess Biosyst. Eng.*, **29(5-6)**, 349-66.