10th International IFAC Symposium on
Computer Applications in Biotechnology
Proceedings Vol.1, June 4-6, 2007, Cancún, Mexico

ASSESSING THE MAIN REACTIONS IN A BIOPROCESS: APPLICATION TO CHEESE RIPENING

Arnaud Hélias * **Olivier Bernard** **

* *UMR782 Génie et Microbiologie des Procédés Alimentaires, INRA, AgroParisTech F-78850 Thiverval-Grignon, France. (arnaud.helias@grignon.inra.fr)*
** *INRIA, COMORE, BP 93, F-06902 Sophia-Antipolis, France. (olivier.bernard@inria.sophia.fr)*

Abstract: The aim of this paper is to select the main reactions of a bioprocess from a set of plausible metabolic pathways provided by expert knowledge. We use a methodology aiming at determining the pseudo-stoichiometric coefficient matrix of a macroscopic mass balance based model. First, the size of the system is identified and a subspace where the bioprocess dynamics lives is established. In a second step, the set of *a priori* plausible reactions is compared with the identified subspace and the most adequate reactions are selected. This approach is applied to cheese ripening experimental data. As the main result the method leads to the identification of a metabolic network that can be the base for dynamical model development.

Keywords: Modelling, Identification, Cheese ripening, Metabolic network.

1. INTRODUCTION

The ripening process is one of the most important steps for many cheese makers. Cheese ripening is a solid substrate fermentation based on a complex ecosystem composed of bacteria, yeast and mould. For example, considering only non lactic acid bacteria, Ogier *et al.* identified 14 to 20 species for different cheese types (Ogier *et al.*, 2004). Development of a ripening microbial consortium, associated with residual lactic acid bacteria activities, leads to organoleptic features of the cheese: (i) rind apparition (mainly composed of *Penicillium Camemberti* in Camembert), (ii) texturization depending on deacidification, proteolysis and liposysis, and (iii) aroma compounds productions.

For cheese ripening, several studies describe the growth characteristics for a given species ,*e.g.* (Aldarf *et al.*, 2006; Riahi *et al.*, 2007; Barba *et al.*, 2001), but to our knowledge, a macroscopic model of cheese ripening does not exist.

Macroscopic modelling can be used to base on-line tools for control and diagnosis of bioprocesses. It is also an interesting way to characterize the main phenomena that take place, especially when a complex ecosystem is used. In the considered approach the system is represented by a limited number of reactions, which are assumed to represent the main mass fluxes throughout the system.

This paper is based on a two step methodology aiming at identifying the structure of the pseudo-stoichiometric (PS) coefficients matrix (Bernard and Bastin, 2005*a*; Bernard and Bastin, 2005*b*). The first step consists in evaluating the number of reactions to be taken into account using a prin-

cipal component analysis. In the second step the unknown coefficients are computed by introducing additional constraints in the PS matrix.

In the present study, the aim is slightly different: A set of realistic theoretical reactions is assumed that may represent the cheese ripening (*i.e.* the metabolic pathways identified for the different species), and we determine among this set those which are mainly triggered. More precisely, the idea consists in comparing each theoretical reaction (represented by a vector of $\mathbf{R}^n$) to the vectorial subspace identified together with the PS matrix structure.

This approach is used on data collected along three experimental ripening runs of surface-mould cheese (Camembert-type).

2. MATERIALS AND METHODS

2.1 Cheese production and ripening

Soft cheeses (Camembert-type) were manufactured in a sterile environment as previously described in (Leclercq-Perlat *et al.*, 2004). 45 cheeses per production run were manufactured. After drainage, the cheeses were aseptically transferred to a sterile ripening chamber (volume = 0.99 m^3, regulated at 13°C); this point was considered as the initial time. Ripening duration is 14 days, the cheeses were turned over on day 5. The atmospheric changes were described with CO_2 and O_2 sensors. Since the ripening chamber was used without an input airflow, the variation of these gas concentrations depended only of exchange with products. During the ripening, a cheese was removed daily for analysis of lactose and lactic acid at the rind and at the core level (see (Leclercq-Perlat *et al.*, 2000) for more details). Three runs were realized. They were carried out with a periodically renewed atmosphere: if necessary, the CO_2 concentration was decreased to 2% by a 6 m^3/h flow rate daily air injection. In practice, the atmosphere was not renewed except 30 min per day.

2.2 Determination of the number of reactions

2.2.1. Bioprocess dynamical model The generic model of a multi-compartment bioprocess can be written as follows:

$$\frac{d\xi}{dt} = Kr(t) - v(t) + \phi(t), \qquad (1)$$

where $\xi = (\xi_1, \xi_2, \ldots, \xi_n)^T$ is the set of biochemical concentrations, which describe the bioprocess state. $v(t)$ is the net balance between inflows and outflows and $\phi(t)$ represents the fluxes between the considered compartments. The term $Kr(t)$ represents the transformation phenomena in the bioreactor. $r(t) = (r_1(t), r_2(t), \ldots, r_p(t))^T$ is a vector of the reaction rates; it is supposed to depend on the state ξ and environmental factors. Matrix K is the pseudo-stoichiometric matrix associated with the macroscopic reaction network. The coefficient K_{ij}, $i = 1, \ldots, n$ and $j = 1, \ldots, p$ represents the relationship between the j^{th} reaction and the i^{th} concentration. A positive K_{ij} value is related to product biosynthesis, while substrate consumption is observed when K_{ij} is negative; if $K_{ij} = 0$ this species is not involved in the j^{th} reaction.

2.2.2. Dimension of the reaction network In a macroscopic approach, the aim is to define the smallest number of reactions that can represent concentration dynamics keeping a biological and biochemical meaning. Let us denote

$$u(t) = \frac{d\xi}{dt} + v(t) - \phi(t)$$

From equation (1) we have

$$u(t) = Kr(t)$$

K is assumed to be a full rank matrix, otherwise, it would mean that the same dynamical behaviour could be obtained with a matrix K of lower dimension, by defining appropriate reaction rates.

The determination of the dimension of the $u(t)$ space is a classical problem in statistical analysis corresponding to the principal component analysis. To address this question, $u(t)$ is considered at N time instants with $N > n$ and we gather these vectors in a matrix U. The number of reactions is then determined by counting the number of non zero singular values of UU^T (Bernard and Bastin, 2005*b*).

In practice, with experimental data, there are no zero eigenvalues for the matrix UU^T due to perturbations (*e.g.* measurement noise, numerical approximation of the derivative). But note that the singular values correspond to the variance associated with the corresponding eigenvectors (inertia axis)(Johnson and Wichern, 1992). The method consists thus in selecting the p first principal axes, which represent a cumulated variance larger than a fixed threshold (e.g. 90%).

2.3 Pseudo-stoichiometric matrix identification

Let ρ be the $n \times p$ matrix made of the p first eigenvectors of the $n \times n$ matrix UU^T. ρ is an orthonormal basis of $\mathcal{I}mK$. Therefore, there exists a $p \times p$ matrix G such that

$$K = \rho G$$

To identify G (and thus K), p^2 additional structural constraints from the *a priori* knowledge on