We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



12

Methodology for Bioprocess Analysis: Mass Balances, Yields and Stoichiometries

Farges Bérangère, Poughon Laurent,
Pons Agnès and Dussap Claude-Gilles
Clermont Université, Université Blaise Pascal,
Laboratoire de Génie Chimique et Biochimique,
Clermont-Ferrand
France

1. Introduction

The stoichiometry of a chemical reaction provides basic information about the nature and the quantities of chemical species consumed and produced. It also intrinsically contains all the information on transformation yields. Such information is useful and necessary for the design of any biotechnological process. In the case of microbiological reactions that support microbial growth, this information deals with the carbon and energy sources consumed, the terminal electron acceptor utilized, other metabolic products formed, as well as the quantity of the biomass produced.

In the first part, general mass balances principles/methodology for the stoichiometric analysis of a bioprocess will be presented. These methods will lead to stoichiometric coefficients estimation including statistical analysis and data reconciliation in case of redundant information. Redundant information means more information than the minimum required to calculate all the conversion yields with a suitable approach of mass balances; this concept of minimum information required is presented (degree of freedom or number of unknowns of the system); conversely over-determined systems, in the case of experimental data in excess, are associated to a statistical treatment.

In the second part, the previous stoichiometry principles will be applied to two practical examples. The first examined case is the continuous anaerobic cultures of *Fibrobacter succinogenes*. This strictly anaerobic bacterium was grown in continuous culture in a bioreactor at different dilution rates (0.02 to 0.092 h^{-1}) on a fully synthetic culture medium with glucose as carbon source. Robustness of the experimental information is checked by C and N balances estimations. A detailed overall stoichiometry analysis of the process for each dilution rate examined, including all substrates and products of the culture, is proposed. The mass balances involved in stoichiometric equations were solved using data reconciliation and linear algebra methods in order to take into account errors measurements.

In the last part, a second practical case,i.e. batch aerobic cultures of *Saccharomyces cerevisiae*, is presented. In this example, the bioprocess is analysed using different methodologies: (i) a

simple mass balance with a minimal set of experimental data, (ii) a mass balance with a large set of experimental data including error measurement and (iii) a mass balance with experimental data obtained with 3 repetitions of the batch culture.

2. Main principles and methods of biochemistry stoichiometry

The quantitative description of a microbial growth and product formation has maturated considerably during the past few decades due on one part to major improvements in the controlled technical equipments and on the other part to important theoretical progress thanks to the application of elemental mass balancing methods.

The mass balance of a fermentation has become more and more recognized as a valuable tool for analytical data validation, for detection of measurement errors and/or unnoticed products (Wang and Stephanopoulos, 1983) for estimation of variables for which no direct analytical methods are available (Humphrey, 1974), and for improving the accuracy and reliability of fermentation parameter estimation (Solomon *et al*., 1982, 1984). The mass balancing theory principles were developed in the late seventies by Minkevich, Erickson, Roels and others (Minkevich and Eroshin, 1973; Erickson *et al*., 1978; Minkevich, 1983; Roels, 1980, 1983). The theory uses the formalism of linear algebra to express the relationships between measurable (macroscopic) flows. Practical applications are mostly in the field of numerical procedures to solve a system for unknown flows or to calculate maximum likelihood estimators in case of over-determined systems. Thus, it is now recognized that biological reactions stoichiometry is mandatory for analysing biological processes.

2.1 Mass-balance equations

The universal principle (Lavoisier principle) that matter cannot be created or destroyed (unless there is a nuclear reaction) holds in biochemical systems. In any closed system the total mass of every element, C, N, O, H, P, etc, is constant over time.

Mass balance of a compound

As a first step in analyzing a biochemical reaction system, the system, its boundary and the surroundings must be defined. A physical entity, such as a cell or a bioreactor, is often defined as a system, and material and energy balance is performed on it. In other cases the balance is done on a set of reactions that is a representative subset of the reaction network in the cell. Material balance can be performed on either a compound (a chemical specie) or an element (Nielsen *et al*, 2003). For a chemical species ***i***, the material balance in a system can be written as:

Accumulation of *i* = rate of inputs of *i* – rate of outputs of *i* + net rate of production/consumption of *i* from all reaction.

or

$$\frac{d(i)}{dt} = E_{(i)} - S_{(i)} + \sum_{k=1}^{n} R_{k(i)} \tag{1}$$

with $R_{k(i)}$ being the rate of consumption/production of i in the kth reaction.

Elemental Mass balance

Typically, chemical reactions are written as a single or a system of stoichiometric equations into which elements should be conserved. For a given element, such as C, it can shift from one compound to another due to chemical reaction, but the total amount is conserved. For the system, it is written:

rate of accumulation of elements = rate of input of element - rate of output of element

or

$$\frac{d(element)}{dt} = E_{(element)} - S_{(element)} \tag{2}$$

As a consequence, at steady state, the net rate of accumulation being zero, the net production/consumption of all elements is zero in the system. Moreover, for each reaction, the net balance for all elements must also be equal to 0. This can be written for the different elements as a linear system of equations, a special set of equation being representative of a particular stoichiometric equation:

$$\sum_{i,1}^{n} \alpha_{k,i} [C]_i = 0 \tag{3}$$

$$\sum_{i,1}^{n} \alpha_{k,i} [H]_i = 0$$

where $\alpha_{k,i}$ is the relational coefficient of compound *i* in the reaction k and $[X]_i$ the composition in element [X] in the compound *i*.

Usually one reaction is represented by stoichiometric reaction. With respect of the elemental balance the general expression is:

$$\alpha_1 [C_a H_b O_c N_d]_1 + \alpha_2 [C_a H_b O_c N_d]_2 + ... \longrightarrow \alpha_5 [C_a H_b O_c N_d]_5 + ... \tag{4}$$

In the previous equation, the stoichiometric coefficients (α) have opposite sign for reactants and products. By convention the products are given positive stoichiometric coefficients, while substrates stoichiometric coefficients are negative.

Multiple stoichiometries

When multiple reactions occur in a system, an overall stoichiometric equation can be written as a linear combination of the stoichiometric equations. This is possible if, and only if, the ratios between the rates of all reactions remain at constant values whatever the external conditions. This means that if all the rates remain proportional, the total set of reactions can be replaced by one single stoichiometric equation. This resulting stoichiometric equation is established by summation of the different stoichiometric coefficients of substrates and products with suitable proportionality constants calculated from the ratios between the rates. When dealing with biochemical systems stoichiometric equations are used for describing reactions in biochemical pathways, as well as for depicting complex systems such as the conversion of nutrients into cells or organisms. Considering that biochemical

pathways are proceeding at constant ratios between the biochemical rates (due to metabolic regulation and enzymatic control), the formulation of traditional chemical reactions and biochemical reactions are virtually identical.

Conversely, when within a set of reactions the ratio between the reactions rates do not remain constant, the resulting stoichiometry may vary with the external conditions, *i.e.* the substrates/products concentrations, time, development phases...etc.. This leads to a variable stoichiometry, *i.e.* variable conversion yields. An instantaneous stoichiometry may be calculated as previously as a rates-weighted sum of the different stoichiometric equations leading to account for instantaneous conversion yields. But, in any case, the balance equations for the compounds must be written by using the reaction rates of all equations separately. This is easily understandable by considering the previous Equation 1, noticing that the instantaneous stoichiometry and the instantaneous overall reaction rate that are established from the set of elementary equations cannot be used directly without considering a second-order term for accounting of yields variations.

Pseudo- stoichiometry

When using stoichiometric equations for non-chemically defined compounds, the elemental composition of which being non-completely defined, such as macro-compounds (proteins, biomass), transaction will only consider C, H, O, N, and P in most cases. The other elements participate only in a small fraction of all biochemical reactions and only slightly contributeto the biomass. Therefore, stoichiometric equations can be written for reactions that occur inside a biological system, such as a cell whether they are enzyme, catalysed or not; stoichiometric equation are also used for characterising cell growth and compounds production inside a reactor. Whatever the case, the material balance involves input and output flows and the reaction rates in the system. In the case of a pseudo-stoichiometry it must be kept in mind that the elemental composition accuracy of macro-compounds is limited by experimental measurements errors (generally not exceeding 10^{-3} relative error), generating a systematic inaccuracy in balance equations.

2.2 Method for stoichiometric coefficients estimation and statistical analysis

2.2.1 Mathematical description of the equation system

Considering the following stoichiometric equation, involving n compounds and c elements:

$$\alpha_1 [CHON]_1 + \alpha_2 [CHON]_2 + ... \longrightarrow \alpha_5 [CHON]_3 + ... \quad (5)$$

The elemental balances imply a system of c equations (here c = 4: C, H, O and N) and n unknown stoichiometric coefficients α_i:

$$\text{C} \quad \sum_{i=1}^{n} \alpha_i [C]_i = 0$$

$$\text{H} \quad \sum_{i=1}^{n} \alpha_i [H]_i = 0 \quad (6)$$

$$\text{O} \quad \sum_{i=1}^{n} \alpha_i [O]_i = 0$$

$$\text{N} \quad \sum_{i=1}^{n} \alpha_i [N]_i = 0$$

In order to know if the system can be solved by linear algebra, it is necessary to analyse its **degree of freedom d** at first. Normally the previous equations are linearly independent. The kernel matrix dimension is n – c.

In all cases, one unknown coefficient (α) is fixed. This leads to:

$$d = n - c - 1 \tag{7}$$

In order to solve the system, it is necessary to provide d more independent information resulting from experimental yields determination and/or theoretical assumptions. Three cases are thus possible:

1. The minimum of experiments or theoretical independent information for solving the system is available; that is to say that d experiments or theoretical independent information are available; thus, the resolution of the system can be realised "by hand" thanks to a substitution method.
2. More information than d is available; the system must be solved with the use of a reconciliation method as well as the use of statistical results.
3. With not enough information, the system cannot be solved.

2.2.2 Mathematical method for data reconciliation

Let the resulting linear system of constraints be written as follows:

$$S\,\alpha = K \tag{8}$$

α (n) is the aforementioned column vector of stoichiometric coefficients (n rows, negative for the substrates, positive for the products). **S** (p, n) contains the elemental formula of all compounds (4 rows for 4 elements) and additional constraints (additional rows) such as the stoichiometric coefficient choice, that is fixed to 1, and eventually other constraints equations between the coefficients. Therefore, p is the total number of equations. **K** (p) is the known column vector of resulting constraints, for example 0 for corresponding conservation equations, 1 for the fixed stoichiometric coefficient and 0 for the linear constraints equation between the coefficients (Urrieta-Saltijeral *et al*, 2001).

If a detailed analysis of a process leads to consider a set of n compounds, an overall stoichiometric equation with (n-1) unknown coefficients is established, knowing that one coefficient is arbitrarily fixed to a value of 1. To determine these (n-1) stoichiometric coefficients, experimental data are needed, and it is necessary to keep in mind that calculations must be performed by meeting the constraint of Lavoisier principle of elements conservation.

Let us introduce experimental data in a column vector $\hat{\mathbf{Y}}_{\mathbf{exp}}$ (m) of m mass yields, all calculated with the same compound as reference, glucose for example. Let us consider **α** (n) the column vector of the stoichiometric coefficients, with the relevant value for the reference compound, again glucose for example, being set to 1. Finally, $\mathbf{Y_r}$ (m) is the column vector of the yields values obtained after data reconciliation. The following matricial expression is written:

$$A\,\alpha = Y_r \tag{9}$$

A (m, n) is the matrix enabling to build the mass yields values knowing the stoichiometric coefficients.

The element conservation balances are written once the elemental formula of all compounds are known. This is a constraint for the identification, as well as the need that one stoichiometric coefficient is fixed to 1.

The complete linear system for the stoichiometric coefficients calculation is obtained by concatenation of the previous expressions, leading to:

$$M\,\alpha = H \tag{10}$$

where **M** is a (m + p, n) matrix composed of matrix **A** and **S**, and **H** is a (m + p) column vector formed by concatenation of $\mathbf{Y_r}$ (m) and **K** (p). Similarly, $\hat{\mathbf{H}}_{\mathbf{exp}}$ (m+p) is the column vector formed by concatenation of $\hat{\mathbf{Y}}_{\mathbf{exp}}$ (m,) and **K** (p). A weighted diagonal matrix **W** (m + p, m +p) is also be filled with the inverse of experimental variances for the n experimental data and with 1 for the constraint relationships. **W** enables to account for the difference in experimental errors of the measurement yields. The system **M α = H** can be solved by direct inversion of the matrix **M**, if matrix rank and (m + p) are equal to α (determined system). In such case:

$$\alpha = M^{-1}\,H \tag{11}$$

In the case of redundant information (more independent rows than columns, *i.e.* (m + p) > α), a data reconciliation method must be used to solve the constrained and over-determined system. We propose to use a Lagrange method for solving this problem of optimisation, with the assumption that errors measurements (*i.e* variances) follow a Gaussian law (Himmelblau, 1970). The method is developed as follows.

The variable to be minimised is the least square estimate (Φ) given by:

$$\Phi = (\hat{H}_{exp} - H)^t\,(W)\,(\hat{H} - H) \tag{12}$$

i.e.

$$\Phi = (\hat{H}_{exp} - M\,\alpha)^t\,(W)\,(\hat{H}_{exp} - M\,\alpha) \tag{13}$$

Φ is a dimensionless number if **W** is used.

In this study the experimental values are known, and both the confidence interval (i_y) and the variance (σ_y) of the model are given:

$$\begin{aligned} y &= \hat{y} \pm i_y \\ \Delta_y &= 2\, i_y \\ \Delta_y &= 2\, t_{1-\alpha/2}\, \sigma_y \\ \sigma_y &= i_y / t_{1-\alpha/2} \end{aligned} \tag{14}$$

If the number of points for the estimation is sufficiently large, $t_{0,975} \approx 2$ ($\alpha = 0.05$), and the interval is a confidence interval of 95%. The weights for all measured yields are given by:

$$W_{yj} = \frac{1}{\sigma_{y_j}^2} = \frac{4}{i_{y_j}^2} \tag{15}$$

Estimation of the vector α components can be directly obtained through matrix calculus from the previous equation. However, a direct inversion does not guarantee that the obtained set of stoichiometric coefficients (vector α) will obey the Lavoisier principle (elements conservation). The originality of the proposed method is to add as constraints the first element conservation equations. The constraints to fulfil are given as previously by:

$$S\,\alpha = K \tag{16}$$

Calculation of components of vector α is therefore obtained by the classical method of Lagrange multipliers. The Lagrangian function to be minimized is:

$$L = \Phi - \Lambda\,(S\,\alpha - K) \tag{17}$$

where $\boldsymbol{\Lambda}$ is the unknown row vector (p) of the Lagrange multipliers.

The p Lagrange multipliers and the n stoichiometric coefficients are obtained by solving the (p + n) system of equations formed by:

$$\begin{aligned} &\frac{\partial L}{\partial \alpha} = 0 \quad (\alpha \text{ equations}) \\ &S\,\alpha = K \quad (\text{p equations}) \end{aligned} \tag{18}$$

$$\frac{\partial L}{\partial \alpha} = 0$$

is a matrix derivation which leads to the n following equations :

$$\frac{\partial L}{\partial \alpha} = -2\, M^t\, W\, (\hat{H}_{exp} - M\,\alpha) - S^t\, \Lambda^t = 0 \tag{19}$$

This leads to:

$$2\,(M^t\, W\, M)\,\alpha = 2\, M^t\, W\, \hat{H}_{exp} + S^t\, \Lambda^t \tag{20}$$

then:

$$\alpha = (M^t\ W\ M)^{-1}\ M^t\ W\ \hat{H}_{exp} + 1/2\ (M^t\ W\ M)^{-1}\ S^t\ \Lambda^t \tag{21}$$

and:

$$S\ (M^t\ W\ M)^{-1}\ M^t\ W\ \hat{H}_{exp} + 1/2\ S\ (M^t W\ M)^{-1}\ S^t\ \Lambda^t = K \tag{22}$$

considering that:

$$\Lambda^t = 2\ [S\ (M^t\ W\ M)^{-1}\ S^t]^{-1}\ [K - S\ (M^t W\ M)^{-1}\ M^t\ W\ \hat{H}_{exp}] \tag{23}$$

if we denote $\psi = M^t\ W\ M$, **α** is given by:

$$\alpha = \psi^{-1}\ M^t\ W\ \hat{H}_{exp} + \psi^{-1}\ S^t\ [S\ \psi^{-1}\ S^t]^{-1}\ [K - S\ \psi^{-1}\ M^t\ W\ \hat{H}_{exp}] \tag{24}$$

Summarizing and combining computed matrices, the calculation leads to:

$$\psi = M^t\ W\ M \tag{25}$$

$$\Omega = M^t\ W\ \hat{H}_{exp} \tag{26}$$

The solution is given by:

$$\alpha = \psi^{-1}\ [\Omega + S^t\ (S\ \psi^{-1}\ St)^{-1}\ (K - S\ \psi^{-1}\ \Omega)] \tag{27}$$

The standard deviations of the calculated coefficients are estimated by diagonal elements of matrices of covariances:

$$Covar\ (\hat{\alpha}) = (M^t\ W\ M)^{-1} = \Psi^{-1} \tag{28}$$

(if W is known from variances)

$$Covar\ (\hat{\alpha}) = (M^t\ M)\ \frac{\Phi}{n + p - \alpha} \tag{29}$$

(if W is not known)

Covariance on estimates of the model is:

$$Covar(\hat{H}_{exp}) = M\ Covar\ (\hat{\alpha})\ M^t \tag{30}$$

This set matrix equations determines the stoichiometric coefficients (vector α), the covariance vector of the estimated coefficients and the covariance of the model prediction. The constraint of elements conservation is intrinsically fulfilled, which is a prerequisite of any robust modeling including mass balance.

3. First application: Stoichiometric analysis of *Fibrobacter succinogenes* growth

3.1 Introduction

Fibrobacter succinogenes is one of the main fibrolytic bacteria in the bovine rumen (Hungate, 1950). It's a strictly anaerobic bacterium with enzymatic equipment well adapted to the degradation of vegetable fibers and plants, especially when these are highly branches and lignified. The degradation steps lead to the production of cellobiose and glucose that are further metabolized by the bacterium. The fermentative metabolism of this bacterium has been studied and leads to the production of succinate, acetate and formate. *Fibrobacter succinogenes* is also able to store intracellular glycogen, even in cells of young cultures (Gaudet *et al*, 1992) and it can produce and release oligosaccharides (Nouaille *et al*, 2005).

However, there is little information on the global stoichiometric description of this metabolism. Such quantitative information is necessary for further understanding the growth of *Fibrobacter succinogenes,* for example through a global stoichiometric approach prior to metabolic flux modelling.

The aim of this work was to establish the overall stoichiometry of the*Fibrobacter succinogenes* S85 growth, cultivated in a standardized continuous anaerobic culture process on a fully synthetic culture medium with glucose as carbon source for different dilution rates (Guiavarch *et al*, 2010). Linear algebra and data reconciliation methods previously developed were applied to solve the overdetermined system obtained from the large number of collected experimental data.

3.2 Culture conditions

The strain used was *Fibrobacter succinogenes* S85 (ATCC 19169), and was grown anaerobically under 100% CO_2 in a synthetic medium with glucose as carbon source. The reactor was a B.BRAUN culture unit (Biostat ED, B.BRAUN Germany). The working volume was 5 L and the stirring speed 100 rpm. Temperature was controlled at 39°C and the pH was maintained at 6.3 by automatic addition of Na_2CO_3 (70 $g.L^{-1}$). The culture vessel was fed with fresh medium completed with various glucose concentrations (from 8.2 to 19.1 $g.L^{-1}$) at three volumetric flow rates (99, 255 and 464 $mL.h^{-1}$) corresponding to three dilution rates (D = 0.02, 0.051 and 0.092 h^{-1}). Culture vessel and all tanks were interconnected by a gas system and the pressure was maintained at 0.2 bars above atmospheric pressure. The whole gas system was continuously flushed with 5 sccm of sterile oxygen-free CO_2 during the culture to preserve anaerobic conditions. This flow rate was controlled using a mass flow controller (0-5 sccm, Tylan), while the gas flow at the exit of the reactor was measured with a mass flow meter (0-20 sccm, Brooks).

Samples were taken at regular time intervals during the experiment. Microscopic observations showed that the culture was always axenic. For each sample, the absorbance was measured at 600 nm. HPLC apparatus was used to determine glucose and organic acids concentrations (Agilent 1100 series fitted with two Phenomenex Rezex ROA columns, 7.8 mm diameter and 300 mm length).

Culture supernatants were obtained after centrifugation of an aliquot (10 000*g*, 5 min), and used to perform the colorimetric assays of ammonium ions, proteins and total

carbohydrates. Soluble carbohydrates production was calculated by difference between total carbohydrates and glucose concentrations.

The pellets resulting from centrifugation step were dried in an oven at 100°C for 24 h to obtain biomass dry weight. A typical correlation OD-cell dry weight was established from data collected during the exponential growth phase of a batch experiment carried out in the same culture conditions as:

$$\text{cell dry weight} = 0.482\ (\pm\ 0.034)\ OD_{600\ nm} \tag{31}$$

where the cell dry weight is expressed in $g.L^{-1}$ (correlation coefficient 0.989).

At steady-state, one sample was centrifuged (10 000*g*, 15 min, 5°C), washed with 0.9% NaCl and dried under vacuum at 65°C (48 h) to determine an average biomass formula (CHONSP) by elemental analysis.

Gas at the exit of the bubble column was analyzed by gas chromatography (Hewlett Packard 5890 series II, fitted with a Thermal Conductivity Detector). Two 1.5 m length, 1/8" diameter stainless steel columns (Porapak Q and 5 Å molecular sieves) connected with a 6-port commutation valve were used.

3.3 Experimental data: Biomass formulae and element recoveries

The average molar biomass formulae established after elemental analysis during steady-state are presented in Table 1.

Dilution rate D (h^{-1})	Average biomass formula				
	C	H	O	N	S
0.020	3.688	6.760	2.665	0.247	0.010
0.051	3.656	6.821	2.791	0.255	0.009
0.092	3.799	6.981	2.545	0.380	0.008

Table 1. Average biomass formulae for different dilution rates.

The biomass formula changed weakly with the dilution rate, the most important modification concerning the nitrogen mass fraction that increases with the dilution rate. This modification could be explained by the variation in glycogen to protein ratio that has been evidenced when the growth rate is increased.

To calculate C-recovery, consumptions of sodium carbonate and glucose were both taken into account as carbon sources as well as production of biomass, soluble proteins, succinate, acetate, formate, and soluble carbohydrates. As reported in Table 2, C-balance was between 97 and 101%. Data considered in N-recovery were the nitrogen source consumption (ion ammonium), the nitrogen content in cell dry weight and the soluble proteins measured in the culture supernatant. At low dilution rate (0.02 h^{-1}), N-balance was satisfactory with a value of 98% but at high dilution rates (0.051 and 0.092h^{-1}), N-balances were low with respectively 62 and 82%.

N-recovery was dilution rate dependent and thus growth rate dependent. This dependence has already been shown in *Fibrobacter succinogenes* by Wells and Russell (1996). Growing cultures of *Fibrobacter succinogenes* were reported to assimilate more ammonia than could be

accounted for cellular protein, RNA, or DNA, and released large amounts of non ammonia nitrogen that were not identified and quantified.

Dilution rate D (h^{-1})	C recovery (%)	N recovery (%)	Redox potential (mV)	Dissolved CO_2 partial pressure (mbar)
0.02	101	98	-358	1167
0.051	98	62	-347	1191
0.092	97	82	-345	1193

Table 2. C and N element recoveries, redox potential and dissolved CO_2 partial pressure for different dilution rates.

All these results showed a rather good consistency of the experimental results obtained during the course of the culture, indicating that the data reconciliation technique can be valuably applied.

3.4 Stoichiometric equation analysis and data reconciliation

The following detailed stoichiometric equation was proposed to describe the culture (Guiavarch *et al*, 2008). The chemical dissociation imposed by the pH value (6.3) was taken into account in the elementary formulae of the organic acids. Nevertheless, it could be assumed that, in the range of pH supported by this bacterium (from 6.0 to 7.0), the relevant formulae remained rather identical. The stoichiometric equation included 13 compounds and therefore 12 stoichiometric coefficients had to be determined, the coefficient for glucose being, as already pointed out, set to 1:

$$\begin{aligned} C_6H_{12}O_6 + \alpha_1\ (NH_4)_2SO_4 + \alpha_2\ Na_2CO_3 \rightarrow\ & \alpha_3\ C_{3.69}H_{6.76}O_{2.66}N_{0.25}S_{0.010}\ \text{(biomass, see table 1)} \\ & + \alpha_4\ C_4H_{4.1864}O_4Na_{1.8136}\ \text{(sodium succinate)} \\ & + \alpha_5\ C_2H_3O_2Na\ \text{(sodium acetate)} \\ & + \alpha_6\ CHO_2Na\ \text{(sodium formate)} \\ & + \alpha_7\ C_6H_{10}O_5\ \text{(carbohydrates)} \\ & + \alpha_8\ C_{4.43}H_{7.09}O_{1.29}N_{1.27}S_{0.042}\ \text{(proteins)} \\ & + \alpha_9\ H_2O + \alpha_{10}\ NaHCO_3 \\ & + \alpha_{11}\ CO_2 + \alpha_{12}\ Na_2SO_4 \end{aligned} \quad (32)$$

At least 12 theoretical and/or experimental data were necessary. On-line and off-line parameters measured during the culture provided nine experimental mass yields ($\hat{Y}_{exp}$) related to ammonium sulfate, sodium carbonate, biomass, succinate, acetate, formate, carbohydrates, carbon dioxide and protein, each value being weighted by a standard deviation associated to the measured value (Table 3). For these data, a A (9, 13) matrix and $\hat{Y}_{exp}$ (9) column vector could be built with molar mass of compounds and nine experimental mass yields. Elemental balances on C, H, O, N, S and Na provided 6 linear equations of constraints. Coefficient of glucose was here again fixed to 1, allowing to build a S (7, 13) matrix filled with chemical formulae of compounds and coefficient of glucose, and a K (7) column vector filled with 0 for conservation equations and 1 for fixed stoichiometric coefficient.

The resulting system of 16 linear equations was made of 9 relationships obtained from experimental measurements and 7 constraints relationships resulting from elemental balances (C, H, O, N, S, Na) and glucose coefficient fixed to 1. It was over-determined since there were only 12 unknown coefficients to calculate. The advantage of data reconciliation was to allow the use of all available information to reduce inaccurate data due to experimental errors. Reconciled molar yields (Y_r) were thus estimated from the calculated stoichiometric coefficients.

At first , this linear system composed of 9 weighted relations from experimental measurements and 7 constraints from elemental balances was used to reconcile molar yields obtained at three different dilution rates (Table 3).

Dilution rate D (h^{-1})	Substrat and Product	$\hat{Y}_{exp\ mass}$	$Y_{exp\ molar}$	Standard deviation of $Y_{exp\ molar}$	Y_r	Standard deviation of Y_r	Confidence interval after data reconciliation	Y_r/Y_{exp}
0.02	$(NH_4)_2SO_4$	-0.044	-0.060	0.008	-0.071	0.004	0.008	1.19
	Na_2CO_3	-0.603	-1.024	0.038	-1.025	0.019	0.038	1.00
	Biomass	0.207	0.387	0.047	0.343	0.024	0.048	0.89
	Succinate	0.553	0.630	0.030	0.630	0.015	0.030	1.00
	Acetate	0.147	0.322	0.028	0.322	0.014	0.028	1.00
	Formate	0.004	0.108	0.012	0.108	0.006	0.012	1.00
	Carbohydrates	0.237	0.235	0.139	0.260	0.067	0.134	1.11
	CO_2	0.000	0.000	0.939	0.389	0.423	0.846	-
	Proteins	0.026	0.048	0.004	0.047	0.002	0.004	0.97
0.051	$(NH_4)_2SO_4$	-0.140	-0.191	0.017	-0.124	0.009	0.018	0.65
	Na_2CO_3	-0.553	-0.939	0.048	-0.940	0.024	0.048	1.00
	Biomass	0.448	0.814	0.041	0.844	0.020	0.040	1.04
	Succinate	0.502	0.572	0.017	0.569	0.009	0.018	0.99
	Acetate	0.141	0.310	0.021	0.308	0.010	0.020	0.99
	Formate	0.055	0.147	0.014	0.146	0.007	0.014	1.00
	Carbohydrates	0.157	0.173	0.131	0.000	0.063	0.126	-
	CO_2	0.007	0.027	1.297	0.537	0.541	1.082	20.0
	Proteins	0.014	0.025	0.003	0.026	0.001	0.002	1.03
0.092	$(NH_4)_2SO_4$	-0.165	-0.225	0.018	-0.159	0.009	0.018	0.71
	Na_2CO_3	-0.589	-0.999	0.027	-1.000	0.024	0.048	1.00
	Biomass	0.496	0.906	0.038	0.751	0.020	0.040	0.83
	Succinate	0.562	0.641	0.018	0.606	0.009	0.018	0.95
	Acetate	0.168	0.369	0.036	0.340	0.010	0.020	0.92
	Formate	0.058	0.152	0.005	0.149	0.007	0.014	0.98
	Carbohydrates	0.038	0.042	0.347	0.000	0.063	0.126	-
	CO_2	0.055	0.227	1.144	0.670	0.494	0.988	2.95
	Proteins	0.014	0.026	0.005	0.025	0.003	0.006	0.98

Table 3. Experimental mass yields values ($\hat{Y}_{exp\ mass}$) (g substrate or product. (g glucose)$^{-1}$) and comparative values of experimental ($Y_{exp\ molar}$) and reconciled (Y_r) average molar yields (mol substrate or product. (mol glucose)$^{-1}$) with the associated variances for different dilution rates.

At a dilution rate of 0.02h^{-1}, reconciled molar yields of acid production and sodium carbonate consumption were equal to experimental yields (Table 3). Biomass and protein yields were slightly decreased which led to an increase of ammonium sulfate yield from -0.060 (σ = 0.008) to -0.071 (σ = 0.004). These variations were of the same order of magnitude than the standard deviation. There was also a slight increase in soluble carbohydrates yield from 0.234 (σ = 0.139) to 0.260 (σ = 0.067) at this dilution rate (0.02 h^{-1}) although the standard deviation associated to this value was high. So, this linear system was sufficient to obtain satisfactory results at a dilution rate of 0.02 h^{-1}.

At the other dilution rates of 0.051 and 0.092 h^{-1}, this linear system did not give satisfactory results for the soluble carbohydrates yield since a negative value was calculated. This result would lead to consider that, under these conditions, soluble carbohydrates were a substrate. This idea could not be considered as realistic since glucose was the sole carbon source in the fresh medium. It was thus necessary to modify the linear system by adding a new constraint on soluble carbohydrates for dilution rates of 0.051 and 0.092 h^{-1}. This supplementary constraint was to fix the coefficient of soluble carbohydrates to zero.

The new system also resulted in a new over-determined linear system of equations, then made of 8 relationships from experimental measurements, 6 constrained equations from elemental balances (C, H, O, N, S, Na) and 2 reference coefficients (glucose set to 1 and carbohydrate set to 0). **A**(8, 13) was the matrix of known constant coefficients and $\hat{\mathbf{Y}}_{\mathbf{exp}}$(8) was the column vector of experimental yields. These data allowed to build a matrix **S**(8, 13) filled with chemical formulae of compounds, and a **K**(8) column vector.

3.5 Discussion

The average reconciled yields and the relevant variances were calculated from this second linear system using data reconciliation. Experimental and reconciled values were compared in Table 3.

This modified linear system gave rather satisfactory results for dilution rates of 0.051 and 0.092 h^{-1} with very close experimental and identified yield values. Particularly Y_r/Y_{exp} ratios were often close to 1 except for carbohydrates that had been set to zero by the supplementary constraint. The same major reconciliations were observed on carbon dioxide and ammonium sulfate yields. The discrepancy between Y_r and Y_{exp} on ammonium sulfate was explained by N-balances that were not satisfactorily assessed at these dilution rates. Y_r took into account only ammonium sulfate used for cellular growth and protein production. During continuous culture, no significant carbon dioxide gas production or consumption had been measured and there was an important standard deviation on these measurements. Carbon dioxide gas production or consumption was calculated by difference between inlet and outlet gas. The whole gas system was flushed with a regular gas flow of 5 sccm of carbon dioxide minimum to preserve correct anaerobic conditions during the continuous culture. However, effluent and medium tank volumes were about 10 times higher than the reactor volume and were not regulated in temperature. Therefore, carbon dioxide solubility was permanently modified by ambient temperature variations that consequently led to unreliable carbon dioxide flow rate at the exit of the culture vessel. As carbon dioxide yield, carbohydrates yield was obtained indirectly by difference between total carbohydrates and glucose concentrations.

These results showed that the stoichiometric equation was dilution rate dependent. When the dilution rate increased from 0.020 to 0.092h^{-1}, biomass and ammonium sulfate yields significantly increased as well. The biomass yield was improved from 0.343 (σ = 0.024) to 0.751 (σ = 0.020) mol biomass (mol glucose)$^{-1}$ (from 0.183 to 0.412 g biomass (g glucose)$^{-1}$). An important result was to notice that no significant variations were observed on succinate, acetate, formate and sodium carbonate yields. In the range of the studied dilution rates, the rates of succinate, acetate, and formate production were proportional to the rate of glucose consumed into the system. This tended to demonstrate that the metabolism of *Fibrobacter succinogenes* was not limited by the production of these acids which were directly linked to energy metabolism.

These results showed that assays used to track products of fermentation, consumption of nitrogen and carbon source were efficient as well as analysis of the biomass. This information was reliable to establish a stoichiometric equation for each dilution rate. It should also be pointed out that soluble carbohydrates production should be measured using a more specific technique. This information must be accurately available to go further in the analysis of *Fibrobacter succinogenes* growth by the use of a metabolic flux model.

4. Application 2: Stoichiometry for an aerated batch culture of *Saccharomyces cerevisiae*

In this second example of application, stoichiometric equations were established using experimental data obtained for the growth of a strain of *Saccharomyces cerevisiae* in an aerated and controlled bioreactor. The method of data reconciliation to establish the stoichiometry is applied here considering 3 approaches for exploiting the experimental results:

1. analysis of a single batch experiment with the minimal experimental data required for the stoichiometry (i.e. simple mass balance approach)
2. analysis of a single batch experiment and use of all experimental data for establishing the stoichiometry, including the experimental uncertainty (i.e. variance) estimation of the experimental results.
3. analysis of several batch experiments (repeatability of experiments) for establishing the stoichiometry, including the statistical analysis of the repetition.

4.1 Batch culture of *Saccharomyces cerevisiae*

4.1.1 Culture conditions, monitoring and analysis

A strain of *Saccharomyces cerevisiae* ATCC 7754 is grown in a controlled 6 liters bioreactor (Biostat A. B-Braun, Germany). *Sc. cerevisiae* is a facultative anaerobic microorganism and can metabolise glucose into ethanol (fermentative metabolism) or/and into CO_2 (oxidative metabolism). It is also a glucose sensitive microorganism, as for glucose concentration above 0.1 g.L^{-1}, the "Crabtree effect" can be observed. The Crabtree effect reflects the respiratory chain saturation, which is the main path for regenerating reduced co-factors, and thus even in an aerated system the alternate route to ethanol is used to regenerating the co-factors.

The bioreactor is operated during 8 hours in liquid batch conditions (4.4 L liquid volume), air flow rate (1.5 NL.min^{-1}) and perfectly mixed (6 blades stirrer at 500rpm). The dissolved

oxygen is monitored (Ingold probe 34-100-3003), and temperature and pH are controlled (respectively at 30°C and pH 5). Bioreactor data (pH, PO2, temperature, stirring rate) are acquired by the digital control unit of the reactor (Micro DCU 300 – B.BRAUN, Germany) as well as the online analysis of the gas output O_2 and CO_2 molar fractions ($y_{O_2}^{out}$ $y_{CO_2}^{out}$) which are measured by an Oxymat 5E (Siemens) and an infrared CO_2 analyser (Schlumberger), respectively.

The growth medium is taken from Kristiansen (1994) (glucose 50 g.L^{-1}, yeast extract 6 g.L^{-1}, KH_2PO_4 5 g.L^{-1}, $(NH_4)_2SO^4$ 2 g.L^{-1}). Organic acids and ethanol products are measured by HPLC (ionic exclusion column Shodex SH1011, 300 x 7.8 mm). Glucose is measured by the 3,5 dinitrosalicylic acid (DNS) method of Summer and Howell (1935). Amino Acids concentrations (yeast extract content) are measured as a leucine equivalent by a colorimetric method after reaction with ninhydrin (Ruhemann's purple read at 570 nm). Ammonium ions (NH_4^+) are measured by the colorimetric method of Patton and Crouch (1977). The biomass growth is followed by optical density at 550 nm and dry mass measurements.

4.1.2 Analysis of a batch culture: Results and theoretical stoichiometry

The main substrates and products identified and expected during the growth of *Sc. cerevisiae* are reported in table 4. It can be noticed that the previously detailed analyses allows quantifying all these compounds during the batch culture. The culture is performed during 8 hours with an initial glucose concentration of about 50 g.L^{-1}, which leads to a growth with a glucose saturated metabolism of *Sc. cerevisiae* (Crabtree effect), *i.e.* a mixed oxidative/fermentative metabolism with ethanol production. It is also considered that the produced metabolites are primary metabolites, which in turns means that all metabolic rates, including biomass synthesis and metabolites production, remain proportional. This justifies that a single stoichiometry approach is applied.

Theoretically, the mass balance on the culture can be expressed with the following stoichiometric equation:

$$\alpha_1\, glucose + \alpha_2\, O_2 + \alpha_3\, NH_3 + \alpha_4\, Amino.Acid$$
$$\rightarrow \alpha_5\, biomass + \alpha_6\, CO_2 + \alpha_7\, ethanol + \alpha_8\, glycerol + \alpha_9\, acetoin + \alpha_{10}\, H_2O \quad (33)$$

Considering the 4 elements (C H O N) balance and that one of the stoichiometric coefficients is arbitrary fixed to 1, the degree of freedom d = n - c - 1 = 10 – 4 - 1 is equal to 5. This means that a minimal set of 5 independent additional information are required to calculate the 10 stoichiometric coefficients. We will obtain these data from the experimental yields calculated from the various experimental data acquired.

All compounds, except water, can be experimentally measured during the culture. An example of the results obtained for the experiment 1 (Exp 1) is reported in table 5 and figure 1. CO_2 production and O_2 consumption have been computed from the integration of the instantaneous CO_2 and O_2 respiration rates (r_{CO2}, r_{O2}) acquired from the online gas balance measurements. The respiratory quotient (RQ = r_{CO2} / r_{O2}) is also computed and its value above 1 is an indicator of the growth fermentative metabolism.

Compound	Formula (CHON)	Molar mass ($g.mol^{-1}$)	% Carbon content	% N content
Biomasse	$C_6H_{1.62}O_{0.52}N_{0.15}P_{0.01}$	24.35	49.28	8.6
Glucose	$C_6H_{12}O_6$	180	40	0
AA	$CH_{2.24}O_{0.48}N_{0.24}$	25.28	47.47	13.3
N-NH3	N	14	0	100
Ethanol	C_2H_6O	46	52.17	0
Glycerol	$C_3H_8O_3$	92	39.13	0
Acetoine	$C_4H_8O_2$	88	54.55	0
CO2	CO_2	44	27.27	0
O2	O_2	32	0	0

Table 4. Compounds involved in the batch culture of *Sc. cerevisiae* and their characteristics. The biomass composition was obtained from elemental analysis of dry sample. The amino acids composition was obtained from average content of amino acids in yeast extract.

Time (h)	Biomass ($g.L^{-1}$)	Glucose ($g.L^{-1}$)	Amino. Ac.($g.L^{-1}$)	N-NH_3 ($g.L^{-1}$)	Ethanol ($g.L^{-1}$)	Glycerol ($g.L^{-1}$)	Acetoin ($g.L^{-1}$)	CO_2 ($g.L^{-1}$)	O_2 ($g.L^{-1}$)	RQ
0.1	0.82	42.88	1.808	0.4226	0.854	0	0.16	0.294	0.049	-
1.1	0.847	42.59	1.936	0.426	0.869	0.045	0.167	0.908	0.118	5.4
2.02	1.033	43.90	1.816	0.4111	1.352	0.06	0.224	2.011	0.212	7.1
3.01	1.486	41.71	1.599	0.394	1.886	0.088	0.242	2.855	0.267	10.2
3.5	1.925	37.65	1.513	0.3931	2.37	0.097	0.269	3.930	0.330	11.7
4	2.235	37.21	1.332	0.3652	3.498	0.115	0.367	5.207	0.404	12.7
4.5	2.61	34.12	1.264	0.3462	2.92	0.11	0.303	6.748	0.498	12.4
5	3.13	31.91	1.095	0.3102	5.015	0.204	0.41	8.588	0.617	11.5
5.51	3.875	27.78	0.544	0.2571	7.46	0.288	0.467	10.766	0.758	11.1
6	4.425	25.93	0.508	0.2129	7.814	0.305	0.426	13.361	0.916	11.4
6.5	5.515	18.90	0.541	0.1363	10.91	0.41	0.491	16.395	1.091	12.2
7.01	6.205	12.85	0.3131	0.0607	11.79	0.51	0.3	0.294	0.049	12.8

Table 5. Results obtained for the growth of *Sc. cerevisiae* - Experiment 1.

A first treatment of the results obtained for the batch culture consists in the calculation of the experimental yields. Considering all products as primary metabolites, the ratios between all production or consumption rates are constant (*i.e* constant yields). Consequently the yield between a compound A and a compound B ($Y_{A/B}$) is calculated by a linear regression (figure 2) using the concentrations obtained in batch:

$$A(t)= \hat{Y}_{A/B}\, B(t) + cst\ ,$$

$$A(t)= \hat{Y}_{A/B}\, B(t), \text{ (when initial values are 0: } A(0)=B(0)=0). \quad (34)$$

The regression line slope is the yield, and calculation of the estimation variance gives the standard deviation. A complete statistical analysis leads to examine the reliability of the linearity assumption between the concentrations of A and B (constant yield and stoichiometry).

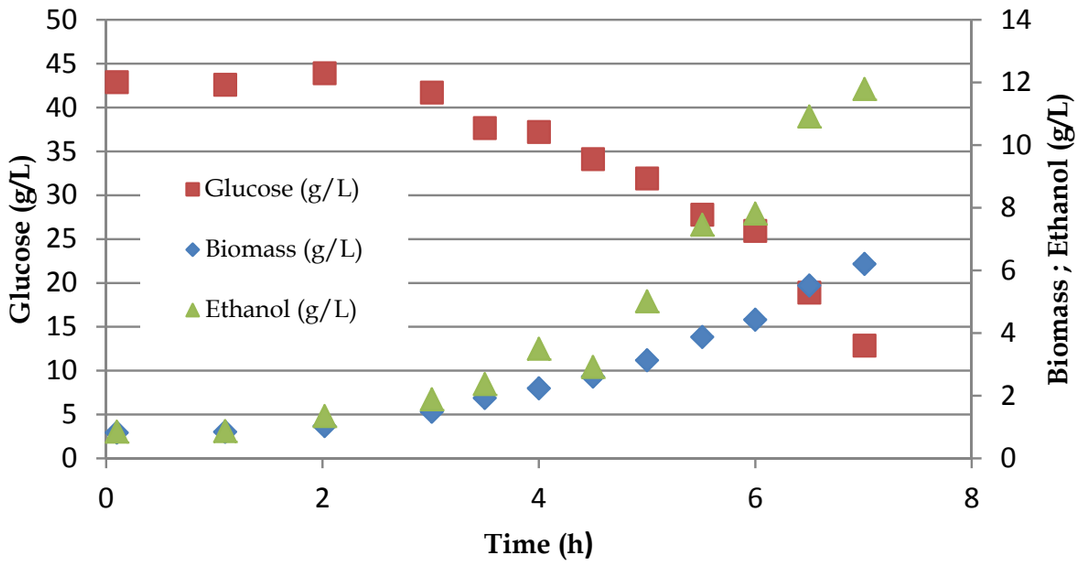


Fig. 1. Glucose consumption, biomass and ethanol production measured for the growth of *Sc. cerevisiae* - Experiment 1.

Some yields calculated for the experiment 1 are reported in table 6 with their standard deviation and correlation coefficient. It is possible to take the minimum of 5 relations required to solve the stoichiometric equation from these results. Besides, the C and N balance can be evaluated, being 0.91 and 0.78 (using $Y_{X/?}$ yields), respectively. These experimental balances may be an indicator of the success for solving the stoichiometric equation, as the theory implies that balances are equal to 1.

	Mass yield ($g.g^{-1}$)	Standard deviation	correlation coefficient	
$Y_{x/glu}$	-0.1806	+/- 3.4%	0.988	
$Y_{AA/glu}$	0.0374	+/- 5.8%	0.965	*
$Y_{N\text{-}NH3/glu}$	0.0096	+/- 2.5%	0.993	*
$Y_{eth/glu}$	-0.3798	+/- 5.6%	0.970	
$Y_{gly/glu}$	-0.0156	+/- 5.7%	0.969	
$Y_{acet/glu}$	-0.0076	+/- 34.2%	0.461	
$Y_{CO2/glu}$	-0.5245	+/- 3.4%	0.990	
$Y_{O2/glu}$	-0.0332	+/- 4.2%	0.984	
$Y_{x/AA}$	-3.0418	+/- 8.1%	0.939	
$Y_{x/N\text{-}NH3}$	-14.9807	+/- 4.8%	0.977	
$Y_{x/Eth}$	0.5498	+/- 4.2%	0.981	*
$Y_{x/gly}$	13.6788	+/- 5.1%	0.973	*
$Y_{x/acet}$	9.2216	+/- 11.8%	0.866	*
$Y_{x/CO2}$	0.4193	+/- 4.8%	0.977	*
$Y_{x/O2}$	6.0254	+/- 2.5%	0.994	*

Table 6. Yields calculated for Exp. 1. * means that a linear regression without intercept was used.

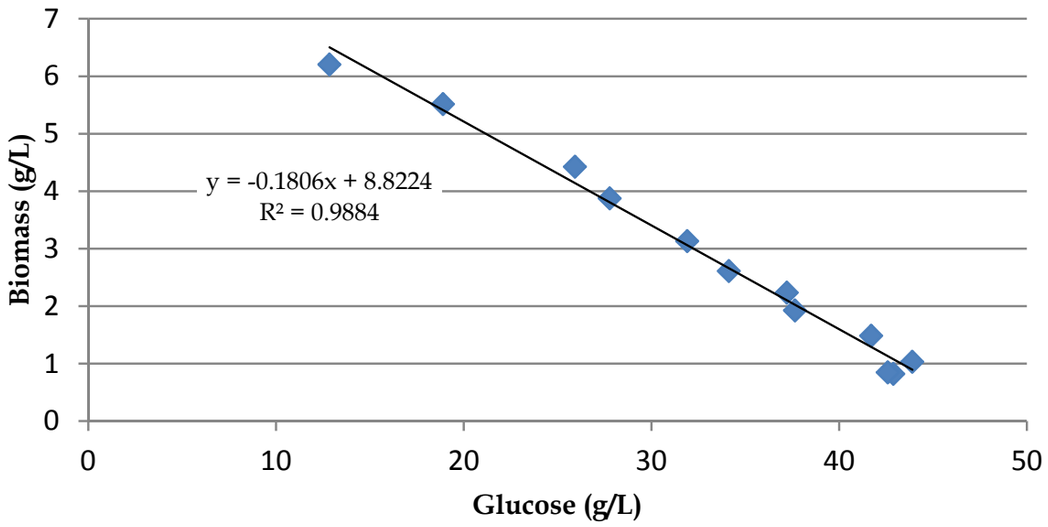


Fig. 2. Graphical representation of the linear regression Biomass(t)= $\hat{Y}_{X/Glucose}$ Glucose(t) + constant.

4.2 Stoichiometry for a minimal set of experimental data (simple elemental balance approach)

If it is chosen to solve the system using the minimal set of 5 experimental data (d=5) and without using reconciliation methods, it is crucial to use strictly independent relations. In the case of the Exp. 1, for instance, it is impossible to select simultaneously the biomass coefficient to 1 and to use the yields $Y_{x/N\text{-}NH3}$ and $Y_{x/AA}$. The 3 information are linked through the N balance (biomass, Amino Acid and NH_3 are the 3 compounds involved in the N-balance) in such a way that only two independent information are really introduced.

The most simple way to solve the stoichiometric equation with the list of yields presented in table 6 is to choose fixing the biomass coefficient $\alpha_5$to 1, considering 4 independent results within the list of the yield $Y_{x/?}$.

Taking for example $Y_{x/Eth}$, we can write :

$$Y_{X/Eth} = \frac{Mass\ Biomass}{Mass\ Ethanol} = \frac{\alpha_5}{\alpha_7}.\frac{Molar\ mass\ biomass}{Molar\ mass\ Ethanol}$$

$$\alpha_7 = \frac{\alpha_5}{Y_{X/Eth}}.\frac{Molar\ mass\ biomass}{Molar\ mass\ Ethanol} = \frac{1}{Y_{X/Eth}}\frac{24.35}{46} = 0.9628 \quad (35)$$

Thus, the ethanol stoichiometric coefficient is experimentally fixed and can be used as a relation to set up one of the system degree of freedom..

In the same way, taking the 3 other yields $Y_{x/AA}$, $Y_{x/gly}$ and $Y_{x/O2}$, the following stoichiometry can be established :

$$0.7492\ glucose + 0.1263\ O_2 + 0.074\ NH_3 + 0.3167\ Amino.Acid \rightarrow 1\ biomass + \\ +1.1305\ CO_2 + 0.9228\ ethanol + 0.0193\ glycerol + 0.1745\ acetoin + 0.7491\ H_2O \quad (36)$$

By taking other experimental data, a different stoichiometric equation is obtained. For example using the experimental yields $Y_{x/CO2}$ instead of $Y_{x/Eth}$ (it can be noticed in table 6 that the two yields were calculated with quite the same accuracy), the stoichiometric equation becomes:

$$0.7492\ glucose + 0.1263\ O_2 + 0.074\ NH_3 + 0.3167\ Amino.Acid \rightarrow 1\ biomass + \\ +1.1305\ CO_2 + 1.1928\ ethanol + 0.0193\ glycerol + 0.0121\ acetoin + 0.4651\ H_2O \quad (37)$$

In this last expression of the stoichiometry, the coefficient for ethanol is 20% higher and the one of acetoin is hundred times lower than in the previous one. This illustrates the influence of the experimental date choice used to establish the stoichiometric equation, as with the minimal set of information necessary to solve the system, all the weight of the constraint for the C H O N conservation remaining on the coefficients that are not given experimentally.

4.3 Stoichiometry using the full set of experimental data

It was previously outlined that even if using only the required experimental data to fulfil the degree of freedom is the easiest methodology, this approach is very sensitive to the experimental data choice and to their accuracy. A way to solve this problem is to use all the available experimental information, with an accuracy evaluation of these data. Such an approach requires the reconciliation method presented in 2.2.2.

With the 8 experimental yields Yx/? of table 6 (and their standard deviation), and by fixing the biomass coefficient $\alpha_5$to 1, a system of 13 equations is built for computing the 10 stoichiometric coefficients of the reaction.

If the variance of $Y_{x/?}$ is computed as:

$$\sigma_Y = (std\ deviation\ \% * Y_{X/?})^2 \quad (38)$$

the variance of $1/ Y_{x/Eth}$ can be estimated to be :

$$\sigma_{1/Y} = \frac{1}{(Y_{X/?})^4} . \sigma_Y \quad (39)$$

Thus, assuming a variance of 0 for the fixed reference stoichiometric coefficient (the one of the biomass), the variance of each stoichiometric coefficient used as a new experimental relation (as detail previously for α_7) can be calculated as

$$\sigma_{\alpha?} = \frac{Molar\ mass\ biomass}{Molar\ mass\ ?} . \sigma_{1/y} \quad (40)$$

The system is computed using the Matlab® (or Octave®) script presented in figure 3.

The resulting stoichiometric equation coefficients are determined, and their confidence interval is calculated (table 7).

The results are consistent with those computed with the previous classical approach. It is noticed that the main uncertainties concern the compounds that have the highest experimental uncertainty (i.e. acetoin and amino acid). This is one of the reconciliation method objectives: to have a high weight on the data that have been obtained with a high accuracy. The standard deviation for the coefficients is lower than the experimental ones but remains in the same order of magnitude.

$$0.7128\ glucose + 0.1267\ O_2 + 0.0980\ NH_3 + 0.2166 Amino.Acid \rightarrow 1\ biomass + 1.2019\ CO_2 + 1.0557\ ethanol + 0.0193\ glycerol + 0.0306\ acetoin + 0.5358\ H_2\ O$$

Compound	Stoichiometric coefficient	Standard deviation
Biomass	1.0000	0.0000
Glucose	-0.7128	+/- 0.0145 (2.03%)
AA	-0.2166	+/- 0.0172 (7.94%)
N-NH3	-0.0980	+/- 0.0041 (4.18%)
Ethanol	1.0557	+/- 0.0287 (2.72%)
Glycerol	0.0193	+/- 0.0010 (5.18%)
Acetoin	0.0306	+/- 0.0036 (11.76%)
CO2	1.2019	+/- 0.0286 (2.38%)
O2	-0.1267	+/- 0.0032 (2.53%)
H2O	0.5358	+/- 0.0096 (1.79%)

Table 7. Stoichiometric equation with reconciliation with all the experimental data "Yx/?" obtained during a batch culture.

4.4 Stoichiometry using several repetition of the batch growth of *Sc. cerevisiae* in bioreactor

It is generally recommended that an experiment is repeated several times in order to check its reproducibility and thus to confirm the results obtained. There are several ways for using the data obtained from the replication of the experiment:

1. By analyzing the results of each experiment independently and comparing the final results. For this purpose, it is necessary to express the results of each experiment (as presented in table 5) into yields (as presented in table 6) and to compute the resulting stoichiometric equations by either simple mass balance or data reconciliation method. In such an approach the reproducibility is checked by comparing the equations obtained for each experiment.
2. By analyzing the results of each experiment independently, but instead of computing one stoichiometric equation for each experiment, the yields calculated are used as a set of information to compute a single average stoichiometry for the repeated experiments. The reproducibility is tested either by comparing the results with the experimental yields or by analyzing the standard deviation computed for the stoichiometric coefficients.
3. By compiling all the results of the experiments as a single set of data and calculating a single average yield from all experiments. The stoichiometric equation should then be computed using a reconciliation data method and the reproducibility is tested with the standard deviation computed for the stoichiometric coefficients.

4.4.1 Experimental yields of 3 repetition of the batch growth of *Sc. cerevisiae*

The results obtained for 3 repetitions of the batch growth of *Sc. cerevisiae* are reported in table 8. The experimental carbon and nitrogen conservation balances reported in the table are calculated using the Yx/? yields (the balance can be different if calculated using the other yields). In experiment 2, the redundant information gives bad results in terms of conservation of elements. This is due to the large standard deviation calculated for the

yields. It can also be observed that within the 3 repetitions there is about 20% of variation in the mass yields calculated.

	Experiment 1 *C balance = 0.92* *N balance = 0.78*			*Experiment 2* *C balance = 0.75* *N balance = 0.71*			*Experiment 3* *C balance = 0.72* ** *N balance = 0.89*		
	Yield ($g.g^{-1}$)	Standard deviation	r^2	Yield ($g.g^{-1}$)	Standard deviation	r^2	Yield ($g.g^{-1}$)	Standard deviation	r^2
$Y_{x/glu}$	**-0.1806**	+/- 3.4%	0.988	**-0.1775**	+/-10.8%	0.887	**-0.1495**	+/-3.9%	0.983
$Y_{AA/glu}$	**0.0374***	+/- 5.8%	0.965	**0.0532 ***	+/-3.8%	0.983	**0.0266 ***	+/-6.1%	0.957
$Y_{N\text{-}NH3/glu}$	**0.0096***	+/- 2.5%	0.993	**0.0104 ***	+/-3.3%	0.987	**0.0079 ***	+/-3.3%	0.987
$Y_{eth/glu}$	**-0.3798**	+/- 5.6%	0.970	**-0.3039**	+/-11.2%	0.878	**-0.3047**	+/-4.5%	0.978
$Y_{gly/glu}$	**-0.0156**	+/- 5.7%	0.969	**-0.0222**	+/-11.3%	0.876	**-**	-	-
$Y_{acet/glu}$	**-0.0076**	+/- 34.2%	0.461	**-0.0130**	+/-14.7%	0.807	**-**	-	-
$Y_{CO2/glu}$	**-0.5245**	+/- 3.4%	0.990	**-0.4288**	+/-12.1%	0.873	**-0.4065**	+/-2.8%	0.992
$Y_{O2/glu}$	**-0.0332**	+/- 4.2%	0.984	**-0.0410**	+/-11.3%	0.888	**-0.0388**	+/-4.0%	0.984
$Y_{x/AA}$	**-3.0418**	+/- 8.1%	0.939	**-2.4486**	+/-7.5%	0.942	**-4.0251**	+/-6.4%	0.957
$Y_{x/N\text{-}NH3}$	**-14.9807***	+/- 4.8%	0.977	**-15.0498 ***	+/-7.1%	0.948	**-15.6974***	+/-9.2%	0.916
$Y_{x/Eth}$	**0.5498***	+/- 4.2%	0.981	**0.6869 ***	+/-3.8%	0.983	**0.5732 ***	+/-4.2%	0.979
$Y_{x/gly}$	**13.6788***	+/- 5.1%	0.973	**10.4657 ***	+/-7.8%	0.932	**-**	-	-
$Y_{x/acet}$	**9.2216***	+/- 11.8%	0.866	**10.6540 ***	+/-6.0%	0.959	**-**	-	-
$Y_{x/CO2}$	**0.4193***	+/- 4.8%	0.977	**0.5549 ***	+/-7.1%	0.947	**0.4441**	+/-4.6%	0.978
$Y_{x/O2}$	**6.0254***	+/- 2.5%	0.994	**5.4944 ***	+/-5.3%	0.970	**4.7570**	+/-5.7%	0.965

Table 8. Yields calculated for 3 repetitions of the batch growth of *Sc. cerevisiae*. . * means that a linear regression without intercept was used. ** balance is obtained without glycerol and acetoin. r^2 is the correlation coefficient of the linear regression used to calculate the yield.

4.4.2 Stoichiometric equation and comparison with experiments

With the 22 experimental yields Yx/? of table 8 (and their standard deviation), and by fixing the biomass coefficient $\alpha_5$to 1, a system of 29 equations is built for computing the 10 stoichiometric coefficients of the reaction. The system is highly over determined and information given is redundant. The solution is computed using the same script (figure 3), by changing the file experimental matrix and imposing the element conservation equations.

The stoichiometric equation computed is reported in table 9. The result obtained remains closed to the one obtained in table 7. It can also be observed that the standard deviations of all coefficients are significantly lower than the results presented in the table 8. In fact, repetitions of the experiment coupled with the over-determined system have increased the computation robustness (on a statistical point of view). Obviously, the stoichiometry remains theoretical and is an average of 3 experiments, but it is representative of the batch growth of *Sc. cerevisiae* and it can be also considered as reproducible for any growth of the strain in the same condition.

$0.6613\ glucose + 0.1322\ O_2 + 0.0946\ NH_3 + 0.2310 Amino.Acid \rightarrow 1\ biomass + 1.109\ CO_2 + 0.9595\ ethanol + 0.0207\ glycerol + 0.0272\ acetoin + 0.5292\ H_2\ O$

Compound	Stoichiometric coefficient	Standard deviation	Ratio table 7 / table 9
Biomass	1.0000	0.0000	100%
Glucose	-0.6613	+/- 0.0083 (1.26%)	92.8%
AA	-0.2310	+/- 0.0098 (4.24%)	106.6%
N-NH3	-0.0946	+/- 0.0024 (2.54%)	96.5%
Ethanol	0.9595	+/- 0.0163 (1.70%)	90.9%
Glycerol	0.0207	+/- 0.0009 (4.35%)	107.3%
Acetoin	0.0272	+/- 0.0015 (5.51%)	88.9%
CO2	1.1090	+/- 0.0164 (1.48%)	92.3%
O2	-0.1322	+/- 0.0028 (2.12%)	104.3%
H2O	0.5292	+/- 0.0053 (1.00%)	98.8%

Table 9. Stoichiometric equation with reconciliation with all the experimental data "Yx/?" obtained with the 3 repetitions of the same batch culture.

	Mass yield ($g.g^{-1}$)	*Standard deviation*
$Y_{x/glu}$	-0.2046	-0.55%
$Y_{x/AA}$	-4.1697	-0.23%
$Y_{x/N\text{-}NH3}$	-18.3857	-0.02%
$Y_{x/Eth}$	0.5517	1.56%
$Y_{x/gly}$	12.7862	0.00%
$Y_{x/acet}$	10.1730	0.00%
$Y_{x/CO2}$	0.4990	1.82%
$Y_{x/O2}$	-5.7560	-0.04%

Table 10. Theoretical yields computed with reconciliation with all the experimental data "Yx/?" obtained with the 3 repetitions of the same batch culture.

One of the main feature of these computation remains in the fact that the computed coefficients satisfy elements conservation; this provides the robustness of the stoichiometric model; on the other hand, it must also be kept in mind that the model, even robust, intrinsically assumes *(i)* that all substrates and products have really been identified and taken in the stoichiometry and *(ii)* that all yields remain constant during the observation time (culture duration) taken for estimation of experimental yields. This is certainly the main difficulty here, both for the estimation of experimental yields from linear regression and for imposing a single stoichiometry model that summarizes three different experiments obtained in batch cultures. Any experimental deviation from the assumption of constant yields creates a lack of adequacy of the model that results in an increase of the variance of the model.

Nevertheless, it can be concluded that the results obtained, within their confidence intervals (table 10), fairly represent the experimental data obtained from the 3 different experiments, knowing that it is not clear if a more complex model (with more coefficients to identify) would be justified by the available degree of experimental information.

```
%=============================================================================================
% Programm for Stoichiometry resolution by data reconciliation and statistical approach.
% AUTHOR : L. POUGHON
%=============================================================================================
% [1] ---> Problem definition <-------------------------------------------------------------
clear all; % clear all previous variables of the workspace
%.....Compounds list. The compounds are ordered from 1 to N
name={'Biomass','Glucose','Amino.Acid.','NH3','Ethanol','Glycerol','Acetoine','CO2','O2','H2O'};
%.....S1 = Compounds composition matrix (CHONSP balance).
S1=[1     6    1     0    2   3   4  1  0 0
           1.62  12   2.24  3    6   8   5  0  0 2
           0.52  6    0.48  0    1   3   2  2  2 1
           0.15  0    0.24  1    0   0   0  0  0 0];
K1=[0 ; 0 ; 0 ; 0];
%....S2 =. Theoreticals relations (typicaly one of the coefficients of the stoichiometry is fixed
S2=[1 0 0 0 0 0 0 0 0 0];
K2=1;
%....A =.Experimentals relations. IF There is No variance then write variance_exp=[]
A =[0       1       0       0       0       0       0       0       0
           0       0       1       0       0       0       0       0       0       0
           0       0       0       1       0       0       0       0       0       0
           0       0       0       0       1       0       0       0       0       0
           0       0       0       0       0       1       0       0       0       0
           0       0       0       0       0       0       1       0       0       0
           0       0       0       0       0       0       0       1       0       0
           0       0       0       0       0       0       0       0       1       0];
Y =[-0.7492 ; -0.3167 ;-0.1161 ; 0.9628 ; 0.0193 ; 0.03 ; 1.3199 ; -0.1263];
variance_exp=[   0.0257^2 ; 0.0256^2 ;   0.0056^2 ; 0.0406^2 ; 0.0010^2 ; 0.0036^2 ; 0.0640^2 ;
0.0032^2];
% [2] ---> Solving problem <----------------------------------------------------------------
%.... Be care for errors (matrix must be homogenous in size)
M=[S1 ; S2 ; A]; R=[K1 ; K2 ; Y];
S=[S1 ; S2 ] ;K=[K1 ; K2] ;
%... matrix for weight
if isempty(variance_exp)
   w=eye(size(A));
else
    w=diag(1./[1e-10*ones(size(K,1),1) ; [variance_exp] ] );
end
%.... Solving the problem
if (size(M,2)==size(M,1) )
   %-- the system is determined (no reconciliations)
   coefficients=M\R;
   covar_coeff=zeros(size(coefficients));
   covar_Acc=zeros(size(R));
   w=zeros(size(R)); % for simplifying the result display
elseif (size(M,2)<size(M,1))
   %-- the system is over determined : reconciliation by lagrangian approach
   V=inv([M]' * w * [M] );
   G=[M]' * w* [R];
   H=inv([ S ] * V * [ S ]' );
   coefficients=V * (G + [S]' * H * ( K - S * V * G ) );
   covar_coeff= [ [R-M*coefficients]' * w * [R-M*coefficients] ] * V ./ ( size(M,1)-size(M,2));
   covar_Acc=[M]*covar_coeff * [M]';
else
   %-- the system cannot be solved.
   disp('System under determined : cannot be solved') ;  return
end
% [3] ---> Display Solution of the problem (coefficients and accumulation vector)<---------------
for i=1:size(coefficients,1)
    display=sprintf('%10s\t   :   %10.4f\t   +/-   %10.4f'   ,char(name(i)),   coefficients(i)   ,
sqrt(covar_coeff(i,i)));  disp(display);
end
Acc=M*coefficients;
disp(' ') ; disp('Accumulation vector')
for i=1:size(Acc,1)
    display=sprintf('line[%2d] : %7.4f\t +/- %7.4f  (theoretical = %7.4f +/- %7.4f)' ,i, Acc(i) ,
sqrt(covar_Acc(i,i)),R(i),sqrt(1./w(i,i)) );  disp(display);
end
```

Fig. 3. Octave/Matlab script used to compute stoichiometric coefficients with the data reconciliation method.

5. Conclusion

Elements conservation laws are a prerequisite to evaluate a solid mass balance model in fermentation technology. As in Chemistry, the perfect vehicle for accounting for elements conservation is the stoichiometric equation. Application of this representation to biochemical systems (and particularly to microbial growth processes) presents some difficulties. The first one is linked to the fact that macromolecules and biomass constituents have seldom a well-determined elemental composition. This is a source of inaccuracy and variability. However general observation of biomass composition clearly shows that it is seldom highly variable (except in the case a product is accumulated into the cells in high quantities). The second obstacle is related to the fact that a single stoichiometry representation intrinsically assumes that the yields are constant. This is certainly a good assumption for continuous processes where biomass metabolism is confined in a constant environment. This is more inaccurate in the case of batch cultures.

Provided a complete statistical analysis is performed, including the calculation of both the coefficients and the model variances, the single stoichiometry approach can be applied for characterizing bioreactions.

We have presented here a data reconciliation technique coupled with the constraint of elements conservation. The main interest of this approach is that the coefficients are obtained with a flexible method applicable with linear algebra techniques, the result being "stoichiometrically" valid.

This technique has been applied to two cases. In the first case, results on the continuous culture of the rumen anaerobic bacteria *Fibrobacter succinogenes* lead to characterize the culture by a stoichiometric equation, slightly depending on the dilution rate. In the second case, batch experiments for the culture of *Saccharomyces cerevisiae* clearly indicate that a single stoichiometry approach is less accurate than for continuous cultures. Nevertheless, a stoichiometric equation has been obtained and realistic mean square deviations have been calculated for the coefficients. The technique has been applied to lump the experimental information from three independent experiments. This shows that this first-order stoichiometric model, including elements balance conservation, is certainly a valuable characterization of the biomass growth and of primary metabolites production. It must also be kept in mind that more complex models would involve more coefficients finally resulting in inaccurate predictions without creating more robustness.

6. References

Erickson, L.E., Minkevich, I.G. & Eroshin, V.K. (1978) Application of mass and energy balance regularities in fermentation. *Biotechnology and Bioengineering*, Vol.20, No.10, pp. 1595-1621.

Gaudet, G., Forano, E., Dauphin, G. & Delort, A.M. (1992). Futile cycle of glycogen in *Fibrobacter succinogenes* as shown by in situ ^{1}H NMR and ^{13}C NMR investigations. *European Journal of Biochemistry*, Vol.107, pp. 155-162.

Guiavarch, E., Pons, A., Creuly, C. & Dussap, C.G. (2008). Application of a data reconciliation method to the stoichiometric analysis of *Fibrobacter succinogenes* growth. *Applied Biochemistry and Biotechnology*, Vol.151, pp.201-210.

Guiavarch, E., Pons, A., Christophe, C., Creuly, C. & Dussap, C.G. (2010). Analysis of a continuous culture of *Fibrobacter succinogenes* S85 on a standardized glucose medium. *Bioprocess and Biosystems Engineering*, Vol.33, pp.417-425.

Himmelblau, D.M. (1970). *Process analysis by statistical methods*, John Wiley Ed, New-York.

Humphrey, A.E. (1974) Current developments in fermentation. *Chemical Engineering*. Vol.81, No.26, pp. 98-112

Hungate, R.E. (1950). The anaerobic mesophilic cellulolytic bacteria. *Bacteriological Review*, Vol.14, No.1, pp.1-49.

Kristiansen B. (1994). Integrated desugn of a fermentation plant. The production of baker's yeast. VCH (Ed.) Weinheim.

Minkevich, I.G. (1983) Mass-energy balance for microbial product synthesis—biochemical and cultural aspects. *Biotechnology and Bioengineering*, Vol.25, No.5, pp. 1267-1293.

Minkevich, I.G. & Eroshin, V.K., (1973) Folio Microbiol, Vol. 18, pp. 376-385.

Nielsen, J., Villadsen, J. & Liden, G. (2003) Bioreaction Engineering Principles, 2nd Ed, Kluwer, Dordrecht.

Nouaille, R., Matulova, M., Delort, A.M. & Forano, E. (2005). Oligosaccharide synthesis in *Fibrobacter succinogenes* S85 and its modulation by the substrate. *FEBS Journal*, Vol. 272, pp.2416-2427.

Patton, C.J., & Crouch, S.R. (1977). Spectrophotometric and kinetics investigation of the Berthelot reaction for the determination of ammonia: *Analytical Chemistry*, Vol.49, pp. 464-469.

Roels, J.A. (1980) Application of macroscopic principles to microbial metabolism. *Biotechnology and Bioengineering*, Vol. 22, No.12, pp.2457-2514.

Roels, J.A. (1983) Energetics and kinetics in Biotechnology, Elsevier Biomedical Press, 117, Amsterdam

Solomon, B.O., Erickson, L.E., Hess, J.E. & Yang, S.S. (1982) Maximum likelihood estimation of growth yields. *Biotechnology and Bioengineering*, Vol. 24, pp.633-649.

Solomon, B.O., Oner, M.D., Erickson, L.E. & Yang, S.S. (1984) Estimation of parameters where dependent observations are related by equality constraints. *AIChE Journal*, Vol.30, pp. 747-757

Summer, J.B. & Howell, S.F. (1935). A method for determination of saccharase activity. *Journal of Biochemical Chemistry*., Vol.108, pp. 51-54.

Urrieta-Saltijeral, J.M., Dussap, C.G., Pons, A., Creuly, C. & Gros, J.B. (2001). Metabolic flux modeling as a tool to analyse the behavior of a genetically modified strain of *Saccharomyces cerevisiae*. *Engineering and Manufacturing for Biotechnology*, pp.143-156.

Wang, N.S. & Stephanopoulos, G. (1983) Application of macroscopic balances to the identification of gross measurement errors. *Biotechnology and Bioengineering*, Vol.25, pp. 2177-2208.

Wells, J.E. & Russel, J.B. (1996). The effect of growth and starvation on the lysis of the ruminal cellulolytic bacterium *Fibrobacter succinogenes*. *Applied and Environmental Microbiology*, Vol. 62, No.4, pp.1342-1346.

IntechOpen

IntechOpen



Stoichiometry and Research - The Importance of Quantity in Biomedicine

Edited by Dr Alessio Innocenti

ISBN 978-953-51-0198-7
Hard cover, 376 pages
Publisher InTech
Published online 07, March, 2012
Published in print edition March, 2012

The aim of this book is to provide an overview of the importance of stoichiometry in the biomedical field. It proposes a collection of selected research articles and reviews which provide up-to-date information related to stoichiometry at various levels. The first section deals with host-guest chemistry, focusing on selected calixarenes, cyclodextrins and crown ethers derivatives. In the second and third sections the book presents some issues concerning stoichiometry of metal complexes and lipids and polymers architecture. The fourth section aims to clarify the role of stoichiometry in the determination of protein interactions, while in the fifth section some selected experimental techniques applied to specific systems are introduced. The last section of the book is an attempt at showing some interesting connections between biomedicine and the environment, introducing the concept of biological stoichiometry. On this basis, the present volume would definitely be an ideal source of scientific information to researchers and scientists involved in biomedicine, biochemistry and other areas involving stoichiometry evaluation.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Farges Bérangère, Poughon Laurent, Pons Agnès and Dussap Claude-Gilles (2012). Methodology for Bioprocess Analysis: Mass Balances, Yields and Stoichiometries, Stoichiometry and Research - The Importance of Quantity in Biomedicine, Dr Alessio Innocenti (Ed.), ISBN: 978-953-51-0198-7, InTech, Available from: http://www.intechopen.com/books/stoichiometry-and-research-the-importance-of-quantity-in-biomedicine/methodology-for-bioprocess-analysis-mass-balances-yields-and-stoichiometries-

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen