KTH Biotechnology

Metabolic reaction network approach for CHO modelling culture

ANTONIO ALIAGA

2010-06-02

Master's Thesis in Bioprocess Department, KTH

Animal Cell Group

Supervisor: Veronique Chotteau

Examiner: Andres Veide

ABSTRACT

Animal cell culture has provided several beneficial improvements in the field of biotechnology. Nowadays, the engineers have focused their work in the optimization of cell cultures techniques. One important tool used is the simulation by computer since it is inexpensive, requires less time than other methods and it is a simple way of understanding the behaviour of the cells in a culture.

This Thesis worked in the design of one simulator that describe the evolution over time of the extracellular metabolites and cell growth of CHO culture. A model that was designed using the concept of metabolic reaction network and the assumption of pseudo-steady state was checked and validated using diverse set of experimental data. In order to trigger different metabolic routes in the cells, these experiments were carried out varying amino acid composition in the medium.

The metabolic reaction network was simplified and consisted in 38 reactions. The set of experimental data was simulated using a graphical interface design in the present Thesis. The model succeeded since the results were very satisfactory. The error was, in general, small and it was checked that the system could also detect the different metabolic pathways that the cells follow when the initial conditions are modified in many metabolites. The model can also be applied without the information of all the essential amino acids, obtaining highly satisfying results as well.

ABBREVATIONS

aa Amino acid
AAA Amino Acid Analysis
AABA Amino Butyric Acid
AcCoA Acetyl Coenzye A
Ala Alanine
Arg Arginine
Asn Asparagine
Asp Aspartate
ATP Adenosine triphosphate
CHO Chinese Hamster Ovary cells
CO_2 Carbon dioxide
CT Centrifuge Tube
Cys Cysteine
DMEM Dulbecco's Modified Eagle's medium
DMEM/F12 Dulbecco's Modified Eagle's Medium: Nutrient Mixture F-12
Eaa Essential amino acids
EFM Elementary Flux Mode
EP Extreme Pathway
FBA Flux-balance analysis
Gln Glutamine
Glu Glutamate
Glc Glucose
Gly Glycine
G6P Glucose-6-Phosphate
His Histidine
HCL Hydrochloric acid
HPLC High Pressure Liquid Chromatography
Ile Isoleucine
Lac Lactate
Leu Leucine
LR Linear Programming
Lys Lysine
MEM Eagle's Minimal Essential medium
Met Methionine
MFA Metabolic flux analysis
MPA Metabolic Pathway Analysis
NaOH Sodium hydroxide
NH_4 Ammonia
O_2 Oxygen
Phe Phenylalanine
Pro Proline
Ser Serine
SF Serum-free
SucCoA Succinyl Coenzyme A
TCA Trichloroacetic acid
Thr Threonine
Trp Tryptophan
Tyr Tyrosine
Val Valine
αKG Alpha-ketoglutarate

NOMENCLATURE

a_{ij}	Elements of the macroscopic stoichiometric matrix
b_j	flux that produce and consume the metabolite in the reaction *j*
C_i	Concentration of the metabolite I [μM]
C_{in}	Inflow concentration of the metabolite [mol/l]
C_{out}	Outflow concentration of the metabolite [mol/l]
D	Dilution factor (dimensionless)
E	Null space
F_{in}	Inflow volume per day [l/h]
F_{out}	Outflow volume per day [l/h]
K_m	Michaelis constant [μM]
MVC	Million viable cells
N	Number of reactions of the metabolic reaction network
NS	Number of samples
p	Number of extracellular metabolites
Q_i	Consumption/production rate of the metabolite I [mol/l·h]
q_i	Specific consumption/production rate of metabolite I [nmol/MVC·day]
$S_{i,j}$	Stoichiometric coefficient of the metabolite *i* in the reaction *j*
t	Number of intracellular metabolites
V	Volume [l]
v_j	Specific flux of the reaction *j*
X_v	Cell density [MVC/mL]
$\mu_{max,j}$	Maximal kinetic rate of the reaction j

Matrices and vectors

$\underline{A}_{ext}$	Extracellular stoichiometric matrix
$\underline{A}_{int}$	Intracellular stoichiometric matrix
$\underline{A}_{mac}$	Macroscopic stoichiometric matrix
$\underline{A}_{red}$	Stoichiometric matrix of the reduced system
$\underline{b}$	Vector of the reaction fluxes

$\underline{C}$ Concentration of the metabolites

$\underline{Q}$ Vector of the consumption/production rate of the metabolites

$\underline{q_{est}}$ Specific consumption/production rate of the extracellular metabolites

$\underline{q_{int}}$ Specific consumption/production rate of the intracellular metabolites

$\underline{\underline{S}}$ Stoichiometric matrix

$\underline{v}$ Vector of specific fluxes

$\underline{w}$ Vector of the macroscopic specific rate fluxes

$\underline{\varepsilon}$ Vector of the extracellular concentration

$\underline{\mu_{max}}$ Vector of the maximal kinetic rates of the reactions

$\underline{\underline{\xi}}$ Matrix elements q_{ij}

TABLE OF CONTENTS

1. INTRODUCTION

Animal cell culture has a special interest in the industry of biotechnology since the ability to grow animal cells in vitro has provided different advances in the fields of biology and medicine [1]. They have been used to study the physiological and biochemical properties of the cells as well as to test the effects of drugs and vaccines, to produce artificial tissue for implantation and to synthesize valuable products from large-scale cultures [2].

Due to their importance, the engineers have focused their efforts in the study and optimization of cell cultivation techniques. There are different methodologies that engineers use to accomplish this goal. However, most of the techniques employed were expensive and time-consuming. Hence the introduction of simulation by computer has helped the work of engineers.

These simulators use mathematical models to describe the cells and the interaction of the different components in cell cultures. They reproduce the behaviour of cell culture in a more efficient way that in real experiments, since it is faster and it does not require a high cost. Moreover, new devices and techniques were designed thanks to the introduction of the simulators. One example is the software sensors, which are more reliable than the hardware ones.

1.1. Aim of the Thesis

The aim of this Thesis is the study of the performance of a dynamical model already done, which can predict the behaviour of CHO cells in a culture over time. The model was studied, validated and verified using different experiments of CHO culture under different conditions (Fig. 1.1).

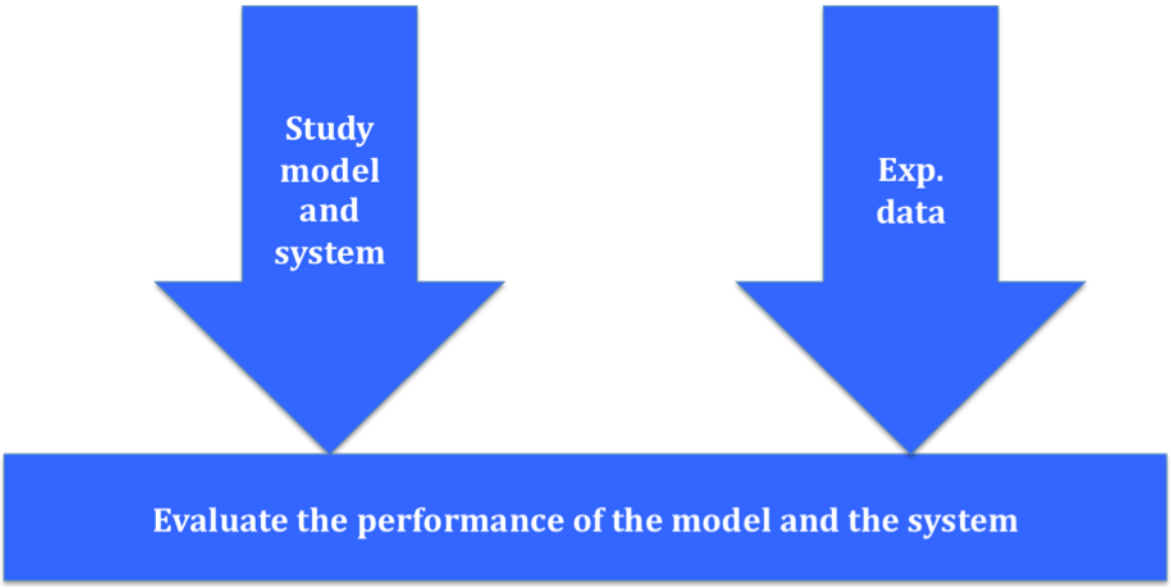


Figure 1.1. Strategy of the Thesis

The concept of metabolic reaction network and the assumption of pseudo-steady state are used in the design of the model. The variables of the model are the major energy source, glucose and glutamine, amino acid and cell density, which are the only parameters that are measured in a cell culture.

Experimental data will be obtained doing different cell culture with different environmental conditions.

2. ANIMAL CELL CULTURE

Ross Harrison was the first person that succeeded in the culture of animal cells in 1907 [3]. But the scientists did not start to use them as an important tool until the 50's and its commercialization still took nearly two decades to carry out [4]. Since then, several advances have been made within different applications and fields like the development of new vaccines and drugs [3].

Animal cell culture consists in the growth and proliferation of cells outside the tissue using a mixture of different components. The mixture is called medium and its composition is essential in the good performance of the culture [4].

A medium has to contain all the components necessary for the nutrition of the cells: vitamins, amino acids, lipids, nucleic acid precursors, carbohydrates, trace elements, salts, bulk ions and often growth factors and hormones. Components that ensure constant pH levels are also necessary [5]. Historically, the media were designed using a base medium supplemented with several factors [4]. These factors included serum or other blood products lipids as well as embryo extracts or yeast extracts. The most common one was the serum. It contained the necessary concentration factors to provide growth and proliferation of the cells like proteins, trace elements, growth factors, vitamins, etc. [5]. Eagle's minimal essential medium (Eagle's MEM or MEM) and MEM modified by Dulbecco (Dulbecco's Modified Eagle's Medium, DMEM) were typical base medium. They are still used nowadays to keep primary cell cultures and cell lines [4]

However, the use of serum as a supplement in the base media carried a high cost and several technical drawbacks. These technical disadvantages include the high risk of contamination (e.g. viruses), the undefined nature of serum and variations in the serum composition due to seasonal and continental variations. [6]. The latter case produces batch-to-batch variations, which causes phenotypical differences in

the cell cultures and as a consequence the results could differ. Moreover, some ethical issues, such as the unnecessary suffering of the animals during serum extraction, have created difficulties in the utilization of such media [4].

As a consequence, nowadays the efforts of the scientists are focused on the creation of serum-free media. The new design uses the base media described above with supplementation of different components. Examples of such supplements are: hormones, growth factors, attachment factors, lipids, protease inhibitors, protein hydrolysates and proteins. It is also common to add some amino acids. Generally, the base media contain the essential amino acids[1] and then additions of nonessential amino acids[2] are done. The difference between essential and nonessential amino acids is that the cells cannot biosynthetise the essential amino acids and the cells need them to grow [4].

An important example is the Ham's F12 medium, which contains lipids, nucleic acid derivatives, vitamins, non-essential amino acids and small concentrations of the essential amino acids and sugar [5].

The use of serum-free media has provided other advantages such as an increased in the cell growth and its productivity, more consistent performance, precise evaluations of cellular function and a better control over physiological responsiveness [7].

There are some aspects that must be taken into account in the design of serum-free media. The first one is that it does not exist a universal media that can be used in all cell types; so different serum-free media have been designed for each cell type [4]. And second the cells need some processes to adapt to the new serum-free media and the whole process must be monitored and checked carefully since small

[1] Arginine, Histidine, Isoleucine, Leucine, Lysine, Methionine, Phenylalanine, Threonine, Tryptophan and Valine
[2] Alanine, Asparagine, Aspartate, Cysteine, Glutamate, Glutamine, Glycine, Proline, Serine and Tyrosine

undesired changes in the culture conditions, may produce alterations in cellular functions [6].

Other possible media are animal-free media and protein-free media. The former case is similar to serum-free media but the components are derived from non-animal sources. And the latter case does not use proteins [8].

2.1. Mammalian cells: Chinese Hamster Ovaries (CHO)

Nowadays, the mammalian cells are one of the most used cell types in biology and medicine. Mammalian cells occupy 60% of the current market [9]. Inside this type of cells, Chinese Hamster Ovary cells are the most important ones. They are widely spread and are used for transfection, expression and large-scale recombinant protein production [10]. Almost 70% of the recombinant protein therapeutics produced today is from CHO [11].

This type of cells is generally grown in incubators under specific conditions: the temperature is kept around 37 °C with a controlled humidified gas mixture of 95% O_2 and 5% CO_2 [4].

3. METABOLIC REACTION NETWORK

Cell metabolism involves thousand of biochemical reactions where the metabolic substrates are transformed either in energy or other components [12]. It is graphically represented by a metabolic network, which illustrates the different biochemical reactions that occur within the cell, together with the reactions that happen with the environment that surrounds the cell [13].

The metabolic network consists of nodes, which symbolize the metabolites, and edges, which represents the metabolic reactions (Fig. 3.1). The inputs and outputs are, respectively, the substrates and the products [14].

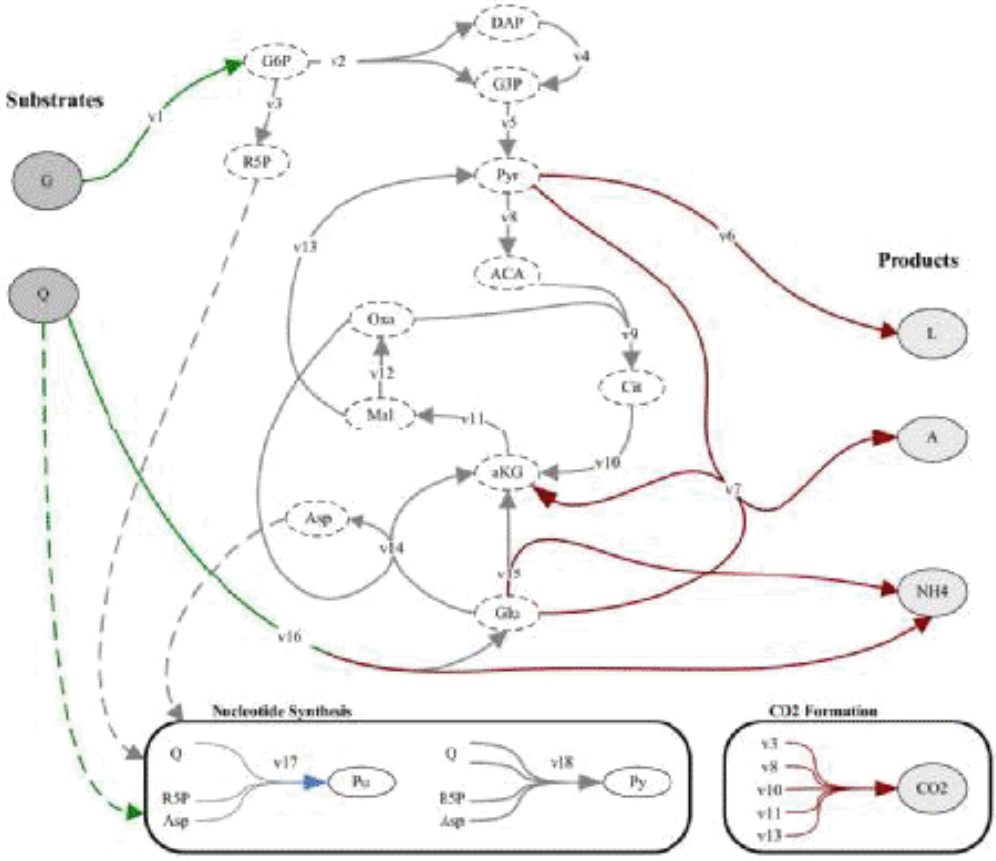


Figure 3.1. Metabolic reaction network [15]

The cells grow as a consequence of the coordinated action of different groups of reactions, usually called pathways. Furthermore, each reaction of each metabolic pathway occurs according to a given rate, called metabolic flux [16]. The metabolic flux of one reaction depends on the concentration of the metabolites involve on it. A mass balance can be applied for a given component [17]. It follows from the physical law that matter or mass can neither be created nor destroyed for every component [18]:

$$Mass_{in} + Mass_{produced} = Mass_{out} + Mass_{accumulated} \quad (1)$$

Since the mass is equal to the multiplication of the concentration times the volume, Eq. 1 for one metabolite can be rewritten as follows:

$$C_{in} \cdot F_{in} + Q_i \cdot V = C_{out} \cdot F_{out} + V \cdot \frac{dC}{dt} \quad (2)$$

where

C_{in} is the inflow concentration of the metabolite [mol/l]

F_{in} is the inflow volume per day [l/h]

Q_i is the consumption/production rate of the metabolite [mol/l·h]

V is the volume of the system (in this case it is assumed that it is constant) [l]

C_{out} is the outflow concentration of the metabolite [mol/l]

F_{out} is the outflow volume per day [l/h]

C is the concentration of the metabolite [mol/l]

The consumption/production rate of a component is given by the sum of the consumption/production rate of all the pathways in which this component is involved weighted by the stoichiometric coefficients S_{ij}. The stoichiometric matrix contains the stoichiometric coefficient S_{ij}. The rows of S corresponds to the metabolites and the column to the reaction kinetics [19,20]:

$$Q_i = S_{i1} \cdot b_1 + S_{i1} \cdot b_2 + S_{i1} \cdot b_3 + \cdots = \sum_j S_{i,j} \cdot b_j \Rightarrow \underline{Q} = \underline{\underline{S}} \cdot \underline{b} \quad (3)$$

where,

$S_{i,j}$ is the stoichiometric coefficient of the metabolite i in the reaction j and determines the number of moles of metabolite i formed in the reaction j

b_j is the flux that produce and consume the metabolite in the reaction j

Q is a vector that contains the consumption/production rate of each metabolite

$S_{i,j}$ is negative when the metabolite i is a substrate and positive when it is a product. And in the same way, Q_i is positive when the metabolite is produced and negative when it is consumed [19]. It is common to use the cell specific consumption/production rate q_i instead of the Q_i:

$$q_i = \frac{Q_i}{X_v} \quad (4)$$

where X_v is the viable cell density.

In case the culture volume is constant F_{in}=F_{out} and the dilution factor D can be defined as:

$$D = \frac{F_{in}}{V} = \frac{F_{out}}{V} \quad (5)$$

It follows, then, from Eq. 2 to Eq. 5 that:

$$\underline{Q} = \underline{q} \cdot X_v = \underline{\underline{S}} \cdot \underline{v} \cdot X_v = \frac{d\underline{C}}{dt} + D(\underline{C_{out}} - \underline{C_{in}}) \quad (6)$$

It can be seen that in this case $\underline{v}$ is the vector of specific fluxes of the biochemical reactions and depends only on the concentrations of the metabolites involve in each reaction ($b = \frac{v}{X_v}$).

Eq. 6 can be particularized for different kind of systems:

- **Batch process**: in this case D is equal to 0:

$$\begin{pmatrix} q_1 \\ \vdots \\ q_M \end{pmatrix} = \begin{pmatrix} S_{1,1} & \cdots & S_{1,N} \\ \vdots & \ddots & \vdots \\ S_{M,1} & \cdots & S_{M,N} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_M \end{pmatrix} = \frac{1}{X_v} \cdot \frac{d\underline{C}}{dt} \tag{7}$$

where M is the number of reactions of the system.

- **Perfusion process**: in this case the variation of the concentrations over time is 0, i.e. $\frac{d\underline{C}}{dt} = \underline{0}$, it follows:

$$\begin{pmatrix} q_1 \\ \vdots \\ q_M \end{pmatrix} = \begin{pmatrix} S_{1,1} & \cdots & S_{1,N} \\ \vdots & \ddots & \vdots \\ S_{M,1} & \cdots & S_{M,N} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_M \end{pmatrix} = \frac{D(\underline{C_{out}} - \underline{C_{in}})}{X_v} \tag{8}$$

3.1. Mathematical cell metabolism models

There are diverse techniques that employ mathematical models to describe the cell metabolism. Despite each methodology has different aims, uses distinct mathematical procedures and is based on different assumptions, all of them exploit the stoichiometric matrix to describe the models and assume the quasi-steady state [21].

The study of the intracellular metabolic fluxes is very important to understand better the different pathways interactions and their impact in the whole metabolic process [22]. The knowledge of the fundamental metabolism of cells in culture under different environmental conditions is indispensable in control strategies, media formulations and for the design of bioreactors [23].

One of the most used methodologies is the called Metabolic Flux Analysis (MFA) and has an important importance in metabolic engineering [24,25]. MFA uses the available data, which is mainly the extracellular fluxes, to determine the fluxes that are not possible to measure, which corresponds to the intracellular fluxes [25]. It is also used to estimate the major metabolic pathway fluxes, using material balancing [26, 27].

Even though MFA has been a widely employed methodology, it has some limitations [21]. The first one comes directly from the methodology itself since MFA requires a lot of measurements to be able to provide results. Unfortunately, the number of external measurements is not sufficient and the obtained system is underdetermined [28]. Therefore, the solution for the system is not unique. A second drawback is that the measurements have a significant error due to the imprecision and the insufficiency of the set of available measurements.

Some changes have been made in order to use MFA in either a determined or overdetermined system. Among them, there are the addition of metabolic theoretical constraints [28], the simplification of the metabolic network by the synthesis of a group of reactions into one simple reaction [23,29], making the system simpler, or the use of linear algebra or convex analysis to get only the solutions that are positive and possible in the system [27, 30].

Nowadays, it is emerging a new technique called flux-spectrum approach, which is employed to obtain the metabolic fluxes over time [31,32]. This technique uses the a priori knowledge to introduce reversibility constraints and takes into account the

uncertainty of the measured fluxes [32]. This approach can also estimate the non-measured fluxes even when the system is undetermined and the obtained results are more reliable in determined system than in other techniques [33].

Another methodology that is used when the metabolic flux distribution is undetermined is the flux-balance analysis (FBA). It is based on linear programming (LP), which optimize an objective function using the most effective and efficient paths through the network [34]. Some examples of the objective function are the maximization of biomass and ATP production [28,34]. The advantage of such technique is that it is possible to obtain quantitative measurements of the behaviour of the network with a minimum biological knowledge and a small amount of data [28].

Metabolic Pathway Analysis (MPA) is another flux-bases analysis method. It works with the concepts of elementary flux modes (EFM) and extreme pathways (EPs) [36-38]. Both concepts are similar since EPs are part of the EFM. MPA studies the properties that the stoichiometric gives, observing all the feasible biochemical network states in the optimal solution space [39].

The concept of EFM is a very important concept since it can be used to reduce a complex system into a simpler network, with a minimum number of reactions, as it was done in [27]. For that reason, in this Thesis it was decided to follow the methodology made by Joan Gonzalez Thesis [40]. It uses elementary flux analysing to calculate a macroscopic reaction. Following, the calculated macroscopic reaction is used to create a simplified system that will be used to predict the production/consumption rates.

3.2. Pseudo-steady state assumption

The common assumption in all the studies, the pseudo-steady state, declares that the intracellular metabolites in growing cells are in quasi-steady state. Therefore, the net sum of the production and consumption fluxes of these intracellular metabolites weighted by their stoichiometric matrix coefficients are zero [16]. It means that the kinetic reactions inside the cells are so much faster, reaching the steady state in a shorter period of time, than the reactions that involve extracellular metabolites [27].

An algebraic relation expresses this assumption:

$$\underline{\underline{A}}_{\text{int}} \cdot \underline{v} = \underline{0} \tag{9}$$

where $\underline{\underline{A}}_{int}$ corresponds to the stoichiometric matrix of the intracellular metabolites.

This expression can be introduced in Eq. 3:

$$\begin{pmatrix} q_1 \\ \vdots \\ q_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} S_{1,1} & S_{1,2} & \cdots & S_{1,M} \\ \vdots & \ddots & \cdots & \vdots \\ S_{p,1} & S_{p,2} & \cdots & S_{p,M} \\ S_{p+1,1} & S_{p+1,1} & \cdots & S_{p+1,1} \\ \vdots & \ddots & \cdots & \vdots \\ S_{p+t,1} & S_{p+t,1} & \cdots & S_{p+t,1} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_p \\ v_{p+1} \\ \vdots \\ v_{p+t} \end{pmatrix} \tag{10}$$

where p is the number of extracellular metabolites and t the number of intracellular metabolites.

Or,

$$\underline{\underline{S}} \cdot \underline{v} = \begin{pmatrix} q_{\text{int}} \\ q_{ext} \end{pmatrix} = \begin{pmatrix} 0 \\ q_{ext} \end{pmatrix} = \begin{bmatrix} A_{\text{int}} \\ A_{ext} \end{bmatrix} \cdot \underline{v} \qquad (11)$$

where $\underline{q_{int}}$ and $\underline{q_{ext}}$ are the vectors of the specific consumption/production rates of the intracellular and extracellular metabolites respectively and $\underline{\underline{A_{int}}}$ and $\underline{\underline{A_{ext}}}$ are the stoichiometric matrices of the intracellular and extracellular components.

The subdivision of the intracellular and extracellular systems gives the following equations:

$$\underline{q}_{\text{int}} = \underline{\underline{A}}_{\text{int}} \cdot \underline{v} = \underline{0} \qquad \underline{q}_{ext} = \underline{\underline{A}}_{ext} \cdot \underline{v} \qquad (12)$$

3.3. Macroscopic matrix

A macroscopic model is used when the only available data, besides the cell density, are the measurements of the extracellular metabolites. Thus, the input data of the system is:

- **Cell density** (*Xv*): number of alive cells in the culture [MVC/mL]. In the case of the cell density, the concentration is computed multiplying the *Xv* by 1000. In that case the units are 10^3 cells/mL.
- **Viability**: rate of living cells to total number of cells [%].
- C_i: concentration of the extracellular metabolite i [μM]
- q_i: cell specific consumption/production rate of the extracellular metabolite i [nmol/MVC·day].

Since, the way that the substrates are converted into products is unknown, the system is considered as a black box that catalyses the conversion of substrates into products [14]:

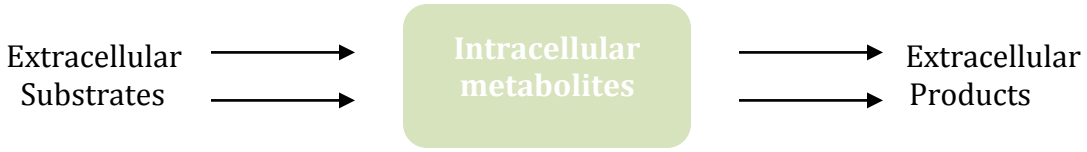


Figure 3.2. Metabolic system seen in a macroscopic point of view

The first stage of this approach is the calculation of the EFM, set of biochemical reactions that starts in one or several substrates and ends in one or several products [27]. In other words, an EFM is a vector that fulfils Eq. 8.

According to linear algebra, the set of EFM can be expressed using the concept of null space ($\underline{\underline{E}}$). All the possible solutions and their corresponding flux distributions can be expressed using the null space. The null space is a basis, a minimum number of linear independent vectors that fulfils Eq. 8 and such that:

$$\underline{\underline{A}}_{\text{int}} \cdot \underline{\underline{E}} = \underline{0} \tag{13}$$

where dimension of $\underline{\underline{E}}$ is obtained using the Rank Theorem of linear algebra [39]:

$$\dim(\underline{\underline{E}}) = N - Rank(\underline{\underline{A}}_{\text{int}}) \tag{14}$$

where N is the number of reactions of the system.

In order to compute the EFM, the Metatool can be used. It is a very simple tool and it is widely used in the computation of metabolic flux analysis [36,41]. The only parameters that Metatool needs is the stoichiometric matrix of the intracellular metabolites ($\underline{\underline{A}}_{int}$) and a vector of length equal to the number of reactions, which expresses if the reaction is reversible or not.

Nevertheless, it has been tested that Metatool has a problem when it has to compute the EFM of the reactions that involve only extracellular reactions. These reactions are expressed in $\underline{\underline{A_{int}}}$ as a column of 0 in all the positions. Normally, the corresponding EFM is expressed as a vector containing 0 in all the positions, except the position of the reaction, which contains a 1. The problem arises in the case that the reaction is reversible where an EFM should contain a -1 in the position of the corresponding reaction. It was decided to add them manually in the computation of the null space.

Once the null space is calculated, the macroscopic reaction can be computed in a matrix called $\underline{\underline{Amac}}$ as:

$$\underline{\underline{Amac}} = \underline{\underline{Aext}} \cdot \underline{\underline{E}} \tag{15}$$

The dynamical model of the extracellular reactions can be written, using Eq. 7 for the extracellular metabolites, as:

$$\frac{d\underline{\varepsilon}}{dt} = \underline{\underline{Amac}} \cdot \underline{w} \cdot Xv \tag{16}$$

where $\underline{\varepsilon}$ are the extracellular metabolite concentrations and w corresponds to the vector of the specific rates of the macroscopic reactions. From Eq. 11, Eq. 15 and Eq. 16 it can be found the relation between $\underline{v}$ fluxes and the macroscopic specific rates $\underline{w}$:

$$\frac{d\underline{\varepsilon}}{dt} = \underline{\underline{Amac}} \cdot \underline{w} \cdot Xv = \underline{\underline{Aext}} \cdot \underline{\underline{E}} \cdot \underline{w} \cdot Xv = \underline{\underline{Aext}} \cdot \underline{v} \cdot Xv \Rightarrow \underline{v} = \underline{\underline{E}} \cdot \underline{w} \tag{17}$$

And introducing this relation in Eq. 12:

$$\underline{q_{ext}} = \underline{\underline{A_{ext}}} \cdot \underline{v} = \underline{\underline{A_{ext}}} \cdot \underline{\underline{E}} \cdot \underline{w} = \underline{\underline{A_{mac}}} \cdot \underline{w} \tag{18}$$

The specific rates of the reactions can be obtained using kinetics. In this case, the Michaelis and Menten model was applied [42]. The model shows the relation between the substrate and the specific rate of the reaction in an enzyme using the following formula:

$$v = \frac{\mu_{\max} \cdot [S]}{K_m + [S]} \tag{19}$$

where μ_{max} corresponds to the maximal kinetic rate when the enzyme is saturated of substrate, *[S]* is the concentration of the substrate of the enzyme and K_m is the called Michaelis constant or half saturation constant. K_m is equal to the value of the substrate concentration where the reaction is half the maximal kinetic rate that the enzyme can reach.

One important aspect that has to be mentioned is that the Michaelis and Menten kinetics assumes that all the reactions are irreversible. So, the information of reversibility will be in the EFM, i.e. two reactions are considered one in each direction.

Eq. 19 can be applied when there is more than one substrate involve in the enzyme, as [16]:

$$v = \frac{\mu_{\max} \cdot \prod_i [S]_i^{\alpha}}{\prod_i (K_{m,i} + [S]_i^{\alpha})} \tag{20}$$

where α is the corresponding stoichiometric value of the substrate i in the reaction in question.

The $\underline{w}$ rates can be computed, applying Eq. 20 and taking into account that in this case, the reactions are seen from a macroscopic point of view:

$$w_j = \mu_j \cdot \prod_i \frac{c_i^{a'_{i,j}}}{(K_i + c_i)^{a'_{i,j}}} \tag{21}$$

where c_i corresponds to the extracellular substrates of the macroreaction j.

Now all the parameters are known except the maximal kinetic rates μ_j. According to that, it can be created a linear system that related the specific consumption/production rates with the maximal kinetic rates, introducing Eq. 21 in Eq. 18:

$$\begin{pmatrix} q_1 \\ \vdots \\ q_p \end{pmatrix} = \begin{pmatrix} a'_{1,1} \cdot \prod_i \frac{c_i^{a'_{i,1}}}{(K_i + c_i)^{a'_{i,1}}} & \cdots & a'_{1,1} \cdot \prod_i \frac{c_i^{a'_{i,N}}}{(K_i + c_i)^{a'_{i,N}}} \\ \vdots & \ddots & \vdots \\ a'_{p,1} \cdot \prod_i \frac{c_i^{a'_{i,1}}}{(K_i + c_i)^{a'_{i,1}}} & \cdots & a'_{p,N} \cdot \prod_i \frac{c_i^{a'_{i,N}}}{(K_i + c_i)^{a'N}} \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix} \tag{22}$$

where a_{ij} are the elements of $\underline{\underline{Amac.}}$

The above system is applied, then, to each sample of the experiment that is been analysing. Consequently and after applying the above system to all the samples, the final system can be written as:

$$\begin{pmatrix} q_{1,1} \\ \vdots \\ q_{p,1} \\ q_{1,2} \\ \vdots \\ q_{1,NS} \\ \vdots \\ q_{p,NS} \end{pmatrix} = \begin{bmatrix} B_1 \\ \vdots \\ B_{NS} \end{bmatrix} \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix} \Rightarrow \underline{\underline{\xi}} = \underline{\underline{B}} \cdot \underline{\mu} \tag{23}$$

where NS is the number of samples of the experiment, $\underline{\underline{\xi}}$ is the matrix of the elements q_{ij} and $\underline{\underline{B_i}}$ is the matrix of the linear system in Eq. 20 applied for each sample.

In order to calculate the maximal kinetic rates μ, the first algorithm that was chosen was the non-negative least-squares where the next function had to be optimize:

$$z = \min_{\mu} \left\| \underline{\underline{B}} \cdot \underline{\mu} - \underline{\underline{\xi}} \right\|_2^2 \quad where \ \mu \geq 0 \tag{24}$$

However, this gave a significant error between the calculated values and the obtained ones using Eq. 24. For that reason, a second function was chosen to obtain better results:

$$z = \min_{\mu} \left\| \underline{\underline{B_{norm}}} \cdot \underline{\mu} - \underline{\underline{\xi_{norm}}} \right\|_2^2 \quad where \ \mu \geq 0 \tag{25}$$

where B_{norm} and Q_{norm} are the normalization of B and Q by the average of the set of consumption/production rates of each metabolite. Moreover, each row of B and Q, which corresponds to specific metabolites, are divided by the average of the consumption/production rates obtained in all the samples of this specific metabolite.

Using the non-negative least-square, it was checked from [40] that in order to assure an optimal solution for μ, the system has to be overdetermined. This is a typical condition that has to be satisfied in least squares methodologies [43]. In this case, to fulfil the condition, the matrix B has to have more rows than columns:

$$\underset{(p)}{N.\ extracellular\ metabolites} \cdot \ \underset{(NS)}{N.\ of\ samples} > \underset{(N)}{N.\ of\ reactions} \tag{26}$$

It could be observable that normally there were several maximal kinetic rates μ_j that were zero and these extracellular reactions do not affect the estimated rates of the model. For that reason, it was decided to design a reduced system where only the reactions with a maximal kinetic rate above zero are taken. The new matrices obtained are $\underline{A_{red}}$ and $\underline{\mu_{red}}$ and they will be used in the estimation of the parameters in the model above described.

4. MATERIALS AND METHODS

4.1. The system

The metabolic network and its corresponding stoichiometric matrix used in this project were taken from [40]. This model took as a starting system the one made by [8], which used the experimental data from different experiments [29,44] to construct the biochemical network, considering the most relevant metabolic routes for animal cells cultured in vitro. Following, some simplifications, corrections and adjustments were made to arrive to a simpler model:

Table 4.1. Biochemical network of CHO cells[3]

R_1 Glc → G6P
R_2 G6P → 2·3phosphaglycerate
R_3 3phosphoglycerate →Pyr
R_4 Pyr → Lac
R_5 Pyr → AcCoA + CO_2
R_6 AcCoA + Oxal → Cit
R_7 Cit → αKG + CO_2
R_8 αKG → SucCoA + CO_2
R_9 SucCoA → Suc
R_{10} Suc → Mal
R_{11} Mal → Oxal
R_{12} Mal → Pyr + CO_2
R_{13} Thr → Gly + AcCoA
R_{14} Trp → Ala + NH_4 +2·AcCoA
R_{15} Lys → NH4 + αKG
R_{16} Ile → Glu + AcCoA + SucCoA
R_{17} Leu → 2·AcCoA + 2·CO_2
R_{18} Tyr → Mal + Oxal + CO_2
R_{19} Ser + Met → Cys + NH_4
R_{20} Val → SucCoA + αKG
R_{21} Glu + Oxal → Asp + αKG
R_{22} Glu → αKG + NH4
R_{23} Glu + Pyr → Ala + αKG
R_{24} Cys → Pyr
R_{25} Ser → NH_4 + Pyr
R_{26} Gly → NH_4 + CO_2
R_{27} Ser + Thr → SucCoA
R_{28} Glu + 3phosphoglycerate → Ser + αKG
R_{29} Ser → Gly
R_{30} Phe → Tyr
R_{31} Asn → Asp + NH_4
R_{32} Gln → Glu + NH_4
R_{33} Arg → Glu
R_{34} Glu → Pro
R_{35} His → Glu + NH_4
R_{36} Gln + Asp → Glu + Asn

R_{37} 0.0208·Glc + 0.0377·Gln + 0.0006·Glu + 0.007·Arg + 0.003·Hist + 0.0084·Ile + 0.0133·Leu + 0.0101·Lys + 0.0033·Met + 0.0055·Phe + 0.008·Thr + 0.004·Trp + 0.0096·Val + 0.0133·Ala + 0.026·Asp + 0.0004·Cys + 0.0165·Gly + 0.0081·Pro + 0.0099·Ser + 0.0077·Tyr → Biomass

[3] R_4, R_{21}, R_{22}, R_{23}, R_{29}, R_{30}, R_{31}, R_{32} and R_{36} are considered reversible

These modifications can be summarized in the following statements [41]:

- Some reactions and metabolites were discarded. This affected, mainly, the co-metabolites (ATP, ADP, NAD…) and the mitochondrial transport, considering that the metabolites inside and outside the mitochondria were the same metabolite.

- Some reactions were modified or corrected and others were added by our group since they occur in mammals and were missing in Altamirano model. These adjustments affected the reversibility of some reactions as well as the addition and elimination of some metabolites in some reactions.

- The biomass was described as only one reaction, using the same procedure as in [23].

- CO_2 was considered as an external metabolites and it is not used in the mathematical model since it is a final product and therefore does not affect the other reactions.

The metabolites of the system can be classified as:

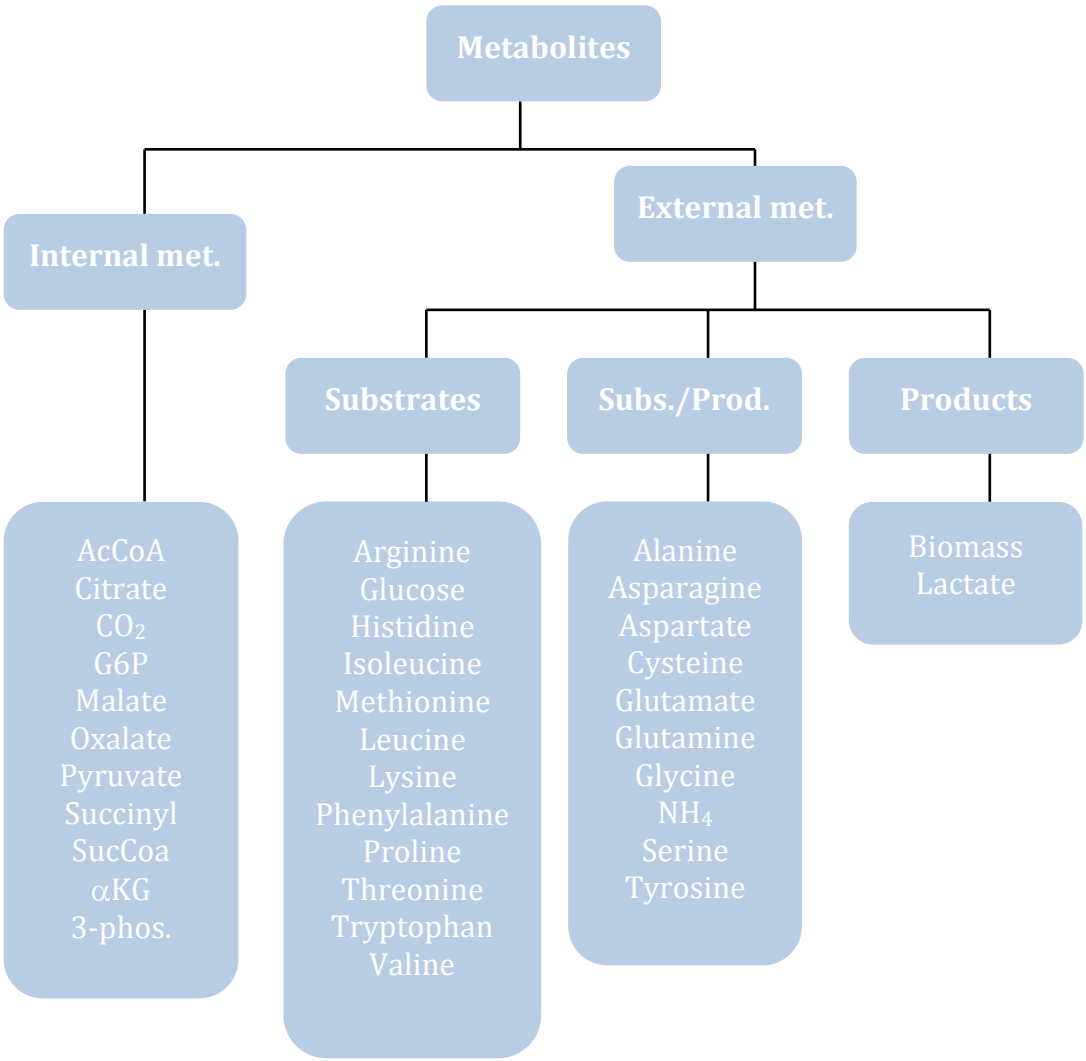


Figure 4.1. Classification of the metabolites in the system

The above system did not consider the state where the cells died and it was only possible to use it in the growing phase of the cells. For that reason, a new reaction was added that took also into account the death phase of the cells (C_d).

$$Cv \longrightarrow Cd$$

where Cv is the cell density [MVC/mL].

Cd can be calculated using the definition of the Viability:

$$Viability = \frac{Cv}{Ctot} = \frac{Cv}{Cv + Cd} \quad \Rightarrow \quad Cd = \frac{Cv(1 - Viability)}{Viability} \tag{27}$$

where C_{tot}=total number of cells (MVC/mL) and C_d has the same units as C_v [MVC/mL].

For computation of the concentration of the cell death, the methodology is the same as in the case of cell density. So, this concentration is obtained multiplying the corresponding cell density by 1000 [10^3 cells/mL].

Regarding the Michaelis constants (K_m), in absence of their value, the KTH Division of Bioprocess suggested the following values based on the order of magnitude of the concentration of these components:

Table 4.2. Michaelis constants of extracellular metabolites

Km [μM]			
Alanine	100	**NH_4**	300
Arginine	300	**Lactate**	300
Asparagine	100	**Leucine**	100
Aspartate	100	**Lysine**	100
Cysteine	100	**Phenylalanine**	100
Glucose	300	**Proline**	100
Glutamate	100	**Serine**	300
Glutamine	100	**Threonine**	100
Glycine	100	**Tryptophan**	100
Histidine	100	**Tyrosine**	150
Isoleucine	100	**Valine**	100
Methionine	100	**Bio./C. Death**	100

4.2. Experiment

The experiment consisted in cell culture of suspended GFP K4 CHO cells, which were carried out in 50 mL filtered cap centrifuge tubes with a working volume of 5 mL. The cells were given by Gemma Ruiz.

A set of different centrifuge tubes, where each tube had different concentrations of amino acids, was performed. Thus, the different behaviour of the system could be analyzed by looking at the distinct reaction pathways in the metabolic network in different conditions.

The cultivation of each centrifuge tube was carried out during several days until the cells attained steady state. The cells needed some days to adapt to the new conditions. In normal situations, the steady state was reached after two or three days of the beginning of the experiment.

Since the Amino Acid Analysis (AAA) was performed during steady state, the experiment could last as long as desired to obtain satisfying results and in a same run the results taken every day were (in principle) repeat of the results obtained the other day after the steady state was reached. Increasing the number of data points per run allowed to improve the quality of the information. The duration of the experiment depended also on how the cells grew.

4.3. Media preparation

The media, which were used in this project, had been developed in a project in the KTH Division of Bioprocess by Gemma Ruiz. The main objective of this project was to develop a medium whose composition was perfectly known and had good performance, meaning that the cells would grow twice every day.

The project in question took as a base medium the composition of the known medium DMEM/F12. Then, additions of components already present or new components were done in order to improve the performance of the original medium. The additions were suggested and tested by Gemma Ruiz before they were accepted.

A first experiment, called RMB03.2, was done using a medium called SF10 (see Appendix B for its composition). The results, as it can be seen in the following section, were not as good as expected, so a second medium called SF14 that has a different composition than SF10 (see Appendix B) was used.

The main differences between these two medias are the following ones:

- The media contained different kind of insulin. SF10 contains pancreatic bovine insulin, while SF14 contains recombinant human insulin.
- SF14 contained Hypoxanthine and Thymidine, which are compounds that help to the duplication of the cells DNA.
- And other components whose concentrations were changed between the first and second media.

Table 4.3. Differences between media

	Concentration (mg/L)	
	SF10 media	SF14 media
Ferric citrate	-	0.25
Hypoxanthine	-	20
Insulin	15	5
Myo-inositol	12.6	70.6
Nicotinamide	2.02	6.02
Putrescine	0.081	1.031
Thymidine	-	5

Andreas Andersson performed the first experiment during the realization of his Thesis. The second experiment was performed during the present Thesis.

Once the final composition of the medium was decided the following step was the preparation of the medium itself. The mixture of the different components had some problems because not all the components could be added directly to the volume established either because the concentration of the different components were impossible to weight with the scales available in the department or because they were not all soluble in the same bases.

Hence, in order to avoid these problems most of the components were prepare in stock solutions before they were added to the final medium in dH_2O, HCL 1 mM, NaOH 1 mM or ethanol 99%. A small amount of components were added directly to the medium. So, as explained before, all the components were added to the final volume except the amino acids. Following this, the base medium was aliquoted in small volumes and the specific concentrations of the amino acids were added.

During the preparation of the medium, the pH and the osmolarity were controlled due to the fact that the cells are only able to grow in specific conditions. In the present case, the range of the pH was between 6.9 and 7.1 and for the osmolarity between 290 and 340 mOsm/Kg. The pH could vary significantly due to the addition of basic stock solution or acidic stock solutions. The medium contained phenol red, pH indicator. So when the medium turned yellowish the pH level was low and when it became pink/purple the pH was too high. In the first case the pH was increased adding some drops of HCL 1mM and in the second case adding some drops of NaOH 1mM.

4.4. Methodology

The main idea of the experiment was to inoculate the cells at a specific cell density (Cv) and renew the culture medium every day by withdrawing used medium and adding with fresh medium, while maintaining the same volume and Cv. According to that, the methodology in the first experiment was as follows:

Day 1. The cells were inoculated at cell density 1 MVC/mL in a volume of 5mL medium was added. Then the tubes were put in the incubator (36,5 °C, 200 rpm, 5% CO_2).

Day 2. One sample was taken to determine the cell density and growth rate (in a perfect case the cell density had to be twice the previous day). Following, a calculated volume of cell broth was discarded in order to have a cell number in each tube of 5 MVC/mL. The remaining culture was centrifuged 5 min. at 1000 rpm (100 g.). Finally, the supernatant was aliquoted and stored at -20 and fresh medium was added to obtain a final volume of 5 mL, giving a cell density of 1 MVC/mL.

Following days. The same procedure as day 2.

The Bioprofile was used to determine the cell density and viability as well as the concentration of some metabolites: Glc, Lac, Gln and Glu. The rest of the concentrations were determined using the HPLC.

The second experiment followed almost the same procedure and had the same experimental conditions, but some changes were made in order to improve the performance of the experiment. The temperature was increased to 37 °C because it was checked that inside the tubes the temperature decreases around 0.5 °C and the cells do not grow properly when the temperature is that low.

Furthermore, the working Cv was changed between both experiments. In RMB03.02 1 MVC/mL was the decided working cell density, thinking that it was high enough to be able to see the different behaviour in the metabolism of the cells. But according to the results, these changes were not always distinguished so the Cv was increased, 1.5 MVC/mL, in the second experiment to make them possible to detect.

Taking into account all the previous aspects, the methodology in experiment 2 was as follows:

Day 1. The cells were inoculated at Cv 0.5 MVC/mL and a volume of 5 mL of medium was added. Then the tubes were put in the incubator (37 °C, 200 rpm, 5% CO_2).

Day 2. The cells were let to grow to reach the desire Cv, 1.5 MVC/mL, so the culture was only centrifuged and 5 mL of fresh medium was added.

Day 3. Same procedure as day 2 in experiment 1.

Day 4. One sample was taken to determine the cell density. The culture was centrifuged and 5 mL of fresh medium was added.

Day 5. Same procedure as day 3 in experiment 2 but using a Cv of 0.8 MVC/mL. In some CT only the sample was taken and they were left again in the incubator in order to let the cells recover.

Following days. Same procedure as day 3 in experiment 2.

Finally, both experiments had different set of CT. The variation in the amino acids must be large enough to produce significant changes between the distinct centrifuges tubes, causing the cells to take other pathways in the metabolic reaction network. For that reason, in RMB03.02 some amino acids were not added to the medium since this was the largest change that could be done.

The different compositions of the amino acids can be seen in Table 4.4. When it says 100% it means that the whole amount of the amino acid was added and when it says 0% it means that the amino acid was not added.

Table 4.4. AA composition the media for RMB03.02 in % of their concentration in the medium

RMB03.02							
Components	**CT_1**	**CT_2**	**CT_3**	**CT_6**	**CT_7**	**CT_9**	**CT_{12}**
Alanine	0	100	100	100	100	100	100
Asparagine	100	0	100	100	100	100	100
Aspartate	100	100	0	100	100	100	100
Cysteine	100	100	100	100	100	100	100
Glutamate	100	100	100	0	100	100	100
Glutamine	100	100	100	100	100	100	100
Glycine	100	100	100	100	0	100	100
Proline	100	100	100	100	100	100	100
Serine	100	100	100	100	100	0	100
Tyrosine	100	100	100	100	100	100	100
EAA [4]	100	100	100	100	100	100	100

The same idea was used to determine the set of centrifuges tubes for Experiment 2. But in this case, the effect in the cells of adding only half of the concentration (50%) of the amino acids used in RMB03.02 was also studied, bearing in mind that the cells need all the essential amino acids in the medium to grow. Thus, the final composition of the tubes in that second experiment was:

[4] Arginine, Histidine, Isoleucine, Leucine, Lysine, Methionine, Threonine, Tryptophan and Valine

Table 4.5. AA composition the media for RMB03.02 in % of their concentration in the medium

Experiment 2														
Comp.	**CT_1**	**CT_2**	**CT_3**	**CT_4**	**CT_5**	**CT_6**	**CT_7**	**CT_8**	**CT_9**	**CT_{10}**	**CT_{11}**	**CT_{12}**	**CT_{13}**	**CT_{14}**
Ala	**0**	**50**	100	100	100	100	100	100	100	100	100	100	100	100
Asn	100	100	100	100	100	100	100	100	100	100	**0**	**50**	100	100
Asp	100	100	100	100	**0**	**50**	100	100	100	100	100	100	100	100
Cys	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Glu	100	100	100	100	100	100	**0**	**50**	100	100	100	100	100	100
Gln	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Gly	100	100	100	100	100	100	100	100	**0**	**50**	100	100	100	100
Pro	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Ser	100	100	**0**	**50**	100	100	100	100	100	100	100	100	100	100
Trp	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Tyr	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Other EAA [5]	100	100	100	100	100	100	100	100	100	100	100	100	**50**	100

[5] Arginine, Histidine, Isoleucine, Leucine, Lysine, Methionine, Threonine and Valine

4.5. HPLC

The instrument used in order to determine the concentration of amino acids was the High Pressure Liquid Chromatography (HPLC). The technique followed was the pre-column derivatization and reversed phase.

The system consisted in 3 Water 510 Pumps, a WISP autoinjector, a column heater and one Waters 486 UV-detector. Three elution buffers were used, one for each pump. Eluent A was obtained with 100 mM NaAc and 5.6 mM Triethylamine and followed by adjusting the pH to 5.7 using 50% phosphoric acid. Eluent B was obtained with 100 mM NaAc and 5.6 mM Triethylamine and followed by adjusting the pH to 6.8 using 50% phosphoric acid. Eluent C consisted in MeCN at 100%.

The derivatization of the amino acids was carried out using Waters AccQ-Fluor reagent 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate (AQC) and they were detected using a UV detector at 254 nm and separated on a C18 column.

In this study the proteins and peptides were not object to study. For that reason, the samples were precipitated before they were analyzed in the HPLC to avoid the interference of these components in the chromatogram. The TCA precipitation protocol designed by [45] was used and it is as follows:

- TCA was added to the samples in order to reach after precipitation a TCA concentration of 0.03M.
- The samples were after incubated at room temperature during 20 min.
- The samples were centrifuged at 13K rpm during 10 min. Following, the supernatant was diluted a factor D and the pellet was discarded.

Once the samples were precipitated, they were mixing with AccQ·Fluor Borate Buffer and AccQ·Fluor Reagent. An internal standard, α-aminobutyric acid (AABA)

was added to the samples mixture, which helped in the quantification of the different peaks of the chromatogram.

The calculation of the concentration of the amino acids were made using the areas of the different peaks of the chromatogram, as follows:

$$C_{aa} = D \cdot C_{st} \cdot \frac{A_{sa} \cdot I_{st}}{A_{st} \cdot I_{sa}} \tag{28}$$

where

C_{aa} is the concentration of amino acid in sample

D is the dilution factor (if applicable)

C_{st} is the concentration of amino acid in standard

A_{sa} is the area of amino acid in sample

I_{ist} is the area of internal standard in standard

A_{st} is the area of amino acid in standard

I_{sa} is the area of internal standard in sample

As it can be seen, a standard is required to determine the concentration of the amino acids. In this case, the standard had a concentration of each amino acid of 100 μM.

4.6. Modelling

The modelling was performed using the software Matlab (MathWorks, Version 7.9). The Metatool was used as it was explained in the introduction part. The modelling methodology was described in the Introduction[6]. A user-friendly modelling tool was developed.

[6] Notice however that in RMB03.02 Eq. 4 was used to calculate the q_i while in Experiment 2 Eq. 6 was used more appropriately since a pseudo-perfussin is applied and $\frac{d\underline{C}}{dt} = \underline{0}$

5. RESULTS

5.1. Experiment 2

The experiment was carried out to obtain data to be used in the mathematical model. The culture was done varying the aa concentration in the medium in the different tubes in order to trigger different metabolic pathways.

At the beginning, 1.5 MVC/mL was thought to be the working Cv. However, between day 2 and 4 there were a lack of CO_2, affecting the well growth of the cells. After this unexpected setback, the Cv was changed to 0.8 MVC/mL in order to have the same conditions in the following days of the experiment.

The cells started to resume after the problem with CO_2. The recovery was already seen in the fifth day of the culture. Nevertheless, this effect was not visible in all the tubes and for that reason the medium was not renewal in these CT's. The tubes affected were CT_2, CT_4, CT_8, CT_9, CT_{11} and CT_{13}. Fortunately, the main methodology of these tubes could be continued again after that day.

The seventh day showed that, even the previous day they seemed that they were getting better, CT_9 and CT_{13} did not recover at all, maintaining a very low Viability and not growing. Accordingly, it was decided not to continue with them after this day.

In this case, not all the samples were analyzed in the HPLC because it was known, from previous experience, that some data would not be able to be used by the simulation program. Therefore, a selection of samples was made taking two aspects as requirements. The first one was that the cells grew from the previous day and the second one that the viability was above a reasonable threshold. It was decided that when the cells had a Viability above 75% was reasonable to be analyzed by HPLC.

Therefore, taking into account the two previous requirements, the selection of samples was the following one:

Table 5.1. Decision of the days to make the AAA

	D1	D2	D3	D4	D5	D6	D7	D8	D9
CT_1	NA	NA	NA	NA	NA	A	A	A	
CT_2	NA	NA	NA	NA	NA	A	A	A	A
CT_3	NA	NA	NA	NA	NA	A	A	A	A
CT_4	NA	NA	NA	NA	NA	A	A	A	
CT_5	NA	NA	NA	NA	A	A	A	A	
CT_6	NA	NA	NA	A	A	A	A	A	
CT_7	NA	NA	NA	NA	A	A	A	A	
CT_8	NA	NA	NA	NA	NA	A	A	A	A
CT_9	NA	NA	NA	NA	NA	NA	NA		
CT_{10}	NA	NA	NA	NA	A	A	A	A	
CT_{11}	NA	NA	NA	NA	A	A	A	A	
CT_{12}	NA	NA	NA	NA	NA	NA	A	A	A
CT_{13}	NA	NA	NA	NA	NA	NA	NA		
CT_{14}	NA	NA	NA	A	A	A	A	A	A

NA Not-anal.

A Analyzed

The problems that there were during the AAA of RMB03.02, where different peaks of aa were eluated together, disappeared in this AAA. In this case, and after two runs of the samples, all the concentrations of all the aa were obtained. However, other problems arose.

Once the qi were calculated, it was noticeable, as in RMB03.02, that the concentrations of some eaa were higher than the concentration of them in fresh medium. This is not feasible since the cells could not produce an eaa.

Another important aspect that must be mentioned is that it seemed that the analysis of some aa was altere when the samples of spent media were frozen for store and defrozen to be analyzed. This was, for example, the case of Gln and the values could be compared with the values of the Bioprofile. It can be said that some values of Gln in the HPLC were not reliable, giving very high values comparing with the same

value in Bioprofile. For that reason, the concentration of Gln obtained in the Bioprofile was used. Other example was NH_4 and the same solution was taken.

All of these dilemmas led to think that the AAA done by the HPLC was not totally trustable and this problem is discussed a little bit further in the discussion section.

Despites the problems encountered with the AAA, several days that could be used in the simulation program. In total, twelve samples fulfilled the condition that an eaa is consumed in the system (CTiDj: "sample of centrifuges tube i of day j"):

- **CT1D06** and **CT1D07**: no Ala was added to the medium.
- **CT2D09**: half of the original concentration of Ala was added to the medium.
- **CT4D06**: half of the original concentration of Ser was added to the medium.
- **CT5D06**, **CT5D07** and **CT5D08**: no Asp was added to the medium.
- **CT6D06**, **CT6D07** and **CT6D08**: half of the original concentration of Asp was added to the medium.
- **CT10D05** and **CT10D08**: half of the original concentration of Gly was added to the medium.

All the data and calculations of Experiment 2 can be found in Appendix D.

5.2. Simulation results of Experiment 2

The model was applied using the stoichiometric matrix that can be found in Appendix A and the samples above mentioned. The model determined that the system had 143 EFM.

A first aspect that can be mentioned is that the inequation in Eq. 24 that must be fulfilled in order to get the optimal solution when applying the non-negative least-square was accomplished:

$$p \cdot NS = 25 \cdot 12 = 300 > N = 143$$

In Tables 5.2 and 5.3 the obtained reduced system are presented. Only the reactions of the macroscopic system that have the maximal kinetic rate above zero are presented.

Table 5.2. Maximal kinetic rates of the reduced model

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7	μ_8	μ_9
477.664	2703.740	130.400	63.384	158.107	854.224	55.522	391.418	549.279
μ_{10}	μ_{11}	μ_{12}	μ_{13}	μ_{14}	μ_{15}	μ_{16}	μ_{17}	μ_{18}
477.664	2703.740	130.400	63.384	158.107	854.224	55.522	391.418	549.279
μ_{19}	μ_{20}	μ_{21}	μ_{22}	μ_{23}	μ_{24}	μ_{25}	μ_{26}	μ_{27}
29.200	1182.876	406.646	183.080	291.968	17.192	148.692	661.105	286.323
μ_{28}	μ_{29}	μ_{30}	μ_{31}	μ_{32}				
249.944	105.417	53.959	1704.1	157.198				

Table 5.3. Reduced stoichiometric matrix of the macroscopic system

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24	R25	R26	R27	R28	R29	R30	R31	R32
Glc	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	-0.5	-0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0208	0
Lac	-1	2	3	1	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Gln	-1	0	0	1	0	-1	-2	0.5	0	0	1	0	0	0	0	0	-1.5	-1	-1	1	1	1	1	0	0	1	0	0	1	0	0	0
NH4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	-1	-0.0377	0
Asp	1	0	2	0	0	0	0	1	0	-1	0	1	0	0	-1	0	1	1	1	0	-1	0	0	0	0	1	1	1	1	1	-0.0006	0
Glu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	-0.007	0
Ser	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	-0.0033	0
Asn	0	0	-1	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0084	0
Gly	0	0	0	0	0	0	0	0	-0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0133	0
His	0	0	0	-1	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0101	0
Thr	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	-0.0033	0
Arg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	-0.0055	0
Ala	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.008	0
Pro	0	0	0	0	0	0	0	-0.5	0	0	0	0	0	0	0	0	-0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.004	0
Tyr	0	0	0	0	-1	0	0	0	0	0	0	0	0	-1	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	-0.0096	0
Cys	0	0	-1	0	-1	1	0	0.5	0	1	0	0	0	0	0	0	1.5	0	0	-1	0	0	0	0	0	0	0	0	0	0	-0.0133	0
Val	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Met	0	0	0	0	1	0	2	-1	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-0.026	0
Iso	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	1	0	0	0	0	0	0	0	0	-0.0004	0
Leu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	-0.0165	0
Lys	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	-0.0081	0
Phe	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	-1	0	1	0	0	0	0	0	0	-0.0099	0
Tryp	0	0	-1	0	0	0	-1	0	0	0	0	0	-1	0	0	0	-1	-0.5	0	0	0	0	0	0	1	0	0	0	0	0	-0.0077	0
Biomass	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1
Death	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

The next figures compare the experimental data specific consumption/production rates $q_{exp,}$ (blue line) using Eq. 8 and the estimated ones by the reduced model q_{est} (red line) using Eq. 23. Each plot represents one day of one CT. The consumption/production rates of all the metabolites are represented by one data point each. The metabolites are given in X-axis and each one is associated with a number listed in Table 5.4.

Table 5.4. Number of the metabolites

1.	Glc	**14.**	Thr
2.	Lac	**15.**	Trp
3.	NH_4	**16.**	Val
4.	Gln	**17.**	Ala
5.	Biomass	**18.**	Asn
6.	Glu	**19.**	Asp
7.	Arg	**20.**	Cys
8.	His	**21.**	Gly
9.	Ile	**22.**	Pro
10.	Leu	**23.**	Ser
11.	Lys	**24.**	Tyr
12.	Met	**25.**	Cell death
13.	Phe		

Normally, the first four external metabolites of the system (Glc, Lac, NH_4 and Gln) and also the biomass have a much higher specific consumption/production rates than the rest. Therefore, it was decided to plot them separately from the other external metabolites. In this manner it was easier to evaluate the performance of the model.

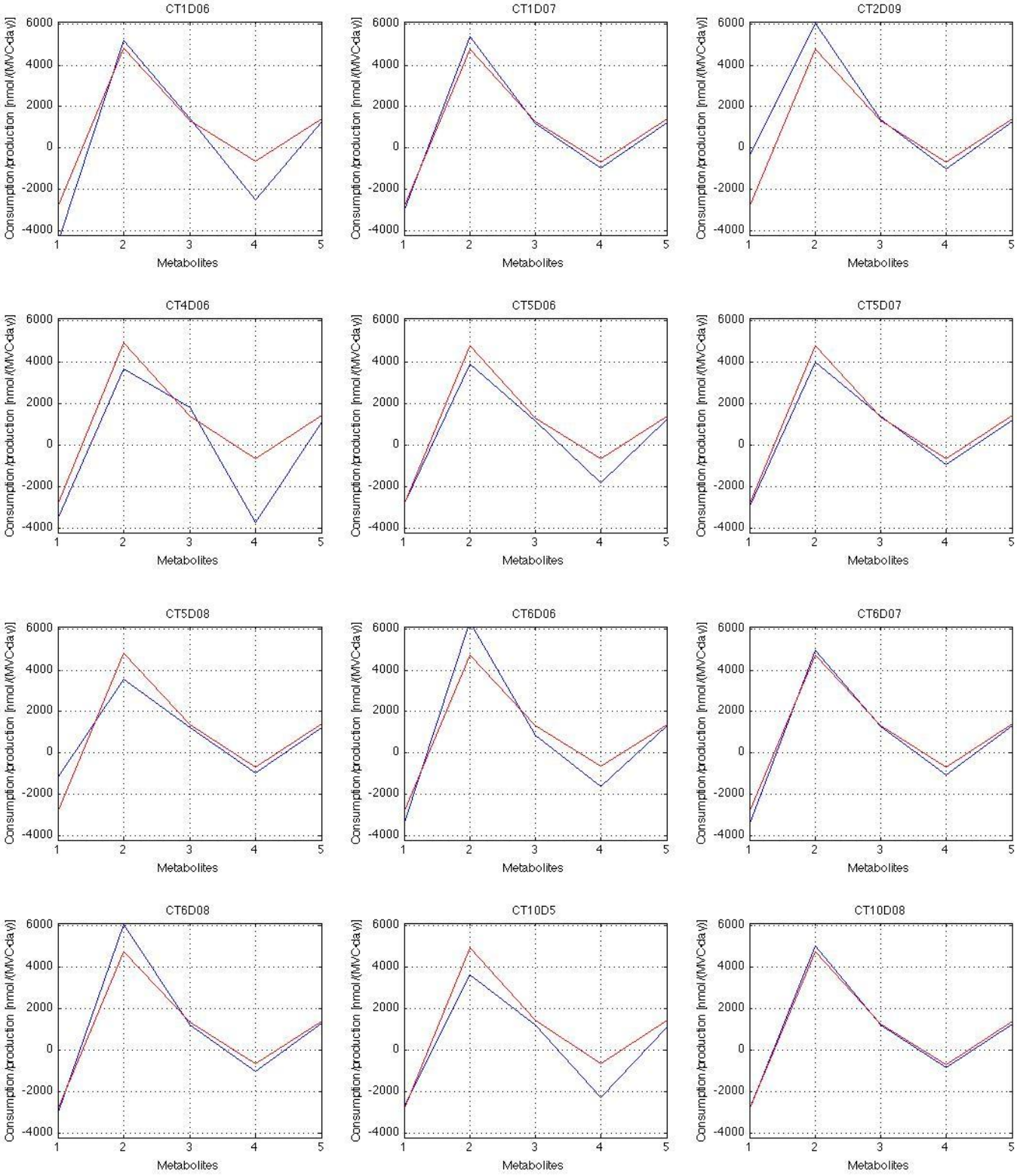


Figure 5.1. Comparison between q_{exp} (Eq. 8) and q_{est} (Eq. 23) for Glc, Lac, NH_4, Gln and Biomass

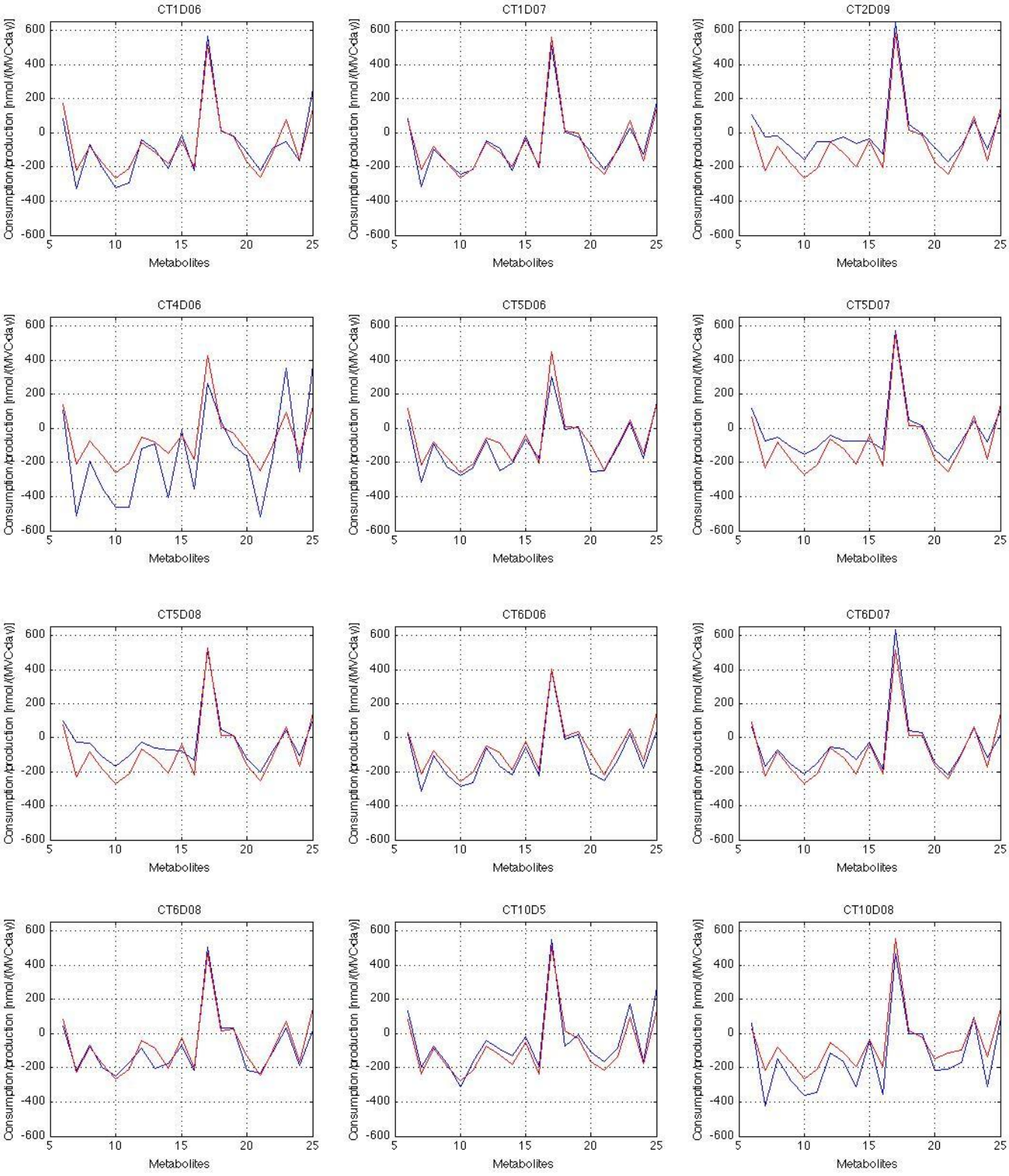


Figure 5.2. Comparison between q_{exp} (Eq. 8) and q_{est} (Eq. 23) for the other external metabolites

These plots gave a qualitative point of view. In order to quantify, the average error of each metabolite between both rates was calculated:

Table 5.5. Average error between q_{exp} (Eq. 8) and q_{ext} (Eq. 23) for all the external metabolites (%)

Glc	Lac	NH_4	Gln	Glu	Arg	His	Ile	Leu	Lys
78	20.4	13.8	46.8	57.4	178.8	71.7	37.7	30.2	64
Met	**Phe**	**Thr**	**Trp**	**Val**	**Ala**	**Asn**	**Asp**	**Cys**	**Gly**
50.4	77	71.1	96	32.2	16.4	153.4	541.3	45.6	23.9
Pro	**Ser**	**Tyr**	**Biomass**		**Cell death**				
30	81.3	41	14.5		151.3				

It can be seen that in general the error between the estimated value and the experimental value is not very high. However, there are 4 metabolites that had a highly unexpected error. They appear in red in Table 5.5 and they are Asn, Asp, cell death and Arg. It was difficult to explain this phenomenon and in a first approach, it could be said that something did not work well in the model for these components. For that reason, the error in each sample was studied. Tables 5.6.a and 5.6.b show these errors.

Table 5.6.a. Error between q_{exp} (Eq. 8) and q_{est} (Eq. 23) in each sample and metabolite (%)

	CT1D06	CT1D07	CT2D09	CT12D06	CT5D06	CT5D07
Glc	39.10	6.71	-677.43	20.27	-0.27	4.45
Lac	7.30	11.11	20.77	-34.36	-22.36	-19.65
NH4	8.45	-10.69	2.80	24.14	-14.13	2.26
Gln	74.09	30.07	32.64	82.50	64.82	28.54
Glu	-103.56	14.39	61.28	-28.27	-148.68	42.36
Arg	32.08	31.59	-826.23	59.91	32.63	-221.08
His	-17.95	16.96	-365.73	60.48	17.12	-69.20
Ile	14.12	-1.12	-111.73	54.34	24.97	-75.69
Leu	16.34	-8.85	-69.55	43.16	5.17	-75.23
Lys	28.95	2.04	-310.24	55.88	9.65	-91.75
Met	-46.49	-30.94	3.15	58.30	20.81	-51.44
Phe	-17.03	-23.15	-390.51	9.84	66.57	-66.63
Thr	14.88	9.41	-211.21	64.10	6.97	-185.88
Trp	-305.00	-105.61	-33.44	-191.46	40.89	45.47
Val	5.79	0.61	-63.49	50.08	-16.81	-81.30
Ala	8.97	-10.17	10.54	-59.65	-48.47	2.48
Asn	-56.67	-77.63	71.95	68.56	239.96	69.94
Asp	-57.90	94.74	-56.65	70.08	56.44	45.85
Cys	-47.29	-52.21	-95.94	21.29	62.36	-35.90
Gly	-16.13	-12.62	-42.59	52.02	1.55	-31.88
Pro	-23.02	3.04	-25.92	42.53	13.85	-47.40
Ser	250.52	-145.74	-24.69	73.63	-50.07	-65.99
Tyr	4.29	-32.05	-69.81	40.18	15.34	-124.24
Biomass	4.87	-19.40	-19.27	-9.81	-20.23	-23.93
Cell_death	-20.88	97.59	96.34	23.17	97.11	96.05

Table 5.6.b. Error between q_{exp} (Eq. 8) and q_{est} (Eq. 23) in each sample and metabolite (%)

	CT5D08	CT6D06	CT6D07	CT6D08	CT10D05	CT10D08
Glc	-139.21	17.83	17.62	6.41	-5.32	-0.79
Lac	-35.13	25.18	3.92	21.32	-37.25	5.60
NH4	-10.86	-53.24	-5.90	-12.01	-19.48	-2.91
Gln	28.99	61.31	36.22	34.53	71.03	21.60
Glu	24.56	-41.73	-44.17	-83.38	37.31	58.48
Arg	-802.15	33.07	-32.99	-4.83	-20.20	48.46
His	-183.08	29.45	-16.08	-20.75	-17.70	45.66
Ile	-64.42	24.08	-21.13	10.15	-10.94	39.36
Leu	-61.26	10.07	-26.70	-6.95	12.04	26.84
Lys	-107.21	23.36	-42.82	-23.07	-33.62	39.21
Met	-162.44	21.29	-10.68	50.90	-93.35	55.26
Phe	-101.88	47.74	-71.00	56.48	-40.90	32.25
Thr	-191.43	14.13	-66.67	-16.53	-34.13	37.47
Trp	55.00	60.67	-35.44	63.26	-213.22	-2.03
Val	-60.59	16.88	-15.26	6.97	-23.10	45.64
Ala	-3.00	-2.09	18.34	6.97	6.79	-19.77
Asn	72.84	211.42	70.38	59.89	122.72	719.28
Asp	-22.39	-101.91	52.96	25.92	-258.49	-5665.12
Cys	-36.38	58.48	-11.43	38.86	-54.84	31.88
Gly	-23.47	14.42	-10.26	-5.12	-29.70	47.08
Pro	-39.00	42.73	-8.97	12.92	-59.10	42.18
Ser	-68.07	-126.78	-8.07	-107.42	44.54	10.00
Tyr	-60.19	24.18	-47.35	11.82	-5.75	56.66
Biomass	-19.75	-11.40	-13.55	-13.87	-6.89	-17.90
Cell_death	95.28	86.25	72.28	75.12	-14.62	94.73

Finally the average error was recalculated using the values that had an error below 100% and they are presented in Table 5.7:

Table 5.7. Average error between q_{exp} (Eq. 8) and q_{est} (Eq. 23) with the error below 100% (%)

Glc	Lac	NH_4	Gln	Glu	Arg	His	Ile	Leu	Lys
11.9	20.3	47.2	43.6	32.9	32.9	31.1	30.1	30.2	35.1
Met	**Phe**	**Thr**	**Trp**	**Val**	**Ala**	**Asn**	**Asp**	**Cys**	**Gly**
40.2	43.2	29.4	42	32.2	16.4	68.5	53.7	45.6	24
Pro	**Ser**	**Tyr**	**Biomass**		**Cell death**				
30.1	58	41	14.5		151.3				

The difference between the estimated values (red line) and the experimental values (blue line) could also be studied using Fig. 5.3, where each plot corresponds to one

metabolite and each plot had all the samples, i.e CT1D06, CT1D07…, order from 1 to 12 in the same order as they were presented before:

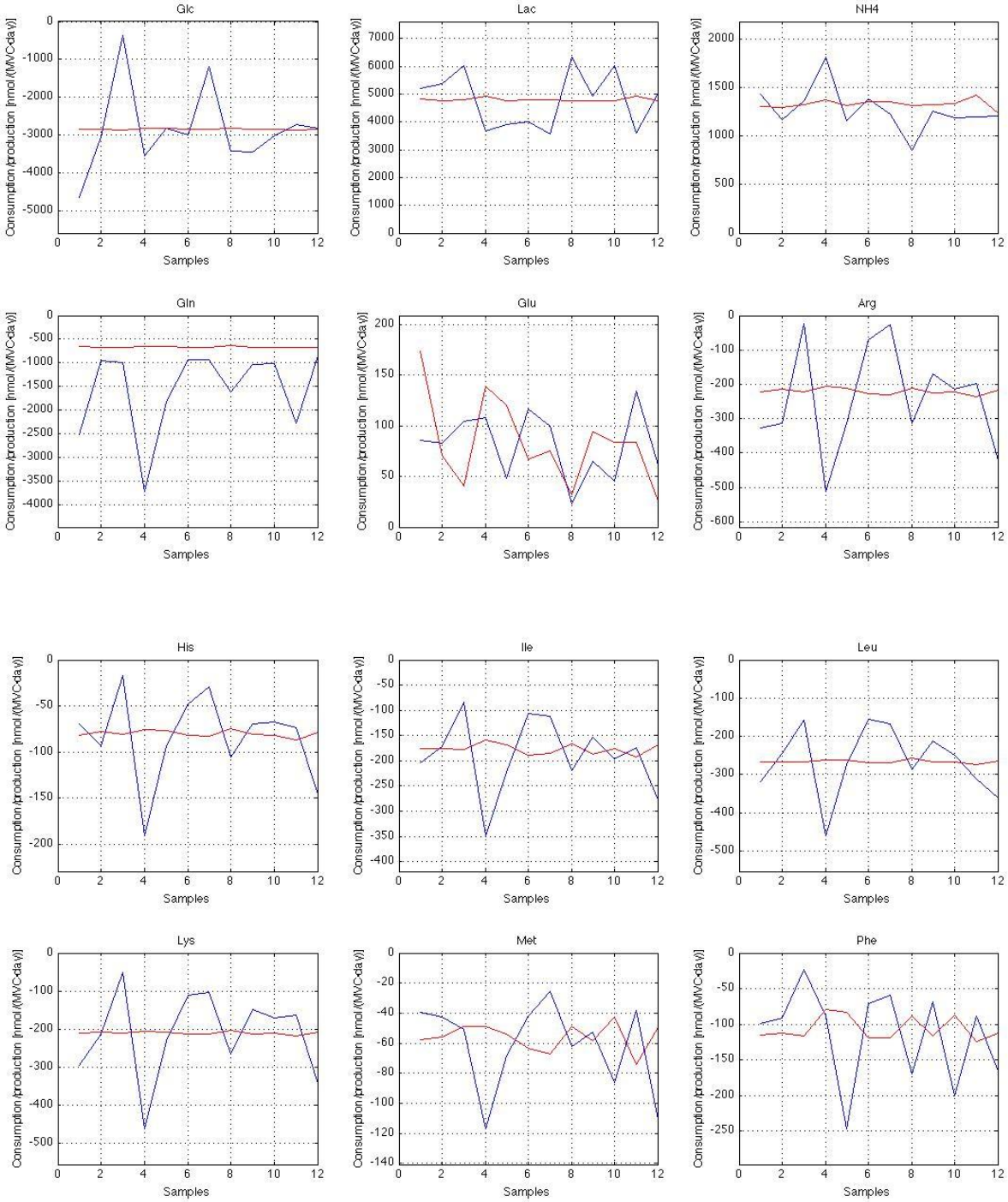


Figure 5.3.a. Comparison between q_{exp} (Eq. 8) and q_{est} (Eq. 23) divided in samples

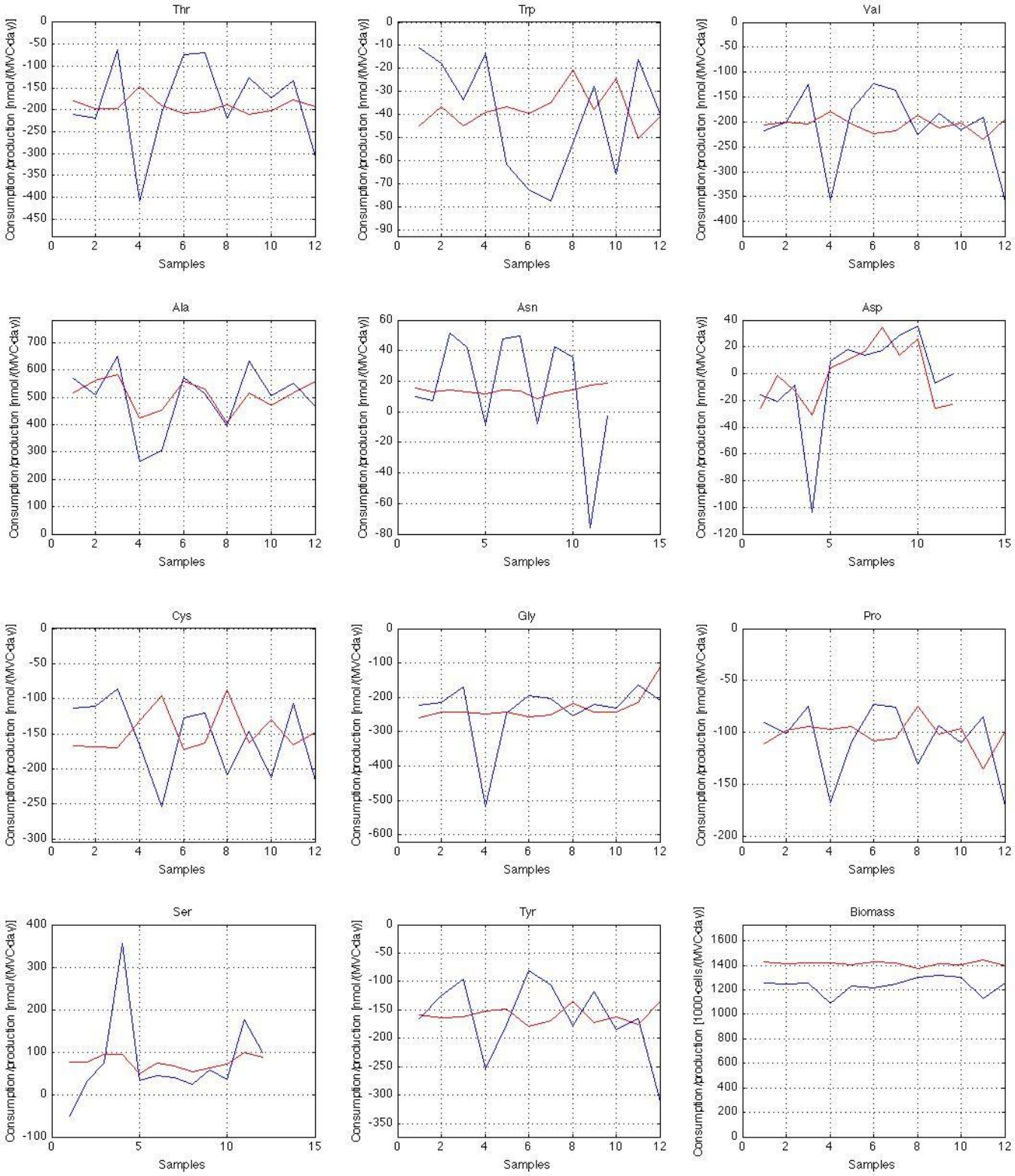


Figure 5.3.b. Comparison between q_{exp} (Eq. 8) and q_{est} (Eq. 23) divided in samples

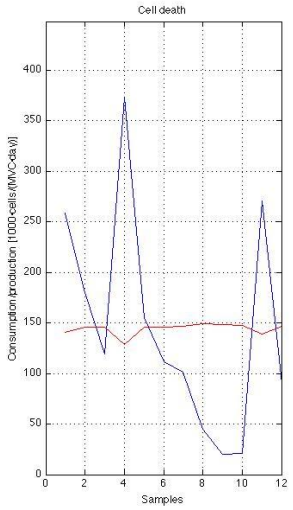


Figure 5.3.c. Comparison between q_{exp} (Eq. 8) and q_{est} (Eq. 23) divided in samples

5.3. Evolution over time prediction

In order to study the quality of the model, a simulation of the concentration in time was done. The main idea of this simulation was to perform one CT culture during one day in time intervals of 0.2 days using Eq. 28. It was decided to take CT_1 and it was simulated from day 6 to day 7. Fig. 5.4 shows the estimated concentrations of the different metabolites (red line) and the experimental data for all the external metabolites (blue line):

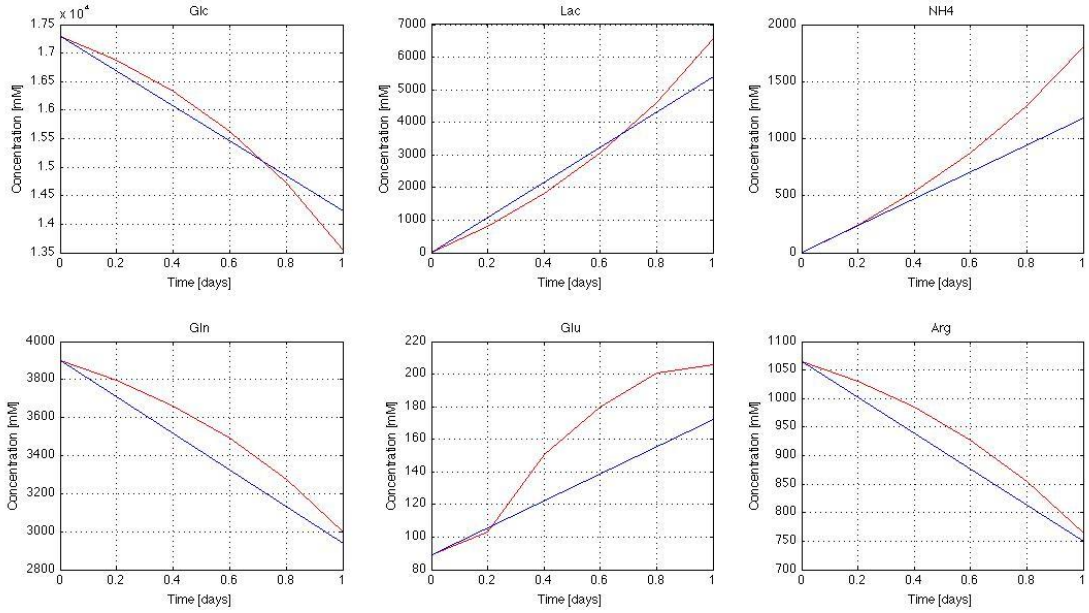


Figure 5.4.a. Time evolution of the concentration of the metabolites from day 6 to 7 in CT_1 (Eq.31)

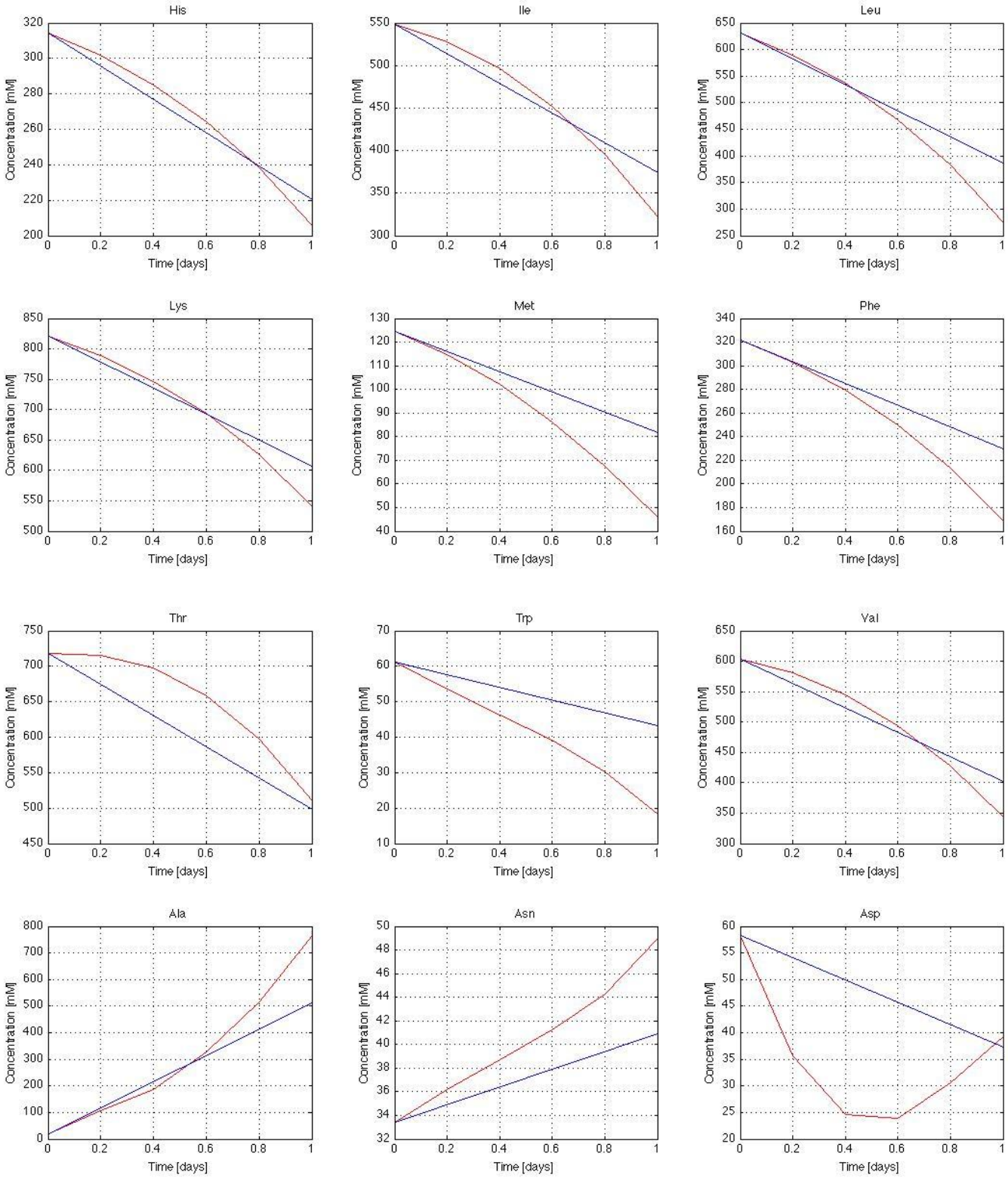


Figure 5.4.b. Time evolution of the concentration of the metabolites from day 6 to 7 in CT_1 (Eq.31)

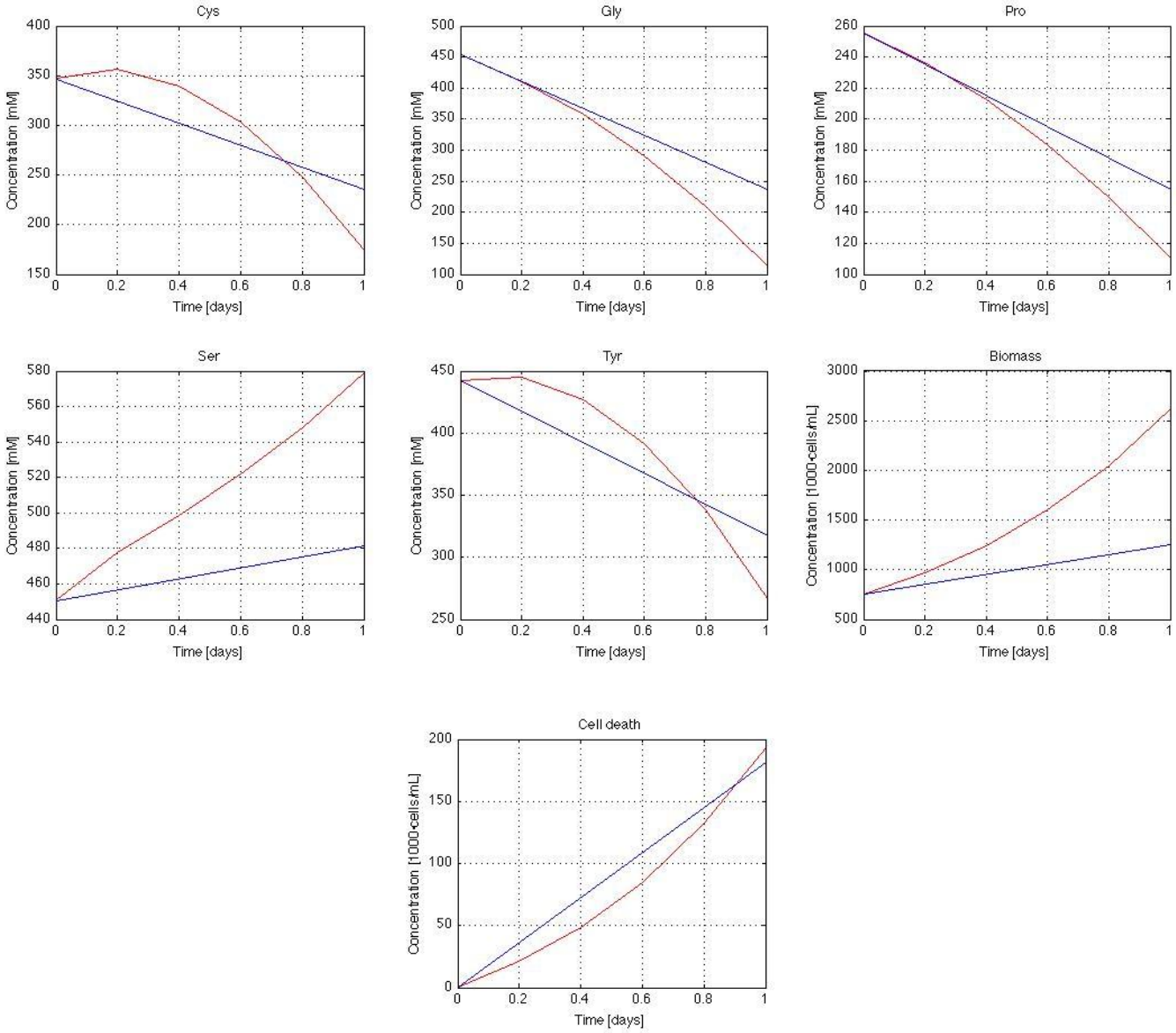


Figure 5.5.c. Time evolution of the concentration of the metabolites from day 6 to 7 in CT_1 (Eq.31)

5.4. Simplified systems

Since the cells cannot synthetize the eaa, these aas are used scarcely by the cells. Consequently, the uptakes of all the eaa are usually very well correlated with each other. In other words, in term of modelling, they bring the same information.

This property was exploited in order to use more experimental data. As a matter of fact some of the experimental data were not used in the complete system since the eaa consumption/production rates were negative. It was tested to discard some eaa.

As a result, it was decided to simplify the model discarding some eaa and be able to use more samples as input parameters. The possibility to use samples of the tubes that were not used in the complete system was the criteria followed to determine these new simplified systems.

Two simplified systems were created and two eaa were discarded in each system:

- Simplified system 1: Trp and Lys were discarded.
- Simplified system 2: Trp and Phe were discarded.

Both systems used the same samples as the complete system as well as other samples that fulfilled the eaa requirement for the eaa that were still in the system:

- Simplified system 1: CT3D07, CT6D04, CT10D06 and CT10D07
- Simplified system 2: CT7D08, CT8D09, CT10D06, CT10D07, CT11D09 and CT12D08.

Table 5.8 and 5.9 show the average error of all the metabolites present in the simplified system 1 and the simplified system 2:

Table 5. 8. Average error between q_{exp} (Eq. 8) and q_{est} (Eq. 23) for the external metabolites in the first simplified system (%)

Glc	Lac	NH_4	Gln	Glu	Arg	His	Ile	Leu	Met
96	19.8	27.6	66.5	91.9	174.0	86.3	39,6	35,3	47,8
Phe	**Thr**	**Val**	**Ala**	**Asn**	**Asp**	**Cys**	**Gly**	**Pro**	**Ser**
121.1	80.0	41.5	16.7	108.1	607.8	44.4	52,7	33,5	97,2
Tyr	**Biomass**		**Cell death**						
47.6	18.5		124						

Table 5.9. Average error between q_{exp} (Eq. 8) and q_{est} (Eq. 23) for the external metabolites in the second simplified system (%)

Glc	Lac	NH_4	Gln	Glu	Arg	His	Ile	Leu	Lys
80.3	18.4	13.2	44.4	66.7	219.4	237.0	44,0	33,7	57,9
Met	**Thr**	**Val**	**Ala**	**Asn**	**Asp**	**Cys**	**Gly**	**Pro**	**Ser**
59.3	97.2	34.0	21.4	140.7	248.9	41.9	36,0	22,6	161,9
Tyr	**Biomass**		**Cell death**						
40.9	21.9		118						

5.5. Graphical interface

A mathematical model that gives the prediction of the evolution of the medium compound was studied in this project. The main objectives was to test the model using different experimental data and check its validity,

The model, which was implemented by Joan Gonzalez [40], was done using Matlab as a programming language. Since previous knowledge of Matlab was required to use it, a Window application was done during the realization of this project. The main purpose of this Window application was to facilitate the user its manipulation, without the need for programming skills.

The former modelling method was divided in two main parts. The first part, called Method Tools, performed the calculation of the model itself. The second part used the parameters calculated by Method Tools to predict the consumption/production rates q_i and represent them graphically. Both parts used the methodology that was presented in section 2.5.

The Window application is a very easy tool that has the same functionality that the former modelling method. Some functions were changed to make the code simpler and more understandable. New functions were also added in order to have more tools to evaluate the behaviour of the system object to study and the validation of the model selected.

When the user opens the program, the user will find a window like the following one:

Figure 5.6. Graphic interface

The main menu that appears in the upper part of the window is divided in 4 submenus:

- Method Tools
- Result Tools
- Network
- Help

5.5.1. Method Tools

This submenu, which is the responsible of the creation of the model, has different tabs.

Clicking in 'See system', a small drop-down menu appears in the upper left corner and the user can choose the input parameter to be see, which are necessary for the program to determine the model:

- **Aint**: stoichiometric matrix of the external metabolites.
- **Aext**: stoichiometric matrix of the external metabolites.
- **Irreversibility**: a vector with length equal to the number of reactions, where each element indicates if the reaction is reversible with a 1 or if it is not with a 0.
- **Michaelis constant**: a vector with length equal to the number of external metabolites, where each element indicates the Michaelis constant of each external metabolite.
- **Concentrations**: matrix where each element $C_{i,j}$ contains the concentration of the metabolite i of the sample j.
- **Q**: matrix where each element $Q_{i,j}$ contains the consumption/production rate of the metabolite i of the sample j.

It is also possible to introduce a new system clicking in 'New system configuration'. The program is made so that the user only can enter the parameters in a predetermined order, without being able to enter the next until the user has introduced the preceding ones. The order is the same as the above description of the parameters.

The parameters have to be introduced directly, typing all the values. It is necessary to introduce some values before defined matrices can be created. These values are the number of reactions and number of samples, as well as the number of internal and external metabolites. The user has to introduce them when the program requires them. An example can be when the user has to introduce the stoichiometric matrix of the internal metabolites (*Aint*), the number of reactions and the number of internal reactions. In the case of *Aint* and *Aext* it is also necessary to introduce the name of the metabolites, in order to establish the order in the stoichiometric matrix.

If they are not introduced the matrix will not be accepted as valid and the user will not be capable to go on in the introduction of the input parameters. The same happens in the introduction of the concentrations, where the sample names have to be inserted.

Each time that a parameter is well introduced, the corresponding button colour turns green.

Finally, the new system can be saved, clicking in 'Save new system'. The user has to enter the path where the system will be saved and each parameter will be saved as an Excel file. They are also saved as .mat, the Matlab format. Thanks to that, they can be used in the future using Matlab.

The last tab of this submenu is called 'Find model' and clicking on it, the model is calculated, introducing first the system that is object to study. The program uses the function called 'findmodel' from [40] to calculate the model, using Eq. 23. The output parameters of the model can be seen, selecting them in the drop-down menu that appears in the upper left corner. A button called 'Save' appears at the bottom of the window and when clicking on it, the model is saved. The model is saved in the same formats as explained above.

The model consists of:

- **Amac**: stoichiometric matrix of the macroscopic reactions, which corresponds to all the elementary flux modes.
- **Maximum rate**: vector of length equal to the number of reactions, whose elements express the maximal kinetic rate in which the reaction can be produced.
- **Elementary modes**: matrix that contains all the elementary flux modes.
- **Number of elementary modes**

- **External reactions**: vector that indicates the reactions that only involve external metabolites.
- **Ared**: the same as Amac but for the reduced model.
- **MUred**: the same as maximum rate but for the reduced model.
- **Error**: error produced when the non-negative least-squares equation is used to determine the maximum rate.

5.5.2. Result Tools

Until the model is calculated, this submenu button remains unable. Once the model is established, the user will be able to interact with this submenu. In it, the user can predict the consumption/production rates and plot the results in very different ways.

The first tab is called 'Prediction of Q' calculates the consumption/production rates Once the user click on 'Prediction of Q', the program shows the estimated q_i in a matrix where each column corresponds to one sample and each row to one metabolite.

After the prediction of the consumption/production rates, the other tabs become enable to be utilized by the user. All of them plot the same but in different forms:

- **Plot Q's samples**: each subplot corresponds to one sample and represents the consumption/production rate of all the metabolites for that sample ordered in the same way as in the stoichiometric matrix.
- **Plot errors**: calculate the error between the consumption/production rates calculated from the experimental data q_{exp} and the estimated ones q_{est}. Each subplot corresponds to a sample. The error is computed using the following formula:

$$error = \frac{q_{\exp} - q_{est}}{q_{average}} \tag{29}$$

where $q_{average}$ is the average of the q_{exp} of the corresponding amino acid in all the samples.

- **Plot absolute errors**: compute the error between the same parameters as plot errors but in another way:

$$Absolute\ error = \frac{q_{\exp} - q_{est}}{q_{\exp}} \tag{30}$$

- **Plot Q's metabolites**: each subplot corresponds to one metabolite and represents the consumption/production rate of all the samples.

In all the cases, the plots are showed in subplots and each figure has at most six subplots. The data from the experiment is presented in blue and all the calculated data in red. At the same time, the values appear in the main window. The different figures are not saved directly by the program, but the user has to do it by itself.

The last tab predicts the evolution in time of the concentration of the different metabolites using the reduced model. It is necessary to introduce some parameters before the program could do the prediction:

- **Co**: vector that contains the initial concentration of the metabolites.
- **Step time:** time increase between two prediction metabolite concentrations.
- **Time of the simulation**: duration of the simulation of the concentration of the different metabolites in days.

And the predictions are done using the following formula:

$$c_{i,k+1} = c_{i,k} + q_{i,k} \cdot \Delta t \cdot X_{v,k} \quad (31)$$

where Δt is the step time and $X_{v,k}$ is the cell density of the system multiplied by 1000.

5.5.3. Network

This tab helps the user to find a metabolite in a system. The user has only to enter the name of the metabolite and the program will show in which reaction it appears, separating it in substrates and products.

5.5.4. Help

Clicking in this submenu, a manual will appear where the user can find a brief description of how the program works.

6. DISCUSSION

6.1. Simulation results Experiment 2

Looking at Fig. 5.1 and 5.2, it can be said that the results are promising since in general the estimated values and the experimental values are very similar or at least they have the same tendency. For example, in CT1D06 and CT1D07 the difference between both rates is practically zero for all the metabolites. But, on the other hand, there are some samples such as CT2D09 and CT5D08 where there are some differences between rates.

As it was said at the beginning, it was thought that the model could not estimate Asn, Asp and Arg correctly since the average error that could be seen in Table 5.5 for these three metabolites were higher than 100%. Nevertheless, when the error was studied separately from all the samples in Table 5.6, it could be seen that this unexpected error did not occur in all the samples. Moreover, the average error is sometimes extremely high due to only one or two samples, when in the rest the error could be acceptable.

CT2D09 and CT5D08 were the samples that had more errors above 100%. It could be seen that the error of some metabolites only occurred in these two samples. These are the cases of Glc, His, Ile, Lys, Met and Phe. It was seen that the experimental rates of most of these metabolites diverge from the rest of the rates for the same CT, being much lower. Furthermore, the rates obtained by the model are more likely to the experimental rates of the other samples than the obtained during the AAA. For all these reasons, it seems reasonable to say that some problems arose in the measurement of these metabolites.

Other two metabolites that had a considerable high error in CT5D08 compared to the other samples are Arg and Thr. However, these two metabolites had also a high

error in CT5D07. This may lead to think that when the Asp concentration in the medium is varied, the cell metabolism takes other pathways that affect the Arg and Thr concentration and the model is not capable to detect them. Nevertheless, these errors are not present in the samples of CT6 where the concentration of Asp in the medium was also modified. Therefore, it is likely to think that this error is caused by other reason. This reason could be the same as explained before, the bad measurement of the concentration of these metabolites.

Besides CT2D09 and CT5D08, there are only a few cases more where the error is above 100%. They seem to be random errors that could occur as a consequence of a bad measurement of the concentrations or to some reactions that are not explained in the model. But, to be able to affirm the second hypothesis more simulations must be done using other set of experimental data. Unfortunately, this is not yet available.

According to all this explanation, it is reasonable to say that the errors above 100% are due to problems in the measurements. For that reason, the error was recalculated using only the errors that were below 100%. This did not affect the cell death.

The results of the new average exposed in Table 5.7 showed that the error diminished for all the samples considerably and in most of the metabolites this error is below 50% which can be consider a well performance of the model. Only four metabolites had an error above 50%: cell death, Asn, Asp and Ser.

According to the results obtained using the average error and Fig. 5.3, there are several aspects that can be mentioned:

- The rate of the biomass is almost all the time very flat; meaning that it does not change too much in the different CT and this is well estimated by the model. The only problem is that the estimated rates of the biomass are always higher than the experimental rates. This could mean that the simplified version of the

biomass reaction works well, but there is something that this reaction could not explain and it happens in the culture of CHO cells.

- The same happens with Gln since the estimated specific consumption/production rates are always higher than the experimental values. The system may lack some reactions to explain this phenomenon.

- Asn rate estimation is one the worse ones and this can be observed in the plot of Asn. The experimental values and the estimated rates diverged in almost all the cases. It means that there are again some effects in the metabolism of the CHO cells that the simplified model cannot explain.

- Even though Asp and Ser have a slightly high error, the tendency that the estimated rates followed is the same as for the experimental rates.

- The cell death was a new metabolite added during the realization of the present Thesis and it was added using a very simple reaction that relates the cell viability with it. Unfortunately, the error and Fig. 6.4 show that there are other aspects that affect the cell death. On the other hand, the system was done without this reaction and the results were very similar.

The trigger of different metabolic pathways in the cell metabolism was accomplished since different consumption/production rates of each metabolite were obtained for the different samples, as it is showed in Fig. 5.3. The capability of the system to detect these changes can be also studied, observing the same figure. The specific consumption/production rate of each eaa should be constant. This is due to the fact that the rates of the eaa depend only on the cell density of the cells and the rates are divided by it, making them constant. This can be seen in Fig. 5.3 for the estimated rates of the eaa, such as Arg and His, but this does not happen in the

experimental values because of the imprecision of the methodology used to calculate them.

In the case of the non-essential aa, the estimated rates are different in different samples and in general the metabolites followed the same trend as the experimental rates. Good examples are Gly and Ala. According to that, it can be said that the system can detect the trigger of the metabolic pathways for the non-essential aa. Unfortunately, this does not happen in the case of Glc, Lac and NH_4 where the corresponding estimated rates remain almost constant for all the samples. This could mean that some reaction are missing in the system that involve these three metabolites or that the system requires more changes in the initial conditions to be able to trigger different pathways for that metabolites.

6.2. Evolution over time

The results shown in Fig.5.4 are very satisfactory since in almost all the cases the simulated values and the experimental values are equal or very similar. The errors of the concentrations after one day are in the range of 10-20% for almost all the metabolites. Moreover, the tendency in all the metabolites is the correct one. The products increase their concentration and the substrates decrease their concentration. Only for Trp and the Biomass the error is around 40%. Concerning these, three aspects can be mentioned.

The first one is the Biomass. The simulated Biomass concentration is so much higher than the experimental one. The cells grew from day 6 to day 7 but not as in theory, where the cells grow twice every day. This could be an explanation of why the Biomass differs from the simulated value.

The other two aspects involve the concentration evolution over time of Trp and Ser. In the case of Trp. it was shown that the simulated value in CT1D07 and the

experimental value were very different. This may be due to an error of the measurement of its concentration. Trp peak is a very small peak in the chromatogram in the HPLC.

Finally, Ser shows a difference between the simulated rate and the experimental value. This, in addition of what was explain before can mean that there is some lack of information in the system that describes the metabolism of Ser.

6.3. Simplified systems

It can be seen from the results obtained in Table 5.8 and 5.9 that both simplified systems had very similar results to the complete system. This shows that some simplifications can be done and the well performance of the model remains. It could be said, that the results could be more reliable since more data were used in the simplified systems.

In the simplified system 2, the His error is very high but it is the same problem observed in the complete system with some metabolites. So, this could also be explained as an error in the measurements of the His concentration.

One important aspect to mention is the case of Ser that was discussed in the previous analysis. In this case, it could be seen that in both cases the error is higher than in the complete system. Since more data is used in these two methods, it is very likely to say that the network does not describe well the metabolic pathways that Ser follows in the cells metabolism, meaning that some reactions are missing in the system that involve Ser. This aspect should be important to study in the future when more data will be available.

In general, the order of magnitude of the error for the rest of the metabolites is the same but there are cases, where the error diminishes. Examples can be Ala and NH_4.

This gives consistency to the method and shows that the network is, for some metabolites, well defined.

6.4. HPLC problems

During the AAA in RMB03.02, it was thought that the fact that the cells did not grow at all could explain all the problems encountered in the eaa. But, as it was said before, these problems did not disappear in Experiment 2 where the cells grew satisfactorily.

Another important aspect that was observable during the AAA, was the difference obtained between the same sample in different runs. It could be reasonable to accept that the error obtained from this difference is below 15-20% since the conditions (temperature. buffers…) from one run to another may differ. On the other hand, in small peaks, and consequently small concentrations, high errors could be also accepted. But in some cases the difference were higher that this acceptable threshold. Diverse hypothesis were thought to explain this phenomenon.

The most likely could be that the precipitation protocol followed with TCA in order to avoid the interferences of proteins and peptides in the AAA did not work properly. One or several non-desired components, which are still unknown, could be eluated together with some aa. This would explain why the concentrations of some eaa's are higher than the concentration of them in fresh media.

This did not happen in the case of fresh medium since these undesired components are not present. Looking at the different runs of the fresh media, the error obtained is, in general, smaller than in the cases of a sample of media that was in contact with cells.

As a consequence of this, some tests are going to be performed in order to check the good performance of the protocol. The main idea is to take one sample with a low Viability, which have a lot of accumulated proteins, and apply different concentration of TCA and time to incubate the sample after precipitation to check the reliability of the method.

7. CONCLUSIONS

The model used in this Thesis was checked and corrected when it was necessary. But after these corrections, the model can be considered a highly satisfying model under a theoretical point of view. A graphical interface was done to facilitate the user use without the necessity of programming skills.

It is difficult to determine the performance of a dynamical model when there is only available one set of data under some defined conditions. A lot of hypothesis can be done using the results obtained but they cannot be tested. In this Thesis the model and the metabolic reaction network used to the modelling succeed in the estimation of the system behaviour under the conditions of changing one aa in the composition of the medium.

In general, the error between the real rates and the estimated rates are small. Only four metabolites have an error above 50%: cell death, Asn, Asp and Ser. It can be said that some reactions are missing in the system that involve these metabolites.

Although the error is not very significant, it was also observable that some reactions that describe the metabolism of Gln and the biomass are missing since the rates of the estimated values are always above the real rates.

The system can follow the difference metabolic pathways that the cells take when one aa in the composition of the media is varied. Only the rates of Glc, Lac and NH_4 are not affected by the changes in the initial conditions of the culture, which is a drawback of the system.

The model also succeeded in the simulation over time of the concentration of the extracellular metabolites. Except Ser and biomass, the errors between the simulated concentration and the real concentrations are below 20%.

Finally, it was observed that it is possible to omit some eaa's in the system without deteriorating the modelling quality.

8. FUTURE WORK

This Thesis was only a start point of the amount of work that must be done in order to validate the model. In the realization of this project, the experiments were carried out changing only the composition of one element in the media, aa. Nevertheless, there are several other experiments that can be done.

According to that, the main work in the future has to be focused in the obtaining of new experimental data under different environmental conditions. These experiments can involve changes in the temperature and the study of the effect of the glutamine in cell culture. It can also be varied more than one aa in the culture media. This data will helped to determine even better than in this project the good performance of the model.

Another important task will be the improvement of the methodology followed in the analysis of the samples since it was seen that some modelling errors were caused by measurements with a high analysis error of the extracellular metabolites.

9. REFERENCES

[1] Gorsfils, A., 2007. ”First principles and black box modelling of biological systems”. <http://theses.ulb.ac.be/ETD-db/collection/available/ULBetd-09122007-105503/> (18 May 2010)

[2] Yang, S. & Basu, S., December 2005 ”Encyclopedia of Chemical Processing”. *Taylor & Francis Group*

[3] Ryan, J.A. (Corning). ”Introduction to Animal Cell Culture”. 2008. <http://www.getter.co.il/biomed/biomed-lab/down/4item_dnlheb.pdf > (19 May 2010)

[4] Van der Valk , J., Brunner, D., De Smet, K., Fex Svennigsen, A., Honegger, P., Knudsen, L, Lindl, T., Noraberg, J., Price, A. Scarino, M.L. & Gstraunthaler, 2010. ”Optimization of chemically defined cell culture media – Replacing fetal bovine serum in mammalian in vitro methods”.

[5] Wolfe, R., Heifetz H., A. & Custer, L.M, 1989.”Basal nutrient medium for cell culture”. <http://www.freepatentsonline.com/5232848.pdf> (19 May 2010)

[6] ”Serum-free media for cell culture”. May 2009.<http://www.focusonalternatives.org.uk/PDFs/FCS-free%20table%20May%2009.pdf> (21 May 2010)

[7]”Serum-free Media and Applications”. <http://tools.invitrogen.com/content/sfs/appendix/Cell_Culture/Serum-Free%20Medium%20And%20Applications.pdf> (19 May 2010)

[8] Sheldon E. Broedel. Jr. ”The case for serum-free media technical brief”. February 2003. <http://www.athenaes.com/tech_brief_serum_free.php> (17 May 2010)

[9] Birch, J., 2004. ”Mammalian cell culture: current status. future prospects”. <http://www.lonzabiologics.com/group/en/company/news/publications_of_lonza.-ParSys-0002-ParSysdownloadlist-0010-DownloadFile.pdf/10_Berlin%20Mammalian%20cell%20culture.pdf> (21 May 2010)

[10]"CHO cells". <http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cell-Culture/Mammalian-Cell-Culture/mammalian-misc1/cho_cells.html> (20 May 2010)

[11]Jayapal, K.P., Wlaschin, K.F. & Hu, W.-S. "Recombinant protein Therapeutics from CHO cells- 20 years and counting" .
<http://hugroup.cems.umn.edu/HuGroup/PDFs/Jayapal_et_al_CHO_Review.pdf> (21 May 2010)

[12] "Metabolism Health Article".
< http://www.healthline.com/galecontent/metabolism> (01 May 2010)

[13] Provost, A., Bastin, G., Agathos, S.N. & Scheider, Y.-J., 2005. "Metabolic Design of Macroscopic Models: Application to CHO cells". *In: Proc. Of the 44th IEEE Conference on Decision and Control, and the European control Conference, Seville, Spain*. p. 2982-2989

[14] Provost, A., Bastin, G., 2006. "Metabolic flux Analysis: An approach for solving non-stationary undetermined systems". *In: Proc. 5th MATHMO, Vienna 2006, Austria*

[15] Biomed Central. "A procedure for the estimation over time of metabolic fluxes in scenarios where measurements are uncertain and/or insufficient". <http://biomedcentral.com/1471-2105/8/421/figure/F4> (14 April 2010)

[16] Zamorano, F., Vande Wouwer, A. & Bastin,. G., 2009. "Metabolic Flux Analysis of an Undetermined Network of CHO Cells"

[17] Segré, D., Zucker, J., Katz, J. Lin X., D'Haeseller P., Rindone, W.P., Kharchenko P., Nguyen, D.H., Wright, M.A. & Church, G.M., 2003. "From Annotated Genomes to Metabolic Flux Models and Kinetic Parameter Fitting". *Journal of Integrative Biology*, Vol. 7(3), p 301-316

[18] "What is mass balance?". <http://www.massbalance.org/about/> (5 May 2010)

[19] Ozturk, S.S. & Palsson., B.O., 1991. "Growth Metabolic and Antibody Production Kinetics of Hybridoma Cell Culture: 1. Analysis of Data from Controlled Batch Reactors". *Biotechnol. Prog.,* Vol. 7, p. 471-480

[20] Altamirano C., Illanes, A., Casablancas, A., Gámez, X., Cairó, J.J. & Gôdia C., 2001. "Analysis of CHO Cells Metabolic Redistribution in a Glutamate-Based Defined Medium in Continuous Culture". *Bitoechnol. Progr.,* Vol. 17, p. 1032-1041

[21] Llaneras, F. & Picó J., 2008. “Stoichiometric Modelling of Cell Metabolism”. *Journal of Bioscience and Bioengineering*, Vol. 105, No. 1, p. 1-11.

[22] Teixera, A.P., Alves, C., Alves, P.M, Carrondo, M. & Oliveira R., 2007. “Hybrid elementary flux analysis/nonparametric modelling: application for bioprocess control”. *BMC Bioinformatics*, Vol. 8, p. 30

[23] Zupke. C. & Stephanopoulos. G., 1995. “Intracellular flux analysis in hybridomas using mass balances and in vitro 13 C NMR”. *Biotechnology and Bioengineering*, Vol. 45, p. 292–303

[24] Stephanopoulos, G. 1999. “Metabolic fluxes and metabolic engineering”. *Metabolic Engineering*, Vol. 1, p. 1–11

[25] Stephanopoulos, G.N., Aristidou, A.A. & Nielsen. J., 1998 “Metabolic engineering: principles and methodologies”. *Academic Press*, San Diego, USA

[26] Nyberg, G. B., Balcarcel, R., Follstad, B. D., Stephanopoulos, G. & Wang, D. I. C.. 1999. “Metabolism of peptide amino acids by Chinese Hamster Ovary cells grown in a complex medium”. *Biotechnology and Bioengineering*, Vol. 62, p. 321–335

[27] Provost, A. & Bastin, G., 2004. “Dynamic metabolic modelling under the balanced growth condition”. *Journal of Process Control*, Vol. 14, p. 717–728

[28] Bonarius, H. P. J., Schmid, G. & Tramper, J., 1997 “Flux analysis of underdetermined metabolic networks: the quest for the missing constraints”. *TIBTECH August*, Vol. 15, p. 308–314

[29] Xie, L. & Wang, D. I., 1996. “Material balance studies on animal cell metabolism using stoichiometrically based reaction network”. *Biotechnology and Bioengineering*, Vol. 52, p. 579–590

[30] Provost, A. & Bastin, G., February 2003. “Dynamical metabolic modelling of biological processes”. *Fourth International Symposium on Mathematical Modelling*

[31] LLaneras, F. & Picó, J., 2007. “An interval approach for dealing with flux distributions and elementary modes activity patterns”. *J. Theor. Biol.*, Vol. 246, p. 290-308

[32] Llaneras, F. & Picó, J., 2007. “A procedure for the estimation over time of metabolic fluxes in scenarios where measurements are uncertain and/or insufficient”. *BMC Bioinformatics,* p. 8-421

[33] Rockafella, R.T., 1996. “Convex analysis”. *Princeton Universtity Press,* Princeton, USA

[34] Varma, A. & Palsson, B.O., 1994. “Metabolic Flux Balancing: Basic Concepts. Scientific and Practical Use”. *Bio-Technology,* Vol. 12, p. 994-998.

[35] Kauffman, K.J., Prakash, P. & Edwards, J.S., 2003 ‘‘Advances in flux balance análisis’’. *Curr. Opin. Biotechnol,* Vol. 14, p. 491–496

[36] Schuster, S., Dandekar, T., & Fell D.A., 1999. “Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering”. *Trends Biotechnology,* Vol. 17, p. 53-60

[37] Schuster S., Fell, D.A. & Dandekar, T., 2000. “A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks”. *Nat Biotechnol*, Vol. 18, p. 326-332

[38] Klamt, S. & Stelling, J., 2003. “Two approaches for metabolic pathway analysis?” *Trends Biotehnol*, Vol. 21, p. 64-69

[39] Schilling, C.H., Schuster, S., Palsson, B.O. & Heinrich, R., 1999. “Metabolic Pathway Analysis: Basic Concepts and Scientific Applications in the Post-genomic Era”. *Biotechnol. Prog.* Vol. 15, p. 296-303

[40] Gonzalez, J., 2009. “Dynamical modelling of a metabolic reaction network”. *Master Thesis in Automatic Control. KTH*

[41] Pfeiffer. T., Sánchez-Valdenebro, I., Nuño, J., Montero, F. & Schuster, S., 2002. “METATOOL: for studying metabolic networks”. *Bioinformatics*, Vol 15, p. 251-257

[42] Westermark, P. 2002. “Physical Biochemistry”. <http://www.csc.kth.se/utbildning/kth/kurser/DD2435/biomod09/forelasnantec kn/F2.pdf> (07 May 2010)

[43] Lawson, C.l. & Hanson, R.J., 1974. “Solving Least-Squares Problems”. *Prentice Hall*, Chapter 23, p. 161

[44] Bonarius, H. P., Hatzimanikatis, V., Meesters, K.P., De Gooijer, CD., Schmid, G. & Tramper, J., 1996. “Metabolic Flux Analysis of Hybridoma Cells in Different Culture Media Using Mass Balances”. *Biotechnol. Gioeng.,* Vol. 50, p. 299-318

[45] Liu, H.-J., 1994. “Determination of amino acids by precolumn derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate and high-performance

liquid chromathography with ultraviolet detection". *J. Chromatography*, 610. p. 59-66

ACKNOWLEDGMENTS

First of all, I would like to thank Veronique for all her guidance, support and patience when I had problems. But I would also like to thank her for giving me the opportunity to carry out this Thesis when I had no knowledge of the subject. At the same time, I would like to acknowledge Andres Veide, my examiner, for his help and comprehension in the last stage of the Thesis.

Many people have contributed to my learning. I would like to thank all of them. Helene Sundström for her support and for introducing me in the lab work. Atefeh Shokri and Ye Zhang for their advices and for helping me when I had doubts. I also would like to thank Jan Kinnander for instructing and aiding me with the HPLC and the methods of amino acid analysis.

But there are two people who without them this would not have been possible. They are Gemma and Andreas. Thanks for being the best partners you can have in the lab and for all the things that I have learned with you.

I would also thank to the rest of the people that I met in the KTH Division of Bioprocess Technology.

We would like to acknowledge IMED AB for the permission of using one of their research cell lines in this study. IMED, located at the Karolinska Institute Science Park, Sweden, develops human monoclonal antibodies for the treatment of diseases associated with deregulation of apoptosis, such as HIV, cancer and transplantation related disorders.

My parents have played an important role in this hard work. Nothing could have been without their unconditional support, advice, dedication and motivation. Thank you very much and I love you.

This has been a long trip and during these seven years, I have met a lot of people. Thank to all of them today I can be writing the final words of my Thesis. I would like to mention everybody but I do not have enough pages to thank you for all your support. My friends from my town, university, DAT, CEET, Best, theatre, German, Erasmus and corridor I can only say to them thank you very much.

APPENDIX A: Metabolic reaction network

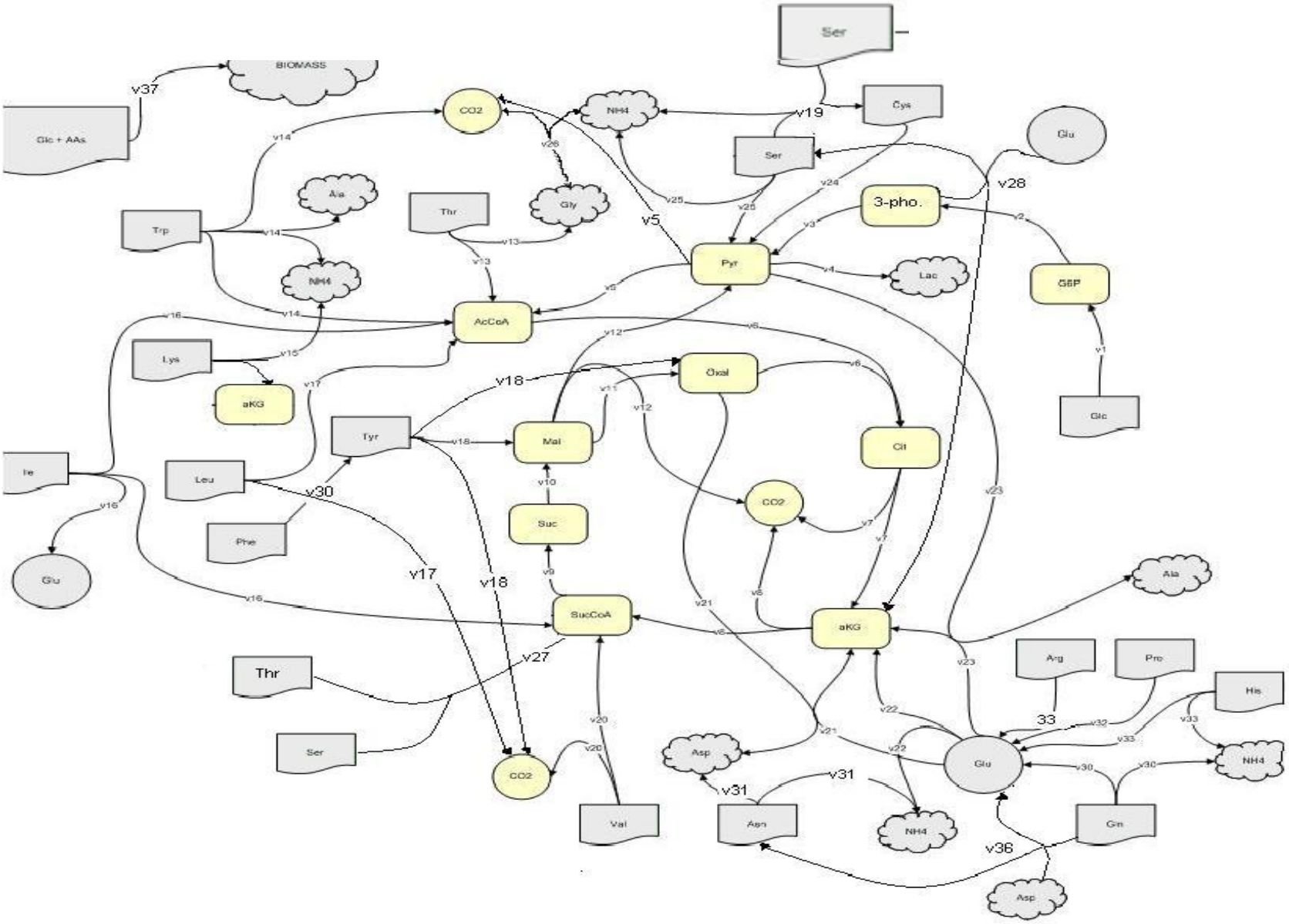


Table A.1. Stoichiometric matrix of a CHO system

		R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}	R_{13}	R_{14}	R_{15}	R_{16}	R_{17}	R_{18}	R_{19}	R_{20}	R_{21}	R_{22}	R_{23}	R_{24}	R_{25}	R_{26}	R_{27}	R_{18}	R_{29}	R_{30}	R_{31}	R_{32}	R_{33}	R_{34}	R_{35}	R_{36}	R_{37}	R_{38}
Aext	Glc	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0208	0
	Lac	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	NH_4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0	1	1	0	0	1	0	0
	Gln	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	-1	-0.0377	0
	Glu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	-1	-1	-1	0	0	0	0	-1	0	0	0	1	1	-1	1	1	-0.0006	0
	Arg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	-0.007	0
	His	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	-0.0033	0
	Ile	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0084	0
	Leu	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0133	0
	Lys	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0101	0
	Met	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0033	0
	Phe	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	-0.0055	0
	Thr	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.008	0
	Trp	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	-0.004	0
	Val	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.0096	0
	Ala	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	-0.0133	0
	Asn	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	1	0	0
	Asp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	-1	-0.026	0
	Cys	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	-0.0004	0
	Gly	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	1	0	0	0	0	0	0	-0.0165	0
	Pro	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+1	0	0	-0.0081	0
	Ser	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	-1	0	0	1	-1	0	0	0	0	0	0	0	-0.0099	0
	Tyr	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	-1	0	0	1	0	0	0	0	0	0	-0.0077	0
	Bio.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1
	C death	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Aint	G6P	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	3-phos.	0	2	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
	Pyr	0	0	1	-1	-1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	AcCoA	0	0	0	0	1	-1	0	0	0	0	0	0	1	2	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Cit	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	KG	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	SucCoa	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Suc	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	Mal	0	0	0	0	0	0	0	0	0	1	-1	-1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Oxal	0	0	0	0	0	-1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX B: Media composition

Table B.1. Composition of SF10 and SF14 medium compared with DMEM/F12

	COMPONENT	DMEM/F12 mg/L	ADD. 1 mg/L	SF10 mg/L	ADD. 2	SF14 mg/L
Amino acids	Alanine	4.45		4.45		4.45
	Arginine-HCl	147.5		147.5		147.5
	Asparagine•H_2O	7.5		7.5		7.5
	Aspartic acid	6.65		6.65		6.65
	Cysteine•HCl•H_2O	28.3		28.3		28.30923275
	Cystine	18.7		18.7		18.74462692
	Glutamate	7.35		7.35		7.35
	Glutamine	365		365		365
	Glycine	18.75		18.75		18.75
	Histidine	31.48		31.48		31.48
	Isoleucine	54.47		54.47		54.47
	Leucine	59.05		59.05		59.05
	Lysine•HCl	91.25		91.25		91.25
	Methionine	17.24		17.24		17.24
	Phenylalanine	35.48		35.48		35.48
	Proline	17.25		17.25		17.25
	Serine	26.25		26.25		26.25
	Threonine	53.45		53.45		53.45
	Tryptophan	9.02		9.02		9.02
	Tyrosine•HCl	55.79		55.79		55.79
	Valine	52.85		52.85		52.85
Organ. Vitamins	Sodium Pyruvate	110		110		110
	Biotin	0.0035		0.0035		0.0035
	Calcium pantothenate	2.24		2.24		2.24
	Choline chloride	8.98		8.98		8.98
	Folic acid	2.66		2.66		2.66
	Myo-Inositol	12.6		12.6	60	72.6
	Nicotinamide	2.02		2.02	4	6.02
	Pyridoxine•HCl	2.031		2.031		2.031
	Riboflavin	0.219		0.219		0.219
	Thiamine•HCl	2.17		2.17		2.17
	Vitamin B12	0.68		0.68		0.68
Salts/metals/trace	$CaCl_2$	116.6632653		116.6632653		116.6632653
	KCl	311.8		311.8		311.8
	Na_2HPO_4	71.02		71.02		71.02
	NaCl	6996		6996		6996
	$NaH_2PO_4•H_2O$	62.445		62.445		62.445
	$MgSO_4$	48.84		48.84		48.84
	$MgCl_2$	28.68844073		28.68844073		28.68844073
	$NaHCO_3$		2.200	2200	2200	2200

	Sodium Selenite	0.00519		0.00519		0.00519
	Ferric Citrate			0	0.25	0.25
	$CuSO_4 \bullet 5H_2O$	0.0013		0.0013		0.0013
	$FeSO_4 \bullet 7H_2O$	0.417		0.417		0.417
	ZnCl	0.432		0.432		0.432
	$MnSO_4 \bullet H_2O$	0.000169		0.000169		0.000169
	$Na_2SiO_3 \bullet 9H_2O$	0.0142		0.0142		0.0142
	$(NH_4)\ 6Mo_7O_{24} \bullet 4H_2O$	0.015525366		0.015525366		0.015525366
	$NiSO_4 \bullet 6H_2O$	0.00022		0.00022		0.00022
	$SnCl_2 \bullet 2H_2O$	7.85714E-05		7.85714E-05		7.85714E-05
Miscellaneous	Putrescine•2HCl	0.081		0.081	0.95	1.031
	Linoleic acid	0.042		0.042		0.042
	Lipoic acid	0.105		0.105		0.105
	Pluronic F-68		1000	1000	1000	1000
	Insulin		15	15	5	5
	Hipoxantine			20	20	20
	Thymidine			5.0	5	5
	Glucose		3150	3.150.0	3150	3150
	Phenol red		8.63	8.6	8.63	8.63

Table B.2. Concentration stock Solution of the components used

	COMPONENT[7]	Dissolution in	Concentration Stock Solution mg/L
Amino acids	Alanine	Water	15000
	Arginine-HCl	Water	49125
	Asparagine•H_2O	HCL 1N	7500
	Aspartic acid	Water	6650
	Cysteine•HCl•H_2O	Water	24000
	Cystine	HCL 2N	9370
	Glutamate	Water	7350
	Glycine	Water	18750
	Histidine	Water	25000
	Isoleucine	HCL 1N	27235
	Leucine	HCL 1N	29525
	Lysine•HCl	Water	91250
	Methionine	Water	17240
	Phenylalanine	HCL 1N	35480
	Proline	Water	17250
	Serine	Water	26250
	Threonine	Water	53450
	Tryptophan	HCL 2N	9020
	Tyrosine•HCl	HCL 1N	14000
	Valine	Water	10570
Organ. Vitamins	Sodium Pyruvate	Water	55000
	Biotin	NaOH 2M	4.17E+03
	Calcium pantothenate	Water	336
	Choline chloride	Water	25000
	Folic acid	Water	25000
	Myo-Inositol	Water	25000
	Nicotinamide	Water	2000
	Pyridoxine•HCl	Water	2000
	Riboflavin	NaOH 0.1M	4.33E+03
	Thiamine•HCl	Water	2000
	Vitamin B12	Water	10200
Salts/metals/trace	$CaCl_2$	Water	100000
	KCl	Water	44000
	Na_2HPO_4	Water	33000
	$NaH_2PO_4•H_2O$	Water	100000
	$MgSO_4$	Water	50000
	$MgCl_2$	Water	72000
	Sodium Selenite	Water	1.07
	Ferric Citrate	Water	500
	$CuSO_4 •5H_2O$	Water	1000
	$FeSO_4 •7H_2O$	HCL 1N	500

[7] The components not listed were added directly in the medium

	ZnCl	Water	300
	$MnSO_4 \bullet H_2O$	Water	0.33
	$Na_2SiO_3 \bullet 9H_2O$	Water	50
	$(NH_4)\ 6Mo_7O_{24} \bullet 4H_2O$	Water	11700
	$NiSO_4 \bullet 6H_2O$	Water	66
	$SnCl_2 \bullet 2H_2O$	Water	6.07E+03
Miscella-neaous	Putrescine•2HCl	Water	250
	Linoleic acid	Water	6300
	Lipoic acid	Ethanol	400
	Insulin	Water	5000000

APPENDIX C: RMB03.02 experiment

Experiment RMB03.02 was carried out varying the aa concentration in the medium in the different tubes in order to trigger different metabolic pathways.

The data were analyzed for modelling. Unfortunately, the cells did not grow significantly or not at all in SF10 medium. The aa variations were small and several problems were encountered.

The two methods used in the AAA, AccQTag and AccQTagComplete, did not separate properly all the aa peaks. In the first method, two pairs of aa were eluated together: Asn with Ser and His with Gln. In contrast, in the second method, only NH_3 and His were eluated together. This affected the calculation of the concentration of the aa.

In general, the concentrations of the aa were obtained doing the average of the concentrations of each method. In the case of Asn, Ser and Gln the concentration was obtained only using the concentration of the method where the peak was correctly separated instead.

The Histidine concentration was more complicated to solve. At first, three possible solutions were thought. These solutions used the concentration values where the Histidine was involved and the concentration of the components that eluated together with it:

Table C.1. Possible solutions to determine His concentration

1.	$His = (His + G\ln)_{AccQTag} - (G\ln)_{AccQTagComplete}$
2.	$His = (His + G\ln)_{AccQTag} - (G\ln)_{Bioprofile}$
3.	$His = (His + NH_3)_{AccQTagComplete} - (NH_3)_{AccQTag}$

Unfortunately, none of the described solutions were good enough since in many cases the concentration of Histidine was negative which is not possible. HPLC is not a totally accurate technique. Furthermore, the values that come from the Bioprofile and HPLC were not completely correlated.

Finally, the first option was decided since it was the case where less Histidine concentrations became negative. However, too many values of His were negative, so this aa was not taken into account in the simulation.

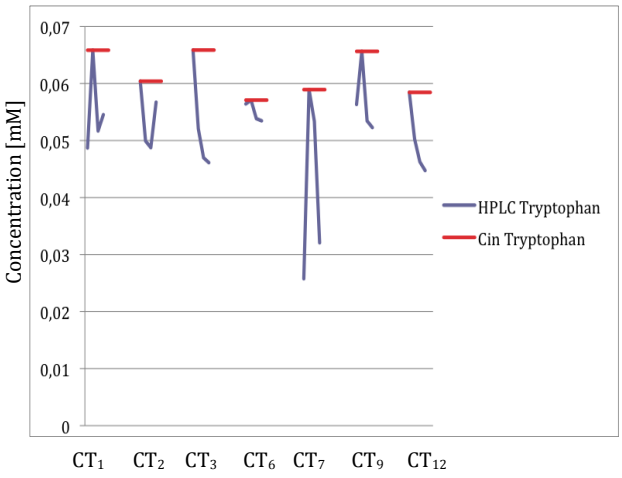
A second problem appeared when the q_i, consumption/production rates, were calculated using Eq. 6. So, according to this formula, when the value of q_i is negative, it means that the amino acid is produced and when it is positive it means that the amino acid is consumed. According to that, the q_i of the essential amino acids were supposed to be positive but, unfortunately, it did not always occur.

From the data it seemed that some essential amino acids were produced since some values of the culture concentrations were above the concentration in fresh medium. This was due to the low consumption of the aa. It suggested that the measurements by HPLC of the concentrations of the essential amino acids in the medium were not accurate. In order to be able to use these data as input in the simulation program, one possible solution was to consider that the measurements of the concentrations of the eaa in the medium were erroneous and to identify better values for these concentrations.

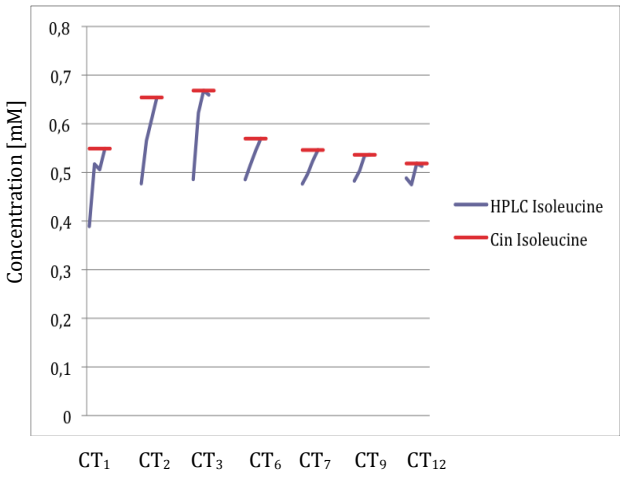
A first approach was to take the average of the values of day 3, which could be also considered the first day of the steady state, of all the experiments. Nevertheless, this approach could not been used in all the essential aa. In some cases the concentrations from day 3 were very different and the average was not a good solution. This was the case of Trp (Fig. C.1.a). In others, the evolution of the concentration of the essential aa over time increased and the problem remained

after the application of the average of day 3. This was the case of Ile (Fig. C.1.b). In this cases the criteria decided was to take the maximum value of the concentrations.

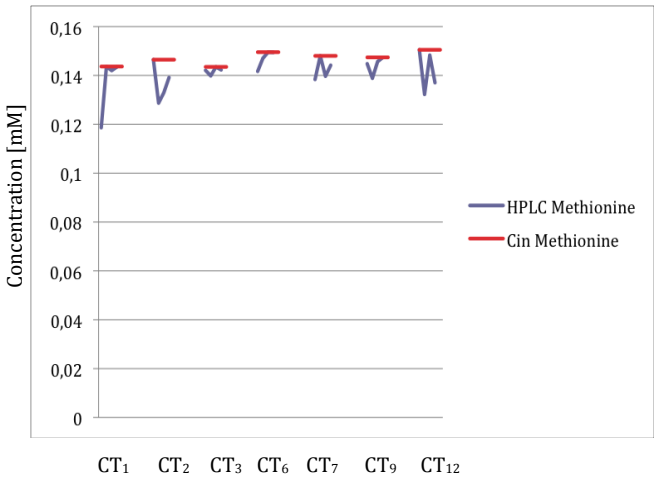
The graphics in Fig. C.1, where the evolution of each essential amino acid and the maximum value appear, helped to arrive to a final decision. In that case the graphics in Fig. C.1 show the final input concentration.



(a)Tryptophan



(b) Isoleucine



(c) Methionine

Figure C.1. Graphics used to obtain the criteria for the first alternative

The final criteria decided were:

Table C.2. Criteria for first alternative

Arginine	Average: CT_2, CT_3, CT_6, CT_7, CT_9, CT_{12} CT_1: maximum value
Isoleucine	Maximum value
Leucine	Maximum value
Lysine	Maximum value
Methionine	Average: CT_2, CT_3, CT_6, CT_7, CT_9, CT_{12} CT_1: first value
Phenylalanine	Average: CT_2, CT_3, CT_6, CT_7, CT_9 CT_2 and CT_{12}: maximum value
Threonine	Average: CT_2, CT_3, CT_6, CT_7, CT_9, CT_{12} CT_1: first value
Tryptophan	Maximum value
Valine	Maximum value

According to the results, there were still consumption/production rates that neglected the definition of an essential amino acid and they are showed in Table C.3.

Table C.3. Eaa which specific consumption/production rate was above zero after the application of the first criteria

	Day 4	**Day 5**	**Day 6**
CT_1	Arg, Ile, Leu, Lys, Met, Phe, Thr, Trp, Val	Lys	Ile, Leu, Lys, Met, Phe, Val
CT_2	Ile	Ile, Lys	Ile, Leu, Lys, Met, Phe, Thr, Trp, Val
CT_3	Ile, Thr	Ile, Leu, Lys, Met, Val	
CT_6	Arg, Met, Phe, Thr, Trp	Ile, Leu, Lys, Met, Val	Arg, Ile, Leu, Lys, Met, Phe, Thr, Val
CT_7	Arg, Met, Phe, Thr, Trp	Ile, Leu, Lys	Ile, Leu, Lys, Met, Phe, Val
CT_9	Trp	Ile, Leu, Lys, Met, Thr, Val	Ile, Met, Phe
CT_{12}		Ile, Leu, Lys, Met, Val, Thr	

As can be seen in Table C.3 only three days could be used, but they were not enough since the inequation in Eq. 26 would not be fulfilled. For that reason, a second alternative was needed. This second criterion, which was a more dramatic change,

used directly as a concentration input the maximum value of the concentrations of each experiment. The next table shows, which essential amino acids continue to be negative after the changing described above.

Table C.4. EAA that have specific consumption/production negative rate after application second criteria

	Day 4	Day 5	Day 6
CT_1	Arg, Ile, Leu, Lys, Met, Phe, Thr, Trp, Val	Lys	Ile, Leu, Lys, Met, Phe, Val
CT_2	Ile	Ile, Lys	Ile, Leu, Lys, ~~Met~~, ~~Phe~~, Thr, Trp, Val
CT_3	Ile, Thr	Ile, Leu, Lys, Met, Phe, Val	
CT_6	Arg, Met, Phe, Thr, Trp	Ile, Leu, Lys, Met, Val	~~Arg~~, Ile, Leu, Lys, ~~Met~~, ~~Phe~~, Thr, Val
CT_7	Arg, Met, Phe, Thr, Trp	Ile, Leu, Lys	Ile, Leu, Lys, Met, ~~Phe~~, Val
CT_9	Trp	Ile, Leu, Lys, Met, Thr, Val	Ile, Met, Phe
CT_{12}		Ile, Leu, Lys, Met, Val, Thr	

Despite that this alternative was the most dramatic change that could be done with the concentration of the fresh medium, there were still some consumption/production rates that were negative.

As a consequence of that, the first attempt could not be directly used in the simulation program and the only possibility to use it was to make some simplifications of the system. These simplifications discarded some eaa where the consumption/production rates were negative in order to be able to use some samples as input in the simulation program.

Several options and systems were used but, unfortunately, the results obtained were not good. Thus, it was decided not to continue with the analysis of the results of this experiment.

APPENDIX D: Experiment 2

Table D.1. Cv and Viabilite of CT_1 and CT_2

Exp CT_1 Ala 0%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 10:59	2	1.675	99.6
12/04/10 01:03	3	2.041	93.4
13/04/10 01:28	4	0.634	66
14/04/10 01:58	5	0.56	65.9
15/04/10 02:48	6	0.842	82.9
16/04/10 01:05	7	1.256	87.4
17/04/10 01:01	8	1.1	89

Exp CT_2 Ala 50%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 11:06	2	1.362	98.5
12/04/10 02:02	3	1.778	93.7
13/04/10 01:33	4	0.581	71
14/04/10 02:06	5	0.44	59.9
15/04/10 11:52	6	0.516	66.7
16/04/10 01:20	7	0.867	85.4
17/04/10 01:59	8	0.959	87.8
18/04/10 12:15	9	1.265	91.3

Table D.2. Cv and Viabilite of CT_3 and CT_4

Exp CT_3 Ser 0%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 11:14	2	1.449	98.7
12/04/10 02:09	3	1.782	95.7
13/04/10 01:46	4	0.608	68.1
14/04/10 03:48	5	0.462	0.638
15/04/10 02:54	6	0.61	82.7
16/04/10 03:22	7	0.937	92
17/04/10 02:08	8	1.173	92.1
18/04/10 12:22	9	1.512	96.5

Exp CT_4 Ser 50%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 11:26	2	1.507	98.1
12/04/10 02:21	3	1.825	93.7
13/04/10 01:52	4	0.47	66.1
14/04/10 03:00	5	0.44	65.9
15/04/10 04:47	6	0.456	74.5
16/04/10 03:29	7	0.602	80.7
17/04/10 01:10	8	0.931	85.2

Table D.3. Cv and Viabilite of CT_5 and CT_6

Exp CT_5 Asp 0%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
10/04/10 01:00	1	-	-
11/04/10 11:45	2	1.515	99
12/04/10 02:28	3	1.665	95.2
13/04/10 03:25	4	0.86	75
14/04/10 05:00	5	0.914	84.5
15/04/10 03:05	6	1.3	88.8
16/04/10 03:36	7	1.271	91.6
17/04/10 01:21	8	1.387	92.5

Exp CT_6 Asp 50%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
10/04/10 01:00	1	-	-
11/04/10 11:52	2	1.515	99.2
12/04/10 02:34	3	1.329	93
13/04/10 03:31	4	0.975	80.5
14/04/10 02:35	5	1.144	91
15/04/10 03:11	6	1.78	96.7
16/04/10 03:52	7	1.613	98.5
17/04/10 01:30	8	1.529	98.4

Table D.4. Cv and Viabilite of CT_7 and CT_8

Exp CT_7 Glu 0%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 11:58	2	1.757	99.5
12/04/10 02:43	3	1.92	93.6
13/04/10 03:40	4	0.929	77.8
14/04/10 03:04	5	0.97	82.8
15/04/10 03:53	6	1.784	89.6
16/04/10 03:59	7	1.239	94.9
17/04/10 01:37	8	1.515	95.3

Exp CT_8 Glu 50%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 12:04	2	1.395	99.2
12/04/10 02:55	3	1.631	89.9
13/04/10 03:46	4	0.47	49.6
14/04/10 03:10	5	0.426	56.7
15/04/10 11:38	6	0.646	67.8
16/04/10 04:20	7	0.896	81
17/04/10 02:21	8	1.026	91.6
18/04/10 12:31	9	1.418	94.4

Table D.5. Cv and Viabilite of CT_9 and CT_{10}

Exp CT_9 Gly 0%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
10/04/10 01:00	1	-	-
11/04/10 12:11	2	1.592	99.4
12/04/10 03:07	3	1.565	91.6
13/04/10 03:52	4	0.914	73.6
14/04/10 04:54	5	0.73	79
15/04/10 11:54	6	0.72	75.9
16/04/10 04:29	7	0.27	58.7

Exp CT_{10} Gly 50%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
10/04/10 01:00	1	-	-
11/04/10 02:00	2	1.75	96.1
12/04/10 03:14	3	1.503	94
13/04/10 05:14	4	0.746	68.1
14/04/10 05:02	5	0.776	80.7
15/04/10 04:04	6	1.288	89.1
16/04/10 04:59	7	1.308	92.2
17/04/10 01:44	8	1.372	93.3

Table D.6. Cv and Viabilite of CT_{11} and CT_{12}

Exp CT_{11} Asn 0%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 02:20	2	1.67	94.1
12/04/10 03:35	3	1.548	94.5
13/04/10 05:30	4	0.711	64.8
14/04/10 05:08	5	0.874	78.1
15/04/10 04:11	6	1.49	89.1
16/04/10 05:07	7	1.375	93.2
17/04/10 01:51	8	1.181	94.5

Exp CT_{12} Asn 50%			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
11/04/10 02:35	2	1.77	97.2
12/04/10 03:28	3	1.861	93
13/04/10 05:38	4	0.516	56.9
14/04/10 05:06	5	0.403	59.8
15/04/10 04:17	6	0.608	62.7
16/04/10 05:14	7	0.624	63.1
17/04/10 02:29	8	0.725	78.7
18/04/10 12:37	9	0.73	81.9

Table D.7. Cv and Viabilite of CT_{13} and CT_{14}

Exp CT_{13} EAA 50%			
Sampling time day	Cult. time Days	Cv MVC/mL	Viability %
09/04/10 01:30	-	0.5	-
10/04/10 01:00	2	1.48	97.6
12/04/10 03:36	3	1.834	92
13/04/10 05:44	4	0.418	48.7
14/04/10 03:30	5	0.28	42.5
15/04/10 04:22	6	0.572	73.8
16/04/10 05:20	7	0.401	54.2

Exp CT_{14} ctrl			
Sampling time day	Cult. time days	Cv MVC/mL	Viability %
	#¡REF!	#¡REF!	#¡REF!
09/04/10 01:30	-	0.5	-
11/04/10 03:15	2	1.48	96
12/04/10 03:49	3	1.264	92.8
13/04/10 02:02	4	0.837	84.7
14/04/10 11:35	5	0.916	86.1
15/04/10 12:45	6	1.104	90.4
16/04/10 05:30	7	1.486	93.1
17/04/10 02:36	8	1.429	95.7
18/04/10 12:45	9	1.284	96.5

Table D.8. Concentration of the external metabolites of the samples used in the complete system

Component	CT1D06	CT1D07	CT2D08	CT4D06	CT5D06	CT5D07	CT5D08	CT6D06	CT6D07	CT6D08	CT10D05
Glc	14.18	14.24	16.93	15.82	14.33	14.18	15.98	12.63	13.08	13.75	15.44
Lac	3.50	5.40	6.08	1.53	4.12	4.19	3.94	8.65	6.06	7.08	2.47
Gln	2.21	2.94	2.90	2.34	1.98	2.92	2.86	1.70	2.63	2.71	2.34
NH_4	0.96	1.18	1.38	0.76	1.22	1.45	1.36	1.17	1.54	1.40	0.82
Asp	0.05	0.04	0.04	0.04	0.03	0.04	0.04	0.02	0.04	0.04	0.06
Glu	0.15	0.17	0.19	0.12	0.13	0.20	0.19	0.13	0.18	0.15	0.20
Ser	0.42	0.48	0.43	0.41	0.48	0.49	0.49	0.45	0.49	0.46	0.46
Asn	0.04	0.04	0.08	0.02	0.03	0.09	0.10	0.04	0.10	0.09	0.04
Gly	0.30	0.24	0.23	0.26	0.23	0.29	0.27	0.16	0.24	0.24	0.16
His	0.27	0.22	0.25	0.20	0.22	0.26	0.28	0.20	0.26	0.26	0.35
Thr	0.58	0.50	0.53	0.43	0.44	0.58	0.58	0.40	0.54	0.49	0.69
Arg	0.85	0.75	0.84	0.64	0.70	0.96	1.01	0.69	0.91	0.87	1.10
Ala	0.38	0.51	0.67	0.29	0.48	0.76	0.73	0.70	0.94	0.75	0.55
Pro	0.20	0.15	0.15	0.15	0.15	0.18	0.18	0.10	0.17	0.15	0.29
Tyr	0.33	0.32	0.32	0.25	0.28	0.38	0.35	0.26	0.36	0.29	0.45
Cys	0.27	0.24	0.20	0.18	0.09	0.22	0.22	0.08	0.18	0.11	0.34
Val	0.46	0.40	0.41	0.35	0.45	0.51	0.49	0.35	0.43	0.40	0.63
Met	0.10	0.08	0.07	0.07	0.08	0.11	0.12	0.07	0.09	0.05	0.16
Iso	0.41	0.37	0.39	0.32	0.33	0.45	0.44	0.31	0.42	0.38	0.54
Leu	0.42	0.39	0.39	0.34	0.33	0.46	0.44	0.29	0.42	0.39	0.55
Lys	0.62	0.61	0.65	0.53	0.58	0.71	0.71	0.51	0.69	0.67	0.85
Phe	0.26	0.23	0.26	0.09	0.10	0.29	0.30	0.12	0.26	0.11	0.35
Tryp	0.05	0.04	0.07	0.05	0.06	0.04	0.03	0.02	0.05	0.01	0.06
Biomass	842.00	1256.00	1265.00	456.00	1300.00	1271.00	1387.00	1780.00	1613.00	1529.00	776.00
Death	173.68	181.07	120.54	156.08	163.96	116.55	112.46	60.74	24.56	24.86	185.59

Table D.9. Specific consumption/production rates of the samples used in the complete system

Component	CT1D06	CT1D07	CT2D08	CT4D06	CT5D06	CT5D07	CT5D08	CT6D06	CT6D07	CT6D08	CT10D05
Glc	-4.65	-3.04	-0.37	-3.54	-2.81	-2.98	-1.19	-3.42	-3.44	-3.02	-2.71
Lac	5.21	5.37	6.03	3.66	3.90	4.00	3.55	6.34	4.94	6.03	3.60
Gln	-2.52	-0.95	-0.99	-3.73	-1.82	-0.94	-0.94	-1.61	-1.04	-1.01	-2.27
NH_4	1.43	1.17	1.37	1.82	1.16	1.38	1.23	0.86	1.26	1.19	1.19
Asp	-0.02	-0.02	-0.01	-0.10	0.01	0.02	0.01	0.02	0.03	0.04	-0.01
Glu	0.09	0.08	0.10	0.11	0.05	0.12	0.10	0.02	0.07	0.05	0.13
Ser	-0.05	0.03	0.07	0.36	0.03	0.04	0.04	0.02	0.06	0.03	0.18
Asn	0.01	0.01	0.05	0.04	-0.01	0.05	0.05	-0.01	0.04	0.04	-0.08
Gly	-0.22	-0.22	-0.17	-0.52	-0.25	-0.19	-0.20	-0.25	-0.22	-0.23	-0.16
His	-0.07	-0.09	-0.02	-0.19	-0.09	-0.05	-0.03	-0.11	-0.07	-0.07	-0.07
Thr	-0.21	-0.22	-0.06	-0.41	-0.21	-0.07	-0.07	-0.22	-0.13	-0.17	-0.13
Arg	-0.33	-0.31	-0.02	-0.51	-0.31	-0.07	-0.03	-0.31	-0.17	-0.21	-0.20
Ala	0.57	0.51	0.65	0.27	0.30	0.57	0.51	0.40	0.63	0.51	0.55
Pro	-0.09	-0.10	-0.08	-0.17	-0.11	-0.07	-0.08	-0.13	-0.09	-0.11	-0.08
Tyr	-0.17	-0.12	-0.09	-0.25	-0.18	-0.08	-0.11	-0.18	-0.12	-0.18	-0.17
Cys	-0.11	-0.11	-0.09	-0.17	-0.25	-0.13	-0.12	-0.21	-0.15	-0.21	-0.11
Val	-0.22	-0.20	-0.12	-0.36	-0.18	-0.12	-0.14	-0.23	-0.18	-0.22	-0.19
Met	-0.04	-0.04	-0.05	-0.12	-0.07	-0.04	-0.03	-0.06	-0.05	-0.09	-0.04
Iso	-0.20	-0.17	-0.08	-0.35	-0.22	-0.11	-0.11	-0.22	-0.15	-0.20	-0.17
Leu	-0.32	-0.24	-0.16	-0.46	-0.28	-0.15	-0.17	-0.29	-0.21	-0.25	-0.31
Lys	-0.29	-0.21	-0.05	-0.46	-0.23	-0.11	-0.10	-0.26	-0.15	-0.17	-0.16
Phe	-0.10	-0.09	-0.02	-0.09	-0.25	-0.07	-0.06	-0.17	-0.07	-0.20	-0.09
Tryp	-0.01	-0.02	-0.03	-0.01	-0.06	-0.07	-0.08	-0.05	-0.03	-0.07	-0.02
Biomass	1154.97	795.86	793.88	1005.84	758.16	764.32	722.22	586.41	652.67	681.59	1041.28
Death	260.02	180.62	120.15	375.60	155.85	111.87	101.97	44.83	20.43	21.61	271.68

APPENDIX E: Code

- **calcAmac**: calculates stoichiometric matrix of the macroscopic reactions

```
function
[Amac,EM,nmodes,externalreactions]=calcAma
c(Aext,Aint,irrev)

[MET,J]=size(Aext);

%Calc AEM (Stoichiometric matrix of
internal metabolites) and AeEM
%(Stoichiometric matrix of external
metabolites)
ne=0;
ni=0;
for j=1:J
    if norm(Aint(:,j))==0
        ne=ne+1;
        AeEM.ext(:,ne)=Aext(:,j);
        AeEM.irrev_react(ne)=irrev(:,j);
        externalreactions(ne)=j;
    else
        ni=ni+1;
        AEM.ext(:,ni)=Aext(:,j);
        AEM.st(:,ni)=Aint(:,j);
        AEM.irrev_react(ni)=irrev(:,j);
    end
end
if ne==0
   AeEM.ext=0;
else
end
if ni==0
   AEM.ext=0;
   AEM.st=0;
   AEM.irrev_react=0;
else
end
%Calc internal elementary modes using
metatool
AEM= metatool(AEM);
AEM.ems= AEM.sub' * AEM.rd_ems;

%Calc macro reactions of the internal
elementary modes

AmacEM=AEM.ext*AEM.ems;

[aux,numbermodesi]=size(AmacEM);

%Calc number of external irrev reaction
nei=0;
for j=1:ne
    if AeEM.irrev_react(j)==0
        nei=nei+1;
    else
    end
end

%Calc external elementary modes
AeEM.ems=zeros(ne,nei+ne);
eem=0;
for j=1:ne
        eem=eem+1;
        AeEM.ems(j,eem)=1;
```

```
        if AeEM.irrev_react(j)==0;
            eem=eem+1;
            AeEM.ems(j,eem)=-1;
        else
        end
end

%Calc macro reactions of external elem
modes
AmaceEM=AeEM.ext*AeEM.ems;

%Calc Amac
if ni==0
    Amac=AmaceEM;
else
    Amac=[AmacEM AmaceEM];
end

%Calc number of EMs
nmodes=numbermodesi+ne+nei;

%Calc EMs from the internal and external
elementary modes matrixes and set
%the reactions order of the beginning
if ni==0
    EM=AeEM.ems;
    nmodes=nmodes-1;
else

    EM=zeros(j,nmodes);
    nee=1;
    nii=1;
    for j=1:J
        if externalreactions(nee)==j
            EM(j,:)=[zeros(1,numbermodesi)
```

```
AeEM.ems(nee,:)];
            nee=nee+1;
        else
            EM(j,:)=[AEM.ems(nii,:)
zeros(1,ne+nei)];
            nii=nii+1;
        end
    end
end
```

- **calcmus**: calculates the values of the maximal kinetic rates μ_j

```
function [MU,error]=calcmus(A,Q,C,K)

[MET,J]=size(A);
[MET,NS]=size(Q);

%Calculate the average value for every
metabolite
AVERAGE=zeros(MET,1);
for met=1:MET
    AVERAGE(met)=norm(Q(met,:),1)/NS;
end

%Create R
R=zeros(1,J*NS);

%Create B
B=zeros(MET*NS,J);

%Calculate Qnorm
```

```
Qnorm=zeros(MET,NS);
for met=1:MET
    Qnorm(met,:)=Q(met,:)/AVERAGE(met);
end

%Create Qt and Ct
Qt=zeros(MET*NS,1);
Ct=zeros(MET*NS,1);
for ns=1:NS
    Qt(MET*(ns-1)+1:MET*ns)=Qnorm(:,ns);
    Ct(MET*(ns-1)+1:MET*ns)=C(:,ns);
end
 %Calculate B
for ns=1:NS
    %Calculate R's such
averagei*bij=aij*rij;
rij=multj((S^aij/(Ki+S)^aij)
    for j=1:J
        R(j+J*(ns-1))=1;
        for met=1:MET


            if A(met,j)<0
                R(j+J*(ns-1))=R(j+J*(ns-
1))*Ct(met+MET*(ns-1))^(-
A(met,j))/(K(met)+Ct(met+MET*(ns-1)))^(-
A(met,j));
            else
            end
        end
    end

    %Calculate B
    for met=1:MET
        for j=1:J
            B(met+MET*(ns-1),j)=R(j+J*(ns-
1))*A(met,j)/AVERAGE(met);
        end
    end
end
for i=1:NS
    t(i,:)=B(17+(i-1)*MET,:);
end
t
%Calculate MU using lsqnonneg
%optional, if is wanted to force a
starting value for the iterations:
x0=0*ones(1,J);
[MU,error] = lsqnonneg(B,Qt,x0);

%Calculate error=norm
error=sqrt(error);
```

- **calcredsys**: calculates the reduced system using the stoichiometric matrix of the macroscopic reactions and the maximal kinetic rates

```
function[Ared,MUred]=calcredsys(A,MU)
[MET,NR]=size(A);
%Count the number of reactions to not
create variables that grow inside a
%loop
nrred=0;
for nr=1:NR
    if MU(nr)==0
    else
        nrred=nrred+1;
    end
```

```
end
%Calculate Ared and MUred
MUred=zeros(nrred,1);
Ared=zeros(MET,nrred);
nrred=0;
for nr=1:NR
    if MU(nr)==0
    else
        nrred=nrred+1;
        Ared(:,nrred)=A(:,nr);
        MUred(nrred)=MU(nr);
    end
end
```

- **findmodel**: combination of the three above functions

```
function[Ared,MUred,error,Amac,MU,EM,nmodes,externalreactions]=findmodel(Aext,Aint,irrev,Q,C,K)
%calculate macroscopic reactions
[Amac,EM,nmodes,externalreactions]=calcAmac(Aext,Aint,irrev);



%Check the cidnition MET*NS>J;
metabolites*samples>elementary modes
J=size(Amac,2);
[MET,NS]=size(Q);

%if MET*NS<J
%    msgBox('Warning: the number of metabolites*the number of samples < the number of elementary modes, solution can be non-optimal*/','Error');
%end

%calc kinetic rates
[MU,error]=calcmus(Amac,Q,C,K);
%calc reduced system
[Ared,MUred]=calcredsys(Amac,MU);
```

- **calcrates**: predicts the specific consumption/production rates of all the external metabolites using the reduced system

```
function [Qcalc]=calcrates(A,MU,C,K,Q)

NS=size(C,2);
[MET,J]=size(A);
Qcalc=zeros(MET,NS);
%Create B*MU
B=zeros(MET,J);
%Create R
R=zeros(1,J);
for ns=1:NS
    %Calculate R's
    for j=1:J
        R(j)=1;
        for met=1:MET
            if A(met,j)<0  %if it is substract
                R(j)=R(j)*C(met,ns)^(-A(met,j))/(K(met)+C(met,ns))^(-A(met,j));
            end
        end
    end
```

```
    %Calculate B*MU
    for met=1:MET
        for j=1:J
            B(met,j)=R(j)*A(met,j);
        end
    end
    %Calc Qcalc
    Qcalc(:,ns)=B*MU;
    if ns==6
        B(6,:)';
    end
end
```

- **plotQ**: plots q_{exp} and q_{est} in the same plot. Each plot represents one day of one CT for the metabolites that have higher q's: Glc, Lac, NH_4, Gln and Biomass

```
function
plotQ(Q,Qcalc,experiments,plotcol)
[MET,J]=size(Q);
figure
plotrows=ceil(J/plotcol);
for ns=1:J
    subplot(plotrows,plotcol,ns)
    %The position of these metabolites can
change between systems
    a1=Q(1:4,ns)'; %Glc, Lac, NH4 and Gln
    a2=Q(24,ns); %Biomass
    a=[a1,a2];
    plot((1:5),a,'b');
    hold
    b1=Qcalc(1:4,ns)';
    b2=Qcalc(24,ns);
    b=[b1,b2];
    plot((1:5),b,'r');
    grid on
    title(experiments(ns));
    xlabel('Metabolites');
    ylabel('Consumption/production
[nmol/(MVC∑day)]');
    axis ([1 5 -4200 6100])
end
```

- **plotQ2**: the same as before but for the rest of the metabolites

```
function
plotQ2(Q,Qcalc,experiments,plotcol)
[MET,J]=size(Q);
figure
plotrows=ceil(J/plotcol);
for ns=1:J
    subplot(plotrows,plotcol,ns)
    a1=Q(5:23,ns)'; %Rest of the
metabolites
    a2=Q(MET,ns); %Cell death
    a=[a1,a2];
    plot((6:MET),a,'b');
    hold
    b1=Qcalc(5:23,ns)';
    b2=Qcalc(MET,ns);
    b=[b1,b2];
    plot((6:MET),b,'r');
    grid on
```

```
    title(experiments(ns));
    xlabel('Metabolites');
    ylabel('Consumption/production
[nmol/(MVC∑day)]');
    axis([5 25 -600 650])
end
```

- **calculate_errorabsolute**: calculates the error between experimental q and estimated q in % for each samples and metablite and the average for all the metabolites

```
function
[ERROR,error_metabolite]=calculate_errorab
solut(Q,Qcalc)

[MET,J]=size(Q);
ERROR=zeros(MET,J);
error_metabolite=zeros(1,MET);
for met=1:MET
    ERROR(met,:)=(Q(met,:)-
Qcalc(met,:))./Q(met,:);
end
for met=1:MET

error_metabolite(met)=mean(abs((ERROR(met,
:))));
end
error_metabolite(:)=error_metabolite(:)*10
0;
```

- **PlotQmet**: plot the same as before but each plot represents one metabolite

```
function
plotQmet(Q,Qcalc,metabolites,plotcol)
[MET,J]=size(Q);
figure
plotrows=ceil(MET/plotcol);
for met=1:MET
    subplot(plotrows,plotcol,met)
    plot((1:J),Q(met,:),'b');
    hold
    plot((1:J),Qcalc(met,:),'r');
    grid on
    title(metabolites(met));
    xlabel('Samples');

    if strcmp(metabolites(met),'Biomass')
|| strcmp(metabolites(met),'Cell death')
        ylabel('Consumption/production
[1000∑cells/(MVC∑day)]');
    else
        ylabel('Consumption/production
[nmol/(MVC∑day)]');
    end
    YMAX=max([Q(met,:) Qcalc(met,:)]);
    YMIN=min([Q(met,:) Qcalc(met,:)]);
    if YMIN>0
        AXIS([0 J 0 YMAX+YMAX*0.2]);
    elseif YMAX<0
        AXIS([0 J YMIN+YMIN*0.2 0]);
    else

    end
```

```
end
```

- **Plotevery** and **ploteverymet** use plotQ, plotQ2 and plotQmet are used in the program in order to decide the number of plots of each figure.

- **calctimeline**: estimated the concentration over time using Eq. 31 and plot this evolution over time for each metabolite

```
function
[Ccalc,time]=calctimeline(step,days,C0,Are
d,MUred,K,nbio,Q,samples,maxplots,plotcol,
C)

nmetab=length(C0);
nstep=ceil(days/step);
time=zeros(1,nstep+1);
Ccalc=zeros(nmetab,nstep+1);
NP=ceil(nmetab/maxplots);
plotrow=ceil(maxplots/plotcol);
%Calcualtes all the steps excepting the
last one using Ccalc
Ccalc(:,1)=C0;
for i=1:nstep-1

Qcalc=calcrates(Ared,MUred,Ccalc(:,i),K,Q(
:,1));

Ccalc(:,i+1)=Ccalc(:,i)+Qcalc*Ccalc(nbio,i
)*step/(1000);
    time(i+1)=i*step;
end
%Calculates the final step which can be
shorter than step
finalstep=days-time(i+1);
i=i+1;
Qcalc=calcrates(Ared,MUred,Ccalc(:,i),K,Q(
:,i));
Ccalc(:,nstep+1)=Ccalc(:,i)+Qcalc*Ccalc(nb
io,i)*finalstep/(1000);
time(i+1)=time(i)+finalstep;
time1=[0,1];
C1=[C0 C];
for np=1:NP-1
    figure
    for nsubplot=1:maxplots
        subplot(plotrow,plotcol,nsubplot);
        plot(time,Ccalc(nsubplot+(np-
1)*maxplots,:),'r');
        grid on
        hold
        plot(time1,C1(nsubplot+(np-
1)*maxplots,:),'b');
        title(samples(nsubplot+(np-
1)*maxplots));
        xlabel('Time [days]');
        if (strcmp(samples(nsubplot+(np-
1)*maxplots),'Biomass')) ||
(strcmp(samples(nsubplot+(np-
1)*maxplots),'Cell death'))
            ylabel('Concentration
[1000∑cells/mL]');
        else
            ylabel('Concentration [mM]');
        end
```

```
    end
end
%Plot the last figures which can have less
graphs
figure

nsubplot_total=mod(nmetab,maxplots);
for nsubplot=1:nsubplot_total
    subplot(plotrow,plotcol,nsubplot);
    plot(time,Ccalc((NP-
1)*maxplots+nsubplot,:),'r');
    grid on
    hold
    plot(time1,C1((NP-
1)*maxplots+nsubplot,:),'b');
    title(samples((NP-
1)*maxplots+nsubplot));
    xlabel('Time [days]');
    if (strcmp(samples((NP-
1)*maxplots+nsubplot),'Biomass')) ||
(strcmp(samples((NP-
1)*maxplots+nsubplot),'Cell death'))
        ylabel('Concentration
[1000∑cells/mL]');
    else
        ylabel('Concentration [mM]');
    end
end
```