Mauro Guilherme Augusto Luís

Bachelor of Science in Chemical and Biochemical Engineering

Development of Principle Culture Medium Formulations for Chinese Hamster Ovary (CHO) Cells

Dissertation to obtain the degree of master in
Chemical and Biochemical Engineering

Supervisor: Rui Manuel Freitas Oliveira, Associate Professor, FCT/UNL
Co-supervisor: Nuno Eduardo Buxo Carinhas, PhD, iBET

Jury:

President: Prof. Dr. Maria da Ascensão Carvalho Fernandes Miranda Reis
Examiner: Dr. Ana Margarida Palma Teixeira

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

September 2013

Mauro Guilherme Augusto Luís

Bachelor of Science in Chemical and Biochemical Engineering

Development of Principle Culture Medium Formulations for Chinese Hamster Ovary (CHO) Cells

Dissertation to obtain the degree of master in
Chemical and Biochemical Engineering

Supervisor: Rui Manuel Freitas Oliveira, Associate Professor, FCT/UNL
Co-supervisor: Nuno Eduardo Buxo Carinhas, PhD, iBET

Jury:

President: Prof. Dr. Maria da Ascensão Carvalho Fernandes Miranda Reis
Examiner: Dr. Ana Margarida Palma Teixeira

September 2013

Development of Principle Culture Medium Formulations for Chinese Hamster Ovary (CHO) Cells

Acknowledgments

This thesis would not have been possible without the support of many people. In few lines I would like to express my gratitude to them.

To Prof. Dr. Rui Oliveira and Dr. Nuno Carinhas, for all the guidance, support and patient;

To all the elements of Systems Biology and Engineering group, for your fellowship;

To my family, especially my parents to whom I owe what I am today;

And at last, but not least, to Jorge, Joana and Tânia that made this journey easier.

To all of you,

Muito Obrigado!

Abstract

Chinese hamster ovary (CHO) cells are the most widely used mammalian cell line for the production of therapeutic proteins. To ensure their maintenance and growth *in vitro*, a complex combination of nutrients have to be supply through culture medium. Thus, in the last decades, efforts have been made to develop better culture media, free of serum. However, the methods still used for media development are too reliant on empirical knowledge and can be expensive and time-consuming. On the other hand, the method used in this thesis relies on the influence of the environment where a cell is inserted in its metabolic performance, i.e. metabolic engineering by medium manipulation.

To study the different physiological states of CHO cells, a metabolic network from the literature was adapted. In that network CHO cells produce Immunoglobulin G (IgG). Using the Metatool 5.1, an initial set of 23240 elementary flux modes (EFMs) were calculated and then divided into: 108 EFMs producing biomass, 16296 EFMs producing IgG, 6658 EFMs carrying out essentially cellular respiration and 178 EFMs assumed as non-feasible.

Each group was reduced to 20 EFMs after normalization and subsequent clustering. One medium formulation was purposed for each group of EFMs; however, since the metabolic network had to contain transport reactions of ATP, NADH, NADPH and $FADH_2$ in order to allow the EFMs calculation, the individual result of each group was not realistic because, for example, it was assumed that ATP could be transported from the "exterior" and vice-versa, according to the needs of each cellular physiological state. To consider the pathways where ATP is highly produced and those where it is required, a formulation was purposed regarding a combination of the results from the three groups.

The application of these tools may allow media development in a faster and more economical way.

Keywords

Cell culture media formulation; Chinese Hamster Ovary (CHO) cells; Elementary mode analysis; Metabolic engineering

Resumo

As células de ovário do hamster chinês (células CHO) são a linha celular de mamífero mais utilizadas para a produção de proteínas terapêuticas. Para garantir a sua manutenção e crescimento *in vitro*, uma complexa combinação de nutrientes tem de ser fornecida através do meio de cultura. Assim, nas últimas décadas, têm sido feitos esforços para desenvolver melhores meios de cultura, sem soro. Contudo, os métodos usados para o desenvolvimento de meios dependem bastante do conhecimento empírico e podem ser dispendiosos em termos de tempo e dinheiro. Por sua vez, o método utilizado nesta tese assenta na influência do ambiente onde uma célula está inserida no seu desempenho metabólico, i.e. engenharia metabólica através de manipulação do meio.

Para estudar os diferentes estados fisiológicos das células CHO, uma rede metabólica publicada foi adaptada. Nessa rede as células CHO produzem Imunoglobulina G (IgG). Usando o Metatool 5.1, obteve-se um conjunto inicial de 23240 modos de fluxo elementares (MFEs) posteriormente dividido em: 108 MFEs que produzem biomassa, 16296 MFEs que produzem IgG, 6658 MFEs que desempenham essencialmente respiração celular e 178 MFEs que foram assumidos como não viáveis.

Cada grupo foi reduzido a 20 MFEs após normalização e agrupamento de dados. Para cada conjunto foi proposta uma formulação de meio; mas, como foi necessário acrescentar à rede metabólica equações de transporte para ATP, NADH, NADPH e $FADH_2$ de forma a permitir o cálculo dos MFEs, o resultado individual de cada grupo não é realístico pois assume que, por exemplo, o ATP pode ser transportado do "exterior" e vice-versa consoante as necessidades de cada estado celular. Para considerar os caminhos metabólicos em que o ATP é altamente produzido e aqueles onde é utilizado, foi proposta uma formulação resultante da combinação das três iniciais.

A utilização deste método pode assim permitir o desenvolvimento de meios de cultura de uma forma mais rápida e económica.

Palavras-chave

Formulação de meios de cultura celular; Células de ovário do hamster chinês (células CHO); Análise de modos elementares; Engenharia metabólica

Contents

List of Figures

List of Tables

List of abbreviations

Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
Cit	Citrate
CO_2	Carbon dioxide
EMA	Elementary mode analysis
EMF	Elementary flux mode
FBA	Flux Balance Analysis
Glc	Glucose
Glu	Glutamate
Gln	Glutamine
Glyc3P	Glycero-3-phosphocholine
Glyc	Glycerol
Gly	Glycine
His	Histidine
Isobut	Isobutyrate
IgG	Immunoglobulin G
Ile	Isoleucine
Isoval	Isovalerate
Lac	Lactate
Leu	Leucine
Lys	Lysine
Mal	Malate
Met	Methionine
MFA	Metabolic Flux Analysis
Phe	Phenylalanine
Pcholine	Phosphocholine
Pro	Proline
Pyr	Pyruvate
Ser	Serine
SQ_{res}	Square of the sum of all the distances between cluster centres and each point
SQ_{tot}	Square of the sum of all points analysed
Suc	Succinate
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

1. Introduction

Production of complex therapeutic proteins is dominantly performed in mammalian cell lines because of their capacity for post-translational modifications and human like protein structure assembly [1]. Chinese Hamster Ovary (CHO) cells have become the most widely used expression system for this type of applications. Since 1987, dozens of biologics have been approved to be produced using CHO cell lines, including monoclonal antibodies used to treat cancer and immunological disorders. In 2007, nearly 70% of all recombinant protein therapeutics produced were made in CHO cells, exceeding US$30 billion worldwide of annual sales [2].

To achieve an effective *in vitro* maintenance and growth, animal cells need culture conditions similar to those found *in vivo* regarding temperature, oxygen and carbon dioxide concentrations, pH, osmolality, and nutrients. The delivery of that complex combination of nutrients to animal cells, which *in vivo* is performed through blood circulation, requires an equivalent supply performed *in vitro* through the addition of culture medium. That medium has to provide appropriate pH and osmolality for cell survival and multiplication, as well as all chemical substances required by the cells that they are unable to synthesize themselves; thus its composition is one of the most important factors regarding culture of animal cells [3].

In the formulation of new culture media, information about the cell metabolism is extremely important. In fact, as stated by Kell *et al.* [4], the physiological state inside a cell is a function of the external metabolite levels, which can be measured. This suggests an opportunity for a new method of media design; based in the "role" of each metabolite in the metabolic network and its impact on metabolic performance. In this context, metabolic modelling may provide essential tools to analyse the available information concerning the metabolic network and may return valuable clues about new culture medium components. Therefore, media development and metabolic modelling will be the main topics of this introduction.

1.1. Media Development

For more than 50 years the development of culture media for mammalian cells has been studied. As previously mentioned, the culture of animal cells requires the supply of a complex combination of nutrients equivalent to those found *in vivo*. For this reason, the first attempts were performed using media based entirely on biological fluids, such as serum and other blood or tissue extracts [5]. However, these media, which are chemically undefined, present some concerns, such as: batch to batch variation – causing inconsistency in growth-promoting properties; high protein content – hindering product purification; risk of contaminants (e.g.

viruses, mycoplasma, prions) that can also contaminate the end products; limited availability and high costs. Therefore media prepared with highly purified compounds and with known composition have been preferentially developed. These media are particularly attractive for biopharmaceutical production since they are less vulnerable to contamination and quality control is easier [3].

To be effective, a culture medium has to contain the nutrients needed for the synthesis of new cells and the substrates necessary for the metabolism, as well as the compounds that act as cofactors. Therefore, a culture medium must have : micronutrients (inorganic salts and trace elements), sugars (carbon and energy sources), amino acids (nitrogen sources), vitamins, hormones and other elements (lipids, organic acids, proteins) and water, besides cell-specific substances [3]. The functions of the main components of a typical culture medium will be briefly explained next.

1.1.1. Main media components

Amino acids

Animal cells in culture require essential amino acids (i.e., those that are not synthesized in animal tissues), plus others depending on the specific requirements of the individual cell line. Nonessential amino acids are often added to alleviate metabolic burden and achieve better growth characteristics. The limitation of an amino acid may limit the maximum cell concentration attainable and may reduce the growth rate [5, 6]. Generally, each amino acid is added to culture medium in a concentration from 0,1 to 1 mM.

According to Freshney, glutamine, methionine and serine are growth limiting amino acids [7]. In the case of glutamine, this amino acid is usually added at high concentrations (from 1 to 5 mM) and acts as source of nitrogen, carbon, and energy [3]. However, glutamine metabolism generates ammonium which has toxic effects and is a cell growth inhibitor when accumulated in large quantities [8]. Glutamate has been shown as a successful substitute for glutamine, as shown in the experiments performed by Altamirano *et al.* regarding CHO cell culture medium [9]. The results of this study will be mentioned below, as, besides glutamine, the substitution of glucose was also tested.

Glucose

Glucose is included in most media as the main source of carbon and energy. It is metabolized mainly by glycolysis, forming pyruvate which can be converted to lactate or acetoacetate. The latter can enter the citric acid cycle (TCA) and be oxidized to form carbon dioxide and water. Typically, there is a high accumulation of lactate in culture medium implying that the citric acid cycle is not functioning *in vitro* as it does *in vivo*. Furthermore,

accumulation of lactate can be a problem since it will decrease external pH and, due to its toxicity to the cells, inhibit cell growth [6], similarly to ammonium.

As previously mentioned, Altamirano *et al.* substituted glucose and glutamine in a CHO Cell Culture Medium Formulation [9]. Glucose was replaced by galactose and good results were achieved. The use of galactose and glutamate simultaneously, in place of glucose and glutamine, allowed the maintenance of cells with lower consumption of nutrients and lower generation of lactate and ammonium; however, the obtained cell growth was also lower. Interestingly, maximum cell growth was obtained using glucose and glutamate, even higher than when glucose and glutamine were used but with very low ammonium accumulation and reduced lactate levels.

Glucose is usually added to the culture medium in concentrations ranging from 5 to 25 mM (0,9 – 4,5 g/L), however it may be up to 56 mM (10g/L) [3].

Inorganic salts and trace elements

The salts are the major components contributing to the osmolality of the medium. The most used are Na^+, K^+, Mg^{2+}, Ca^{2+}, Cl^-, SO_4^{2-}, PO_4^{3-}, and HCO_3^-. In brief, the main functions of these ions are [6]:

- Ca^{2+} - required by some cell adhesion molecules; intermediary in signal transduction; can influence whether cells will proliferate or differentiate;
- Na^+, K^+, Cl^- - regulation of the cellular membrane potential;
- SO_4^{2-}, PO_4^{3-}, HCO_3^- - macromolecular synthesis; regulation of intracellular charge.

To avoid cell aggregation and adhesion in suspension cultures, the concentration of calcium and magnesium must be kept low [3]

Other inorganic elements, which can be found in serum, are usually added to the culture medium at reduced concentrations. These include elements such as iron, selenium, zinc and copper. According to the information provided by *Sigma-Aldrich Media Expert™* [10], the main functions of these elements are:

- Fe – cell respiration and energy management;
- Se – protection of cells from oxidative damage; it can act as an antioxidant or a pro-oxidant;
- Zn – important in amino and nucleic acids metabolism; roles in transcription and as an antioxidant;
- Cu – homeostasis of iron.

Other media components

Other components may be, for example, vitamins and hormones, which are present at relatively low concentrations and are utilized as essential metabolic cofactors and growth regulators, respectively.

Water used for media preparation is simultaneously one of the basic and one of the most critical components since mammalian cells are extremely sensitive to water quality [5]. To ensure the physicochemical and microbiological standards required by pharmacopoeia [11] specialized water purification systems have to be used. This process can involve four stages, which are reverse osmosis (or distillation), charcoal filtration, deionization, and micropore filtration [5].

The components in each media formulation will vary depending on the cell line for which it was designed.

1.1.2. Previous studies and media formulations

As mentioned before, the use of serum in a culture medium presents many disadvantages.

Several commercial serum-free media are available for CHO cells; however, they are expensive and their compositions are proprietary, which makes further improvement difficult.

Thus, there have been lots of efforts to develop serum-free media for the production of therapeutic proteins by mammalian cells, as well as protein-free media, which facilitates the downstream purification of recombinant proteins secreted by the cells. However, the media development efforts carried out are essentially based on experimental tests and statistical methods, such as Design of experiments. These techniques can be expensive and time-consuming, and are reliant on empirical knowledge. Some examples are briefly described below.

In 1999, Lee *et al.* used a statistical optimization approach in order to develop a serum-free medium for the production of erythropoietin (EPO) by suspension culture of recombinant CHO cells [12]. The medium was intended to be used in suspension cultures, since this type of manufacturing process has become the method of large-scale, commercial production of therapeutically important proteins from rCHO cells [13]. The approach was based on a Plackett-Burman design, where every ingredient is tested at high and low concentrations in order to identify those that are important for cell growth and/or EPO production. That group was formed by glutamate, serine, methionine, phosphatidylcholine, hydrocortisone and pluronic F68.

The starting point of this formulation was the Iscove's modified Dulbecco's medium (IMDM), which is a basal formulation of known composition (see for instance, the *life technologies*^{TM} website). The components mentioned above were added to IMDM at the highest

concentrations tested. The developed culture medium was compared with serum-supplemented IMDM. Although cell growth and EPO production where lower using the serum-free medium, the productivity per cell was higher, and the advantages of using a serum-free medium can compensate economically and qualitatively.

According to the authors of the mentioned study and Castro *et al.* [14], who performed an identical study, the classical approach of changing one medium component at a time is impractical because it is time-consuming and has the risk of neglecting interactions among the supplements.

Kim *et al.* also developed a serum-free medium formulation based upon IMDM, using, in their own words, a simple but efficient strategy [15]. This study was based in experimental work using a recombinant CHO cell line expressing a chimeric antibody against the S surface antigen of Hepatitis B virus. The effects of the addition of different supplements to the initial medium were analysed, concluding that yeast extract was the best substitute for serum in nine experimental groups established. Identical analyses were performed regarding the cell aggregation and proliferation and additional vitamins. The final formulation achieved good results, with higher viable cell density and higher productivity. The comparison included a serum-supplemented medium.

The media formulation considered in this thesis as baseline formulation was developed by Schröder *et al.* and it will be described in the next section.

In another study, the efficacy of seven commercially available serum-free media (EX-CELL, ISF-I, CD CHO, CDM4CHO, CHO-III-A, Octomed and HybridoMed) was tested [16]. It was used a transfected CHO-K1 cell line and the evaluation included cell growth and monoclonal antibody (mAb) production. In the two last mentioned media, the cell adaptation was not successful. The most recommended were EX-CELL and particularly CDM4CHO.

1.2. Metabolic modelling of bioprocesses

For more than three decades, computational tools to analyse cellular metabolism have been developed. Metabolic fluxes can be determined, being defined by the rates of enzyme-catalysed reactions from a metabolic network, and each metabolic flux vector (or distribution) defines cellular phenotype under a given growth condition [17]. There are essentially three main groups of metabolic modelling tools that are related as they are developed from the same mathematical principle but with differences in the problem formulation:

i. Metabolic Flux Analysis;
ii. Flux Balance Analysis;
iii. Metabolic Pathway Analysis.

1.2.1. Theory behind metabolic network analysis

Generally, a biological system consists of a single cell or a cell compartment that contains metabolites that are interconverted in an intricate metabolic network of enzyme-catalyzed reactions [18]. Usually the reactions that transform metabolites within the system are classified as internal reactions, whereas the reactions involving the transport of metabolites in and out of the system are considered as exchange reactions [19].

The analysis of metabolic networks is based on the first principle of mass conservation of internal metabolites within a system [20-22]. In a system of defined volume, that principle can be generally described by the equation:

$$\frac{d}{dt}C = S \cdot r - \mu . C. \quad (1)$$

where C (mol/L) is the concentration vector of m internal metabolites, r (mol/L/h) is the reaction rate (flux) vector of n reactions that convert metabolites, S is the stoichiometry matrix of dimension $m_x n$ whose elements s_{ij} represents the stoichiometry coefficient of the metabolite i involved in reaction j, μ (h^{-1}) is the specific growth rate associated with the change in volume of the system. The later can be considered as negligible since the growth rate is much slower than the reaction rates thus not changing significantly the concentrations of metabolites. At steady state, there is no accumulation of internal metabolites in the system so Equation (1) can be simplified to:

$$S \cdot r = 0. \quad (2)$$

which, in the case of cellular metabolism, is typically an undertermined system with as many unknowns as the number of reactions.

On the other hand, thermodynamic constraints imply reactions have to proceed in the appropriate direction thus irreversible reactions have positive or null flux values [17]
:

$$r_i \geq 0. \quad (3)$$

1.2.2. Approaches

The use of Metabolic Flux Analysis (MFA) to solve this system implies the knowledge of some experimentally measured fluxes. Thus, the flux vector is portioned into an unmeasured flux vector (r_u) and an measured flux vector (r_m) resulting in the following equation:

$$S_u \cdot r_u = -S_m \cdot r_m \quad (4)$$

With a sufficient number of experimentally determined fluxes, matrix S_u is invertible and r_u can be calculated however, that demonstrates how reliant on experimental data this technique is. The solution obtained is a single metabolic flux vector (r) for the measured fluxes considered, hence different growth conditions will result in a different metabolic determination.

When the known measured fluxes are not enough or the matrix S_u is not invertible, Flux Balance Analysis (FBA) can be used. With this tool, the metabolic flux vector is determined by imposing an objective function (e.g. maximize the product or the specific growth rate) and some constraints (e.g. substrate uptake rate and thermodynamic constraints). Like the previous mentioned technique, FBA only determines a metabolic flux vector under given growth conditions, which does not exclude the existence of alternative optimal solutions or suboptimal solutions, and although this tool does not require so many experimental data as MFA, more measured fluxes will lead to a more accurately determined flux vector.

Unlike the previous methods, metabolic pathway analysis does not require measured fluxes neither the imposition of an objective function, and it allows the identification of all metabolic flux vectors possible for a metabolic network. Network-based elementary flux modes and extreme pathways define uniquely and mathematically biochemical pathways directly from the structure of the metabolic network characterized by its reactions and metabolites [23].

As Schuster *et al.* mentioned an elementary mode is defined as a minimal set of enzymes that could operate at steady state, and any elementary mode has to use the irreversible reactions in the appropriate direction [24]. Therefore, these pathways have to be non-decomposable which means if an active flux is restricted to zero then the flux through the entire pathway must be zero.

Geometrically, these pathways can be visualized as vectors in a space with as many dimensions as the number of reactions in the metabolic network, where the numerical value on a given axis is the flux level in the corresponding reaction [23]. The allowed flux distributions are within the admissible flux space, kwon as convex polyhedral cone, illustrated in Figure 1.1 [25].

The edges of the convex cone correspond to the unique extreme pathways that mathematically describe the minimal set of convex basis vectors needed to describe all allowable steady-state flux distributions through the metabolic network [26]. Extreme pathways are obtained when additionally to the constraints considered for elementary mode analysis (EMA) systemic independence is required, which means no extreme pathways can be represented by non-negative linear combinations of other extreme pathways. Therefore, extreme pathways are a minimal subset of the elementary flux modes. When all the reactions in a metabolic network are irreversible the extreme pathways and the elementary flux modes result in the same set of pathways.

Pathway analysis can be useful for predictions of a minimal medium, identifying which substrates are required for which biomass constituents. For example, that was made by Schilling

et al. for *Haemophilus influenza* and *Helicobacter pylori* with computed compositions consistent with experimental results [27, 28].

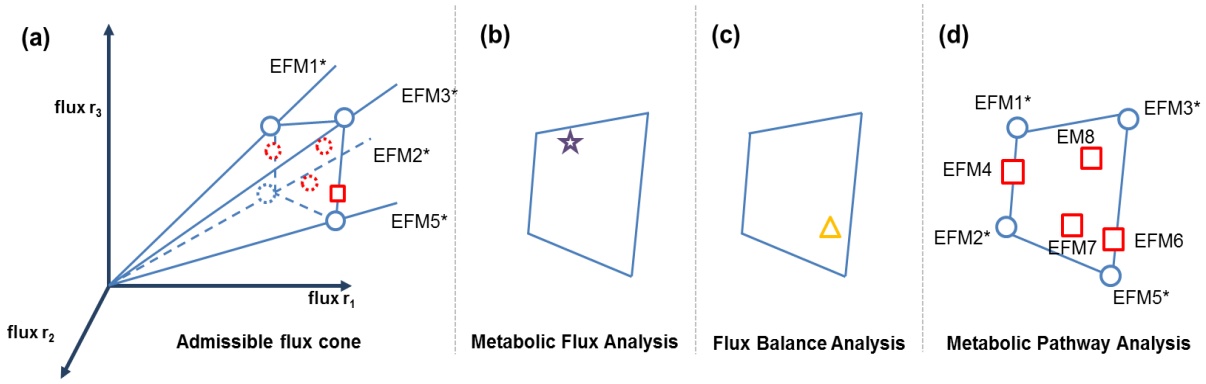

Figure 1.1 – Geometric interpretation of a metabolic network analysis. (a) All possible flux distributions of a metabolic network lie within the admissible flux cone. Each axis corresponds to the flux through a given reaction. The edges of the cone are represented by the extreme pathways (represented by ○ and by the elementary flux modes marked with an asterisk). (b) Metabolic Flux Analysis identifies only a combination of fluxes that lies anywhere in the cone (☆); (c) as well as Flux Balance Analysis that represents only a combination of fluxes that lies also anywhere in the cone (△) and satisfies the defined objective function. (d) The remaining elementary flux modes lie on the face and inside the cone (□). Adapted from [17].

1.3. Objectives

As shown in the first part of this introduction, even using statistical tools to enhance data analysis, the methods still used for media development are too reliant on empirical knowledge; thus, that can be expensive and time-consuming.

It has been shown that the environment where a cell is inserted can influence its performance; however, metabolic engineering by medium manipulation has not been used.

Therefore, the main objective of this thesis is to leverage the knowledge of the metabolic network of CHO cells to develop principle culture medium formulations that cover their distinct metabolic states. The ultimate aim is to establish a culture medium which allows higher yields of cell growth and product. The main work can be split into the following tasks:

1. Implementation of a "parsimonious" metabolic network for CHO cells which includes a biomass synthesis reaction and a product synthesis reaction;
2. Computation of elementary flux modes in order to identify all possible biochemical states of the network; the complexity of the network cannot be too large for this purpose;

3. Clustering of elementary flux modes to obtain a representative and reasonable group of biochemical states;
4. Formulation of culture medium based on a conjugation between bibliography review and the results obtained in the *in silico* analysis.

2. Materials and methods

2.1. Metabolic network

The metabolic network considered in this thesis results from an adaptation of the network of central metabolic pathways of glutamine-synthetase (GS)-CHO cells, producing Immunoglobulin G (IgG), developed by Carinhas *et al.* [29], which was based on published models [30-32] and complemented with reactions taken from biochemistry textbooks in agreement with the results provided by ^{1}H-NMR exometabolomic analysis. This network is composed by 117 reactions, involving the pathways of glycolysis, pentose phosphate pathway (PPP), TCA cycle, glutaminolysis, amino acids metabolism, by-product formation and extracellular transport fluxes. The cell biomass composition was taken from the literature [33].

Due to the high number of reversible reactions in the original metabolic network it was not possible to calculate all the elementary flux modes, since the number of modes grows exponentially with the size of the network. This created the need to simplify the network. Thus, for most of reversible reactions it was defined a single flux direction; and some of the internal (balanced) metabolites were left out. Since the studied CHO cells were generated by transfection with a construct containing glutamine synthetase (GS), which is the enzyme responsible for the biosynthesis of glutamine from glutamate and ammonia, they can grow in a glutamine-free medium. Thus, glutamine transport flux was omitted. The extracellular transport fluxes of ATP, NADH, NADPH and FADH2 were added to the network, as these were considered as non-balanced metabolites. It should be noticed that these molecules are not added to culture medium as a supplement; they are produced and used by the cell. However, it was necessary to add these reactions to the metabolic network in order to allow the calculation of elementary flux modes.

The resulting adapted metabolic network is formed by 119 reactions (9 reversible), with 126 metabolites (41 external or non-balanced). The metabolites with suffix "ext" were defined as external, or non-balanced. The list of reactions is shown in Table 6.1 (Appendix A) and a simplified representation with the main reactions is shown in Figure 2.1.

In the reaction of biomass production (r_{69}), the coefficients represent the amount of each metabolite in nmol required to produce 10^6 cells; and in the reaction of IgG production (r_{78}), the coefficients are also in nmol and represent the amount required to produce 1 mg of IgG.

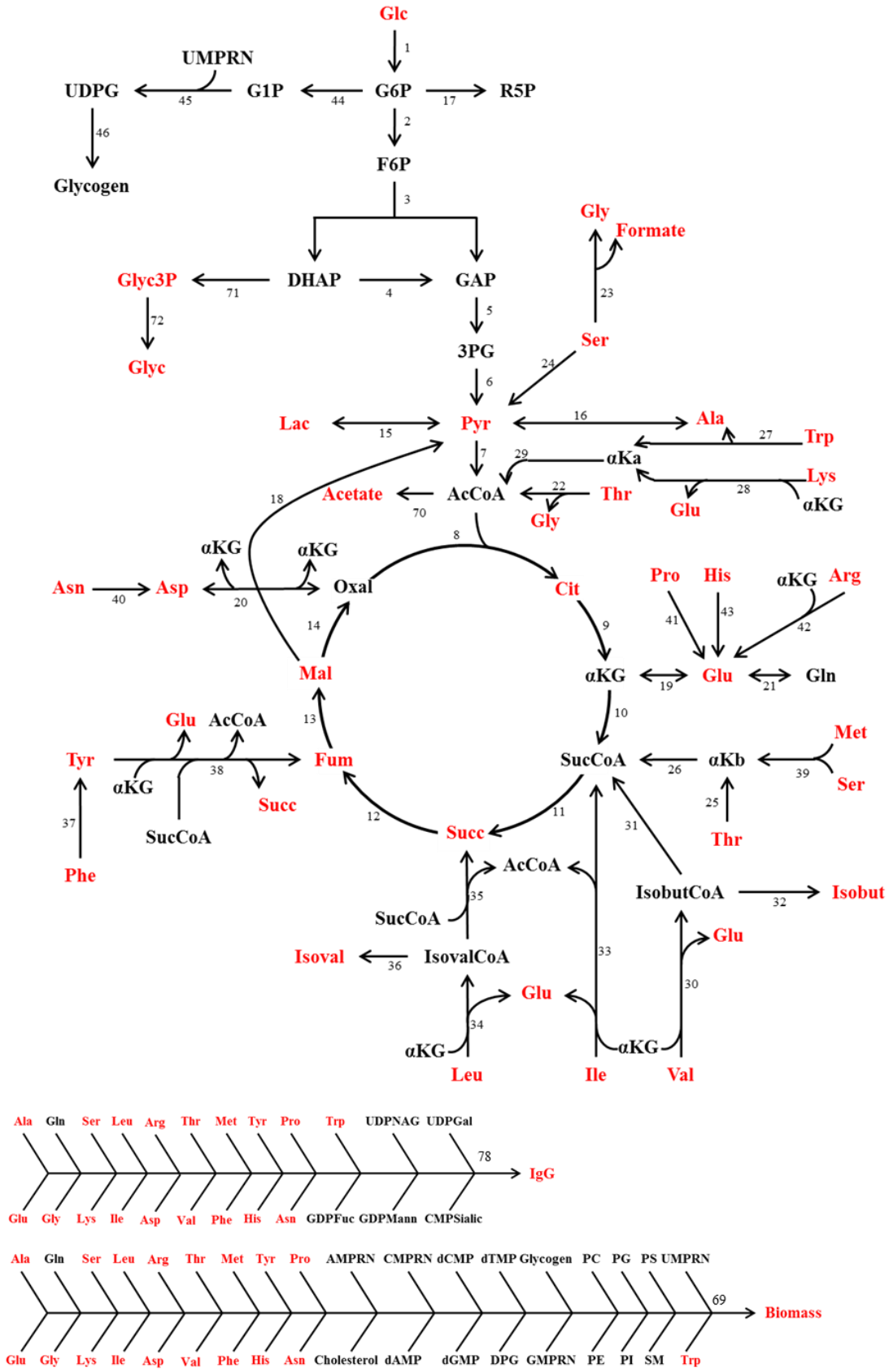

Figure 2.1 - Metabolic network of CHO cells (adapted from [29]). Nucleotide synthesis, lipid synthesis, IgG glycosylation and transport reactions were omitted. Metabolites marked in red were implemented as external.

2.2.Elementary mode analysis

In this study, metabolic modelling was performed using elementary mode analysis. This tool can decompose the intricate metabolic network comprised of highly interconnected reactions into uniquely organized pathways which, as previously mentioned, consist of a minimal set of enzymes that can support steady state operation and each one representing an independent cellular physiological state [17]. Furthermore, elementary flux modes (EFMs) have to be independent of each other, i.e. the reactions involved in one EFM must not be a subset of the reactions of any other EFM.

EFMs were calculated using the program Metatool 5.1 [34, 35], freely available on http://pinguin.biologie.uni-jena.de/bioinformatik/networks/; interfaced with MATLAB (version 8.0.0.783 64-bit) in a 3,6 GHz PC with 4 GB RAM. To carry out this analysis, the metabolites classified as external were listed and the 119 reactions from the metabolic network were implemented.

Results from the Metatool are provided as a MATLAB structure, composed by several matrices, including the reduced system matrix ("rd") and the matrix with the elementary flux modes of the reduced system ("rd_ems"). These two matrices can be used to examine the numerical quality of the result. Multiplying the first one by the second, the maximum absolute value in the result must be zero if all the stoichiometric coefficients are integers, or at least a very small number.

Inside that structure is the matrix "ems", whith rows corresponding to metabolites and columns to EFMs. However, for this study, which regards media formulation, external metabolites are the main interest. The matrix "ext", also in the structure, represents all EFMs (columns) versus external metabolites (rows). Thus, multiplying "ems" by "ext" allows obtaining a matrix that correlates EFMs (columns) with the external metabolites (rows).

2.3.Data normalization and clustering

Each elementary flux mode represents an independent physiological state of the cell, from which a particular medium formulation could be developed. However, it would not be reasonable obtaining thousands of formulations for all EFMs, and many EFMs can be similar so they can be grouped in the same set. Thus, by performing a cluster analysis one can take a small and representative subset of EFMs.

Before cluster analysis some steps had to be implemented. The first was the normalization of the matrix containing the EFMs versus external metabolites. Normalization was done in two ways. First, all the stoichiometric coefficients of each EFM (i.e. each column of the matrix) were divided by the norm of the vector which represents that EFM. This step

aims to prevent two EFMs that are linearly dependent to be in different clusters; hence, after this operation, those two EFMs will have the same coefficients and will be grouped in the same set.

The second normalization was performed by row. The coefficients for each metabolite were divided by the maximum absolute value of the corresponding row in order to ensure that all metabolites have coefficients in the same range and their contribution to the cluster analysis will not be biased. Importantly, it must be taken into account that this second normalization can only be used for the clustering step, since each coefficient in the same EFM will be divided by a different value and in consequence the resulting EFM will not be the same. Therefore, this normalization is reversed after clustering; multiplying the coefficients of each metabolite by the same value they were previously divided by. After normalization and before the cluster analysis, EFMs were divided into three groups according to their contribution to the production of: only biomass; only IgG; only CO_2 with no production of biomass or IgG. Clustering was applied to these groups.

The cluster analysis was implemented in MATLAB using the *k-means* algorithm, a partitional clustering technique that is one of the simplest and most commonly used algorithms employing a squared error criterion [36]. Using *k-means* implies the predefinition of the number of clusters intended. To determine *k* clusters, the algorithm randomly chooses *k* cluster centres or centroids, then assign each EFM to the closest cluster centre, and finally after all the EFMs are assigned recalculates the centre of each cluster. In this study, 10 clusters were calculated for EFMs subsets. The choice of the number of clusters was made considering an acceptable explained variance of the data:

$$Explained\ Variance = 1 - \frac{\mathrm{SQ_{res}}}{SQ_{tot}} \tag{5}$$

where SQ_{res} is the square of the sum of all the distances from each point (representing each EFM) to the respective cluster centre; and SQ_{tot} is the square of the sum of all points analysed.

Finally, the randomness of the initial choice can be a limitation when it is not done properly. To avoid it, the algorithm was set to repeat the operation 5 times, each with a new initial set of centres, and chose the set with the lowest SQ_{res} value. Additionally, the maximum number of iterations allowed was increased to 1000.

2.4.Medium formulation

As mentioned in the introduction of this thesis, several serum-free media for CHO cells have been developed and are commercially available. However, their formulations are proprietary and cannot be found in the literature. From the published formulations found, the

one proposed by Schröder *et al.* [37] seemed to be the most complete. This formulation was used to compare with the results obtained.

Table 2.1 - Composition of serum-free medium CHO-T1-SF [37].

Component	Specific concentration (mg/l)	Molar concentration (mM)
Inorganic salts		
$CaCl_2$	211,288	1,9
$CuSO_4 \cdot 5H_2O$	0,0012	$4,81 \times 10^{-6}$
$FeSO_4 \cdot 7H_2O$	0,3336	$1,20 \times 10^{-3}$
KNO_3	0,0912	$9,02 \times 10^{-4}$
KCl	485,44	6,51
$MgCl_2$	22,888	0,24
$MgSO_4$	117,204	0,576
NaCl	3039,6	52
$NaHCO_3$	4099,2	48,8
Na_2HPO_4	56,816	0,4
$NaH_2PO_4 \cdot H_2O$	150	1,25
$Na_2SeO_3 \cdot 5H_2O$	0,02601	$9,89 \times 10^{-5}$
$ZnSO_4 \cdot 7H_2O$	0,3452	$1,20 \times 10^{-3}$
Carbohydrates		
d-Glucose	6120,8	34
l-Amino acids		
l-Alanine	33,56	0,377
l-Arginine·HCl	185,2	0,879
l-Asparagine·H_2O	40,084	0,267
l-Aspartic acid	41,32	0,31
l-Cysteine·HCl·H_2O	14,048	0,08
l-Cystine·2HCl	109,488	0,349
l-Glutamic acid	95,88	0,652
l-Glutamine	1052,28	7,2
Glycine	39	0,52
l-Histidine·HCl·H_2O	58,784	0,28
l-Isoleucine	127,576	0,973
l-Leucine	131,24	1
l-Lysine·HCl	189,8	1,04
l-Methionine	37,792	0,253
l-Phenylalanine	81,184	0,491
l-Proline	61,8	0,537
l-Serine	54,6	0,52
l-Threonine	118,76	0,997
l-Tryptophan	20,016	0,098
l-Tyrosine, disodium salt	128,164	0,569
l-Valine	117,48	1
Vitamins and miscellaneous compounds		
d-Biotin	0,01852	$7,58 \times 10^{-5}$
dl-Pantothenic acid, calcium salt	4,992	$1,09 \times 10^{-2}$
Choline chloride	10,384	$7,44 \times 10^{-2}$
Ethanolamine	1,22	0,02
Folic acid	5,32	$1,21 \times 10^{-2}$

Table 2.1 (Continued)

Component	Specific concentration (mg/l)	Molar concentration (mM)
Hypoxanthine	13,6	0,1
i-Inositol	15,84	$8{,}79 \times 10^{-2}$
Linoleic acid	0,0336	$1{,}20 \times 10^{-4}$
Lipoic acid	0,084	$4{,}07 \times 10^{-4}$
Methotrexate (MTX)		1×10^{-4} to 0,1
Nicotinamide	48,148	$3{,}94 \times 10^{-2}$
Phenol red	18,48	$5{,}21 \times 10^{-2}$
Pluronic F-68	2000	–
Pyridoxal·HCl	4,8	$2{,}36 \times 10^{-2}$
Pyridoxine·HCl	0,0248	$1{,}21 \times 10^{-4}$
Riboflavin	0,4952	$1{,}32 \times 10^{-3}$
Sodium pyruvate	276	2,51
Thiamine·HCl	4,936	$1{,}46 \times 10^{-2}$
Thymidinea	3,88	0,016
Vitamin B_{12}	0,5596	$4{,}13 \times 10^{-4}$
Peptones and proteins		
Fetuin	10	–
Insulin	5	–
holo-Transferrin	5	–
Casein peptone soybean flour peptone broth	5000	–

2.5. Experimental data

In order to compare the metabolite concentrations achieved from the results of elementary mode analysis and subsequent cluster analysis, experimental data were considered. A set of concentration profiles of 36 external metabolites was analyzed by Carinhas *et al.* [29]. The study carried out regarded cultures of GS-CHO cells expressing IgG_4. The supernatants were analysed using ^{1}H-nuclear magnetic resonance (^{1}H-NMR).

The data used here are the measurements performed using the lower producer (LP) CHO cell clone without butyrate treatment. Biomass and IgG productions were also measured. In the comparisons carried out in this thesis, the initial measured values (i.e. age=0h) were assumed for the consumed metabolites and final concentrations for biomass and IgG.

2.6. Cost estimate

In order to provide information essential for an estimate of the cost of a culture medium formulation, the price of each supplement added was obtained and compiled in Table 6.2 (Appendix B). All the prices are from *Sigma-Aldrich* and are available on the company website (www.sigmaaldrich.com).

Since the range of package sizes varies depending on the product, it was chosen as criterion to present the price for the larger package, which corresponds to the lowest unit price

(i.e. lowest price per gram). Thus, the price estimate for a formulation using the values from this table will correspond to the lowest price achievable.

All the prices were gathered on 2 Aug. 2013.

3. Results and Discussion

3.1. Elementary mode analysis

The calculation of EFMs for the metabolic network listed in the Table 6.1 performed by Metatool 5.1 in MATLAB (version 8.0.0.783 64-bit; shared libraries compiled for the 32-bit were not used in this study) took approximately one hour.

Before the calculation, Metatool pre-analyses the metabolic network. In that pre-analysis it was confirmed that all the balanced metabolites were consumed and produced at least in one reaction each, as well as that none of the metabolites takes part in only one reversible reaction and all the metabolites are used. Therefore, all the reactions in the metabolic network were considered by Metatool.

In total, 23240 EFMs were computed by the algorithm. Figure 3.1 has a graphical representation of the coefficients of each external metabolite for each EFM. The representation was made as an heat map, where the green colour represents metabolites with negative stoichiometric coefficients, i.e. the metabolite is consumed, while metabolites represented in red are produced. All the heat maps shown in this thesis were obtained using MALTAB function *HeatMap*.

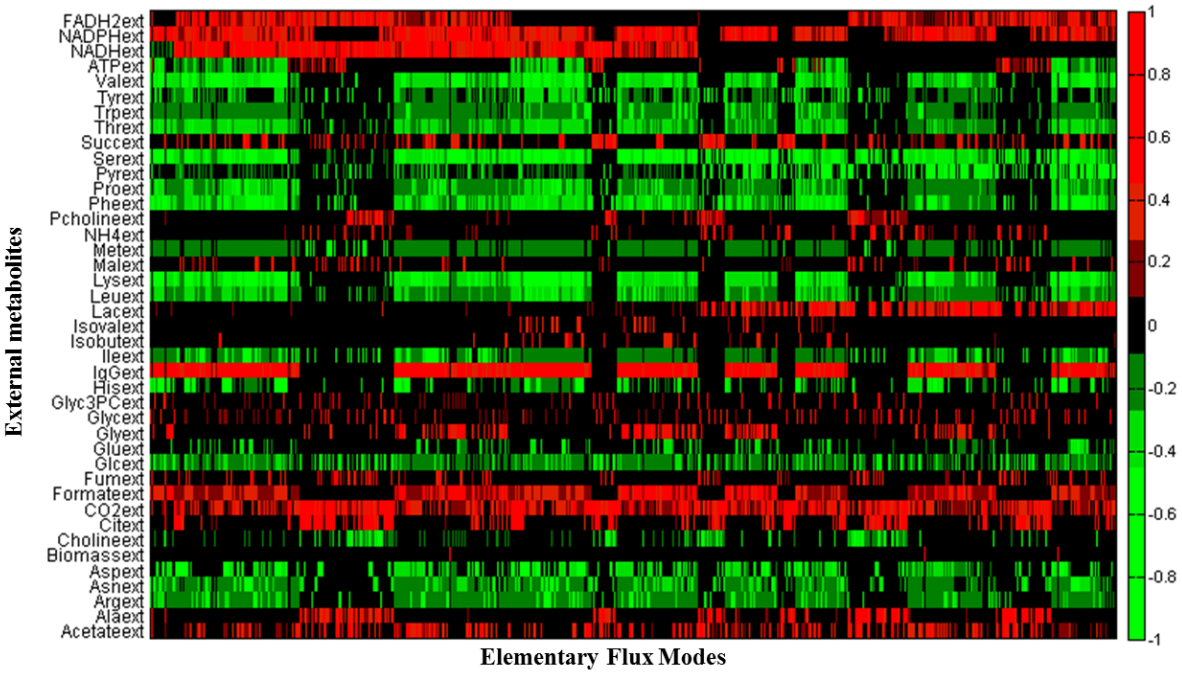

Figure 3.1 – Heat map of all EFMs initially obtained (normalized data). Metabolites in green are consumed and in red are produced.

The numerical quality of the result was examined by multiplying the reduced system matrix by the matrix with elementary flux modes of the reduced system (see Section 2.2). The maximum absolute value obtained was approximately $3{,}27 \times 10^{-8}$, which is in agreement with the required (i.e. a very small number; or even zero when all the stoichiometric coefficients are integers).

Before the construction of the heat map in Figure 3.1, the matrix with the data of external metabolites vs. EFMs was normalized, as described in Section 2.3. This normalization is very important or the graph would have been essentially occupied by a black area, since the scale of colours used depends on the range of the coefficients. Thus, if a metabolite, for instance glucose, has high values compared to the other metabolites, it would be marked on the heat map, but the remaining metabolites with smaller values would lay down in the dark area as their coefficients are not representative when compared to the higher one. After normalization, all the coefficients of all the metabolites range between -1 and 1 and the heat map is more illustrative.

In Figure 3.1 it can be seen that all metabolites, except ATP, have the same behaviour (i.e. are always consumed or produced) through the EFMs represented. ATP will be produced or consumed depending on the energy credit of each EFM, since ATP was considered as an "external" metabolite (i.e. unbalanced). Thus, ATP will be produced if the EFM is related to cellular catabolism or it will be consumed in EFMs related to cellular anabolism. It should be noted that in practice ATP is not available in the culture medium and its transport reaction does not occur. Catabolism and anabolism occur simultaneously in a balanced way. However, as mentioned in Section 2.1, the transport of ATP as well as the transport of NADH, NADPH and $FADH_2$ had to be added to the metabolic network in order to be possible the calculation of EFMs by Metatool.

In the same figure, it can be observed that biomass is produced, or at least it is represented, in few EFMs. In fact, when the EFMs are grouped according to its function (see Section 2.3) one can notice that the pathways producing biomass are a minority. Therefore, the group originally formed by 23240 EFMs was divided into:

- 108 EFMs producing only biomass;
- 16296 EFMs producing only IgG;
- 6658 EFMs producing CO_2, with no production of biomass or IgG;
- 178 remaining EFMs with other functions.

To note that none of the EFMs produced biomass and IgG simultaneously. Each of these groups will be analysed in more detail next in this thesis.

3.2. Biomass production

The EFMs that produce biomass are represented in Figure 3.2. It can be seen that many metabolites are not observable in the heat map graph. This does not mean that they do not

participate in the EFMs, but their coefficients may be too low when compared to the values associated with metabolites highly consumed/produced.

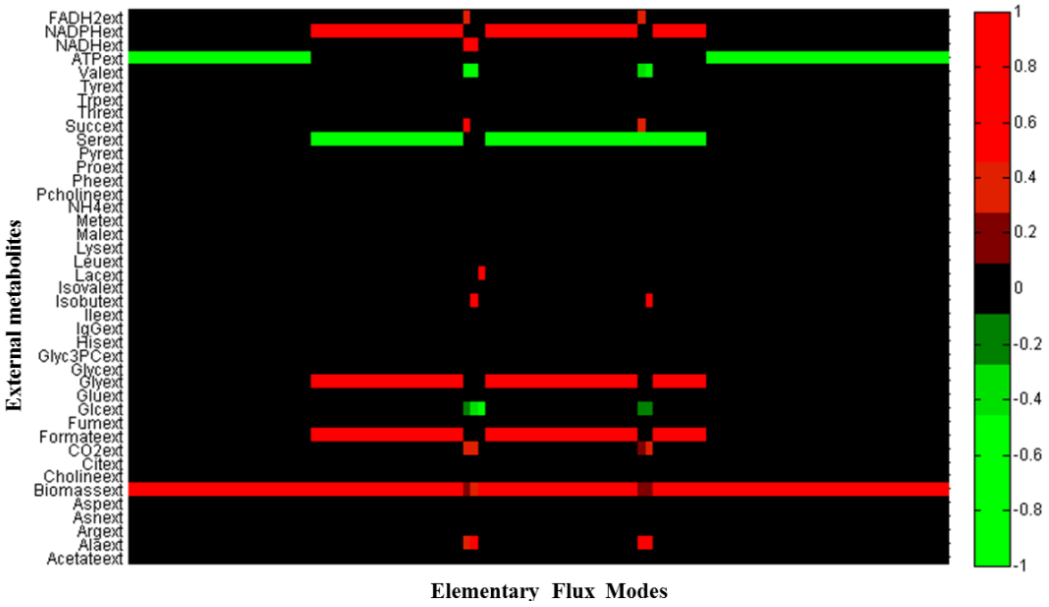

Figure 3.2 - Heat map of all the EFMs which produce biomass (normalized data). Metabolites in green are consumed and in red are produced.

Although this is one of the smallest groups of EFMs obtained, in order to have a representative sample of the cellular physiological states, where biomass is produced, cluster analysis was performed using MATLAB *kmeans* function.

The data were grouped in 10 clusters, with an explained variance of 99,6%.This value of explained variance increases as more clusters are calculated. In this case, it will not be significant since the value is already very high.

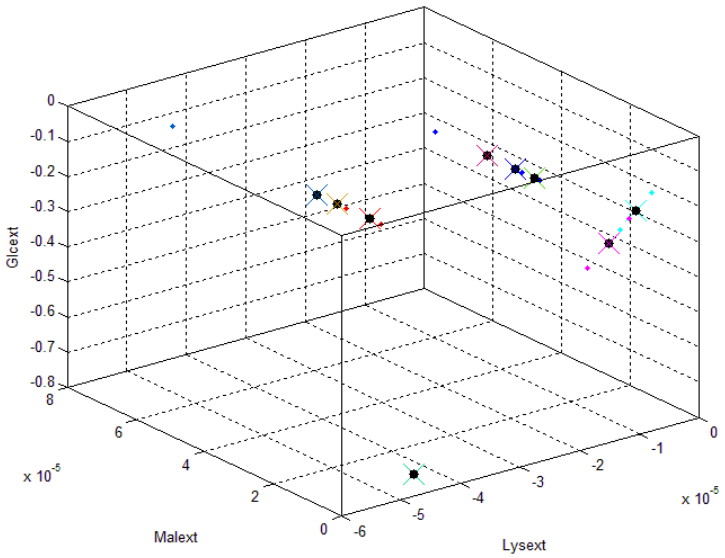

Figure 3.3 – Three-dimensional representation of biomass EFMs clustering. Each coloured dot represents an EFM and each colour a cluster. Black dots are the centroid of each cluster.

To perform the cluster analysis 41 variables have to be considered; one for each external metabolite. It is not possible to graphically represent the data with all the variables. However, in order to provide a simple illustrative example, data were represented considering the coefficients of glucose, malate and lysine in Figure 3.3. Although there are 108 possible EFMs in this group, which correspond to 108 points plotted on the graph, only a few can be visually distinguished. This suggests a possible collinearity of the data, i.e. many EFMs are similar in most of the coefficients but not entirely equal. Thus, it is not surprising the high value of explained variance achieved in the clustering step.

In this discussion it will be shown only the results for the following metabolites: alanine, arginine, asparagine, choline, glucose, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, pyruvate, serine, threonine, tryptophan, tyrosine, and valine; as well as the production of biomass, IgG and carbon dioxide, according to the analysis performed. However, the whole set of results are available in the appendices. The metabolites mentioned were chosen because they are usually consumed and consequently they are interesting regarding media formulation

The coefficients of the 10 clusters (after the second normalization is reversed, see Section 2.3) are shown in Table 6.3 (Appendix C) with the correspondent standard deviations. In this discussion, in order to be more comprehensible and easier to compare with literature data, the values of each metabolite in each cluster (for all the clustering results) will be shown as medium concentrations proportional to an initial glucose availability of 40 mM. This value was the initial concentration measured by Carinhas *et al.* [29]. Literature data comes from the formulation shown in Table 2.1 (page 15).

The 10 cluster centres are also graphically represented as a heat map in the Figure 3.4 and 5 of them were considered for Table 3.1. Cluster 5 was removed because it is identical to the first one; cluster 9 is identical to cluster2; clusters 4, 7 and 10 produce very little amounts of biomass. Cluster 4 produces essentially lactate; and clusters 7 and 10 correspond basically to cellular respiration, since they are the two states from this set that produce more CO_2, but producing a little amount of biomass simultaneously.

Glutamate is not consumed, and also glutamine since it was assumed this was not part of the medium when the transport reaction for this metabolite was omitted in the metabolic network. Consequently, it is not surprising that ammonium is not produced. Aspartate and glycine are not part of the medium formulation as well, according to these results, since they can be produced from asparagine and serine, respectively. In fact, serine is highly consumed in these cellular metabolic states, as also mentioned in the Introduction (Section 1.1) based on literature review. However, the value of serine concentration required for cluster 6 is extremely high. This metabolite is being used to produce energy. As shown in Figure 3.4, when serine is highly consumed, the cell does not need ATP from an external source.

Table 3.1 – Medium composition from clusters 1, 2, 3, 6 and 8 (aiming biomass production) and from published data [29, 37].

Metabolites	Cluster1	Cluster2	Cluster3	Cluster6	Cluster8	Carinhas *et al.*	Schöder *et al.*	Units
Arg	14,69	5,05	7,59	8,00	5,98	2,64	0,88	mM
Asn	62,62	21,53	32,34	34,11	25,48	4,83	0,27	mM
Choline	3,34	27,39	3,70	1,82	1,36	0,60	-	mM
Glc	40,00	40,00	40,00	40,00	40,00	40,04	34,00	mM
Glu	-	-	-	-	-	1,88	0,65	mM
His	5,80	1,99	2,99	3,16	2,36	1,13	0,28	mM
Ile	13,22	4,55	6,83	7,20	5,38	2,60	0,97	mM
Leu	61,61	21,19	31,82	33,55	25,07	3,92	1,00	mM
Lys	23,58	8,11	12,18	12,84	9,59	3,16	1,04	mM
Met	6,40	2,20	3,31	3,49	2,61	0,86	0,25	mM
Phe	10,78	3,71	5,57	5,87	4,39	1,22	0,49	mM
Pro	57,76	19,86	29,83	31,46	23,50	4,71	0,54	mM
Pyr	86,42	-	1,98	19,83	-	1,50	2,51	mM
Ser	67,63	25,07	44,46	$1{,}02 \times 10^{6}$	36,54	5,71	0,52	mM
Thr	15,32	5,27	7,91	8,34	6,23	2,85	1,00	mM
Trp	1,48	0,51	0,76	0,80	0,60	0,98	0,10	mM
Tyr	7,60	2,61	3,92	4,14	3,09	0,90	0,57	mM
Val	55,36	19,04	28,59	30,15	22,53	2,97	1,00	mM
Biomass	159,93	55,00	82,59	87,10	65,08	7	n.a.	10^{6} cells/mL

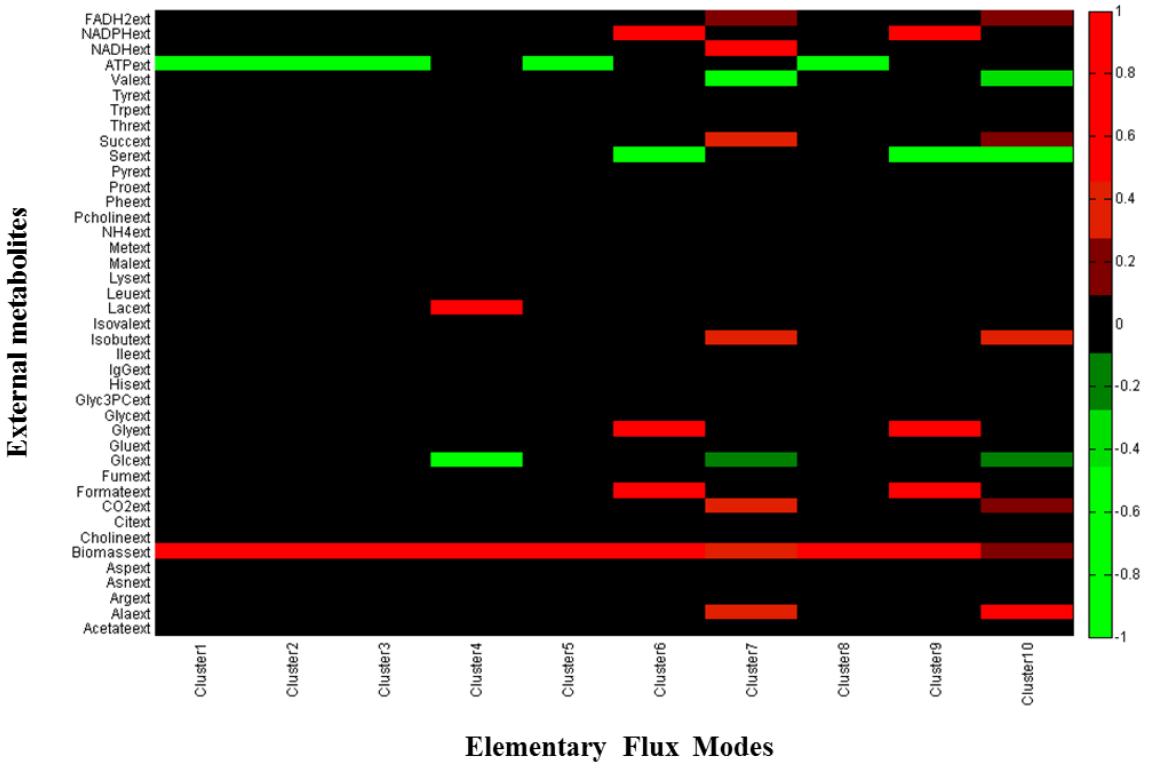

Figure 3.4 - Heat map of pathways obtained after cluster analysis of biomass production EFMs (normalized data). Metabolites in green are consumed and in red are produced.

The amount of biomass produced seems to be too high when compared to the amount achieved by Carinhas *et al.*: 7 millions of cells per millilitre. However, it must be taken into account that, as previously mentioned, due to the four last transport reactions added to the metabolic network (see Section 2.1 and Appendix A) it is not imposed that catabolism and anabolism have to coexist in the same cell. Actually, as it can be seen in Figure 3.4, in several cluster centres ATP is consumed from an external source. In practice that does not happen since the culture medium is not supplied with ATP; then the cell has to produce it in the catabolic process. Thus, in these metabolic states, glucose is being used essentially for biomass production. Actually, the ratio of consumed glucose per produced biomass for cluster 2 is approximately 727 nmol Glc/10^6 cells while in the experiments by Carinhas *et al.* it was near 5714 nmol Glc/10^6 cells.

Cluster 2 seemed to be the metabolic pathway with the most reasonable set of concentrations required in the medium. Figure 3.5 compares the achieved values in that cluster with the values from the literature. It can be seen that most of the metabolites are required at higher concentrations in this formulation, which is justified with the much higher production of biomass. This also happens in the other centroids with high production of biomass.

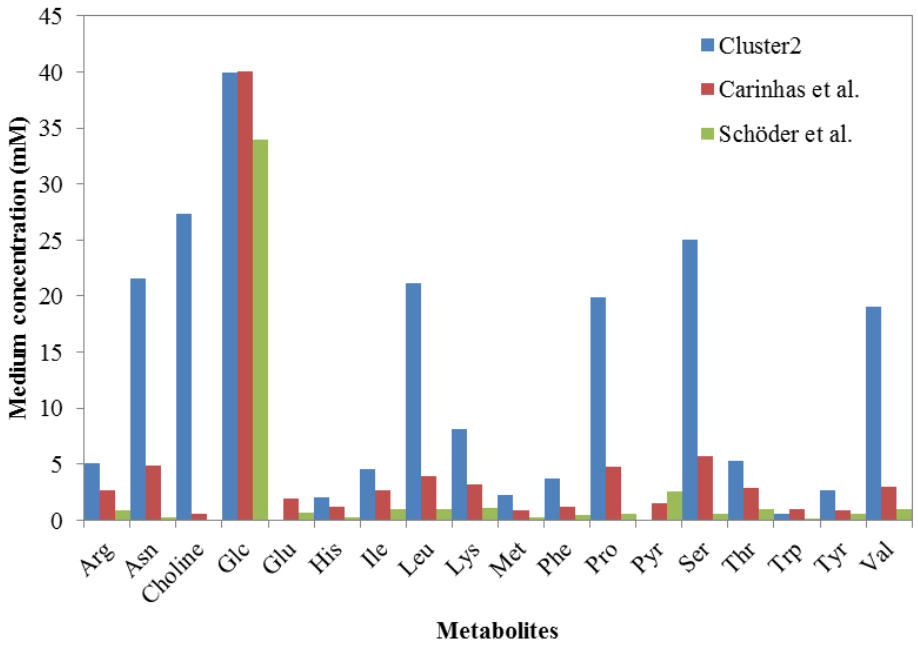

Figure 3.5 – Comparison between the concentrations from the formulation of cluster 2 (biomass production) and from published data [29, 37].

3.3. IgG production

The 16296 EFMs with IgG production are represented in Figure 3.6. As this heat map suggests, and contrary to the results depicted in Figure 3.2 with the EFMs of biomass production, most of the metabolites take part in significant amount in the pathways aiming IgG production. Once more, the behaviour (i.e. consumption or production) of metabolites seems to be preserved regardless of the EFM in which it takes part. The exception is again ATP, which is consumed in most of the EFMs but produced in a few of them.

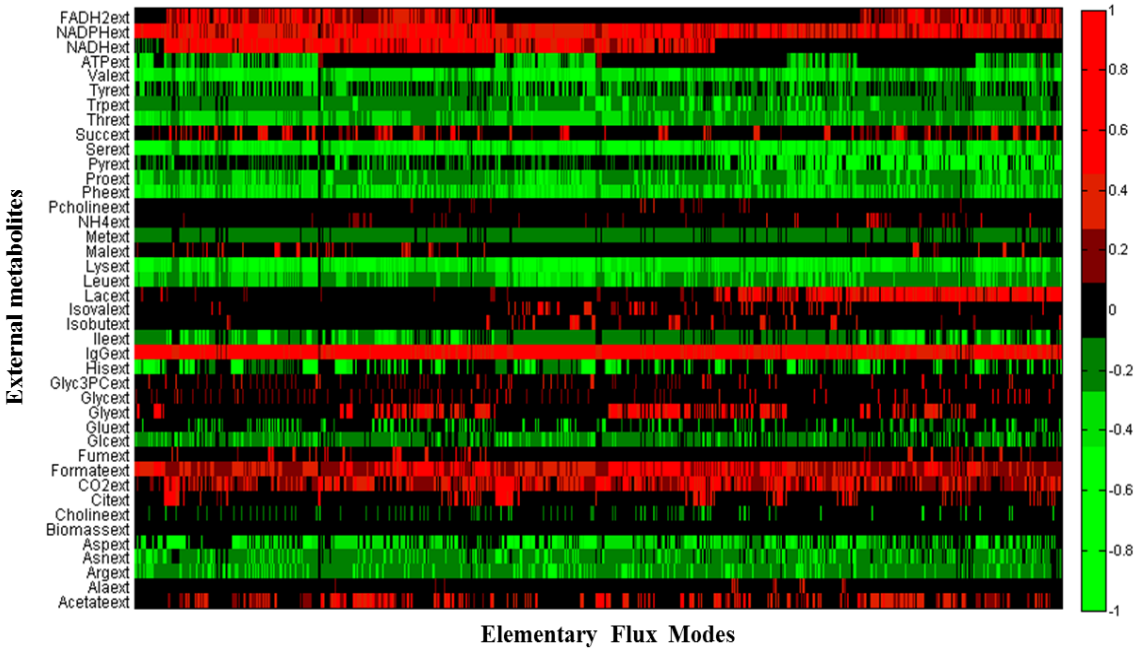

Figure 3.6 – Heat map of all the EFMs with IgG production (normalized data). Metabolites in green are consumed and in red are produced.

Cluster analysis provided a group of 10 clusters, each represented by its centre, with an explained variance of about 83,6%. It was expected a lower value when compared to the previous one since the EFMs are many more in this case, and taking into account that 41 variables are analysed. Visualizing the representation of the EFMs treated in this section considering only three variables (i.e. metabolites) in Figure 3.7, becomes clear that clustering of these data is not so simple as the previous case, even just for three variables since the EFMs are relatively close to each other. The greater the number of clusters is, the greater the explained variance is, as shown in Figure 3.8. Therefore, the chosen number of clusters must be a reasonable number to analyse but the explained variance achievable should not be neglected.

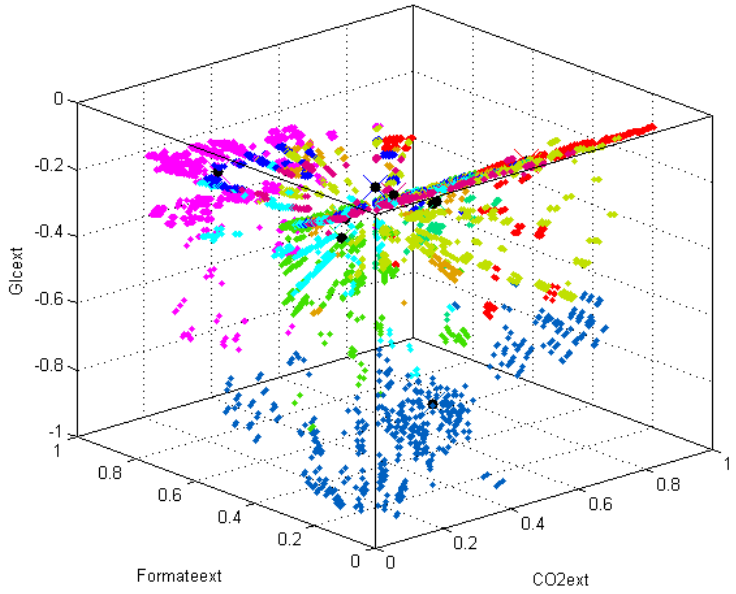

Figure 3.7 - Three-dimensional representation of IgG production EFMs clustering. Each coloured dot represents an EFM and each colour a cluster. Black dots are the centroid of each cluster.

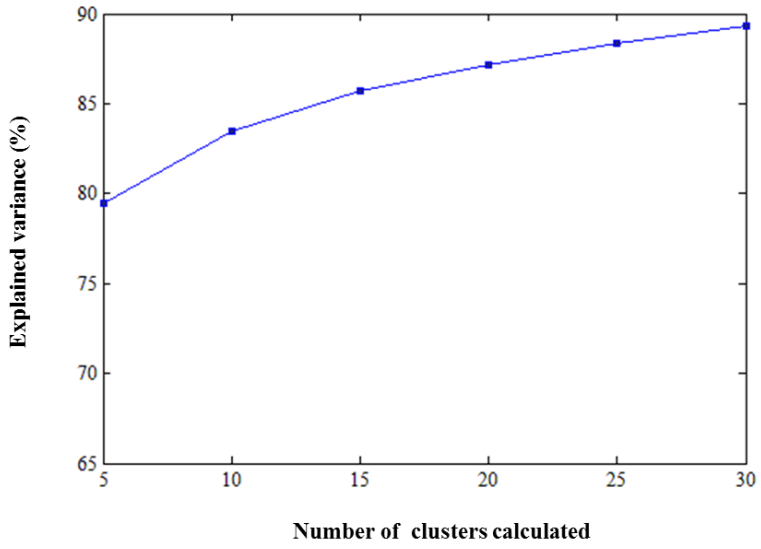

Figure 3.8 – Evolution of explained variance of the results from cluster analysis (of IgG production EFMs) as function of the number of cluster calculated.

The whole set of stoichiometric coefficients of the 10 clusters is listed in Table 6.4 (Appendix) with respective standard deviations, and represented in Figure 3.9 as a normalized heat map. It can be seen that IgG is highly produced in all clusters. ATP in some clusters is obtained by the cell from an external source, but in smaller amounts than for biomass production. This can be explained by the high production of CO_2, also visible in the heat map, which suggests the occurrence of cellular respiration in order to obtain energy. However, NADH, NADPH and $FADH_2$ are transported to the cell exterior, when they should be capitalized to produce more energy.

Table 3.2 shows 5 of the 10 cluster centres. Clusters 2, 3, 5, 6 and 9 have identical values of IgG production to some of those clusters that are represented. In the same table, where the second normalization of the values was already reversed and the data are presented as concentrations corresponding to 40 mM of glucose available in the medium, the high value of IgG produced becomes more evident; especially when compared to the value achieved experimentally. Cluster 10 has the lowest value of IgG produced; however, it is almost 50 times higher than the experimental one. To understand this difference, one must keep in mind that in these EFMs no biomass is produced, which poses a much larger metabolic commitment to the cell. Thus, glucose can be used mostly to produce IgG, with part being used to produce energy as previously said.

Regarding the presented coefficients, cluster 10 seems to be the most biologically feasible. Therefore, this was the one considered in the comparison made with published data, in Figure 3.10. The concentrations required by the pathway represented by cluster 10 are higher than those from the reports, especially in the case of choline, pyruvate and tyrosine. However, the higher demand is comprehensible since more antibody (i.e. IgG) is produced. Furthermore, when compared to Figure 3.5 (concentrations for biomass production) it can be noticed that in this case the difference between calculated and published data is not so remarkable.

Also in contrast to results for biomass production is the consumption of glutamate and production of ammonium (from amino acids metabolism) present in these pathways, although in small amounts. The ratio of ammonium produced per glutamate consumed in cluster 10 is near 0,24 mmol NH_4 / mmol Glu.

The ratio of produced IgG per consumed glucose of cluster 10 is near 321mmol IgG / mmol Glc, while for experimental data is near 6,4 mmol IgG / mmol Glc.

Although in cluster 10 that does not happen, serine is highly consumed in the remaining clusters, in agreement with the previous results.

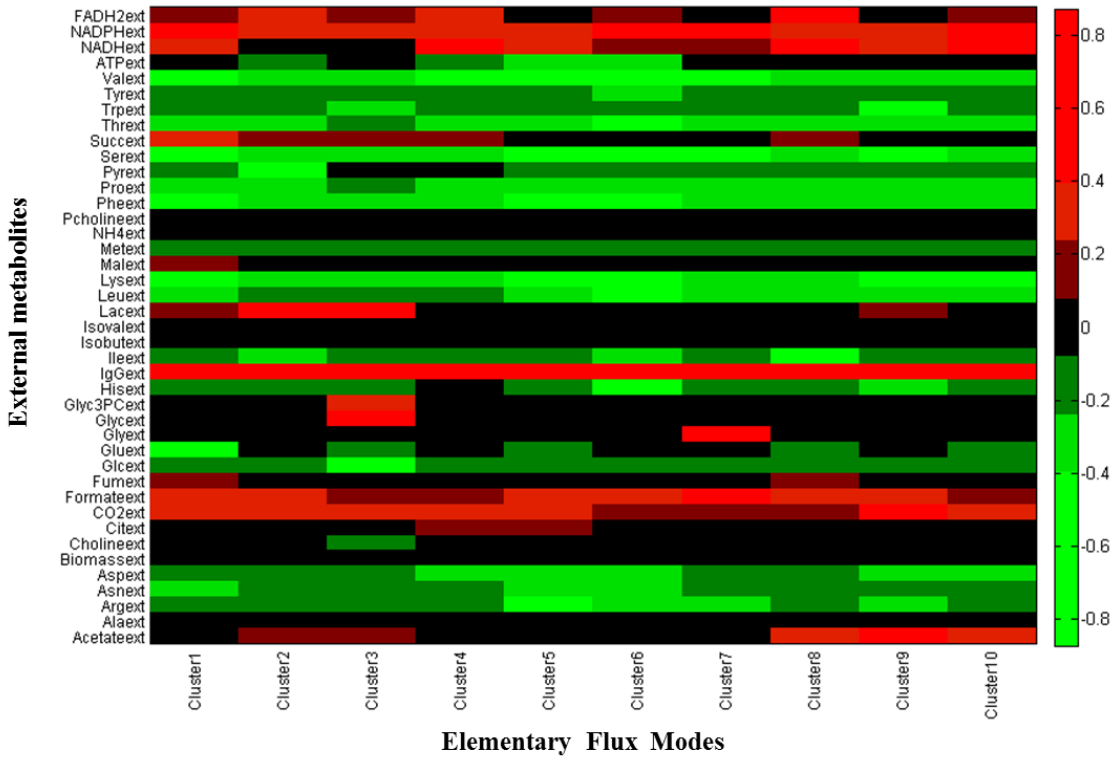

Figure 3.9 - Heat map of pathways obtained after cluster analysis of IgG production EFMs (normalized data). Metabolites in green are consumed and in red are produced.

Table 3.2 - Medium composition from clusters 1, 4, 7, 8 and 10 (aiming IgG production) and from published data [29, 37].

Metabolites	Cluster1	Cluster4	Cluster7	Cluster8	Cluster10	Carinhas *et al.*	Schöder *et al.*	Units
Arg	22,50	56,19	26,93	31,42	5,88	2,64	0,88	mM
Asn	20,59	46,61	30,30	38,45	5,20	4,83	0,27	mM
Asp	29,96	44,65	31,03	31,26	3,97	1,56	0,31	mM
Choline	10,04	1,23	5,02	5,11	11,21	0,6	0	mM
Glc	40,00	40,00	40,00	40,00	40,00	40,04	34	mM
Glu	17,13	30,23	0,21	30,99	4,48	1,88	0,65	mM
His	14,06	33,91	78,02	25,22	3,21	1,13	0,28	mM
Ile	11,44	56,66	22,84	16,21	6,25	2,6	0,97	mM
Leu	36,30	68,75	48,90	46,86	6,80	3,92	1	mM
Lys	30,75	68,41	42,90	91,07	6,35	3,16	1,04	mM
Met	4,29	8,05	5,74	5,98	0,85	0,86	0,25	mM
Phe	26,80	48,81	32,19	35,37	5,01	1,22	0,49	mM
Pro	63,00	89,28	44,24	60,29	9,17	4,71	0,54	mM
Pyr	23,19	418,97	35,65	54,32	20,94	1,5	2,51	mM
Ser	112,00	206,86	134,49	135,58	8,35	5,71	0,52	mM
Thr	47,55	80,53	55,26	57,57	5,52	2,85	1	mM
Trp	9,30	21,20	12,43	12,96	2,38	0,98	0,1	mM
Tyr	10,39	21,17	17,52	16,84	10,55	0,9	0,57	mM
Val	91,18	111,12	71,27	64,80	3,72	2,97	1	mM
IgG	65058,64	122023,81	86964,46	90692,84	12851,37	260	n.a.	mg/L

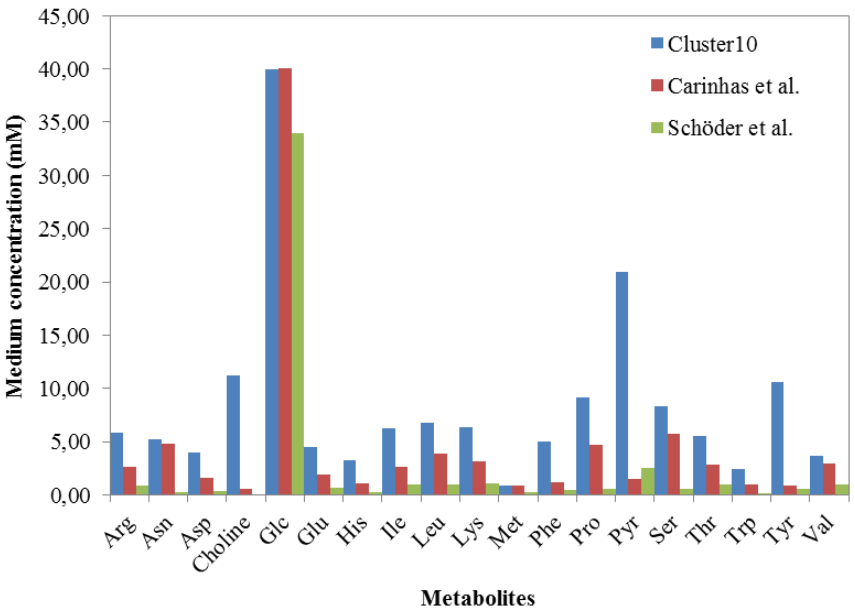

Figure 3.10 - Comparison between the concentrations from the formulation of cluster 10 (IgG production) and from published data [29, 37].

3.4. CO_2 production / Cellular respiration

In this subset are the elementary flux modes that do not produce biomass nor IgG but produce carbon dioxide. Thus, these 6658 EFMs are the ones carrying out exclusively cellular respiration. As mentioned before, due to the addition of the transport reactions of ATP, NADH, NADPH and $FADH_2$, it is implicit that when the cell needs energy it can get it from the "exterior"; but in practice that energy should come from the catabolic process. These pathways are essentially associated with that process.

Figure 3.11 has a general view of all these EFMs. It is clear the high production of CO_2, but also the high production of metabolites such as citrate and succinate (from TCA cycle), alanine and lactate (from pyruvate), as well as acetate as a by-product.

In this case, contrary to the previous two groups of EFMs, ATP is usually produced in excess instead of consumed. This is expectable since the metabolic process here considered has the function to provide energy to the cell. Thus, ATP produced in these pathways should be used by the pathways with biomass and IgG production.

Additionally, NADH, NADPH and $FADH_2$ would also be converted into ATP if the reactions of oxidative phosphorylation were considered in the metabolic network. However, if those were assumed the software would not be able to calculate EFMs due to the complexity of the network.

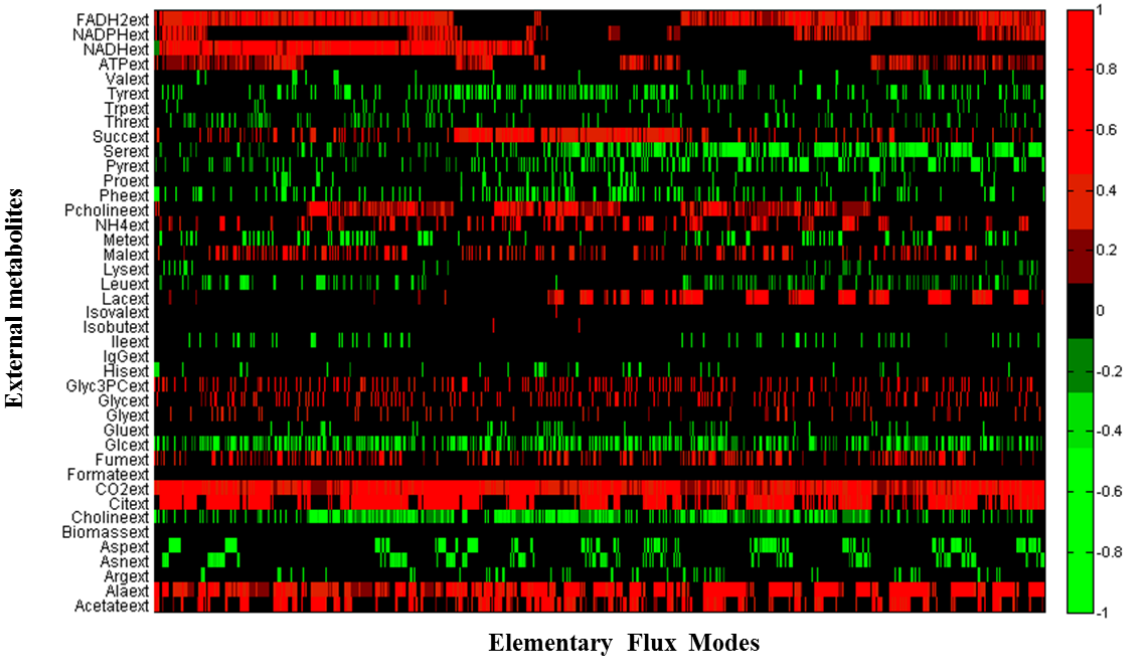

Figure 3.11 - Heat map of all the EFMs with CO_2 production and no biomass or IgG (normalized data). Metabolites in green are consumed and in red are produced.

After cluster analysis, the explained variance of the 10 cluster centres calculated was about 67,7 %. Indeed, as represented in Figure 3.12, it seems to be more complex to identify each cluster when compared, for instance, to Figure 3.7 (page 26), although the regarded metabolites are not the same. Even with 30 clusters the explained variance is not high, as shown in Figure 3.13.

Clusters 4, 6, 7, 8 and 9 were assumed as the most reliable. Despite the experimental data here used for comparison does not include measurements for CO_2, the remaining clusters have coefficients for CO_2 and/or other metabolites that seem to be too high. For that reason, these are the cluster centres shown in Table 3.3; however, all 10 are represented in Figure 3.14 with the respective consumed and produced metabolites, and their stoichiometric coefficients with respective standard deviations are shown in Table 6.5.

In Table 3.3, it is perceptible the increase of ATP production as more CO_2 is produced, which is expected since the production of CO_2 is related to the production of energy from glucose in the catabolic process.

The wide dark area on the heat map of Figure 3.14 shows that several metabolites do not participate or participate but in insignificant amounts in these pathways. That can be explained by the fact that neither biomass nor IgG are produced in these pathways, so the demand for these metabolites is not so high.

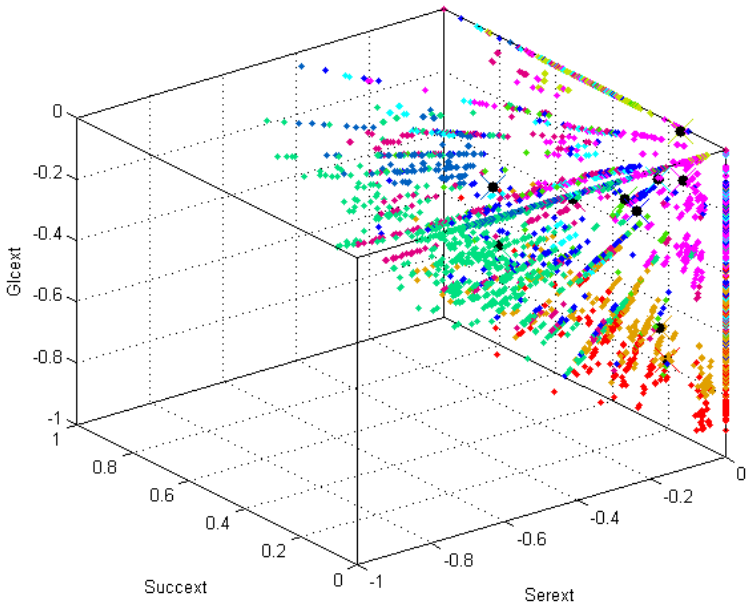

Figure 3.12 - Three-dimensional representation of cellular respiration EFMs clustering. Each coloured dot represents an EFM and each colour a cluster. Black dots are the centroid of each cluster.

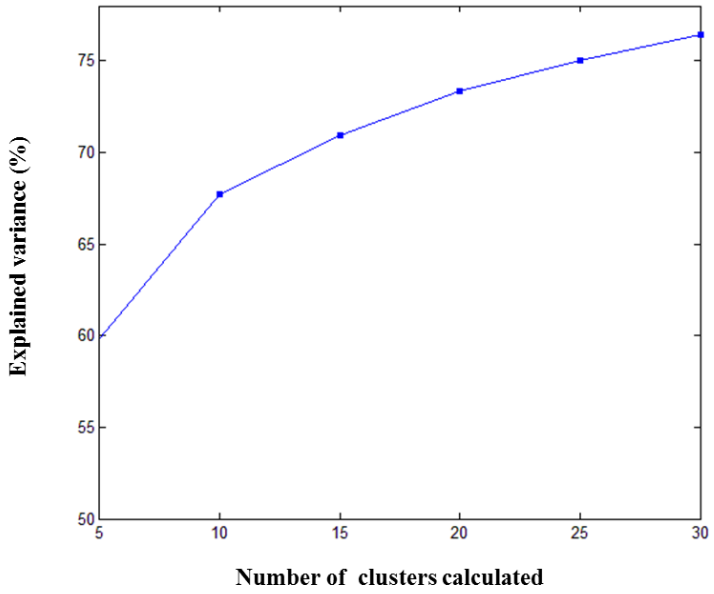

Figure 3.13 - Evolution of explained variance of the results from cluster analysis (of cellular respiration EFMs) as function of the number of cluster calculated.

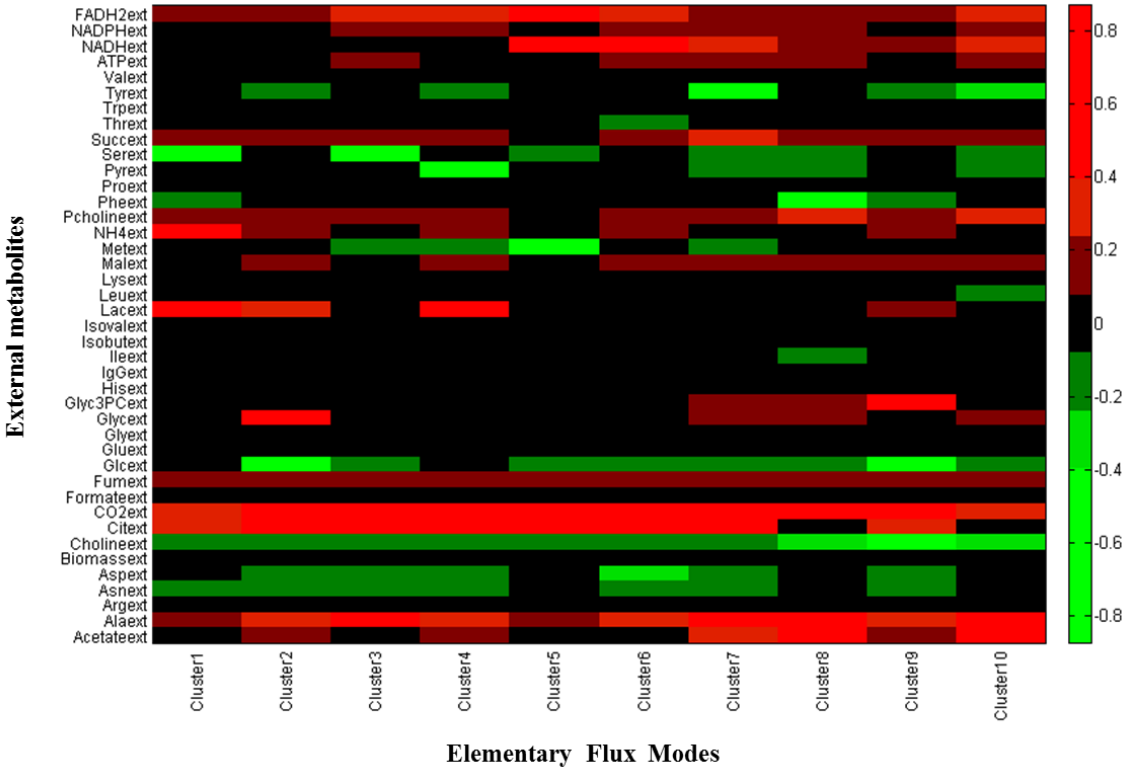

Figure 3.14 - Heat map of pathways obtained after cluster analysis of cellular respiration EFMs (normalized data). Metabolites in green are consumed and in red are produced.

Table 3.3 – Medium composition from clusters 4, 6, 7, 8 and 9 (aiming cellular respiration) and from published data [29, 37]..

Metabolites	Cluster4	Cluster6	Cluster7	Cluster8	Cluster9	Carinhas *et al.*	Schöder *et al.*	Units
Arg	-	1,38	2,67	10,98	7,42	2,64	0,88	mM
Asn	0,88	9,68	8,17	2,65	4,85	4,83	0,27	mM
Asp	151,47	4,27	5,08	0,31	0,87	1,56	0,31	mM
Choline	42,88	4,78	44,60	58,58	51,18	0,6	0	mM
Glc	40,00	40,00	40,00	40,00	40,00	40,04	34	mM
Glu	-	1,23	1,72	14,34	8,65	1,88	0,65	mM
His	-	1,01	2,23	10,31	4,82	1,13	0,28	mM
Ile	8,53	1,55	1,51	4,93	11,43	2,6	0,97	mM
Leu	9,19	2,08	2,48	48,33	0,65	3,92	1	mM
Lys	6,70	1,19	1,15	3,34	-	3,16	1,04	mM
Met	-	2,72	2,09	4,07	11,71	0,86	0,25	mM
Phe	22,60	3,78	6,03	-	-	1,22	0,49	mM
Pro	-	0,89	1,15	11,24	7,34	4,71	0,54	mM
Pyr	17,27	-	-	23,56	24,55	1,5	2,51	mM
Ser	8,52	1,59	3,48	27,50	30,72	5,71	0,52	mM
Thr	10,07	1,35	1,45	1,20	2,96	2,85	1	mM
Trp	7,76	1,02	1,00	1,86	-	0,98	0,1	mM
Tyr	29,50	5,27	6,97	0,98	119,60	0,9	0,57	mM
Val	-	0,54	0,47	8,25	5,86	2,97	1	mM
CO_2	154,96	23,53	24,59	54,11	87,51	n.a.	n.a.	mM
ATP	21,02	3,37	4,04	25,13	25,56	n.a.	n.a.	mM

Due to its reasonable coefficients (i.e. not too high), cluster 6 was assumed as the most feasible, and it was compared to the published concentrations in Figure 3.15. In this case, the concentrations are not very different. In fact, this is the achieved formulation most close to experimental data from Carinhas *et al.* and to the formulation proposed by Schöder *et al.*

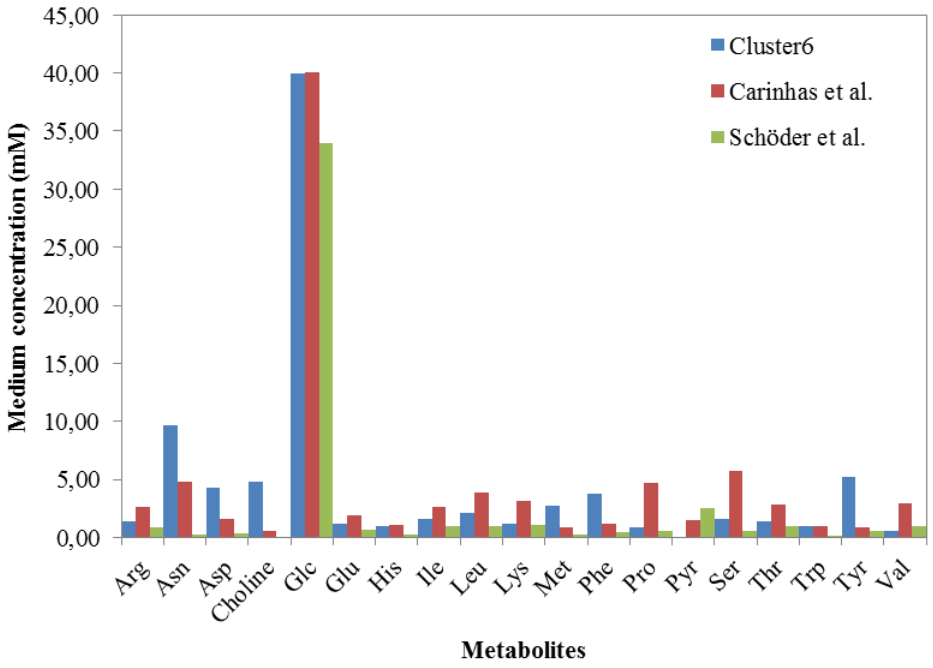

Figure 3.15 – Graphical comparison between the concentrations from the formulation of cluster 6 (cellular respiration) and from published data [29, 37].

3.5. Remaining EFMs

178 elementary flux modes were not in any of the previous categories, since they do not produce biomass, IgG nor they carry out cellular respiration. Among these EFMs are the following examples:

$$0{,}71\ Ser + 0{,}66\ NADH \rightarrow 0{,}77\ Ala \tag{6}$$

$$0{,}54\ Pyr + 0{,}91\ Thr \rightarrow 0{,}68\ Acetate + 0{,}82\ Gly + 0{,}46\ Lac + 0{,}41\ ATP \tag{7}$$

As seen in these examples, each provides little information aiming media formulation, and the physiological states they represent can be very different (as seen in Figure 3.16), so cluster analysis may not be successful.

Thus, this small group of EFMs was considered not important to perform this study and was neglected.

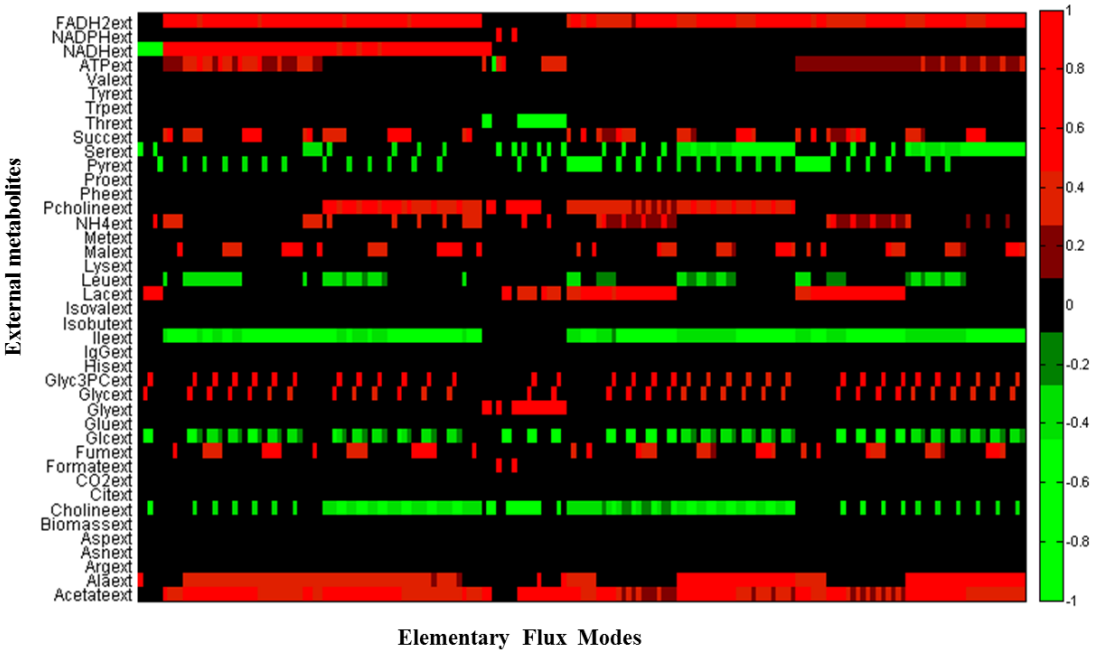

Figure 3.16 - Heat map of all the EFMs with no production of biomass, IgG and CO_2 (normalized data). Metabolites in green are consumed and in red are produced.

3.6.Medium formulation

As previously said, due to the last four transport reactions in the metabolic network (see Appendix A), which were required in order to allow the EFMs calculation, it is not implicit that the cell has to carry out catabolic and anabolic processes, since there is the option of obtaining energy from the unbalanced ATP, NADH and NADPH. In practice, ATP is not supplemented to culture medium so the cell has to produce its own energy from the nutrients available.

In the previous analysis, the calculated physiological states were grouped according to their main function: biomass production, IgG production and cellular respiration. However, when used industrially, CHO cells are intended to perform these three processes simultaneously. Thus, culture medium must stimulate cells to carry them out with the best performance, providing the required supplements.

For this reason it is not straightforward to contemplate a formulation from only one subset of EFMs, since it would be designed aiming essentially one cellular function, and it would be assumed that energy could be "imported" from culture medium. Therefore, a combination of three formulations (one for each cellular function) can be tested.

To provide an example, the formulations of cluster 2 from biomass production, cluster 10 from IgG production and cluster 6 from cellular respiration were combined. It was assumed the next proportion for the final formulation: 60% of cellular respiration pathway, 20% of

biomass and 20% of IgG production. The results are shown in Table 3.4. Assuming the same percentages for production, it would be possible to reach 11 million of cells per mL and nearly 2570 mg/L of IgG. *A priori*, the latter still is high when compared to experimental results. This proportion is only to provide an example; to find the best combination it is necessary to test it experimentally with a cell line.

Table 3.4 - Medium composition from the clusters selected from each subset, final formulation and compositions from published data [29, 37].

Metabolites	Biomass production	IgG production	Cellular respiration	Final formulation	Carinhas *et al.*	Schöder *et al.*	Units
Arg	5,05	5,88	1,38	3,01	2,64	0,88	mM
Asn	21,53	5,20	9,68	11,16	4,83	0,27	mM
Asp		3,97	4,27	3,36	1,56	0,31	mM
Choline	27,39	11,21	4,78	10,59	0,6	0	mM
Glc	40	40,00	40,00	40,00	40,04	34	mM
Glu	-	4,48	1,23	1,63	1,88	0,65	mM
His	1,99	3,21	1,01	1,65	1,13	0,28	mM
Ile	4,55	6,25	1,55	3,09	2,6	0,97	mM
Leu	21,19	6,80	2,08	6,85	3,92	1	mM
Lys	8,11	6,35	1,19	3,61	3,16	1,04	mM
Met	2,2	0,85	2,72	2,24	0,86	0,25	mM
Phe	3,71	5,01	3,78	4,01	1,22	0,49	mM
Pro	19,86	9,17	0,89	6,34	4,71	0,54	mM
Pyr	-	20,94	-	4,19	1,5	2,51	mM
Ser	25,07	8,35	1,59	7,64	5,71	0,52	mM
Thr	5,27	5,52	1,35	2,97	2,85	1	mM
Trp	0,51	2,38	1,02	1,19	0,98	0,1	mM
Tyr	2,61	10,55	5,27	5,79	0,9	0,57	mM
Val	19,04	3,72	0,54	4,87	2,97	1	mM

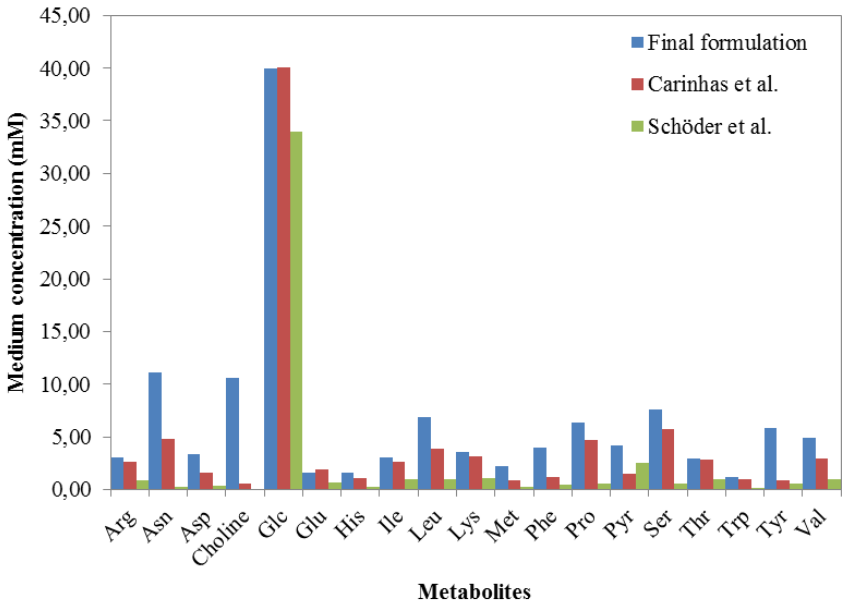

Figure 3.17 - Graphical comparison of metabolites concentration between the final formulation and published data [29, 37]

In Figure 3.17 the achieved formulation is compared to published results. At the first sight, most of the metabolites are present at higher concentration in the proposed medium, which is rational since it is intended to achieve higher production rates. The major difference is the concentration of choline, due to its high demand in biomass production elementary modes. However, in general the differences are not very large so the concentrations proposed should not be toxic for cellular growth.

It should also be taken into account that besides these amino acids and glucose, medium formulation should also have the other elements described in Table 2.1 (Section 2.4), such as: inorganic salts, vitamins and miscellaneous compounds.

Proteins are not desirable in the media formulation in order to simplify the product purification process. Therefore, Schröder *et al.* [37] suggested the omission of fetuin and coating of tissue culture dishes with fibronectin, the replacement of *holo*-transferrin with tropolone and ferric ammonium citrate, and the progressive decrease of insulin concentration. According to the authors, the results of this technique were satisfactory.

4. Conclusion

In this thesis, elementary mode analysis was implemented using a metabolic network of CHO cells expressing IgG with 119 reactions and 126 metabolites (41 assumed as external), with the objective to design culture medium formulations that enhance the cellular performance when used to produce biopharmaceuticals. The adopted network includes transport reactions for ATP, NADH, NADPH and $FADH_2$ in order to allow the Metatool 5.1 algorithm to perform the analysis; otherwise, the calculation would become too complex and RAM memory available would not be enough. However, it must be highlighted that with these reactions it is assumed by the algorithm that energy can be obtained from an external source, which will influence the results.

Therefore, the 23240 EFMs using Metatool were divided into: 108 producing biomass, 16296 producing IgG, 6658 carrying out essentially cellular respiration and 178 were assumed as non-feasible. Cluster analysis was employed to calculate a representative subset of 10 pathways for each group. None of the EFMs had production of biomass and IgG simultaneously.

In the subset of pathways regarding cellular growth, biomass production presented values much higher than those obtained experimentally in other studies. The reason is that in these pathways energy (ATP) is transported into the cell so glucose is essentially used for the production of biomass. Thus, these EFMs would have to be combined with another one having energy production. On the other hand, some EFMs have very low biomass production and a low transport of ATP from the outside, which emphasizes the fact that high cellular growth is related to glucose being used for biomass production and not for energy.

Concerning IgG, very high production rates were also achieved. In fact, the presented values are too high when compared to experimental data, but this results from the high uptake of required metabolites which are not fuelled to the more demanding biomass formation. Additionally, ATP comes from the "exterior" in the studied pathways too. However, it must be kept in mind that too high metabolite concentrations can have a toxic effect on the cellular growth, so some formulations as they were presented may not be feasible.

Finally, the pathways with production of carbon dioxide and no production of biomass or IgG were analysed. As expected, ATP is produced, as well as NADH, NADPH and $FADH_2$ which could also be converted into ATP if the oxidative phosphorylation was considered in the metabolic network.

The medium compositions from the three previous analyses only regarded one main cellular function. However, when CHO cells are used to produce biopharmaceuticals it is

necessary that all these three process occur inside the cell. Therefore, a medium composition was estimated combining one EFM from each group, in order to ensure that the three cellular functions occur. It was assumed that 60% of the available glucose was used in cellular respiration, 20% in IgG production and 20% in cellular growth. The final medium composition has concentrations closer to those experimentally observed and, theoretically, would allow higher production rates of biomass and product (IgG).

In the future, these results should be tested *in vitro* and, according to the reached performance, the percentage of each formulation can be fine-tuned, aiming the main objective.

A more simple metabolic network could also be tested to allow for catabolic and anabolic processes to coexist during elementary mode analysis.

Another approach is to test elementary mode analysis using the present network without the four last transport reactions and accounting for oxidative phosphorylation using a computer with more RAM memory; but this can be time-consuming.

To conclude, with this method it was possible to design a medium composition in a costless way (excluding the investment in computer equipment) and using less time than empirical methods. However, experimental results are required to ensure if this method is successful.

5. References

[1] J. W. Zhu, "Mammalian cell protein expression for biopharmaceutical production," *Biotechnology Advances,* vol. 30, pp. 1158-1170, Sep-Oct 2012.

[2] K. R. Jayapal, K. F. Wlaschin, W. S. Hu, and M. G. S. Yap, "Recombinant protein therapeutics from CHO cells - 20 years and counting," *Chemical Engineering Progress,* vol. 103, pp. 40-47, Oct 2007.

[3] L. d. R. Castilho, *Animal cell technology : from biopharmaceuticals to gene therapy*. London: Taylor & Francis, 2008.

[4] D. B. Kell, M. Brown, H. M. Davey, W. B. Dunn, I. Spasic, and S. G. Oliver, "Metabolic footprinting and systems biology: the medium is the message," *Nat Rev Microbiol,* vol. 3, pp. 557-65, Jul 2005.

[5] S. S. Ozturk and W.-S. Hu, *Cell culture technology for pharmaceutical and cell-based therapies*. New York ; London: Taylor & Francis, 2006.

[6] R. I. Freshney, *Culture of animal cells : a manual of basic technique*, 5th ed. ed. Hoboken, N.J.: Wiley-Liss, Chichester : John Wiley [distributor], 2005.

[7] R. I. Freshney, *Animal cell culture : a practical approach*, 2nd ed. ed.: IRL Press at Oxford University Press, 1992.

[8] M. Butler and A. Christie, "ADAPTATION OF MAMMALIAN-CELLS TO NONAMMONIAGENIC MEDIA," *Cytotechnology,* vol. 15, pp. 87-94, 1994.

[9] C. Altamirano, C. Paredes, J. J. Cairo, and F. Godia, "Improvement of CHO cell culture medium formulation: Simultaneous substitution of glucose and glutamine," *Biotechnology Progress,* vol. 16, pp. 69-75, Jan-Feb 2000.

[10] D. R. Conrad. (2/9/2013). *Sigma-Aldrich Media Expert*. Available: http://www.sigmaaldrich.com/life-science/cell-culture/learning-center/media-expert/sa-media-expert.html

[11] R. O. Slabicky, K. G. Chapman, B. W. Daugherty, T. F. Genova, C. R. Kern, L. E. Kovach, *et al.*, "UPDATING REQUIREMENTS FOR PHARMACEUTICAL GRADES OF WATER - PROPOSED REVISIONS," *Pharmacopeial Forum,* vol. 20, pp. 7526-7544, May-Jun 1994.

[12] G. M. Lee, E. J. Kim, N. S. Kim, S. K. Yoon, Y. H. Ahn, and J. Y. Song, "Development of a serum-free medium for the production of erythropoietin by suspension culture of recombinant Chinese hamster ovary cells using a statistical design," *Journal of Biotechnology,* vol. 69, pp. 85-93, Apr 1999.

[13] M. S. Sinacore, T. S. Charlebois, S. Harrison, S. Brennan, T. Richards, M. Hamilton, *et al.*, "CHO DUKX cell lineages preadapted to growth in serum-free suspension culture enable rapid development of cell culture processes for the manufacture of recombinant proteins," *Biotechnology and Bioengineering,* vol. 52, pp. 518-528, Nov 1996.

[14] P. M. L. Castro, P. M. Hayter, A. P. Ison, and A. T. Bull, "APPLICATION OF A STATISTICAL DESIGN TO THE OPTIMIZATION OF CULTURE-MEDIUM FOR RECOMBINANT INTERFERON-GAMMA PRODUCTION BY CHINESE-HAMSTER OVARY CELLS," *Applied Microbiology and Biotechnology,* vol. 38, pp. 84-90, Oct 1992.

[15] D. Y. Kim, J. C. Lee, H. N. Chang, and D. J. Oh, "Development of serum-free media for a recombinant CHO cell line producing recombinant antibody," *Enzyme and Microbial Technology,* vol. 39, pp. 426-433, Jul 2006.

[16] M. E. Rodrigues, A. R. Costa, M. Henriques, J. Azeredo, and R. Oliveira, "Comparison of commercial serum-free media for CHO-K1 cell growth and monoclonal antibody production," *International Journal of Pharmaceutics,* vol. 437, pp. 303-305, Nov 2012.

[17] C. T. Trinh, A. Wlaschin, and F. Srienc, "Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism," *Applied Microbiology and Biotechnology*, vol. 81, pp. 813-826, Jan 2009.

[18] J. A. Roels, *Energetics and kinetics in biotechnology*. Amsterdam ; Oxford: Elsevier Biomedical, 1983.

[19] S. Schuster, C. Hilgetag, J. H. Woods, and D. A. Fell, "Reaction routes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism," *Journal of Mathematical Biology*, vol. 45, pp. 153-181, Aug 2002.

[20] C. Reder, "METABOLIC CONTROL-THEORY - A STRUCTURAL APPROACH," *Journal of Theoretical Biology*, vol. 135, pp. 175-201, Nov 21 1988.

[21] B. L. Clarke, "COMPLETE SET OF STEADY-STATES FOR THE GENERAL STOICHIOMETRIC DYNAMICAL SYSTEM," *Journal of Chemical Physics*, vol. 75, pp. 4970-4979, 1981 1981.

[22] R. Schuster and S. Schuster, "REFINED ALGORITHM AND COMPUTER-PROGRAM FOR CALCULATING ALL NONNEGATIVE FLUXES ADMISSIBLE IN STEADY-STATES OF BIOCHEMICAL REACTION SYSTEMS WITH OR WITHOUT SOME FLUX RATES FIXED," *Computer Applications in the Biosciences*, vol. 9, pp. 79-85, Feb 1993.

[23] J. A. Papin, N. D. Price, S. J. Wiback, D. A. Fell, and B. O. Palsson, "Metabolic pathways in the post-genome era," *Trends in Biochemical Sciences*, vol. 28, pp. 250-258, May 2003.

[24] S. Schuster, D. A. Fell, and T. Dandekar, "A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks," *Nature Biotechnology*, vol. 18, pp. 326-332, Mar 2000.

[25] R. T. Rockafellar, *Convex analysis*. Princeton, New Jersey: Princeton University Press, 1970.

[26] C. H. Schilling, D. Letscher, and B. O. Palsson, "Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from? A pathway-oriented perspective," *Journal of Theoretical Biology*, vol. 203, pp. 229-248, Apr 7 2000.

[27] C. H. Schilling and B. O. Palsson, "Assessment of the metabolic capabilities of Haemophilus influenzae Rd through a genome-scale pathway analysis," *Journal of Theoretical Biology*, vol. 203, pp. 249-283, Apr 7 2000.

[28] C. H. Schilling, M. W. Covert, I. Famili, G. M. Church, J. S. Edwards, and B. O. Palsson, "Genome-scale metabolic model of Helicobacter pylori 26695," *Journal of Bacteriology*, vol. 184, pp. 4582-4593, Aug 2002.

[29] N. Carinhas, T. M. Duarte, L. C. Barreiro, M. J. Carrondo, P. M. Alves, and A. P. Teixeira, "Metabolic signatures of GS-CHO cell clones associated with butyrate treatment and culture phase transition," *Biotechnol Bioeng*, Jun 2013.

[30] F. Zamorano, A. Wouwer, and G. Bastin, "A detailed metabolic flux analysis of an underdetermined network of CHO cells," *Journal of Biotechnology*, vol. 150, pp. 497-508, DEC 2010 2010.

[31] L. Quek, S. Dietmair, J. Kromer, and L. Nielsen, "Metabolic flux analysis in mammalian cell culture," *Metabolic Engineering*, vol. 12, pp. 161-171, MAR 2010 2010.

[32] N. Sengupta, S. T. Rose, and J. A. Morgan, "Metabolic flux analysis of CHO cell metabolism in the late non-growth phase," *Biotechnol Bioeng*, vol. 108, pp. 82-92, Jan 2011.

[33] K. Sheikh, J. Förster, and L. K. Nielsen, "Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of Mus musculus," *Biotechnol Prog*, vol. 21, pp. 112-21, 2005 Jan-Feb 2005.

[34] T. Pfeiffer, I. Sanchez-Valdenebro, J. C. Nuno, F. Montero, and S. Schuster, "METATOOL: for studying metabolic networks," *Bioinformatics*, vol. 15, pp. 251-257, Mar 1999.

[35] A. von Kamp and S. Schuster, "Metatool 5.0: fast and flexible elementary modes analysis," *Bioinformatics*, vol. 22, pp. 1930-1931, Aug 1 2006.

[36] A. Jain, M. Murty, and P. Flynn, "Data clustering: A review," *Acm Computing Surveys*, vol. 31, pp. 264-323, SEP 1999 1999.

[37] M. Schroder, K. Matischak, and P. Friedl, "Serum- and protein-free media formulations for the Chinese hamster ovary cell line DUKXB11," *Journal of Biotechnology*, vol. 108, pp. 279-292, Mar 2004.

6. Appendices

6.1. Appendix A

Table 6.1 - Metabolic network for CHO cells.

Flux	Reaction
	Glycolysis
$\mathbf{r_1}$	Glc + ATP → G6P
$\mathbf{r_2}$	G6P → F6P
$\mathbf{r_3}$	F6P + ATP → DHAP + GAP
$\mathbf{r_4}$	DHAP → GAP
$\mathbf{r_5}$	GAP → 3PG + NADH + ATP
$\mathbf{r_6}$	3PG → Pyr + ATP
	TCA cycle
$\mathbf{r_7}$	Pyr → AcCoA + CO2 + NADH
$\mathbf{r_8}$	AcCoA + Oxal → Cit
$\mathbf{r_9}$	Cit → alfaKG + CO2 + NADH
$\mathbf{r_{10}}$	alfaKG → SucCoA + CO2 + NADH
$\mathbf{r_{11}}$	SucCoA → Succ + ATP
$\mathbf{r_{12}}$	Succ → Fum + FADH2
$\mathbf{r_{13}}$	Fum → Mal
$\mathbf{r_{14}}$	Mal → Oxal + NADH
	Pyruvate fates
$\mathbf{r_{15}}$	Pyr + NADH ⇆ Lac
$\mathbf{r_{16}}$	Pyr + Glu ⇆ Ala + alfaKG
	Pentose Phosphate Pathway
$\mathbf{r_{17}}$	3 G6P → 3 CO2 + 3 R5P + 6 NADPH
	Anaplerotic Reaction
$\mathbf{r_{18}}$	Mal → Pyr + CO2 + NADPH
	Amino Acid Metabolism
$\mathbf{r_{19}}$	Glu ⇆ alfaKG + NH4 + NADH
$\mathbf{r_{20}}$	Oxal + Glu ⇆ Asp + alfaKG
$\mathbf{r_{21}}$	Gln ⇆ Glu + ATP + NH4
$\mathbf{r_{22}}$	Thr → Gly + NADH + AcCoA
$\mathbf{r_{23}}$	Ser → Gly + ATP + Formate + NADPH
$\mathbf{r_{24}}$	Ser → Pyr + NH4
$\mathbf{r_{25}}$	Thr → alfaKb + NH4
$\mathbf{r_{26}}$	alfaKb + ATP → SucCoA + NADH
$\mathbf{r_{27}}$	Trp → Ala + 2 CO2 + alfaKa
$\mathbf{r_{28}}$	Lys + 2 alfaKG → alfaKa + 2 Glu + 3 NADPH + FADH2
$\mathbf{r_{29}}$	alfaKa → 2 AcCoA + 2 NADH + 2 CO2
$\mathbf{r_{30}}$	Val + alfaKG → IsobutCoA + Glu + CO2 + NADH
$\mathbf{r_{31}}$	IsobutCoA + ATP → SucCoA + FADH2 + 2 NADH
$\mathbf{r_{32}}$	IsobutCoA → Isobut

Table 6.1 (Continued)

Flux	Reaction
r_{33}	Ile + alfaKG + ATP → AcCoA + SucCoA + Glu + 2 NADH + FADH2
r_{34}	Leu + alfaKG → IsovalCoA + Glu + CO2 + NADH
r_{35}	IsovalCoA + ATP + CO2 + SucCoA → 3 AcCoA + Succ + FADH2
r_{36}	IsovalCoA → Isoval
r_{37}	Phe + NADH → Tyr
r_{38}	Tyr + alfaKG + SucCoA → Fum + 2 AcCoA + Succ + Glu + CO2
r_{39}	Met + Ser + ATP → alfaKb + NH4
r_{40}	Asn → Asp + NH4
r_{41}	Pro → Glu + NADH
r_{42}	Arg + alfaKG → 2 Glu + NADH
r_{43}	His → Glu + NH4
	Glycogen Synthesis
r_{44}	G6P → G1P
r_{45}	G1P + UMPRN + 2 ATP → UDPG
r_{46}	UDPG → Glycogen
	Nucleotide Synthesis
r_{47}	R5P + ATP → PRPP
r_{48}	PRPP + 2 Gln + Gly + Asp + 5 ATP + CO2 → IMP + 2 Glu + Fum
r_{49}	IMP + Asp + ATP → AMPRN + Fum
r_{50}	IMP + Gln + ATP → GMPRN + Glu + NADH
r_{51}	CO2 + NH4 + Asp + 2 ATP → Orotate + NADH
r_{52}	Orotate + PRPP → UMPRN + CO2
r_{53}	UMPRN + Gln + ATP → CMPRN + Glu
r_{54}	AMPRN → dAMP
r_{55}	GMPRN → dGMP
r_{56}	CMPRN → dCMP
r_{57}	UMPRN → dTMP
	Lipid Synthesis
r_{58}	Choline + ATP → Pcholine
r_{59}	Pcholine + 18 AcCoA + Glyc3P + 22 ATP + 33 NADH → PC
r_{60}	PC + Ser → PS + Choline
r_{61}	PS → PE + CO2
r_{62}	Choline + Glyc3P → Glyc3PC
r_{63}	G6P → Inositol
r_{64}	Inositol + 18 AcCoA + Glyc3P + 22 ATP + 33 NADH → PI
r_{65}	18 AcCoA + 2 Glyc3P + 22 ATP + 33 NADH → PG
r_{66}	2 PG → DPG + Glyc
r_{67}	16 AcCoA + Ser + Choline + 16 ATP + 29 NADPH → SM + 2 CO2
r_{68}	18 AcCoA + 18 ATP + 14 NADPH → Cholesterol + 9 CO2
	Biomass Formation
r_{69}	160.1015 Ala + 235.2056 Glu + 70.3787 Gln + 174.6799 Gly + 114.9787 Ser + 147.4132 Lys + 157.4070 Leu + 82.6648 Ile + 91.8543 Arg + 169.4920 Asp + 95.7754 Thr + 118.3569 Val + 40.0354 Met + 67.4027 Phe + 47.4956 Tyr + 36.2551 His + 55.5590 Pro + 70.3787 Asn + 11694463 ATP + 8.943 AMPRN + 4.878 Cholesterol + 14.9321 CMPRN + 4.0108 dAMP + 2.6829 dCMP + 2.6829 dGMP + 4.0108 dTMP + 0.8130 DPG + 75.6090 Glycogen + 16.9104 GMPRN + 18.6990 PC + 7.046 PE + 0.271 PG + 2.71 PI + 0.813 PS + 2.168 SM + 8.943

Table 6.1 (Continued)

Flux	Reaction
	UMPRN + 9.2297 Trp → 1 Biomass
	Other by-products
$\mathbf{r_{70}}$	AcCoA → Acetate + ATP
$\mathbf{r_{71}}$	DHAP + NADH → Glyc3P
$\mathbf{r_{72}}$	Glyc3P → Glyc
	IgG Glycosylation
$\mathbf{r_{73}}$	UDPG → UDPGal
$\mathbf{r_{74}}$	Glc + 3 ATP + GMPRN → GDPMann
$\mathbf{r_{75}}$	F6P + Gln + AcCoA + UMPRN + 2 ATP → UDPNAG + Glu
$\mathbf{r_{76}}$	UDPNAG + 3 ATP + 3PG + CMPRN → CMPSialic
$\mathbf{r_{77}}$	GDPMann + NADPH → GDPFuc
	IgG Formation
$\mathbf{r_{78}}$	428.7 Ala + 362.75 Glu + 351.76 Gln + 516.64 Gly + 934.36 Ser + 472.67 Lys + 516.64 Leu + 175.88 Ile + 307.79 Arg + 296.8 Asp + 626.57 Thr + 714.51 Val + 65.954 Met + 285.8 Phe + 285.8 Tyr + 164.89 His + 505.65 Pro + 263.82 Asn + 142.9 Trp + 10.992 GDPFuc + 54.962 UDPNAG + 32.977 GDPMann + 21.985 UDPGal + 21.985 CMPSialic → 1 IgG
	Transport Reactions
$\mathbf{r_{79}}$	Aspext → Asp
$\mathbf{r_{80}}$	Gly → Glyext
$\mathbf{r_{81}}$	Serext → Ser
$\mathbf{r_{82}}$	Gluext → Glu
$\mathbf{r_{83}}$	Tyrext → Tyr
$\mathbf{r_{84}}$	Ala → Alaext
$\mathbf{r_{85}}$	Argext → Arg
$\mathbf{r_{86}}$	Asnext → Asn
$\mathbf{r_{87}}$	Hisext → His
$\mathbf{r_{88}}$	Ileext → Ile
$\mathbf{r_{89}}$	Leuext → Leu
$\mathbf{r_{90}}$	Lysext → Lys
$\mathbf{r_{91}}$	Metext → Met
$\mathbf{r_{92}}$	Pheext → Phe
$\mathbf{r_{93}}$	Proext → Pro
$\mathbf{r_{94}}$	Thrext → Thr
$\mathbf{r_{95}}$	Trpext → Trp
$\mathbf{r_{96}}$	Valext → Val
$\mathbf{r_{97}}$	Cholineext → Choline
$\mathbf{r_{98}}$	NH4 → NH4ext
$\mathbf{r_{99}}$	CO2 → CO2ext
$\mathbf{r_{100}}$	Cit → Citext
$\mathbf{r_{101}}$	Fum → Fumext
$\mathbf{r_{102}}$	Pyrext → Pyr
$\mathbf{r_{103}}$	Succ → Succext
$\mathbf{r_{104}}$	Mal → Malext
$\mathbf{r_{105}}$	Glyc → Glycext
$\mathbf{r_{106}}$	Pcholine → Pcholineext

Flux	Reaction
$\mathbf{r}_{107}$	Glyc3PC → Glyc3PCext
$_{108}$	Formate → Formateext
$\mathbf{r}_{109}$	Acetate → Acetateext
$\mathbf{r}_{110}$	Isobut → Isobutext
$\mathbf{r}_{111}$	Isoval → Isovalext
$\mathbf{r}_{112}$	Biomass → Biomassext
$\mathbf{r}_{113}$	Glcext → Glc
$\mathbf{r}_{114}$	Lac → Lacext
$\mathbf{r}_{115}$	IgG → IgGext
$\mathbf{r}_{116}$	ATP ⇆ ATPext
$\mathbf{r}_{117}$	NADH ⇆ NADHext
$\mathbf{r}_{118}$	NADPH ⇆ NADPHext
$\mathbf{r}_{119}$	FADH2 ⇆ FADH2ext

6.2. Appendix B

Table 6.2 - Price list of supplements present in the medium composition proposed by Schröder *et al.* [37]. Prices and references were gathered from *Sigma-Aldrich* website on 2 Aug. 2013, aiming reagents suitable for cell culture.

Component	Stock Keeping Unit (SKU)	Package size (g)	Price (€)	Price per gram (€/g)
Inorganic salts				
$CaCl_2$	C7902	1000	116,5	0,117
$CuSO_4 \cdot 5H_2O$	C8027	1000	114	0,114
$FeSO_4 \cdot 7H_2O$	F7002	1000	122	0,122
KNO_3	P8291	25000	1530	0,061
KCl	P5405	1000	78,5	0,079
$MgCl_2$	M2393	500	56	0,112
$MgSO_4$	M2643	1000	196	0,196
NaCl	S5886	10000	293,5	0,029
$NaHCO_3$	S5761	5000	111,5	0,022
Na_2HPO_4	S5136	5000	467	0,093
$Na_2SeO_3 \cdot 5H_2O$	S5261	100	133	1,330
$ZnSO_4 \cdot 7H_2O$	Z0251	500	103	0,206
Carbohydrates				
d-Glucose	G7021	10000	187	0,019
l-Amino acids				
l-Alanine	A7469	1000	690	0,690
l-Arginine	A8094	1000	412	0,412
l-Asparagine	A4159	500	350	0,700
l-Aspartic acid	A7219	1000	154	0,154
l-Cysteine·HCl·H_2O	C7602	1000	870	0,870
l-Cystine·2HCl	C2526	1000	14500	14,500
l-Glutamic acid	G8415	1000	141,5	0,142
l-Glutamine	G7513	100 (mL)	16,8	0,168 (€/mL)
Glycine	G8790	1000	123,5	0,124
l-Histidine·HCl·H_2O	H5659	1000	646	0,646
l-Isoleucine	I7403	1000	1605	1,605
l-Leucine	L8912	1000	403	0,403
l-Lysine·HCl	L8662	1000	244,5	0,245
l-Methionine	M5308	1000	685	0,685
l-Phenylalanine	P5482	1000	787	0,787
l-Proline	P5607	1000	906	0,906
l-Serine	S4311	1000	784	0,784
l-Threonine	T8441	1000	1860	1,860
l-Tryptophan	T8941	1000	1535	1,535
l-Tyrosine, disodium salt	T1145	500	668	1,336
l-Valine	V0513	1000	822	0,822
Vitamins and miscellaneous compounds				
d-Biotin	B4639	5	498	99,600
dl-Pantothenic acid, calcium salt	P5155	500	338,5	0,677

Table 6.2 (Continued)

Component	Stock Keeping Unit (SKU)	Package size (g)	Price (€)	Price per gram (€/g)
Choline chloride	C7527	1000	163	0,163
Ethanolamine	E0135	500 (mL)	49,3	0,099 (€/mL)
Folic acid	F8758	25	120	4,800
Hypoxanthine	H9636	25	128,5	5,140
i-Inositol	I7508	1000	457,5	0,458
Linoleic acid	L1012	5	85,2	17,040
Lipoic acid	T1395	5	116,5	23,300
Methotrexate (MTX)	M8407	0,5	821	1642,000
Nicotinamide	N0636	500	97,2	0,194
Phenol red	P3532	25	74,5	2,980
Pluronic F-68	P5556	100 (mL)	29,7	0,297 (€/mL)
Pyridoxal·HCl	P6155	25	336,5	13,460
Pyridoxine·HCl	P6280	100	220,5	2,205
Riboflavin	R9504	100	134	1,340
Sodium pyruvate	P5280	500	597	1,194
Thiamine·HCl	T1270	100	117	1,170
Thymidinea	T1895	25	599	23,960
Vitamin B_{12}	V6629	5	780	156,000

6.3. Appendix C

In this appendix, the coefficients of each cluster, from each clustering performed will be shown, with the respective standard deviation. The metabolites with negative values are consumed.

Table 6.3 - Coefficients of the 10 clusters (after the second normalization is reversed) from EFMs producing biomass.

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5		
Acetate	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	1,45E-01	0,00E+00	±	0,00E+00
Ala	3,66E-01	±	1,16E-01	3,95E-01	±	1,21E-01	4,65E-01	±	2,20E-01	2,58E-01	±	1,87E-01	0	±	1,13E-01
Arg	-7,85E-06	±	5,65E-15	-7,85E-01	±	9,56E-15	-7,85E-06	±	1,37E-01	-7,24E-06	±	8,42E-01	-7,85E-06	±	5,87E-15
Asn	-3,35E-01	±	2,45E-14	-3,35E-01	±	4,75E-14	-3,35E-01	±	5,84E-01	-2,99E-01	±	9,12E-01	-3,35E-01	±	2,57E-14
Asp	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	6,97E-01	0,00E+00	±	0,00E+00
Biomass	8,55E-08	±	6,14E-17	8,55E-08	±	1,48E-16	8,55E-08	±	1,49E-12	7,65E-08	±	2,55E-08	0,00E+00	±	6,39E-17
Choline	-1,78E-06	±	1,28E-15	-4,26E-01	±	6,27E-06	-3,83E-06	±	5,59E-01	-1,60E-06	±	1,57E-01	-1,78E-06	±	1,33E-15
Cit	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	2,63E-01	0,00E+00	±	0,00E+00
CO2	3,46E-01	±	1,95E-06	3,43E-01	±	1,82E-06	4,50E-01	±	1,27E-01	3,32E-01	±	2,00E-01	0,00E+00	±	8,83E-06
Formate	1,77E-01	±	1,27E-14	1,77E-01	±	2,16E-14	1,77E-01	±	3,88E-01	1,58E-01	±	1,22E-01	1,77E-01	±	1,32E-14
Fum	1,30E-06	±	1,95E-06	4,96E-06	±	9,23E-06	5,84E-06	±	1,45E-01	0,00E+00	±	9,25E-01	1,15E-01	±	1,25E-01
Glc	-2,14E-01	±	1,54E-14	-6,22E-01	±	6,27E-06	-4,15E-01	±	1,34E-01	-4,47E-01	±	1,46E-01	-2,14E-01	±	1,60E-14
Glu	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	1,11E-01	0,00E+00	±	0,00E+00
Gly	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	1,22E-01	0,00E+00	±	0,00E+00
Glyc	6,96E-08	±	4,99E-17	6,96E-08	±	8,46E-17	1,25E-01	±	1,74E-01	6,22E-08	±	1,43E-01	0,00E+00	±	5,19E-17
Glyc3PC	0,00E+00	±	0,00E+00	4,80E-01	±	6,27E-06	2,47E-06	±	5,59E-06	0,00E+00	±	1,18E-01	0,00E+00	±	0,00E+00
His	-4,00E-06	±	2,23E-15	-4,00E-01	±	3,77E-15	-4,00E-06	±	5,43E-11	-2,77E-06	±	9,73E-01	0,00E+00	±	2,32E-15
IgG	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	8,63E-01	0,00E+00	±	0,00E+00
Ile	-7,67E-06	±	5,77E-15	-7,67E-06	±	8,64E-15	-7,67E-06	±	1,23E-01	-6,32E-06	±	7,25E-01	0,00E+00	±	5,28E-15
Isobut	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	9,81E-01	0,00E+00	±	0,00E+00
Isoval	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	9,81E-01	0,00E+00	±	0,00E+00
Lac	7,75E-06	±	1,16E-01	2,11E-06	±	7,84E-06	0,00E+00	±	0,00E+00	8,94E-01	±	2,15E-01	0,00E+00	±	1,24E-01
Leu	-3,29E-01	±	2,37E-14	-3,29E-01	±	4,94E-14	-3,29E-01	±	5,75E-01	-2,95E-01	±	9,81E-01	-3,29E-01	±	2,46E-14
Lys	-1,26E-01	±	9,53E-15	-1,26E-01	±	1,53E-14	-1,26E-01	±	2,20E-01	-1,13E-01	±	5,68E-01	-1,26E-01	±	9,41E-15
Mal	1,30E-06	±	1,95E-06	4,96E-06	±	9,23E-06	5,84E-06	±	1,45E-01	0,00E+00	±	9,25E-01	1,15E-01	±	1,25E-01
Met	-3,42E-06	±	2,46E-15	-3,42E-06	±	4,17E-15	-3,42E-01	±	5,97E-11	-3,62E-06	±	2,46E-01	-3,42E-06	±	2,56E-15
NH4	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	1,82E-01	0,00E+00	±	0,00E+00
Pcholine	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	1,39E-01	0,00E+00	±	0,00E+00
Phe	-5,76E-06	±	4,14E-15	-5,76E-01	±	7,15E-15	-5,76E-06	±	1,45E-01	-5,15E-06	±	4,91E-01	0,00E+00	±	4,35E-15
Pro	-3,88E-01	±	2,22E-14	-3,88E-01	±	3,76E-14	-3,88E-01	±	5,38E-01	-2,76E-01	±	1,12E-01	-3,88E-01	±	2,37E-14
Pyr	-4,62E-01	±	1,18E-01	0,00E+00	±	0,00E+00	-2,47E-06	±	5,59E-06	0,00E+00	±	1,82E-01	-4,64E-01	±	1,14E-01
Ser	-3,62E-01	±	1,16E-01	-3,90E-01	±	1,21E-01	-4,62E-01	±	2,20E-01	-2,55E-01	±	1,95E-01	-3,55E-01	±	1,13E-01
Succ	1,95E-01	±	1,40E-14	1,24E-01	±	9,83E-06	0,00E+00	±	0,00E+00	1,74E-01	±	9,84E-01	1,29E-21	±	0
Thr	-8,19E-06	±	5,88E-15	-8,19E-06	±	9,97E-15	-8,19E-06	±	1,43E-01	-7,32E-06	±	1,75E-01	-8,19E-06	±	6,12E-15
Trp	-7,90E-07	±	5,67E-16	-7,90E-07	±	9,66E-16	-7,90E-07	±	1,38E-11	-7,58E-07	±	4,84E-01	-7,90E-07	±	5,89E-16
Tyr	-4,66E-06	±	2,92E-15	-4,66E-06	±	4,94E-15	-4,67E-06	±	7,78E-11	-3,63E-06	±	5,63E-01	-4,66E-06	±	3,33E-15
Val	-2,96E-01	±	2,13E-14	-2,96E-01	±	3,63E-14	-2,96E-01	±	5,16E-01	-2,65E-01	±	1,54E-01	-2,96E-01	±	2,22E-14
ATP	-1,00E+00	±	7,18E-01	-1,00E+00	±	1,22E-09	-1,00E+00	±	1,95E-09	0,00E+00	±	2,43E-01	-1,00E+00	±	7,47E-01
NADH	7,75E-06	±	1,16E-01	1,57E-01	±	1,21E-01	1,76E-01	±	2,20E-01	5,29E-01	±	2,98E-01	6,64E-06	±	1,13E-01
NADPH	3,15E-01	±	1,95E-01	3,13E-01	±	1,82E-06	4,20E-01	±	1,27E-01	3,51E-01	±	1,44E-01	3,36E-01	±	8,83E-06
FADH2	3,90E-01	±	2,80E-14	4,64E-01	±	9,83E-06	5,84E-01	±	1,19E-09	3,48E-01	±	1,28E-01	5,84E-01	±	4,37E-14

Table 6.3 (Continued)

	Cluster 6			Cluster 7			Cluster 8			Cluster 9			Cluster 10		
Acetate	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Ala	2,20E-01	±	1,88E-01	2,83E-01	±	9,28E-01	4,85E-01	±	2,60E-01	1,97E-01	±	6,69E-06	7,44E-01	±	6,47E-01
Arg	-3,93E-06	±	6,27E-11	-2,22E-06	±	7,29E-07	-7,85E-06	±	1,57E-01	-3,93E-06	±	1,18E-11	-1,56E-06	±	5,52E-07
Asn	-1,67E-01	±	2,67E-01	-9,46E-06	±	3,16E-01	-3,35E-01	±	6,69E-01	-1,67E-01	±	5,23E-11	-6,64E-06	±	2,35E-06
Asp	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Biomass	4,27E-08	±	6,83E-13	2,42E-08	±	7,93E-09	8,55E-08	±	1,79E-12	4,27E-08	±	1,28E-13	1,70E-08	±	6,97E-09
Choline	-8,92E-07	±	1,42E-11	-5,47E-07	±	1,66E-07	-1,78E-06	±	3,57E-11	-2,13E-01	±	3,13E-06	-3,54E-07	±	1,25E-07
Cit	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
CO2	1,77E-01	±	8,27E-07	2,83E-01	±	9,28E-01	3,48E-01	±	1,97E-06	1,72E-01	±	9,86E-07	1,98E-01	±	7,29E-01
Formate	5,00E-01	±	2,72E-06	5,59E-06	±	1,64E-06	1,77E-01	±	3,54E-01	5,00E-01	±	1,54E-01	3,51E-01	±	1,25E-06
Fum	2,65E-06	±	4,56E-06	0,00E+00	±	0,00E+00	1,20E-06	±	1,87E-06	2,48E-06	±	4,62E-06	0,00E+00	±	0,00E+00
Glc	-1,96E-01	±	8,86E-06	-1,41E-01	±	4,64E-01	-5,25E-01	±	1,28E-01	-3,19E-01	±	3,13E-01	-9,92E-01	±	3,51E-01
Glu	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Gly	0,499988469	±	2,72E-06	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,499989272	±	0,153963523	0,00E+00	±	0,00E+00
Glyc	6,27E-06	±	9,68E-06	1,96E-08	±	6,45E-09	2,24E-01	±	2,22E-01	3,48E-08	±	1,43E-13	1,38E-08	±	4,89E-09
Glyc3PC	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	2,40E-01	±	3,13E-06	0,00E+00	±	0,00E+00
His	-1,55E-06	±	2,48E-11	-8,76E-07	±	2,88E-07	-4,00E-06	±	6,20E-11	-1,55E-06	±	4,66E-12	-6,15E-07	±	2,18E-07
IgG	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Ile	-3,53E-01	±	5,64E-11	-2,00E-01	±	6,56E-07	-7,68E-06	±	1,41E-01	-3,53E-06	±	1,64E-11	-1,42E-06	±	4,97E-07
Isobut	0,00E+00	±	0,00E+00	0,174727287	±	0,246176138	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,124372433	±	0,175459325
Isoval	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Lac	1,94E-06	±	4,39E-06	0,00E+00	±	0,00E+00	1,79E-01	±	6,45E-06	1,57E-01	±	3,54E-06	0,00E+00	±	0,00E+00
Leu	-1,65E-01	±	2,64E-01	-9,36E-06	±	3,56E-06	-3,29E-01	±	6,58E-01	-1,65E-01	±	4,94E-11	-6,53E-01	±	2,32E-06
Lys	-6,31E-06	±	1,64E-01	-3,57E-06	±	1,17E-06	-1,26E-01	±	2,52E-01	-6,31E-06	±	1,90E-11	-2,58E-06	±	8,86E-07
Mal	2,65E-06	±	4,56E-06	0,00E+00	±	0,00E+00	1,20E-06	±	1,87E-06	2,48E-06	±	4,62E-06	0,00E+00	±	0,00E+00
Met	-1,71E-06	±	2,73E-11	-9,67E-07	±	3,18E-07	-3,42E-06	±	6,84E-11	-1,71E-06	±	5,14E-12	-6,79E-07	±	2,46E-07
NH4	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Pcholine	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Phe	-2,88E-06	±	4,62E-11	-1,63E-06	±	5,35E-07	-5,76E-06	±	1,15E-01	-2,88E-06	±	8,65E-12	-1,14E-06	±	4,57E-07
Pro	-1,54E-01	±	2,47E-01	-8,72E-06	±	2,86E-06	-3,88E-01	±	6,17E-01	-1,54E-01	±	4,63E-11	-6,12E-01	±	2,17E-06
Pyr	-9,73E-06	±	1,22E-01	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Ser	-5,12E-01	±	8,17E-06	-8,25E-06	±	2,64E-06	-4,86E-01	±	2,60E-01	-5,85E-01	±	4,59E-01	-5,45E-01	±	6,96E-01
Succ	5,95E-06	±	4,81E-06	1,85E-01	±	1,53E-01	1,95E-01	±	3,89E-01	6,20E-01	±	4,91E-06	7,43E-01	±	1,51E-01
Thr	-4,95E-06	±	6,54E-11	-2,31E-06	±	7,60E-07	-8,19E-06	±	1,64E-01	-4,94E-06	±	1,23E-11	-1,62E-01	±	5,76E-07
Trp	-3,95E-07	±	6,31E-12	-2,23E-07	±	7,32E-08	-7,89E-07	±	1,58E-11	-3,95E-07	±	1,18E-12	-1,57E-07	±	5,55E-08
Tyr	-2,33E-06	±	3,24E-11	-1,15E-06	±	3,77E-07	-4,67E-06	±	8,12E-11	-2,33E-06	±	6,93E-12	-8,55E-07	±	2,85E-07
Val	-1,48E-01	±	2,36E-01	-2,83E-01	±	9,28E-01	-2,96E-01	±	5,92E-01	-1,48E-01	±	4,45E-11	-1,98E-01	±	7,29E-01
ATP	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	-1,00E+00	±	2,22E-09	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
NADH	7,55E-06	±	1,88E-01	7,82E-01	±	1,21E-01	1,96E-01	±	2,60E-01	5,28E-06	±	6,69E-01	0,00E+00	±	0,00E+00
NADPH	5,45E-01	±	2,82E-06	9,64E-06	±	3,17E-06	3,17E-01	±	1,97E-06	5,46E-01	±	1,75E-01	6,77E-06	±	2,40E-06
FADH2	2,33E-01	±	4,81E-06	1,85E-01	±	1,53E-01	3,90E-01	±	7,79E-01	2,32E-01	±	4,91E-06	7,43E-01	±	1,51E-01

Table 6.4 - Coefficients of the 10 clusters (after the second normalization is reversed) from EFMs producing IgG.

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5		
Acetate	4,47E-03	±	2,00E-02	1,14E-02	±	2,99E-02	2,55E-01	±	6,50E-02	9,80E-02	±	1,01E-01	2,72E-02	±	7,31E-02
Ala	0,00E+00	±	0,00E+00	1,59E-02	±	6,06E-02	1,07E-03	±	1,78E-02	2,74E-03	±	2,14E-02	1,97E-02	±	6,57E-02
Arg	-5,13E-02	±	3,09E-02	-2,80E-01	±	3,92E-02	-9,04E-02	±	6,18E-02	-7,54E-02	±	5,90E-02	-7,45E-02	±	9,60E-03
Asn	-4,70E-02	±	3,31E-02	-1,20E-01	±	7,18E-02	-6,79E-02	±	3,56E-02	-6,26E-02	±	4,80E-02	-1,46E-01	±	1,01E-01
Asp	-6,83E-02	±	8,35E-02	-7,29E-02	±	5,99E-02	-8,54E-02	±	4,69E-02	-5,99E-02	±	5,46E-02	-6,65E-02	±	8,13E-02
Biomass	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Choline	-2,29E-02	±	5,38E-02	-1,76E-02	±	4,44E-02	-2,40E-02	±	5,53E-02	-1,65E-03	±	6,20E-03	-2,09E-02	±	5,53E-02
Cit	2,00E-03	±	3,00E-03	1,32E-04	±	6,66E-04	2,81E-04	±	1,00E-03	7,47E-04	±	1,75E-03	1,78E-03	±	3,04E-03
CO2	3,20E-01	±	1,30E-01	1,02E-01	±	5,61E-02	5,43E-01	±	1,15E-01	2,02E-01	±	1,12E-01	2,88E-01	±	8,92E-02
Formate	1,15E-01	±	6,50E-02	1,54E-01	±	1,58E-02	1,39E-01	±	6,60E-02	1,23E-01	±	6,35E-02	1,48E-01	±	4,23E-02
Fum	3,22E-02	±	6,41E-02	7,50E-03	±	2,01E-02	3,09E-03	±	5,97E-03	2,13E-02	±	4,77E-02	1,17E-02	±	4,38E-02
Glc	-9,12E-02	±	6,15E-02	-1,29E-01	±	5,97E-02	-8,32E-02	±	5,77E-02	-5,37E-02	±	1,49E-02	-1,28E-01	±	7,62E-02
Glu	-3,91E-02	±	9,39E-02	-1,79E-03	±	9,52E-03	-2,84E-02	±	5,29E-02	-4,06E-02	±	8,26E-02	-1,52E-01	±	1,59E-01
Gly	3,14E-02	±	6,38E-02	0,00E+00	±	0,00E+00	2,40E-02	±	5,32E-02	3,10E-02	±	6,28E-02	1,24E-02	±	4,04E-02
Glyc	1,79E-02	±	4,87E-02	1,85E-02	±	4,81E-02	8,42E-03	±	3,80E-02	1,23E-03	±	1,79E-03	2,19E-02	±	5,93E-02
Glyc3PC	1,72E-02	±	4,74E-02	1,76E-02	±	4,44E-02	8,26E-03	±	3,71E-02	1,16E-03	±	1,77E-03	2,05E-02	±	5,52E-02
His	-3,21E-02	±	3,48E-02	-4,95E-02	±	2,11E-02	-1,04E-01	±	9,32E-02	-4,55E-02	±	4,98E-02	-4,02E-02	±	8,06E-03
IgG	1,48E-04	±	4,52E-05	2,74E-04	±	2,81E-05	2,05E-04	±	5,14E-05	1,64E-04	±	4,38E-05	2,42E-04	±	2,90E-05
Ile	-2,61E-02	±	7,94E-03	-5,75E-02	±	2,99E-02	-3,92E-02	±	1,75E-02	-7,60E-02	±	7,04E-02	-4,25E-02	±	5,11E-03
Isobut	1,63E-02	±	5,30E-02	6,45E-03	±	2,56E-02	0,00E+00	±	0,00E+00	9,71E-03	±	3,89E-02	2,45E-02	±	6,78E-02
Isoval	6,12E-03	±	3,35E-02	4,69E-03	±	2,23E-02	0,00E+00	±	0,00E+00	7,35E-03	±	3,54E-02	2,90E-02	±	7,39E-02
Lac	3,66E-03	±	1,90E-02	2,58E-02	±	8,35E-02	1,33E-01	±	1,68E-01	4,70E-01	±	7,48E-02	4,57E-02	±	1,03E-01
Leu	-8,28E-02	±	4,27E-02	-1,46E-01	±	2,57E-02	-1,07E-01	±	2,72E-02	-9,23E-02	±	4,36E-02	-1,54E-01	±	7,57E-02
Lys	-7,01E-02	±	2,13E-02	-1,32E-01	±	1,69E-02	-9,84E-02	±	2,53E-02	-9,18E-02	±	4,11E-02	-1,14E-01	±	1,37E-02
Mal	2,37E-02	±	5,63E-02	7,31E-03	±	2,01E-02	2,74E-03	±	4,41E-03	1,31E-02	±	3,91E-02	1,16E-02	±	4,38E-02
Met	-9,79E-03	±	2,98E-03	-1,81E-02	±	1,85E-03	-1,35E-02	±	3,39E-03	-1,08E-02	±	2,89E-03	-1,59E-02	±	1,92E-03
NH4	1,34E-03	±	1,04E-02	1,28E-03	±	4,58E-03	2,77E-02	±	9,39E-02	1,09E-02	±	4,44E-02	1,13E-02	±	5,85E-02
Pcholine	5,67E-03	±	2,69E-02	0,00E+00	±	0,00E+00	1,57E-02	±	4,38E-02	4,84E-04	±	6,02E-03	3,67E-04	±	4,59E-03
Phe	-6,11E-02	±	2,83E-02	-1,12E-01	±	4,12E-02	-7,74E-02	±	3,26E-02	-6,55E-02	±	2,92E-02	-1,02E-01	±	3,63E-02
Pro	-1,44E-01	±	1,22E-01	-1,40E-01	±	1,55E-02	-1,17E-01	±	4,05E-02	-1,20E-01	±	7,36E-02	-1,61E-01	±	8,32E-02
Pyr	-5,29E-02	±	7,50E-02	-8,41E-02	±	1,25E-01	-8,48E-02	±	1,45E-01	-5,62E-01	±	8,43E-02	-8,40E-02	±	1,33E-01
Ser	-2,55E-01	±	8,94E-02	-4,52E-01	±	1,02E-01	-3,52E-01	±	1,21E-01	-2,78E-01	±	8,53E-02	-4,77E-01	±	1,52E-01
Succ	4,90E-02	±	7,38E-02	1,18E-02	±	3,03E-02	2,93E-03	±	1,52E-02	4,10E-02	±	6,42E-02	4,03E-02	±	8,35E-02
Thr	-1,08E-01	±	2,72E-02	-1,75E-01	±	1,71E-02	-1,29E-01	±	3,16E-02	-1,08E-01	±	2,57E-02	-1,54E-01	±	1,92E-02
Trp	-2,12E-02	±	6,45E-03	-3,92E-02	±	4,02E-03	-1,49E-01	±	2,65E-02	-2,85E-02	±	2,12E-02	-3,45E-02	±	4,15E-03
Tyr	-2,37E-02	±	2,34E-02	-4,47E-02	±	3,92E-02	-4,00E-02	±	3,06E-02	-2,84E-02	±	2,55E-02	-3,62E-02	±	3,52E-02
Val	-2,08E-01	±	9,22E-02	-2,12E-01	±	3,25E-02	-1,46E-01	±	3,68E-02	-1,49E-01	±	6,24E-02	-2,04E-01	±	7,46E-02
ATP	-9,52E-02	±	1,24E-01	-3,45E-01	±	7,49E-02	-1,83E-02	±	8,58E-02	-7,80E-02	±	1,11E-01	-2,15E-01	±	1,41E-01
NADH	6,76E-01	±	8,29E-02	4,07E-01	±	2,13E-01	2,82E-01	±	2,26E-01	0,00E+00	±	0,00E+00	2,86E-01	±	2,18E-01
NADPH	1,85E-01	±	8,03E-02	2,57E-01	±	4,23E-02	2,12E-01	±	8,03E-02	2,32E-01	±	1,13E-01	2,35E-01	±	5,59E-02
FADH2	1,99E-01	±	1,15E-01	4,16E-02	±	6,64E-02	7,25E-03	±	2,50E-02	1,40E-01	±	1,11E-01	3,43E-02	±	7,62E-02

Table 6.4 (Continued)

	Cluster 6			Cluster 7			Cluster 8			Cluster 9			Cluster 10		
Acetate	1,19E-02	±	3,19E-02	3,60E-02	±	5,07E-02	2,33E-01	±	6,23E-02	1,98E-01	±	7,48E-02	1,30E-01	±	9,92E-02
Ala	6,95E-03	±	3,35E-02	0,00E+00	±	0,00E+00	9,29E-03	±	4,20E-02	2,70E-03	±	2,11E-02	1,80E-03	±	1,53E-02
Arg	-8,85E-02	±	5,73E-02	-9,47E-02	±	7,85E-03	-7,11E-02	±	3,07E-02	-7,23E-02	±	4,14E-02	-6,63E-02	±	4,94E-02
Asn	-8,41E-02	±	5,55E-02	-1,07E-01	±	5,95E-02	-8,70E-02	±	5,43E-02	-7,49E-02	±	5,18E-02	-5,86E-02	±	4,23E-02
Asp	-6,57E-02	±	5,16E-02	-1,09E-01	±	5,66E-02	-7,07E-02	±	6,21E-02	-5,56E-02	±	6,70E-02	-4,47E-02	±	4,07E-02
Biomass	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Choline	-1,26E-02	±	3,54E-02	-1,76E-02	±	4,74E-02	-1,16E-02	±	2,80E-02	-1,14E-02	±	2,80E-02	-1,26E-01	±	1,75E-01
Cit	5,93E-04	±	1,56E-03	3,49E-05	±	2,64E-04	5,52E-04	±	1,60E-03	7,78E-04	±	1,76E-03	2,96E-04	±	1,02E-03
CO2	1,34E-01	±	8,44E-02	1,38E-01	±	6,46E-02	3,00E-01	±	5,90E-02	1,22E-01	±	1,07E-01	2,81E-01	±	1,80E-01
Formate	3,49E-01	±	3,80E-02	1,71E-01	±	1,31E-02	1,15E-01	±	1,96E-02	1,35E-01	±	6,08E-02	1,01E-01	±	5,40E-02
Fum	8,61E-03	±	2,40E-02	1,29E-02	±	3,26E-02	2,44E-03	±	4,08E-03	3,56E-02	±	6,85E-02	2,20E-03	±	7,99E-03
Glc	-9,09E-02	±	5,26E-02	-1,41E-01	±	6,56E-02	-9,05E-02	±	4,01E-02	-8,64E-02	±	4,64E-02	-4,51E-01	±	6,92E-02
Glu	-2,17E-02	±	5,09E-02	-7,24E-04	±	4,41E-03	-7,01E-02	±	7,93E-02	-4,66E-02	±	7,08E-02	-5,05E-02	±	8,26E-02
Gly	2,42E-01	±	3,45E-02	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	2,56E-02	±	5,62E-02	1,95E-02	±	4,94E-02
Glyc	1,48E-02	±	4,27E-02	2,00E-02	±	5,33E-02	1,21E-02	±	3,12E-02	1,31E-02	±	3,77E-02	2,86E-01	±	2,10E-01
Glyc3PC	1,26E-02	±	3,54E-02	1,76E-02	±	4,74E-02	1,10E-02	±	2,78E-02	9,48E-03	±	2,64E-02	1,20E-01	±	1,72E-01
His	-9,12E-02	±	7,48E-02	-2,74E-01	±	3,85E-02	-5,70E-02	±	5,55E-02	-5,61E-02	±	6,03E-02	-3,62E-02	±	3,90E-02
IgG	1,91E-04	±	2,34E-05	3,06E-04	±	2,34E-05	2,05E-04	±	3,50E-05	1,95E-04	±	4,08E-05	1,45E-04	±	3,59E-05
Ile	-4,37E-02	±	3,18E-02	-8,03E-02	±	5,24E-02	-3,67E-02	±	7,76E-03	-2,19E-01	±	3,92E-02	-7,04E-02	±	6,53E-02
Isobut	2,16E-02	±	5,00E-02	1,41E-02	±	3,98E-02	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	1,41E-03	±	1,40E-02
Isoval	2,08E-02	±	4,97E-02	1,40E-02	±	3,99E-02	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	9,02E-04	±	1,26E-02
Lac	5,82E-02	±	1,07E-01	4,94E-02	±	9,70E-02	1,36E-02	±	6,25E-02	1,14E-02	±	3,84E-02	4,10E-01	±	6,45E-02
Leu	-1,20E-01	±	4,91E-02	-1,72E-01	±	4,12E-02	-1,06E-01	±	1,81E-02	-1,01E-01	±	2,11E-02	-7,66E-02	±	2,39E-02
Lys	-9,05E-02	±	1,10E-02	-1,51E-01	±	2,22E-02	-2,06E-01	±	2,24E-02	-9,21E-02	±	1,93E-02	-7,15E-02	±	2,36E-02
Mal	8,16E-03	±	2,38E-02	1,20E-02	±	3,23E-02	2,43E-03	±	4,07E-03	2,80E-02	±	6,36E-02	8,46E-04	±	4,25E-03
Met	-1,26E-02	±	1,54E-03	-2,02E-02	±	1,54E-03	-1,35E-02	±	2,31E-03	-1,29E-02	±	2,69E-03	-9,55E-03	±	2,37E-03
NH4	8,23E-03	±	4,12E-02	9,27E-03	±	4,62E-02	3,98E-02	±	6,37E-02	5,74E-03	±	2,56E-02	1,19E-02	±	3,03E-02
Pcholine	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	5,76E-04	±	4,35E-03	1,90E-03	±	1,11E-02	6,11E-03	±	2,37E-02
Phe	-7,54E-02	±	2,69E-02	-1,13E-01	±	4,11E-02	-8,00E-02	±	3,18E-02	-8,04E-02	±	3,31E-02	-5,65E-02	±	2,44E-02
Pro	-1,13E-01	±	4,82E-02	-1,56E-01	±	1,27E-02	-1,36E-01	±	6,41E-02	-1,37E-01	±	6,55E-02	-1,03E-01	±	5,97E-02
Pyr	-1,06E-01	±	1,36E-01	-1,25E-01	±	1,34E-01	-1,23E-01	±	1,26E-01	-8,12E-02	±	1,06E-01	0,00E+00	±	0,00E+00
Ser	-5,65E-01	±	7,89E-02	-4,73E-01	±	5,38E-02	-3,07E-01	±	5,24E-02	-3,17E-01	±	8,33E-02	-2,36E-01	±	7,43E-02
Succ	1,80E-02	±	3,70E-02	2,53E-02	±	4,86E-02	4,47E-04	±	1,45E-03	8,00E-02	±	9,32E-02	3,64E-02	±	5,68E-02
Thr	-1,21E-01	±	1,46E-02	-1,94E-01	±	1,90E-02	-1,30E-01	±	2,26E-02	-1,33E-01	±	2,60E-02	-9,42E-02	±	2,18E-02
Trp	-2,74E-02	±	3,34E-03	-4,37E-02	±	3,34E-03	-2,93E-02	±	5,01E-03	-2,88E-02	±	1,04E-02	-6,23E-02	±	5,61E-02
Tyr	-3,40E-02	±	2,75E-02	-6,16E-02	±	4,01E-02	-3,81E-02	±	3,01E-02	-3,19E-02	±	3,09E-02	-2,69E-02	±	2,24E-02
Val	-1,75E-01	±	5,28E-02	-2,51E-01	±	4,59E-02	-1,47E-01	±	2,50E-02	-1,39E-01	±	2,92E-02	-1,19E-01	±	4,31E-02
ATP	0,00E+00	±	0,00E+00	-3,65E-01	±	6,73E-02	-4,65E-02	±	7,89E-02	-4,37E-02	±	7,81E-02	-4,19E-02	±	8,76E-02
NADH	1,85E-01	±	1,87E-01	1,39E-01	±	1,81E-01	4,54E-01	±	1,36E-01	6,12E-01	±	1,04E-01	1,48E-03	±	2,53E-02
NADPH	4,13E-01	±	4,20E-02	3,00E-01	±	6,47E-02	5,10E-01	±	6,84E-02	2,30E-01	±	8,00E-02	1,95E-01	±	8,09E-02
FADH2	4,08E-02	±	6,52E-02	8,11E-02	±	9,15E-02	1,11E-01	±	3,09E-02	3,01E-01	±	1,19E-01	1,07E-01	±	1,21E-01

Table 6.5 – Coefficients of the 10 clusters (after the second normalization is reversed) from EFMs producing only CO_2.

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5		
Acetate	6,69E-02	±	2,00E-02	1,00E-01	±	2,99E-02	1,12E-02	±	6,50E-02	5,82E-02	±	1,01E-01	6,59E-02	±	7,31E-02
Ala	8,53E-01	±	0,00E+00	2,47E-01	±	6,06E-02	1,83E-01	±	1,78E-02	3,17E-01	±	2,14E-02	1,08E-01	±	6,57E-02
Arg	-1,38E-02	±	3,09E-02	-2,00E-02	±	3,92E-02	-5,34E-03	±	6,18E-02	0,00E+00	±	5,90E-02	-1,97E-02	±	9,60E-03
Asn	-8,14E-02	±	3,31E-02	-1,34E-01	±	7,18E-02	-1,02E-01	±	3,56E-02	-3,77E-03	±	4,80E-02	-1,17E-01	±	1,01E-01
Asp	-1,66E-01	±	8,35E-02	-1,35E-01	±	5,99E-02	-3,38E-03	±	4,69E-02	-6,47E-01	±	5,46E-02	-7,14E-02	±	8,13E-02
Biomass	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Choline	-1,45E-01	±	5,38E-02	-8,76E-02	±	4,44E-02	-1,30E-01	±	5,53E-02	-1,83E-01	±	6,20E-03	-8,55E-02	±	5,53E-02
Cit	4,77E-01	±	3,00E-03	4,80E-01	±	6,66E-04	6,66E-01	±	1,00E-03	7,22E-01	±	1,75E-03	3,23E-01	±	3,04E-03
CO2	4,80E-01	±	1,30E-01	4,96E-01	±	5,61E-02	6,02E-01	±	1,15E-01	6,61E-01	±	1,12E-01	3,74E-01	±	8,92E-02
Formate	1,68E-04	±	6,50E-02	1,78E-04	±	1,58E-02	4,21E-04	±	6,60E-02	0,00E+00	±	6,35E-02	1,84E-04	±	4,23E-02
Fum	1,01E-01	±	6,41E-02	1,28E-01	±	2,01E-02	9,35E-02	±	5,97E-03	1,55E-01	±	4,77E-02	1,28E-01	±	4,38E-02
Glc	-1,57E-01	±	6,15E-02	-3,52E-05	±	5,97E-02	-1,14E-01	±	5,77E-02	-1,71E-01	±	1,49E-02	-3,02E-03	±	7,62E-02
Glu	-8,93E-03	±	9,39E-02	-1,60E-02	±	9,52E-03	-2,69E-03	±	5,29E-02	0,00E+00	±	8,26E-02	-2,27E-02	±	1,59E-01
Gly	2,70E-02	±	6,38E-02	2,39E-02	±	0,00E+00	4,46E-02	±	5,32E-02	3,81E-02	±	6,28E-02	2,03E-02	±	4,04E-02
Glyc	6,86E-02	±	4,87E-02	0,00E+00	±	4,81E-02	4,08E-02	±	3,80E-02	8,43E-02	±	1,79E-03	2,03E-03	±	5,93E-02
Glyc3PC	5,64E-02	±	4,74E-02	0,00E+00	±	4,44E-02	4,79E-02	±	3,71E-02	6,66E-02	±	1,77E-03	1,22E-03	±	5,52E-02
His	-3,39E-03	±	3,48E-02	-1,08E-02	±	2,11E-02	-6,54E-04	±	9,32E-02	0,00E+00	±	4,98E-02	-6,59E-02	±	8,06E-03
IgG	0,00E+00	±	4,52E-05	0,00E+00	±	2,81E-05	0,00E+00	±	5,14E-05	0,00E+00	±	4,38E-05	0,00E+00	±	2,90E-05
Ile	-3,07E-02	±	7,94E-03	-3,08E-02	±	2,99E-02	-5,31E-02	±	1,75E-02	-3,64E-02	±	7,04E-02	-3,11E-02	±	5,11E-03
Isobut	2,51E-03	±	5,30E-02	2,21E-03	±	2,56E-02	8,90E-04	±	0,00E+00	0,00E+00	±	3,89E-02	5,59E-03	±	6,78E-02
Isoval	2,51E-03	±	3,35E-02	2,21E-03	±	2,23E-02	8,90E-04	±	0,00E+00	0,00E+00	±	3,54E-02	3,19E-03	±	7,39E-02
Lac	0,00E+00	±	1,90E-02	4,80E-01	±	8,35E-02	1,01E-02	±	1,68E-01	9,86E-03	±	7,48E-02	4,15E-01	±	1,03E-01
Leu	-4,48E-02	±	4,27E-02	-4,64E-02	±	2,57E-02	-3,07E-02	±	2,72E-02	-3,92E-02	±	4,36E-02	-2,75E-02	±	7,57E-02
Lys	-2,11E-02	±	2,13E-02	-2,37E-02	±	1,69E-02	-2,59E-02	±	2,53E-02	-2,86E-02	±	4,11E-02	-1,96E-02	±	1,37E-02
Mal	6,74E-02	±	5,63E-02	9,20E-02	±	2,01E-02	7,86E-02	±	4,41E-03	1,42E-01	±	3,91E-02	7,34E-02	±	4,38E-02
Met	-9,95E-02	±	2,98E-03	-9,18E-02	±	1,85E-03	-3,12E-01	±	3,39E-03	0,00E+00	±	2,89E-03	-6,97E-02	±	1,92E-03
NH4	2,33E-02	±	1,04E-02	1,27E-01	±	4,58E-03	1,17E-01	±	9,39E-02	1,03E-01	±	4,44E-02	8,03E-01	±	5,85E-02
Pcholine	1,17E-01	±	2,69E-02	9,93E-02	±	0,00E+00	1,07E-01	±	4,38E-02	1,51E-01	±	6,02E-03	9,59E-02	±	4,59E-03
Phe	-4,55E-02	±	2,83E-02	-5,88E-02	±	4,12E-02	-5,81E-02	±	3,26E-02	-9,65E-02	±	2,92E-02	-8,09E-02	±	3,63E-02
Pro	-1,23E-02	±	1,22E-01	-1,56E-02	±	1,55E-02	-5,81E-03	±	4,05E-02	0,00E+00	±	7,36E-02	-1,71E-02	±	8,32E-02
Pyr	-5,30E-02	±	7,50E-02	-7,73E-01	±	1,25E-01	-4,29E-02	±	1,45E-01	-7,37E-02	±	8,43E-02	-1,64E-03	±	1,33E-01
Ser	-5,42E-01	±	8,94E-02	-1,20E-02	±	1,02E-01	-5,27E-02	±	1,21E-01	-3,63E-02	±	8,53E-02	-5,26E-01	±	1,52E-01
Succ	1,02E-01	±	7,38E-02	1,47E-01	±	3,03E-02	8,78E-02	±	1,52E-02	1,98E-01	±	6,42E-02	1,46E-01	±	8,35E-02
Thr	-4,73E-02	±	2,72E-02	-3,69E-02	±	1,71E-02	-9,56E-02	±	3,16E-02	-4,30E-02	±	2,57E-02	-2,36E-02	±	1,92E-02
Trp	-2,98E-02	±	6,45E-03	-2,24E-02	±	4,02E-03	-4,00E-02	±	2,65E-02	-3,31E-02	±	2,12E-02	-4,08E-03	±	4,15E-03
Tyr	-9,58E-02	±	2,34E-02	-1,16E-01	±	3,92E-02	-5,27E-02	±	3,06E-02	-1,26E-01	±	2,55E-02	-6,28E-02	±	3,52E-02
Val	-9,50E-03	±	9,22E-02	-1,50E-02	±	3,25E-02	-8,85E-03	±	3,68E-02	0,00E+00	±	6,24E-02	-1,07E-02	±	7,46E-02
ATP	8,69E-02	±	1,24E-01	6,54E-02	±	7,49E-02	8,17E-02	±	8,58E-02	8,97E-02	±	1,11E-01	5,96E-02	±	1,41E-01
NADH	1,62E-03	±	8,29E-02	1,93E-02	±	2,13E-01	7,75E-01	±	2,26E-01	6,06E-01	±	0,00E+00	4,56E-02	±	2,18E-01
NADPH	1,07E-01	±	8,03E-02	9,30E-02	±	4,23E-02	8,93E-02	±	8,03E-02	1,21E-01	±	1,13E-01	7,66E-02	±	5,59E-02
FADH2	3,04E-01	±	1,15E-01	2,96E-01	±	6,64E-02	4,67E-01	±	2,50E-02	2,83E-01	±	1,11E-01	2,31E-01	±	7,62E-02

Table 6.5 (Continued)

	Cluster 6			Cluster 7			Cluster 8			Cluster 9			Cluster 10		
Acetate	8,95E-02	±	3,19E-02	1,47E-01	±	5,07E-02	5,14E-01	±	6,23E-02	4,74E-01	±	7,48E-02	5,61E-01	±	9,92E-02
Ala	2,54E-01	±	3,35E-02	3,19E-01	±	0,00E+00	5,09E-01	±	4,20E-02	5,71E-01	±	2,11E-02	5,49E-01	±	1,53E-02
Arg	-2,58E-02	±	5,73E-02	-4,17E-02	±	7,85E-03	-5,67E-02	±	3,07E-02	-4,10E-02	±	4,14E-02	-6,04E-02	±	4,94E-02
Asn	-1,81E-01	±	5,55E-02	-1,28E-01	±	5,95E-02	-1,37E-02	±	5,43E-02	-2,68E-02	±	5,18E-02	-9,59E-03	±	4,23E-02
Asp	-8,01E-02	±	5,16E-02	-7,94E-02	±	5,66E-02	-1,58E-03	±	6,21E-02	-4,83E-03	±	6,70E-02	-2,96E-03	±	4,07E-02
Biomass	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	0,00E+00
Choline	-8,96E-02	±	3,54E-02	-6,97E-01	±	4,74E-02	-3,03E-01	±	2,80E-02	-2,83E-01	±	2,80E-02	-3,11E-01	±	1,75E-01
Cit	4,10E-01	±	1,56E-03	3,11E-01	±	2,64E-04	7,00E-02	±	1,60E-03	1,81E-01	±	1,76E-03	1,71E-02	±	1,02E-03
CO2	4,41E-01	±	8,44E-02	3,84E-01	±	6,46E-02	2,79E-01	±	5,90E-02	4,84E-01	±	1,07E-01	4,47E-01	±	1,80E-01
Formate	1,13E-04	±	3,80E-02	7,22E-05	±	1,31E-02	0,00E+00	±	1,96E-02	0,00E+00	±	6,08E-02	0,00E+00	±	5,40E-02
Fum	1,21E-01	±	2,40E-02	1,35E-01	±	3,26E-02	8,86E-02	±	4,08E-03	1,72E-01	±	6,85E-02	2,11E-01	±	7,99E-03
Glc	-7,50E-01	±	5,26E-02	-6,25E-01	±	6,56E-02	-2,07E-01	±	4,01E-02	-2,21E-01	±	4,64E-02	-2,32E-01	±	6,92E-02
Glu	-2,30E-02	±	5,09E-02	-2,69E-02	±	4,41E-03	-7,41E-02	±	7,93E-02	-4,78E-02	±	7,08E-02	-3,39E-02	±	8,26E-02
Gly	2,09E-02	±	3,45E-02	1,92E-02	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	5,62E-02	1,17E-02	±	4,94E-02
Glyc	7,63E-01	±	4,27E-02	0,00E+00	±	5,33E-02	9,89E-02	±	3,12E-02	1,10E-01	±	3,77E-02	1,24E-01	±	2,10E-01
Glyc3PC	0,00E+00	±	3,54E-02	7,79E-01	±	4,74E-02	6,88E-02	±	2,78E-02	9,03E-02	±	2,64E-02	8,30E-02	±	1,72E-01
His	-1,89E-02	±	7,48E-02	-3,49E-02	±	3,85E-02	-5,32E-02	±	5,55E-02	-2,66E-02	±	6,03E-02	-7,31E-02	±	3,90E-02
IgG	0,00E+00	±	2,34E-05	0,00E+00	±	2,34E-05	0,00E+00	±	3,50E-05	0,00E+00	±	4,08E-05	0,00E+00	±	3,59E-05
Ile	-2,91E-02	±	3,18E-02	-2,36E-02	±	5,24E-02	-2,55E-02	±	7,76E-03	-6,32E-02	±	3,92E-02	-8,79E-02	±	6,53E-02
Isobut	4,95E-03	±	5,00E-02	4,42E-03	±	3,98E-02	4,51E-03	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	1,40E-02
Isoval	4,95E-03	±	4,97E-02	4,42E-03	±	3,99E-02	6,13E-03	±	0,00E+00	0,00E+00	±	0,00E+00	0,00E+00	±	1,26E-02
Lac	3,48E-01	±	1,07E-01	2,21E-01	±	9,70E-02	1,30E-02	±	6,25E-02	2,34E-02	±	3,84E-02	3,31E-02	±	6,45E-02
Leu	-3,90E-02	±	4,91E-02	-3,88E-02	±	4,12E-02	-2,50E-01	±	1,81E-02	-3,61E-03	±	2,11E-02	-1,33E-03	±	2,39E-02
Lys	-2,23E-02	±	1,10E-02	-1,80E-02	±	2,22E-02	-1,72E-02	±	2,24E-02	0,00E+00	±	1,93E-02	-2,26E-03	±	2,36E-02
Mal	9,11E-02	±	2,38E-02	1,10E-01	±	3,23E-02	9,01E-02	±	4,07E-03	1,47E-01	±	6,36E-02	1,77E-01	±	4,25E-03
Met	-5,10E-02	±	1,54E-03	-3,27E-02	±	1,54E-03	-2,10E-02	±	2,31E-03	-6,47E-02	±	2,69E-03	0,00E+00	±	2,37E-03
NH4	1,69E-01	±	4,12E-02	1,13E-01	±	4,62E-02	3,95E-02	±	6,37E-02	2,39E-02	±	2,56E-02	3,38E-02	±	3,03E-02
Pcholine	1,02E-01	±	0,00E+00	1,33E-01	±	0,00E+00	2,85E-01	±	4,35E-03	2,45E-01	±	1,11E-02	2,83E-01	±	2,37E-02
Phe	-7,08E-02	±	2,69E-02	-9,43E-02	±	4,11E-02	0,00E+00	±	3,18E-02	0,00E+00	±	3,31E-02	-7,31E-01	±	2,44E-02
Pro	-1,67E-02	±	4,82E-02	-1,80E-02	±	1,27E-02	-5,80E-02	±	6,41E-02	-4,06E-02	±	6,55E-02	-6,03E-02	±	5,97E-02
Pyr	0,00E+00	±	1,36E-01	0,00E+00	±	1,34E-01	-1,22E-01	±	1,26E-01	-1,36E-01	±	1,06E-01	-1,76E-01	±	0,00E+00
Ser	-2,99E-02	±	7,89E-02	-5,44E-02	±	5,38E-02	-1,42E-01	±	5,24E-02	-1,70E-01	±	8,33E-02	-9,84E-02	±	7,43E-02
Succ	1,69E-01	±	3,70E-02	1,55E-01	±	4,86E-02	8,47E-02	±	1,45E-03	2,57E-01	±	9,32E-02	2,31E-01	±	5,68E-02
Thr	-2,53E-02	±	1,46E-02	-2,27E-02	±	1,90E-02	-6,22E-03	±	2,26E-02	-1,63E-02	±	2,60E-02	-1,31E-02	±	2,18E-02
Trp	-1,90E-02	±	3,34E-03	-1,56E-02	±	3,34E-03	-9,62E-03	±	5,01E-03	0,00E+00	±	1,04E-02	-1,46E-02	±	5,61E-02
Tyr	-9,87E-02	±	2,75E-02	-1,09E-01	±	4,01E-02	-5,08E-03	±	3,01E-02	-6,61E-01	±	3,09E-02	-4,99E-05	±	2,24E-02
Val	-1,01E-02	±	5,28E-02	-7,31E-03	±	4,59E-02	-4,26E-02	±	2,50E-02	-3,24E-02	±	2,92E-02	-5,61E-02	±	4,31E-02
ATP	6,32E-02	±	0,00E+00	6,31E-02	±	6,73E-02	1,30E-01	±	7,89E-02	1,41E-01	±	7,81E-02	1,76E-01	±	8,76E-02
NADH	6,76E-02	±	1,87E-01	8,13E-02	±	1,81E-01	3,80E-01	±	1,36E-01	2,32E-01	±	1,04E-01	9,34E-02	±	2,53E-02
NADPH	8,00E-02	±	4,20E-02	5,78E-02	±	6,47E-02	8,33E-02	±	6,84E-02	9,74E-02	±	8,00E-02	1,08E-01	±	8,09E-02
FADH2	2,28E-01	±	6,52E-02	2,03E-01	±	9,15E-02	4,60E-01	±	3,09E-02	1,80E-01	±	1,19E-01	2,15E-01	±	1,21E-01