João Rodrigues Correia Ramos

Licenciado em Biotecnologia

Analysis of Metabolic Flux Distributions in Relation to the Extracellular Environment in Avian Cells

Dissertação para obtenção do Grau de Mestre em Biotecnologia

Orientador: Dr. Moritz von Stosch, investigator Post-doc, FCT-UNL

Co-orientadores: Dr. Rui M. Freitas Oliveira, Professor Associado, FCT-UNL

Júri:

Presidente: Dr. Pedro Miguel Calado Simões

Arguentes: Dra. Ana Margarida Palma Teixeira

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

June, 2015

Analysis of metabolic flux distribution in relation to extracellular environment in avian cells

To my father…

Acknowledgments

This leap in to an excitement future would not be possible without all the emotional and scientific support of many people who have accompanied me through these last years. This a like a dream coming true and I have many people to thank for it.

First, I would like to thank my advisor, Dr. Moritz Von Stosch, for the opportunity to work in this exciting field. I am very grateful for his encouragement, guidance and knowledge, which allowed me to fulfill my goals. Also, I am thankful to my co-advisor Dr. Rui Oliveira for all the help and insight during this work.

I would like to thank Fundação Lapa do Lobo, especially Dr. Carlos Torres, for believing in me all this time and of course for the precious scholarship that got me this far.

I am very thankful to my family and to my girlfriend, whose love and the continuous support during all these years, made all things possible in my life.

Finally yet importantly, a warm thanks to the people from Max Planck Institute for the data that made this thesis possible.

Having the data is not enough. I have to show it in ways people both enjoy and understand.

Hans Rosling

Abstract

Continuous cell lines that proliferate in chemically defined and simple media have been highly regarded as suitable alternatives for vaccine production. One such cell line is the AG1.CR.pIX avian cell line developed by PROBIOGEN. This cell line can be cultivated in a fully scalable suspension culture and adapted to grow in chemically defined, calf serum free, medium [1]–[5]. The medium composition and cultivation strategy are important factors for reaching high virus titers.

In this project, a series of computational methods was used to simulate the cell's response to different environments. The study is based on the metabolic model of the central metabolism proposed in [1]. In a first step, Metabolic Flux Analysis (MFA) was used along with measured uptake and secretion fluxes to estimate intracellular flux values. The network and data were found to be consistent. In a second step, Flux Balance Analysis (FBA) was performed to access the cell's biological objective. The objective that resulted in the best predicted results fit to the experimental data was the minimization of oxidative phosphorylation. Employing this objective, in the next step Flux Variability Analysis (FVA) was used to characterize the flux solution space. Furthermore, various scenarios, where a reaction deletion (elimination of the compound from the media) was simulated, were performed and the flux solution space for each scenario was calculated. Growth restrictions caused by essential and non-essential amino acids were accurately predicted. Fluxes related to the essential amino acids uptake and catabolism, the lipid synthesis and ATP production via TCA were found to be essential to exponential growth. Finally, the data gathered during the previous steps were analyzed using principal component analysis (PCA), in order to assess potential changes in the physiological state of the cell. Three metabolic states were found, which correspond to zero, partial and maximum biomass growth rate. Elimination of non-essential amino acids or pyruvate from the media showed no impact on the cell's assumed normal metabolic state.

Keywords: Avian cells, AG1.CR.pIX, Metabolic Flux Analysis, Flux Balance Analysis, Flux Variability Analysis, Principal Component Analysis.

Resumo

Culturas de células contínuas, capazes de proliferar em meios simples e definidos, são vistas como possíveis alternativa para produção de vacinas. Uma tal alternativa é a linhagem celular aviaria AG1.CR.pIX, recentemente desenvolvido pela PROBIOGEN. Estas células, além de crescer em suspensão, crescem num meio simples sem derivados de animais.

Neste projeto, vários métodos computacionais foram usados para simular a resposta destas células a diferentes meios. Este estudo é baseado no modelo de metabolismo central proposto em [1]. Numa primeira abordagem, *Metabolic Flux Analysis* (MFA) com os fluxos de consumo e de secreção foi aplicado para estimar os fluxos intracelular. A rede e os dados revelaram ser consistentes. Numa segunda fase, *Flux Balance Analysis* (FBA) foi implementado para aferir o objetivo biológico das células. O objetivo para o qual foi obtido uma melhor correlação entre os fluxos previstos com os experimentais foi a minimização da fosforilação oxidativa. Usando este objetivo, *Flux Variability Analysis* (FVA) foi implementado para obter a variabilidade dos fluxos. Além disso, este método foi aplicado a vários cenários onde a eliminação de uma reação (equivalente a eliminação de compostos do meio) foram simulados. As restrições causadas por aminoácidos essências e não essenciais foram corretamente previstos. Os fluxos relacionados com consumo e catabolismo de aminoácidos essenciais, síntese lipídica e produção de ATP via TCA revelaram-se como essências durante o crescimento exponencial. Por fim, os dados obtidos na etapa anterior foram analisados usando o *Principal Component Analysis* (PCA), para aferir sobre possíveis mudanças no estado fisiológico das células. Foram encontrados três estados metabólicos, correspondentes a zero, parcial e máximo crescimento celular. A eliminação de aminoácidos não essenciais ou do piruvato do meio não mostrou nenhum impacto no estado metabólico assumido como o normal para estas células.

Palavras-chave: células aviárias, AG1.CR.pIX, *Metabolic Flux Analysis*, *Flux Balance Analysis*, *Flux Variability Analysis*, *Principal Component Analysis*.

Contents

List of Figures

List of Tables

Acronyms

FBA → Flux Balance Analysis.

MFA → Metabolic Flux Analysis.

PCA → Principal Component Analysis.

mol → unity used in chemistry to express the amount 6.022×10^{23} atoms.

h^{-1} → hours.

µmol → micro moles (10^3 mol).

gDW → gram of dry weight.

1

1 Introduction

1.1 The avian cell line AG1.CR.pIX

1.1.1 Background

Since the 18th century vaccine research and development has been a constant focus of scientists, allowing novel therapeutic options. With the first vaccine, developed for smallpox dated from 1796, vaccine production has had since a great impact in human health and it is considered a turning point in human evolution. Nowadays, the problems with vaccine production are the high/varying demands and the production processes itself, which is cost intensive and has a relatively low efficiency. Science is driven by the search for a greater efficiency, which comprises cost and timeline minimizations for industrial process, such as vaccine production. Despite this fact, several viral vaccines, including the human influenza vaccine, are still being produced in primary cell lines such as embryonated eggs or chicken embryo fibroblasts [1]–[4]. Few continuous cell lines have been developed and are considered safe for the production of vaccines for humans (e.g. MDCK cells and Vero cells) [1]. However, these cell lines have many limitations, such as the low number of passages possible due to genetic instability and scalability [3]. Further, scale up is limited as the cultivation is done in adherent plates and the cells require calf serum supplement [1], [2]. Calf serum is a potential source of contamination and lot-to-lot changes, which leads to heterogeneous products. The heterogeneity, also caused by primary culture use, impairs the quality of the final product. As such, continuous cell lines that proliferate in chemically defined and simple media have been highly regarded as suitable alternatives.

One such potential cell line is a new avian cell line, AG1.CR.pIX, developed from sample tissue of Muscovy duck around 2003 by PROBIOGEN [1], [4]. Duck was chosen because it had been found to have no endogenous avian retrovirus (EAV) or endogenous avian leucosis virus (ELV-E) [4], which is a very important indicator for vaccine safety. Furthermore, duck eggs are available from monitored stocks that are free of pathogen. This new cell line resulted from tissue samples of duck embryo, previously tested for various pathogens. The immortalization was performed with adenovirus type 5 E1 genes as these virus are considered to be ubiquitous and therefore safe for human [2]. The transfected genes were E1A and E1B, originated from adenovirus. The transfected gene products have been shown to promote cell cycle progression and interfere with the p53 transcription factor, which persistent activation had been correlated to apoptosis, therefore allowing immortalization [2]. This immortalization has been shown to be stable after a high number of passage (several years) [3]. Further details can be found in [3], [4].

This avian cell line has been shown to be robbust and support viral replication such as modified ankara virus (MVA) and influenza virus at titers equal or higher than current pharmaceutical processes [5], [1]–[4]. This cell line could not only be a substitute for current viral vaccine production, but also even be adapted to produce future viral outbreak vaccines or the production of viral vectors. Due to its interesting characteristics, detailed studies of this cell line are of importance. Others aspects of interest are that this cell line can be cultivated in a fully scalable suspension culture and that it can be adapted to grow in chemically-defined, calf serum free, medium [5], [1]–[4], i.e.: no growth hormones or complex animal serum are needed. A suitable media composition, at which high cell density and high virus titer has been obtained included glucose, pyruvate and amino acids [5], [1]–[3]. As this cell line proliferates in suspension media, cultivations in stirred tank reactors have been shown to be suitable [5], [1]–[3], though different reactors and operation conditions such as alternating tangential flow perfusion (ATF) have been shown to increment virus titers and facilitate downstream processes [6]. However, the cellular requirements for growth are still largely unknown and detailed studies of the intracellular processes are largely missing. In this thesis, flux distributions are studied in relation to certain extracellular environments, more specifically it is analyzed how eliminations of certain medium compounds impact on biomass growth and the physiological state of the cell. Various methods were used to estimate intracellular flux distributions, assess the cell's objective and predict the physiological state changes in relation to the extracellular environments / compounds in the medium. For these tasks, a metabolic network model is required and the recently published model of Lohr et all [1] was used.

1.1.2 The CR.pIX metabolic model

In this thesis, a metabolic model of central metabolism for the avian cell line CR.pIX, proposed by Lohr et al in [1] was used. A metabolic model is the aggregate of all the reactions assumed or verified to occur in any given cell [7]. A reaction constitutes a sequence of probable and observable steps between a set of input and output of metabolites [7]. The overall cellular reactions result in a conversion of substrates into free energy and a large set of metabolites, including precursor metabolites and building blocks for macromolecular synthesis [7]. This set of metabolites is divided in intracellular metabolites, metabolic product or more complex products such as secondary metabolite [7]. Macromolecular pools usually falls in the category constituents of biomass [7]. In this context, the following formalisms are used:

- Substrate as a compound present in the medium, which is taken up and directly incorporated or further metabolized by the cell. The substrate is a broad category, ranging from carbon, nitrogen and various minerals source, essentials for cell function [7]. In our case, the substrates are glucose, pyruvate and amino acids (found in **Table 7.1**, appendix).
- Metabolic product as a compound produced by cell that is excreted to extracellular medium as a result of primary or secondary metabolism [7]. In the present case, the metabolic products are lactate, ammonia, uric acid and carbon dioxide.
- Intracellular metabolite as all other class of metabolites that is found inside that cell, including intermediary and building blocks [7]. Example of this is glycolytic, TCA cycle and amino acid catabolism intermediates (see list in **Table 7.2**, appendix).
- Flux, as the reaction rate or speed at which a set of reactants is converted into products.

The cells metabolic network is obtained from the integration of the reactions' stoichiometries. Stoichiometry is based on the law of mass conservation, where the total mass of reactants is equal to the total mass of products. Typically, the amount of product and reactants are defined in ratios of positive integers, as follows:

Equation 1: A general chemical reaction.

$$aA + bB + ... \longrightarrow ... + yY + zZ$$

For the reactants A and B the numbers a and b are known as the stoichiometric numbers. The same principle applies to y and z, making them the stoichiometric numbers of products Y and Z, respectively.

As stated before a metabolic model includes a set of reactions, and the stoichiometry of each reactions is very important in modelling. In this case, the metabolic model used comprised 97 reactions and 72 metabolites. The pathways included in the model are TCA cycle, glycolysis, pentose pathway, anaplerosis, amino acid uptake and catabolism, transport reactions, MTHF & uric acid synthesis, lipid synthesis, metabolic product release and biomass production.

In **Figure 1.1**, the metabolic model of central metabolism for the CR.pIX proposed by Lohr et all in [5], is presented.

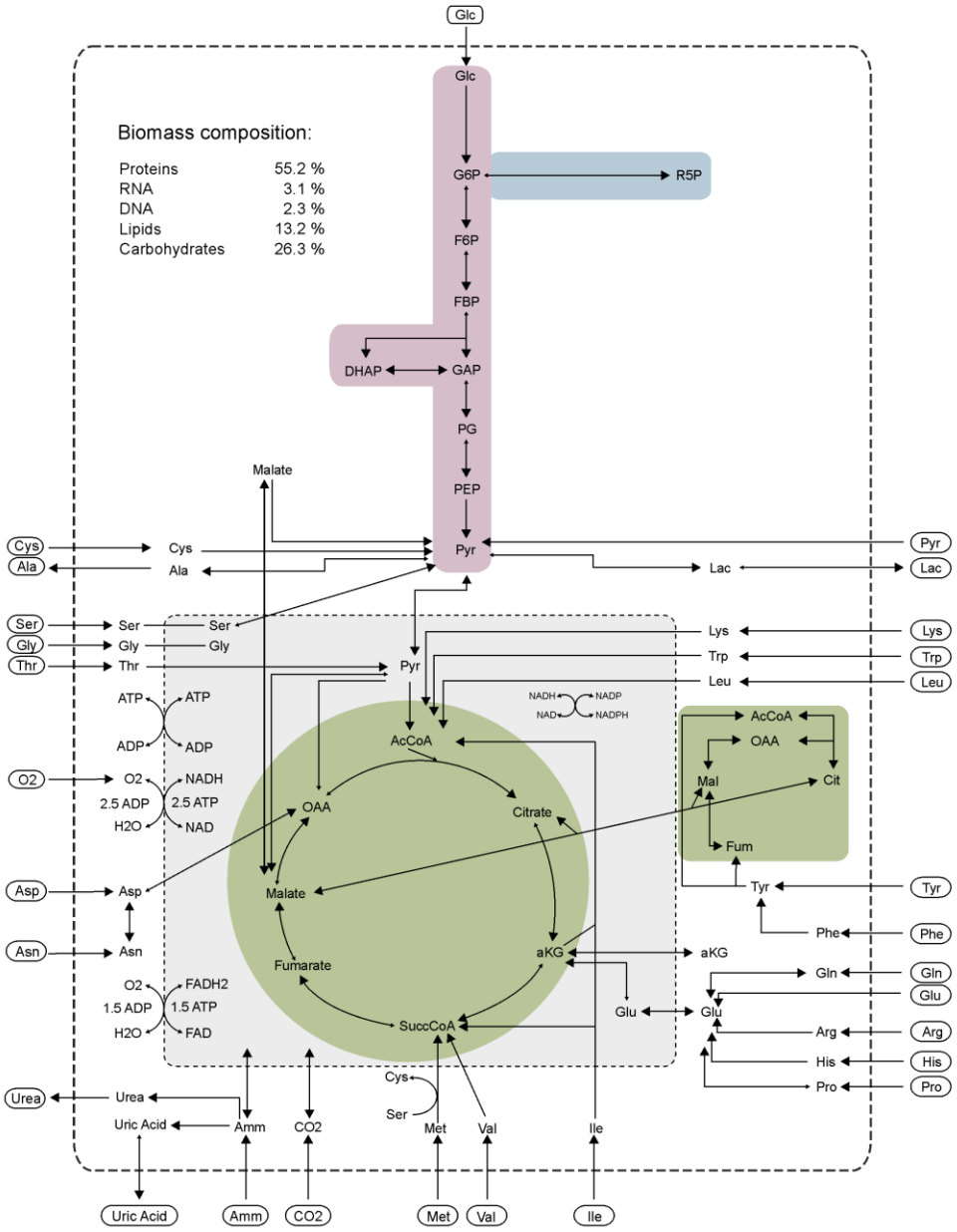

***Figure 1.1**: The CR.pIX central metabolic model, adopted from Lohr et al* [1] .

Detailed reactions included in the metabolic network model are presented in **Table 8.2**, Annex.

1.2 Objectives

1.2.1 General objectives

The main goal of this thesis is to understand which and how certain compounds that are present in the media impact the biomass growth and the physiological state of the cell. Knowledge about this can support the development of optimal media, in which the biomass growth rate is maximized. It can also help to avoid depletion or inhibition when aiming for very high cell densities, a common process goal, since higher biomass concentrations in principle result into higher virus titer.

1.2.2 Specific objectives

The specific objectives in this thesis included:

1. Validate the model of central metabolism
   1.1. Determine the extracellular fluxes and the intracellular fluxes;
   1.2. Monte Carlos Sampling to determine all the fluxes standard deviations;
   1.3. Determine model consistency, using the standard deviations.
2. Use FBA to determine what the objective of the cell is.
3. Apply FVA to study how certain reaction eliminations (tantamount to eliminations of media compounds) affect the flux distribution obtained by FBA, implicitly elucidating the impact on cell growth and the most important reactions related to growth.
4. Use PCA to analyze and compare the cell metabolic states predicted by FVA.

2

2 Methods

2.1 Cell culture and sampling

The experimental data used in this study were obtained by Lohr et al in [1], as a result from a batch experiment in 1 L tank reactor well stirred and aerated. Along the process, triplicate samples were taken for off-line analytic analysis. The medium concentration of various substrates and metabolic product as well a viable cell number were measured. The analytic methods used for each concentration measurement are provided in (**Table 9.1**, annex). The biomass composition and specific dry weight were also determined experimentally. The detailed experimental and analytic procedures are also described in [5], [1].

2.2 Calculation of the uptake and secretion fluxes

2.2.1 Calculation of the fluxes

A differential approach was used to calculate the flux values [8]. Firstly, cubic smoothing spline functions were formulated for the time profiles of the concentrations based on the experimental data for each concentration C_i, including biomass. The smoothing spline is an interpolation method, where a function that simulates a continuous concentration variation is calculated. The function that is minimized to obtain the smoothing spline f is:

Equation 2

$$p\sum_{j=1}^{n} \mathrm{w(j)}\left|C_i(\mathrm{j}) - \mathrm{f}\left(t(\mathrm{j})\right)\right|^2 + (1-p)\int \beta(x)\left|\mathrm{D}^2\ f(x)\right|^2 dx$$

Where p is a parameter called smooth factor and belongs to the interval [0,1], and j = 1,…,n are the number of points at which concentration data are available. The variable t represents the time at of each entry j, while w represent the weight of each point $C_{i,j}$, the default value for the weight is one. The variable β is the piecewise constant weight function, and in the presented cases has a constant value of one. Finally, the D^2f denotes the second derivative of f.

With a smooth factor of one (p=1), a smooth curve which goes through every point in a given data set is calculated, whereas for a smooth factor of zero (p=0), a linear curve fit is calculated.

In the following table, **Table 2.1**, the p values used in this thesis for the creation of the cubic smoothing spline are presented.

***Table 2.1**: Table with p values used for the creation of the cubic smoothing spline for each concentration.*

	p value
r1, Glc	1E-03
r2, Pyr	5E-04
r4, Gln	1E-05
r5, Glu	1E-02
r7,Asp	1E-04
r8,Arg	1E-05
r9,Asn	9E-05
r10, Cys	1E+00
r11,Gly	1E-05
r12,His	1E-04
r13,Ile	1E-05
r14, Leu	1E-04
r15,Lys	1E-04
r16, Val	9E-05
r17, Met	1E-04
r18, Phe	1E-04
r19, Pro	5E-04
r20, Ser	1E-04
r21, Thr	9E-05
r22,Trp	1E-04
r23, Tyr	1E-04
r78, Lac	7E-03
r82, Amm	1E-02
r79, Ala_{out}	1E-03
r80, UricAcid	1E00
r97, Biomass	1E00

After the spline of each concentration has been created, the smoothing spline function was differentiated with respect to the time, whereby functions for the time gradients of substrate uptake and metabolic product releases are obtained. Finally, the homogenous material balances for the reactor were used to calculate the variations of the specific rates (V_i, µmol/gDW/h) along time, where the units of biomass concentration are in dry weight. The glutamine concentration variation was found to be due to chemical decomposition to ammonia and not by cellular uptake as described in [8], [9]. The rate at which this process occurs is a first order reaction with k_{Gln} = 0.0032 h^{-1} obtained from [1]. Based on this, the glutamine and ammonia rates were corrected. The material balance read:

Equation 3

$$\frac{dX}{dt} = \mu X$$
$$\frac{dc_i}{dt} = v_i X$$
$$\frac{dc_{Gln}}{dt} = v_{Gln} X - k_{Gln}.Gln$$
$$\frac{dc_{Amm}}{dt} = v_{Amm} X + k_{Gln}.Gln$$

Where X is the viable cell concentration, t is time and µ is the specific biomass growth rate. The variable v_i is the specific rate for each compound (i).

2.2.2 Monte Carlo Sampling for the calculation of the flux standard deviation

Monte-Carlo sampling, based on the measured concentrations (metabolites and biomass) measurements' standard deviations as boundaries, was performed in order to calculate standard deviations for each flux. The latter was achieved by generating 1000 Monte Carlo sample for each measurement of each concentration. Then 1000 cubic smoothing splines for each concentration were created with the p-values described above. From these smoothing spline 1000 profiles of the uptake, release and biomass growth rates were calculated. The standard deviations for each time-point of each flux were calculated from the 1000 generated flux values, which yields the standard deviation of each flux at each time point.

2.3 Constraint Based Models - Methods for determination and analysis of the cellular flux distribution

Constraint based methods are yield from the homogenous material balances of the intracellular metabolites:

Equation 4

$$\frac{dc_{\text{int}}}{dt} = S.v - \mu.c_{\text{int}}$$

assuming a pseudo-steady-state for all intracellular metabolites, meaning that no metabolite accumulation occurs, results in:

Equation 5

$$0 = S.v - \mu.c_{\text{int}}$$

This assumption is usually valid because enzymatic reactions rates take milliseconds, which is very small compared to the biomass growth rate (up to days) [7]. The second term on the right hand side is typically much smaller than the reaction rate values (v), wherefore it is typically neglected. Both assumptions simplify metabolic processes analysis significantly [7]. Detailed description of this assumption can be found in [7]. Furthermore, no kinetics parameters need to be determined for these model applications [10].

Constraint based models can be used for the determination, estimation, prediction and analysis of flux distributions in metabolic networks. These methods allow comprehensive studies of interactions of the pathways in order to identify factors responsible for the control of the overall metabolism [7]. Due to the complexity of biological system, various mathematical methods have been developed in metabolic engineering [11], [12]. These methods are essential for rational metabolic flux modification in order to achieve a certain goal or other application such as medium optimization.

2.3.1 Metabolic Flux Analysis

Metabolic Flux Analysis (MFA) is a constraint based mathematical method used in metabolic modelling to estimate unknown flux distributions for a given metabolic model or network [7], [13]–[15]. It is widely used to estimate intracellular flux distributions of a metabolic network,

provided that a minimum number of reaction rates are known. MFA does not include kinetic enzymatic parameter of reactions as restrictions. This means incorrect, flux estimations can be observed when compared to in vivo. MFA is based on the pseudo-steady-state assumption, this means that the produced metabolites are consumed at the same rate as they are being produced [7], [10]. This method consists on solving the linear equation problem which consists in the stoichiometric matrix (S) multiplication to the reaction rates vector (v). With pseudo-steady state assumption and neglected growth rate dilution the system reads:

$$0 = S.v$$

It is clear that a system of linear equations must be solved in order to obtain each reaction rate (v_i). The linear equation system is typically underdetermined as there are more unknown variables than the number of linear equations, i.e. of higher number of reactions than the number of metabolites present in the metabolic model [7]. In another words, in order to solve the linear system a minimum of reaction rates must be determined experimentally. This means that MFA is driven by experimental data, and degree of confidence of the results varies according to the reliability of the latter [7].

With some reaction rates previously determined the linear system can be separated into known ($v_{k,i}$) and unknown reaction rates ($v_{n,i}$).

Equation 6

$$0 = S.v$$
$$0 = S_n.v_n + S_k.\mathrm{v}_k$$
$$S_n.v_n = S_k.\mathrm{v}_k$$
$$v_n = -S_n^{-1}.S_k.\mathrm{v}_k \text{ ,if } S_n \textit{ is square}$$
$$v_n = -S_n^{\#}.S_k.\mathrm{v}_k \text{ ,if } S_n \textit{ is not square}$$
$$S_n^{\#} \textit{ is Moore} - \textit{Penrose pseudoinverse}$$
$$v_k \textit{ is the known reactions rates}$$
$$v_n \textit{ is the unknown reactions rates}$$

Depending on the number of known reaction, the resulting linear system can fall into the following categories, as described in [16].

Determinacy:

- Underdetermined: rank(S_n) < u [a], implying there are not enough linearly independent constraints for computing all rates of v_n uniquely.

- Determined: rank(S_n) ≥ u [a], implying that there are enough linearly independent constraints for computing all rates of v_n uniquely.

Redundancy:

- Redundant: rank(S_n) < m [b], some rows of S_n can be expressed as linear combinations of other rows. In this case, methods for consistency check are applied to verify (and eventually improve) the accuracy of the calculated v_n.
- Not redundant: rank(S_n) = m [b], system consistent by itself because no dependent rows exist in S_n.

[a] u = number of unknown rates; [b] m = number of metabolites.

For any given redundant linear system a redundancy matrix can be obtained by the following equation, as described in [7], [16].

Equation 7

$$R = S_k - S_n.S_n^{-1}.S_k \text{ , if } S_n \textit{ is square}$$
$$R = S_k - S_n.S_n^{\#}.S_k \text{ , if } S_n \textit{ is not square}$$

A consistent system fulfils the following conditions, as described in [7], [16],

Equation 8

$$0 = R.v_k$$

While an inconsistent system does not fulfill these conditions. Consistency can be assessed by calculating the ratio between the calculated error of non-balanced metabolites and their estimated variance and comparing the value of this ratio to that of a Chi-square distribution with the same number of degrees of freedom (rank(R)) for a specified confidence value (in this theses .95), for further details see [7]. In case of consistency, detailed procedures to improve the estimated rates (v_n) were applied, as well described in more detail in [7].

2.3.2 Flux Balance Analysis

Flux Balance Analysis (FBA) is a mathematical optimization methodology vastly used to predict gene activity in genome scale network analysis and flux distribution through the pathways of a metabolic model [17]–[20]. This method is based on the pseudo-steady-state assumption and

is widely used since no kinetic parameter or a minimum number of experimental flux values are required. FBA consists on solving a system linear equations problem.

Given a space of possible biological reactions rates (Upper and Lower boundaries), FBA is used to predict the optimal flux distribution through each pathway given an assumed cellular objective function. This concept is based on the idea that the evolutionary pressure enabled the cells to redirect the metabolic flux according to a certain goal or global objective [21]. It has been shown that the maximization of the cellular growth rate in FBA many times yields flux distributions, which are comparable to measured ones, and it is assumed that this may be due to evolutionary pressure that allowed cells to grow as fast as possible [10], [22]. However, since FBA does not include thermodynamic constraints, e.g. enzymatic parameters, the predicted optimal reactions rate for reactions other than biomass growth rate is usually considerable different from experimental data. In eukaryotic cells it is particularly harder to obtain meaningful physiological steady state with this method, given the more complex cellular behavior and the little information that is used in the predictions [23].

A critical point for the application of this method is the determination of the cell's biological objective. The objective for which the best fit to the experimental data is obtained using FBA is usually assumed to be the "true" cellular objective.

The following describes how FBA is applied.

Equation 9

$$Objective,$$

$$\max_{v}(Z) = \sum_{i=1}^{\overrightarrow{q}} w_i.\mathrm{v}_i, \; subject \; to$$

$$0 = S.v$$

$$v_l \le v_i \le v_b$$

$$v_l = 0 \;\; \text{reaction } i \; irreversible$$

$$v_l < 0 \;\; \text{reaction } i \; reversible$$

In objective function, the vector w has the same size as reaction vector (v) and it contains positive or negative entries, according to maximization or minimization of the corresponding reaction. S is the stoichiometric matrix.

In the following Figure, a visual demonstration of FBA for a given metabolic model is provided.

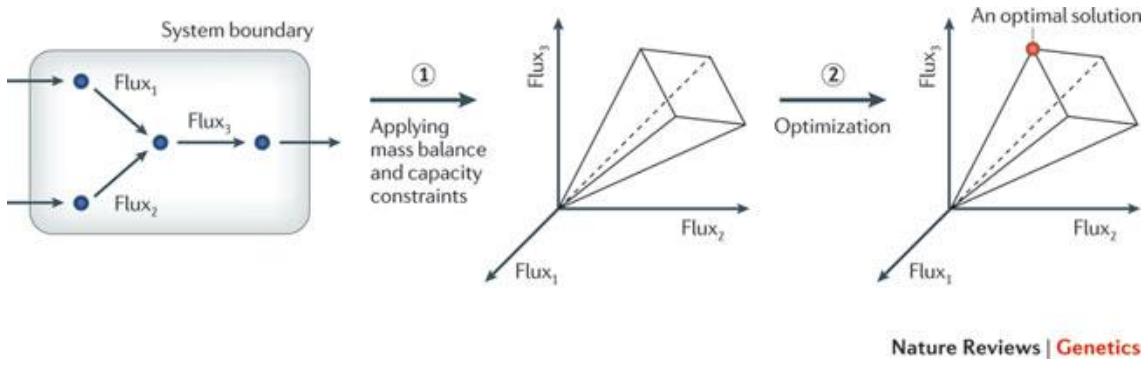

Figure 2.1*: Example of FBA applied to a metabolic network, adapted with permission from Macmillan Publishers Ltd: Systems-biology approaches for predicting genomic evolution from* [23]*, copyright 2011.*

In **Figure 2.1** it is seen that given a set of reactions, a defined space of solution, FBA is then applied to obtain an optimal solution within that space.

2.3.3 Flux Variability Analysis

Flux Variability Analysis is a mathematical optimization methodology used to characterize the solution space of reactions rates in a metabolic model or gene regulatory network, using a set of constraints [17]–[19]. Since the system S*v=0 is typically underdetermined, no unique solution can be obtained by FBA [13], [24]–[26] and FVA can be used to assess the variability of the flux distribution. Other applications include the study of flux distribution under suboptimal conditions, optimization of medium compositions, metabolic flexibility and so on [24].

The following describes the implementation of FVA.

Equation 10

$$\max_v / \min_v \quad \left\{ c^T v \right\}$$
$$0 = S.v$$
$$v_l \le v_i \le v_b$$
$$v_l = 0 \text{ reaction } i \textit{ irreversible}$$
$$v_l < 0 \text{ reaction } i \textit{ reversible}$$

Where **S** is the stoichiometric matrix with m metabolites and n reactions and c is a vector with the linear objective. The variable **v** represents the fluxes.

In metabolic engineering, FVA is essential in the analysis of cellular metabolic flexibility as it allows the determination of the boundaries of each flux for different environmental conditions or mutations [27].

2.3.4 Principal Component Analysis

Principal Component Analysis (PCA) is statistical procedure that uses orthogonal transformation to convert a set of possible correlated data into a set of linearly uncorrelated data [26], [28].

In PCA a matrix of data, X, is decomposed into matrices of loadings W and scores T such that a maximum amount of variance of the data is captured in an underlying latent space for a specified number of latent variables or number of components (p).

Equation 11

$$X = WT;$$
$$X \text{ is } \mathrm{m} \times \mathrm{n}$$
$$W \text{ is } \mathrm{m} \times p$$
$$T \text{ is } 1 \times p$$
$$p \text{ is the number of components}$$

The scores describe the patterns of the data in the underlying (orthogonal) latent space and the loadings described the relation between the latent space and the patterns of the data. The PCA loadings are determined in an iterative procedure from the data, X, such that for each latent variable a maximum of variance of the data can be captured.

This transformation is essential to analyze high dimensional or large experimental data sets and identify correlations [28]. The PCA decomposition highlights the small number of flux whose variability accounts for almost all the other observed flux variability [26]. This means that small numbers of reactions are determinant for a certain cellular metabolic state. In metabolic engineering, PCA has been shown to highlight the most active reactions through a metabolic genomic network [28]. Furthermore, the PCA scores plotting has been shown to cluster similar metabolic states together [28] and comprehensive analysis of the loadings provides clues about regulation mechanism behind the flux variability.

3

3 Results and discussion

3.1 Cell culture

3.1.1 Exponential biomass growth phase

The overall cell metabolism dynamics have an impact on the biomass growth [7]. As such, the quantification of the biomass concentration is essential for metabolic flux determination and analysis [10].

The experimental data used in this thesis were obtained through a batch experiment in 1 L stirred tank reactor well stirred and aerated. Along the process, triplicate samples were taken for off-line analytic analysis. The biomass composition and specific dry weight were determined as described elsewhere [1]. The detailed experimental and analytic procedures are also described in [1]. In **Fig. 3.1**, the measured viable biomass concentration is presented in relation to the time for the exponential growth phase.

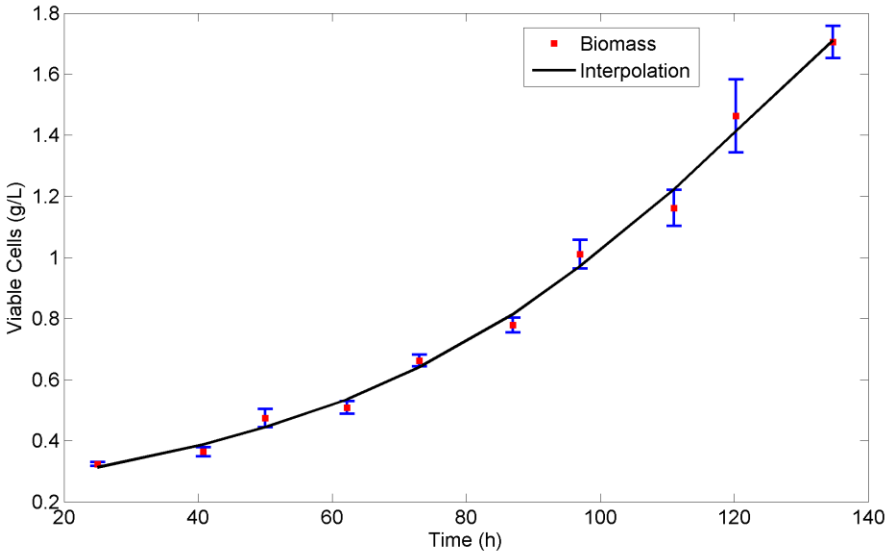

***Figure 3.1**: Biomass exponential growth curve.*

It can be seen that the biomass concentration increases, where biomass duplication takes up to two days.

Metabolic Flux Analysis (MFA) is based on a quasi-steady state assumption, meaning that the intracellular metabolite concentrations are almost constant along time [7]. Due to this assumption, the considered time frame of the biomass exponential growth phase for MFA and flux variability analysis (FVA) is from 41-70 hours (R^2 is 99.999%). During this phase, it is very likely that the quasi steady state assumption holds [1], [29]. Furthermore, this is the typical growth period before the infection in virus production processes [1], [3].

In **Fig. 3.2**, the measurements of the viable biomass concentration are shown in logarithmic scale for the exponential phase as well as a linear regression model. It can be seen that in this phase, the corresponding value of the growth rate is 0.0152 h^{-1}, which corresponds to a biomass duplication time of 45h and 35 minutes.

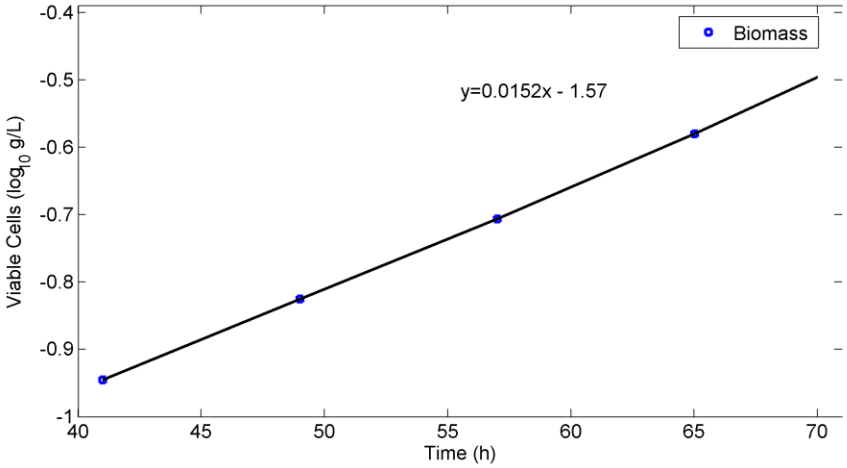

***Figure 3.2**: Measurements and linear regression model of the biomass concentration in logarithmic scale over time for the exponential growth phase.*

3.1.2 Substrates and metabolic products

In order to validate the metabolic model with as much information as possible and to minimize the errors caused by the experimental variations (peaks), data interpolation was performed. The interpolation should on the one side approximate the data well and be within the standard deviation of each measurements, but on the hand be rather smooth without too much low frequency variations. The interpolation is a function that simulates a continuous concentration variation. In this case, the cubic smoothing spline interpolation was performed using the csaps function available in MATLAB®. For each concentration the following functions was defined:

Equation 12

$$F(c) = csaps([t_1\ t_2\ \ldots\ t_n],[c_1\ c_2\ \ldots\ c_n],p)$$

Where t_i is time, c_i is the concentration and p is the smooth factor, and varies between zero and one.

In **Figure 3.3**, typical variation in extracellular substrate and metabolic product concentrations, as were observed in the process, are presented.

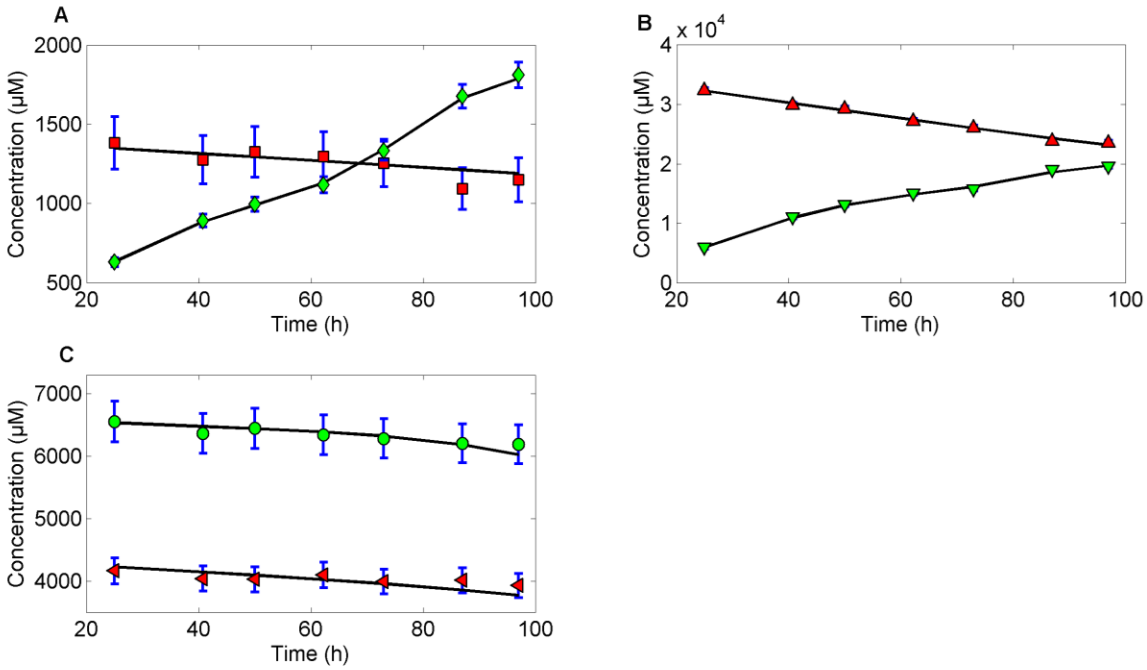

Figure 3.3: *Typical variations in substrate uptake and metabolic product formation for CR.pIX cells cultivation. Black line: interpolation.* ***A****: extracellular concentration of glutamine (■) and ammonia (♦).* ***B****: extracellular concentration of glucose (▲) and lactate (▼).* ***C****: extracellular concentration of serine (●) and glycine (◀).*

From the analysis of these extracellular concentrations' variations important conclusions can be drawn. First, the ammonia concentration is very low (**Fig. 3.3A**) when compared to other cell lines, which is very beneficial as ammonia release inhibits cell growth and has a negative impact in product formation [1]. Secondly, the glutamine uptake is zero and the variation in the media concentration is accounted for by glutamine hydrolyses.

A high glucose consumption can also be observed, in **Fig. 3.3B**, followed by lactate production and release into the medium. Most likely, almost all of the glucose is converted into lactate even in a well-oxygenated reactor. This is usually termed as aerobic glycolysis, and it is frequently found in cancer cells, although it is an inefficient way to produce ATP [30], [31]. In fact a high aerobic glycolysis is often assumed as a requirement for the increment growth rate, although it is not clear which advantageous it confers to the cancer cells [30], [31]. From a biological point of view the CR.pIX is indeed an abnormal or "cancer" cell as it was immortalized [3]. These metabolic changes can be explained by the fact that transformed cells, like CR.pIX, undergo specific metabolic reprogramming in order to support proliferation [32], [33].

Another key difference is glycine uptake by the CR.pIX, which is contrary to other cell lines [1]. For the presented experimental conditions, glycine might be essential for avian cells as reported in [1], [34]. Another explanation may be that glycine is taken up by rapidly proliferating

both, normal and abnormal cells [33]. Another amino acid with such importance to this cells is serine [32], for which the uptake was observed to be above average. In fact glycine and serine uptakes have been found to be biosynthetically linked [32].

3.2 Metabolic Flux Analysis

Metabolic Flux Analysis (MFA) is a constraint based mathematical method used in metabolic modelling to estimate unknown flux rates distribution for a given metabolic model or network [7]. It is based on the pseudo-steady-state.

The first stage of the work consisted in testing the model previously developed by another group in order access weather it can explain the experimental data. To do this MFA was performed to estimate intracellular flux distribution, followed by a consistency check of the results.

3.2.1 The determined Fluxes

The metabolic model tested in this thesis consists of 97 reactions and 72 metabolites, and the stoichiometric matrix rank is 69. A determined system, in this case, requires a number of at least 28 measured fluxes (97-69). In the present case, these constraints are the uptake rates of substrate and metabolic product formation in the avian cell culture. Other constraints used were the growth rate and the pyruvate carboxylase (r46) which can be set to zero because the enzyme was found to be inactive in [1]. Additional constraints rose from the fact that the same reaction cannot be active both ways, i.e. either uptake or release is observed.

As described in the previous chapter, the variation of each flux, i.e. substrate uptake or metabolic product release, was calculated. Furthermore, Monte Carlo sampling was performed for the standard deviation calculation for each calculated flux.

In **Table 3.1**, the determined fluxes mean and standard deviation is presented calculated for the time frame 41-70h.

***Table 3.1**: The determined extracellular fluxes (41-70h) for the avian cells.*

	Fluxes (µmol/gDW/h)
r1, Glc	-270.6 ± 27.9[a]
r2, Pyr	-42.8 ± 2.0
r4, Gln	-4.0 ± 0.6
r5, Glu	-22.0 ± 3.6
r7,Asp	-18.4 ± 4.6
r8,Arg	-7.8 ± 4.2
r9,Asn	-10.7 ± 10.1
r10, Cys	0.0 ± 0.0
r11,Gly	-3.4 ± 4.3
r12,His	-2.2 ± 1.2
r13,Ile	-6.5 ± 3.9
r14, Leu	-7.6 ± 5.8
r15,Lys	-2.1 ± 3.1
r16, Val	-4.0 ± 3.8
r17, Met	-4.9 ± 2.6
r18, Phe	-2.6 ± 1.7
r19, Pro	0.0 ± 0.6
r20, Ser	-7.8 ± 9.6
r21, Thr	-2.8 ± 2.9
r22,Trp	-0.2 ± 0.4
r23, Tyr	-1.7 ± 2.1
r78, Lac	355.4 ± 38.1
r82, Amm	22.2 ± 1.8
r79, Ala_{out}	34.2 ± 17.2
r80, $UricAcid_{out}$	0.0 ± 0.0

[a]Standard deviation calculated from Monte Carlo sampling using experimental measure errors

The analysis of the data presented in **Table 3.1** reveals that generally, the measured reaction rate values partly agree with studies on other cell lines, having the same direction (uptake and release) [1]. The uptake rates depend mainly on two factors: the concentration of the substrate in the environment and the cells metabolic needs. In addition, the uptake rates tend to diminish overtime as the substrate media concentrations diminish. On the other hand, the formation rates of the released product depend on the uptake rates and by default on the cells metabolic state. This means that the intracellular flux distributions are correlated with the uptake and release rates, which might render the identification of the cell system, e.g. via MFA, possible. The cells needs/ state changes over time and environment, which has a strong influence on the cell cycle [35].

3.2.2 Estimated intracellular fluxes

Using equation (6), previously shown, the substrate uptake and product formation rates and the other mentioned constraints, MFA was performed in order to estimate the intracellular flux distribution.

In **Figure 3.4**, a general overview of both intracellular and extracellular reactions rates variation along the chosen time frame is presented.

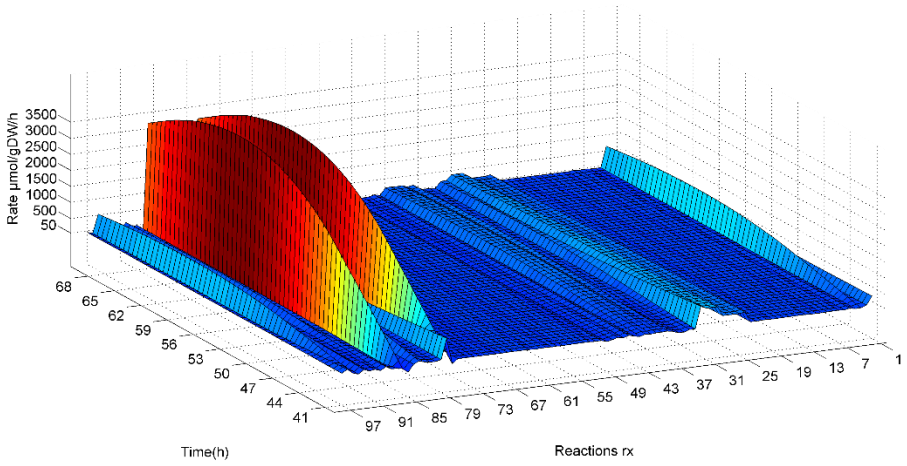

***Figure 3.4**: Metabolic flux distribution in CR.pIX.*

The flux distribution obtained by the MFA method is one solution from a space of feasible solution, given a certain values for a certain number of fluxes. As stated before, each flux value depends of many factors, such as initial concentrations of substrates. What we are interested in is how the intracellular fluxes change overtime and how the relations between those and the uptake/release rates changes. The IDs of the 97 reactions are provided in **Table 7.2** (see appendices). Two of the greatest reaction rate values (see **Figure 3.4**) correspond to ATP transport (r86) and ATP consumption for maintenance and for other futile reactions (r77). Clearly, the model does not consider other ATP consuming reactions. Also other cell lines like MCDK have been observed to have the same or greater consumption of ATP through maintenance or other futile reactions [36].

Another aspect of the CR.pIX metabolism is that the uptaken glucose (via r1) seems to be almost all converted into pyruvate (r31) and then to lactate (r32). The amount of pyruvate, which is estimated to enter into the TCA cycle (r33), is much lower than the flux to lactate, implying that the cell in not performing cellular respiration, although it would have sufficient oxygen to perform this task. The amount of ATP generated in this case of lactate production is lower than with oxidative phosphorylation (2 versus 36), but the rate value of the ATP generation is greater [37]. This is in fact, not something new and for cancer cells it is termed aerobic glycolysis or Warburg effect [30], [31], but could apply also to the analyzed cell as discussed above.

On the other hand, levels of the estimated oxygen consumption (r83) and carbon dioxide release rates are also high. The respiration quotient (RQ= CO_2/ O_2) is 1.06, which is near unity. This is what is expected for a culture supplied with a carbohydrates diet [38], meaning that the flux distribution estimated by MFA might be correct in most of the cases.

3.2.3 Consistency check

The consistency test of a model is defined as the statistical evaluation of the consistency of the data with the assumed biochemistry [7].This tests how consistent the measurements are in relation to the assumed network using the redundancy of the measured fluxes. In case of consistency the accuracy of the intracellular flux estimation can be increased through least square calculation [7]. Since each concentration measurement has an associated standard deviation, Monte Carlo sampling was performed, allowing standard deviation estimation for each flux. In this case, 1000 Monte Carlo samplings were performed for each concentration profile, where each concentration value is varied using the standard deviation of each of the measurements. The chi-square distribution test with a degree of freedom equal to number of redundant measurements can be used to do the consistency check [7]. This test is easy to perform and it can be used to see if there is a significant level of difference between two or more sets of data, given a confidence level. The system we used has two degree of redundancy. The problem then reads $h<X^2(0.95, 2)$. In a consistent model, the value of hypothesis has to be lower than the X^2 test at 95% confidence level for the two degrees of freedom. In **Figure 3.5**, results of the consistency are presented.

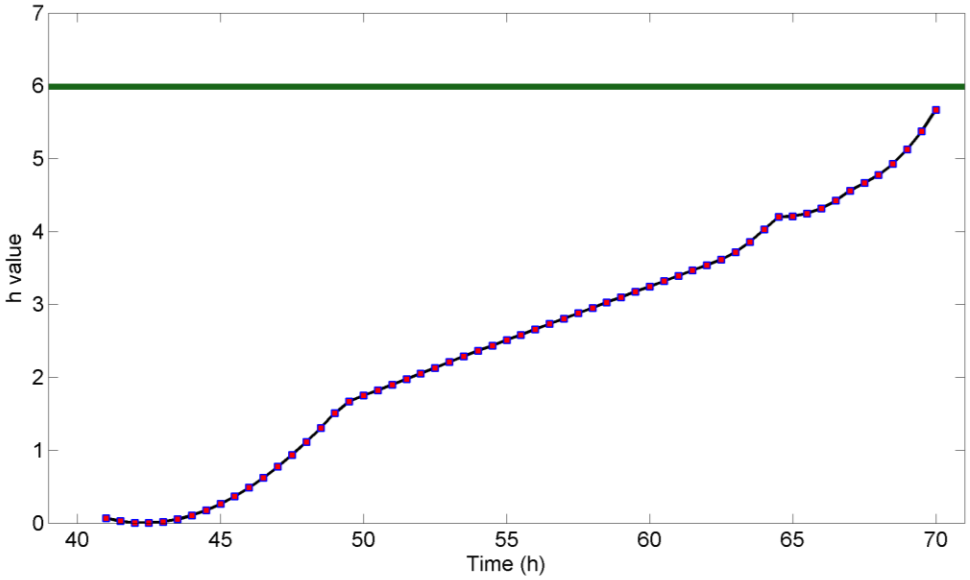

***Figure 3.5**: Consistency check results: h-value over time. Test hypostasis $X^2(0.95, 2)$ (—).*

It can be seen that the consistency (h) of the model estimates along time is lower than the test hypothesis ($X^2(0.95, 2) = 5.99$), in the considered period. This means the model can explain the flux data [1], [7] and the solution can be considered satisfactory or suitable. On the other hand, the consistency diminishes over time. This could probably mean that although the cells are exponentially growing, the steady state assumption might tend not to hold, when higher cellular concentration is reached [31]. This is most likely due to differential metabolic states, caused by cellular agglomerations or imperfect medium mix. Another explanation for this fact might be that the network used might not fully describe the cells metabolic behavior.

3.2.3.1 Influence of the Error measures in the consistency

The consistency check derives from the standard deviation, which is calculated from the measurements errors. The consistency check can allow the identification of the inconsistencies in the model or which standard deviations have the most impact on the overall system consistency [7]. In the presented case, the consistency changes overtime. Therefore, we assessed which influence each standard deviation of each measured rate (known rates) has on the system consistency. In order to calculate the contribution of each standard deviation of the measured compound the following equation was derived:

Equation 13: Contribution of experimentally measured compounds standard deviation to the consistency of the solutions.

$$n.std_i \begin{cases} h_{new} = h_{initial} + \frac{1}{n}.\beta_i.h_{initial} \; ;n =]0,1[\\ h_{new} = h_{initial} - n.\beta_i.h_{initial} \; ;n > 1 \end{cases}$$

$$contribution_i = \beta_i.h_{initial}$$

If we multiply the standard deviation (std) of a specific substrate or metabolic product (i) by a factor (n) and recalculate the system consistency (h_{new}), then we find a coefficient of contribution of each compound standard deviation (β_i) on the overall system consistency. The contribution of each compound is then the coefficient of contribution multiplied by the calculated system consistency.

Since the standard deviation for each compound (**std_i**) and the overall system h-value (**$h_{initial}$**) changes over time, the coefficient of contribution also changes slightly over time.

The mean for each estimated contribution coefficient for substrate uptake or metabolic product formation is presented in **Figure 3.6**.

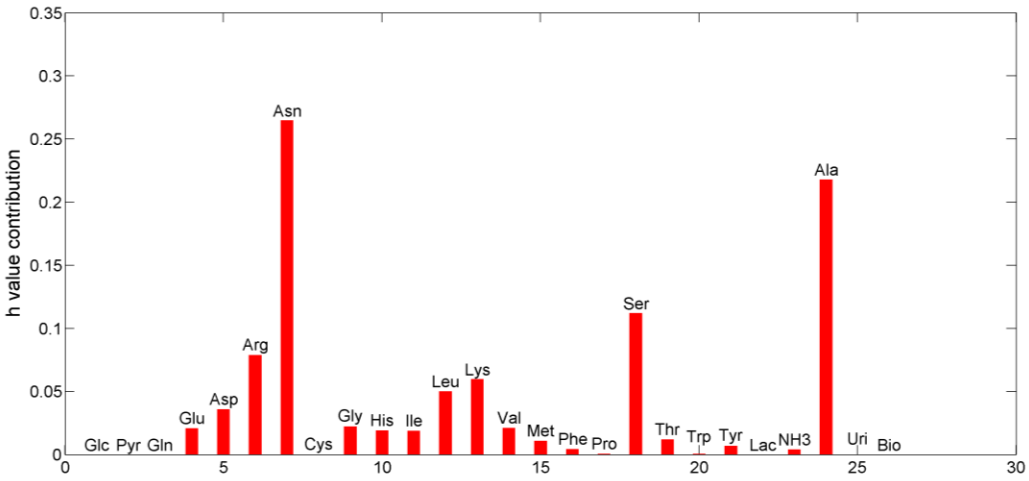

***Figure 3.6**: Mean of the coefficient of contribution for each of experimentally measured compounds on the model consistency.*

Surprisingly, there is a differential effect from each substrate standard deviation on the consistency. This suggests that the errors in measurements have different impacts on the overall system consistency. The glucose, pyruvate, biomass and glutamine flux standard deviations have zero impact on the system consistency. On the other hand, asparagine, alanine and serine have a high influence. The metabolic model presented is then highly sensitive to variations in these compounds. These findings suggest that for some compounds a rigorous quantification is uttermost important in order to improve the MFA results for estimated intracellular fluxes. In the following Figure, **Fig. 3.7**, the coefficient of contribution for compounds that most impact the overall system consistency are presented, over the studied time frame.

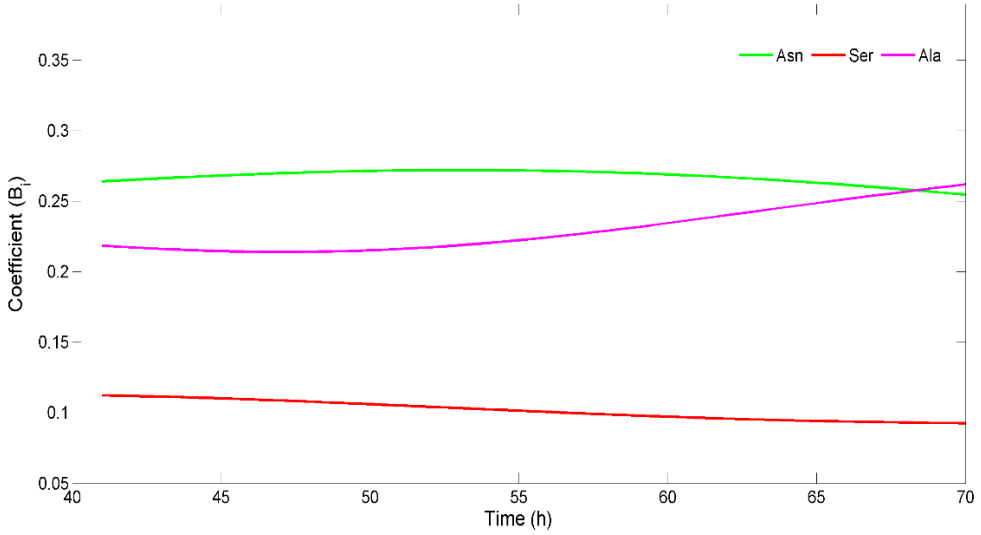

***Figure 3.7**: Coefficient of contribution value for compounds with the most impact on the model consistency over time.*

This Figure reveals that for the compound for which the measurements errors have the higher influence on the consistency, the coefficients of contribution changes narrowly over time. This highlights their predominance as the main source of the inconsistency.

3.3 Flux Balance Analysis

Flux Balance Analysis (FBA) is a mathematical methodology vastly used to predict gene activity in genome scale network analysis and metabolites flow through the pathways of a metabolic model [17]–[19]. A critical point for the application of this method is the determination of the cell's biological objective. This concept is based on the idea that the evolutionary pressure enabled the cells to redirect the metabolic flux according to a certain goal or global objective [21].

It is also worth mentioning that in more complex systems like eukaryotic cells the objective may be a combination of two or more.

In theory, every cell is considered to have a global objective for a given moment, which governs or directs flux distribution through the metabolic pathways. This goal can change from one environment to another and according to the cell life cycle itself [17]. In the exponential biomass growth phase, this objective usually is the maximization of biomass growth, but this is not always the case [18], [21], [39]. In order to encounter the objective of the analyzed cell, multiple combinations of objective functions, which were taken from literature, were tested (e.g. maximization of ATP production, minimization of reductive power and maximization of lactate production) and the obtained flux distribution compared to the experimental data (data not shown). The objective for which the best fit to the experimental data was obtained, was assumed to describe the cells objective best.

The Linear Programming function (linprog) available in MATLAB® was used to for the FBA calculations. The following parameters where defined:

- Upper boundaries (UB)
- Lower boundaries (LB)
- Objective function (Z)

The problem then reads:

Equation 14: FBA optimization parameters.

$$\textit{Objective,}$$

$$\max_{v}(Z) = \max_{v}\left(\sum_{i=1}^{\overrightarrow{q}} w_i.\mathrm{v}_i\right),\ \textit{subject to}$$

$$\frac{dc}{dt} = 0 = S.\mathrm{v} = S_n.\mathrm{v}_n + S_k.\mathrm{v}_k$$

$$0 \le v_i \le 10000 \ \text{ reaction } i \textit{ irreversible}$$

$$-10000 \le v_i \le 10000 \ \text{ reaction } i \textit{ reversible}$$

$$225 \le \mathrm{v}_{o_2} \le 728$$

$$241 \le \mathrm{v}_{Glc} \le 303$$

The use of FBA as a predictive technique is considerable impaired in aerobic conditions, because the exact oxygen flux value is difficult to predict [21]. Tighter upper and lower boundaries for the oxygen levels has been observed to remarkably improve the predictions of FBA, which is in agreement to [21]. Tighter boundary conditions for the carbon source (e.g. glucose in this case) have also been observed to have an impact on the FBA prediction, which also is in agreement to

[21], [40]. Therefore, additional boundary conditions were proposed for the glucose and oxygen flux boundaries. The oxygen and glucose upper boundaries were defined as the sum of the corresponding mean flux value and their standard deviation, calculated from data for the studied phase. The lower boundaries were defined respectively as the sum of the mean flux value minus their corresponding standard deviation. In the studied system, the oxygen boundaries were chosen based on the flux values estimated by the MFA method. Beside the oxygen and carbon source, additional constraints might be needed to avoid futile cycling or unrealistic flux values as shown in [13]. In order to improve the FBA predictability two scenarios were tested, with the same cellular objective and different constraints. The FBA scenarios are described in the following in more detail. In the first, glucose and oxygen boundary conditions were tightened, and in the second scenario, an additional constraint was proposed as a response to the results observed for the first scenario.

3.3.1 The cell objectives

The cellular objective that best describes the experimental data of the CR.pIX cultivation was found to be the minimization of the oxidative phosphorylation occurring in the mitochondria (Z= {Minimization $_{\text{Oxidative Phosphorylation}}$}). This also is in agreement with the observed high lactate production and the lower activity of the TCA cycle found during MFA. This objective was rather unexpected tough the TCA cycle is known to be balanced by the oxidative phosphorylation levels [41] and lower TCA cycle activity means low oxidative phosphorylation.

In the first tested scenario, the predicted specific biomass growth rate value was 0.0083 $h^{-1,}$ which is lower than the experimental value of 0.0152 h^{-1}. In **Figure 3.8** experimental and FBA predicted substrate uptake and metabolic product formation values are presented alongside.

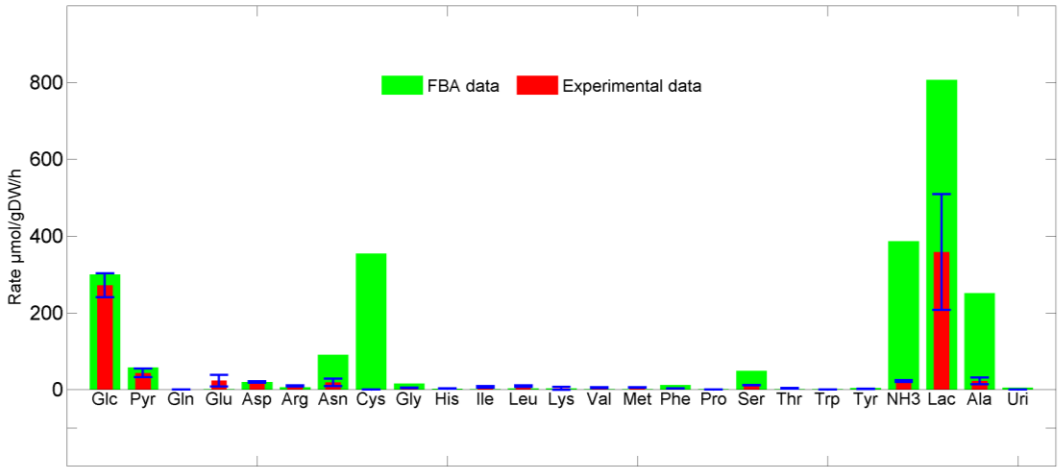

***Figure 3.8**: Experimental substrate uptake and metabolic product formation rates plus standard deviations and the rates predicted by FBA, for the first scenario.*

These results reveal that the worst predicted rate is the cysteine rate (355 µmol/gDW/h compared to zero µmol/gDW/h determined experimentally). This over uptake resulted in higher ammonia and lactate estimations. Since the predicted cysteine uptake rate disagrees the most with the observed (see **Figure 3.8**), in the second tested scenario the cysteine uptake (r_{cys}=0 µmol/gDW/h) was defined as an additional FBA constraint. It has been reported that the inclusion of uptake rates in the FBA constraints can improve FBA predictions, since it limits the possible solution according to substrate media availability [42], [43].

In **Figure 3.9** a comparison between experimental and FBA predicted substrate uptake and metabolic product formation values for the second scenario is presented.

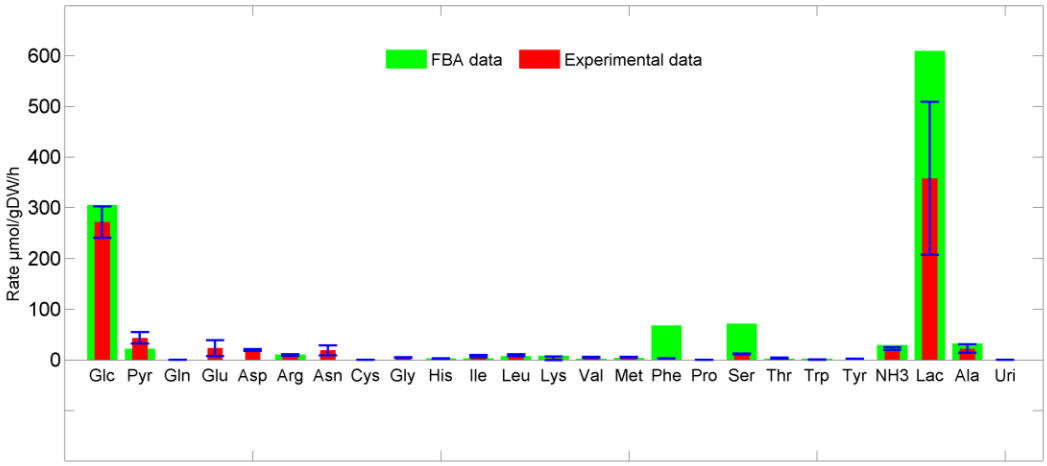

***Figure 3.9**: Experimental substrate uptake and metabolic product formation rates, their standard deviations and the rates predicted by FBA for the second scenario.*

In this scenario, the FBA predicts a biomass growth rate of 0.0188 h^{-1}, which is closer to the experimental value of 0.0152 h^{-1}. For most of the tested objective functions (data no shown), the biomass growth rate was well predicted by FBA, whereas the predicted fluxes differed significantly for those fluxes for which experimental data were available. The fluxes that achieve accurate growth rate predictions are not unique but one solution of a large space of solutions [10]. This is one of the main reasons why FBA predictions poorly represent the experimentally measured rates when the objective function is not adequate [10].

In **Figure 3.9**, it can be seen that most of the predicted flux values are in good agreement with the experimental data, though some fluxes still deviate significantly. Firstly, predicted glutamate and asparagine uptake are both lower than their measured uptake rates. The FBA with this objective compensates these lower fluxes by over uptake of other amino acids like phenylalanine. Phenylalanine catabolism provides the needed glutamine and asparagine and also provides the TCA precursors. FBA with the given objective predicts higher serine flux values than measured. Serine catabolism provides pyruvate that can enter the TCA cycle. It can been seen that estimated ammonia levels are within expected boundaries, which suggests that amino acids catabolism levels are balanced.

The estimated phenylalanine uptake rate is 20 times higher than the experimentally calculated rate, while the serine uptake rate is five times higher (**Figure 3.9**). Most likely, the cell objective includes a subset of other objectives, influenced by the media composition. The observed differences between predicted and measure fluxes in the results are not unforeseen, since cellular objectives in eukaryotic cells are much more complex than in prokaryotic cells [21].

The availability of the amino acids present on the media balances the levels of uptake by the cells. In FBA the lack of the predictability of amino acid uptake is because FBA "does not know" how much of what it should assume as the optimal value, if two amino acids give origin to the same needed precursors. Media availability of the substrates dictates the uptake rates when cells are cultivated. In order to improve FBA performance, the environment composition must be an input constraint, as described in [18]. On the other hand, the wide boundaries chosen as constraints can still condition the solutions, even with further constraints.

The lack of predictability of fluxes, i.e. uptake of amino acids, can be accounted for by different reasons, i.e. lack of thermodynamic constraints [10]. A way to improve rate predictions of FBA could be inclusion of a second optimization such as the minimization of the differences between predicted and experimentally determined fluxes [44]. It is of interest to use as few as possible constraints to describe the cell's objective, since in the next step Flux Variability Analysis (FVA) is used to provide ranges of values for each reaction [10]. FBA is a very good method to estimate growth rate, it also very unreliably in predicting other flux values such as metabolic product formation [10]. With few constraints, the solution space is typically wider and the flux variability analysis provides better results [10]. In this scenario, only two out of these 26 predicted flux deviate significantly from the experimental flux values, which is considered to be only a minor defect and so no further constraints were used.

Concluding, a list of causes can be named for the observed deviations: 1) Though the predicted and measured fluxes largely agree the chosen objective function might not completely describe the complex function of the cell. 2) Insufficient large upper and lower bounds or insufficient thermodynamic constraints [45] might in addition contribute to incorrect or infeasible metabolic states. Nevertheless, the selected cellular objective seems to describe the regulation of the levels of oxidative phosphorylation relatively well.

3.4 Flux Variability Analysis

Flux Variability Analysis (FVA) is a mathematical method used to estimate the range of values for the rates of metabolic pathways, maximizing and minimizing the value of each flux value [10]. In the present case, the cellular objective determined in the prior section for which the best fit to the experimental data was obtained and the prior mentioned constraints were included

into FVA as additional constraints. Various scenarios, where a reaction deletion is simulated, were performed and the solution space for each scenario was calculated using FVA. This task was performed by setting the flux through a given reaction, e.g. amino acid uptake, to zero. Thereupon, FVA is used to obtain the feasible solution space. Important reactions are the most inclined to have lower variability in the flux values and FVA is a promising method for the identification of these reactions [7].

The Linear Programming function (linprog) available in MATLAB® was used to for the FVA calculations. The following parameters where defined:

- Upper boundaries (UB)
- Lower boundaries (LB)
- $\max_v$ and $\min_v$ objectives

The problem then reads,

Equation 15 – The FVA optimization parameters.

$$\max_v / \min_v \left\{ c^T v \right\}$$

$$s.t \quad \frac{dc}{dt} = 0 = S.\mathrm{v}$$

$$v_{phosporylsation} = \min_v(phosporylsation) = \min_v \left(\sum_{i=1}^{\overrightarrow{q}} w_i.\mathrm{v}_i \right), \; s.t \quad 0 = S.\mathrm{v}$$

$$0 \le v_i \le 10000 \text{ reaction } i \textit{ irreversible}$$

$$-10000 \le v_i \le 10000 \text{ reaction } i \textit{ reversible}$$

$$225 \le v_{o_2} \le 728$$

$$241 \le v_{Glc} \le 303$$

$$v_{cys} = 0$$

3.4.1 Flux variability for FBA results with assumed biological objective

FVA was performed to analyze the solution space of the FBA flux distribution, which was calculated with the objective that best described the measured data. In the following Figure, box plots of these results are presented.

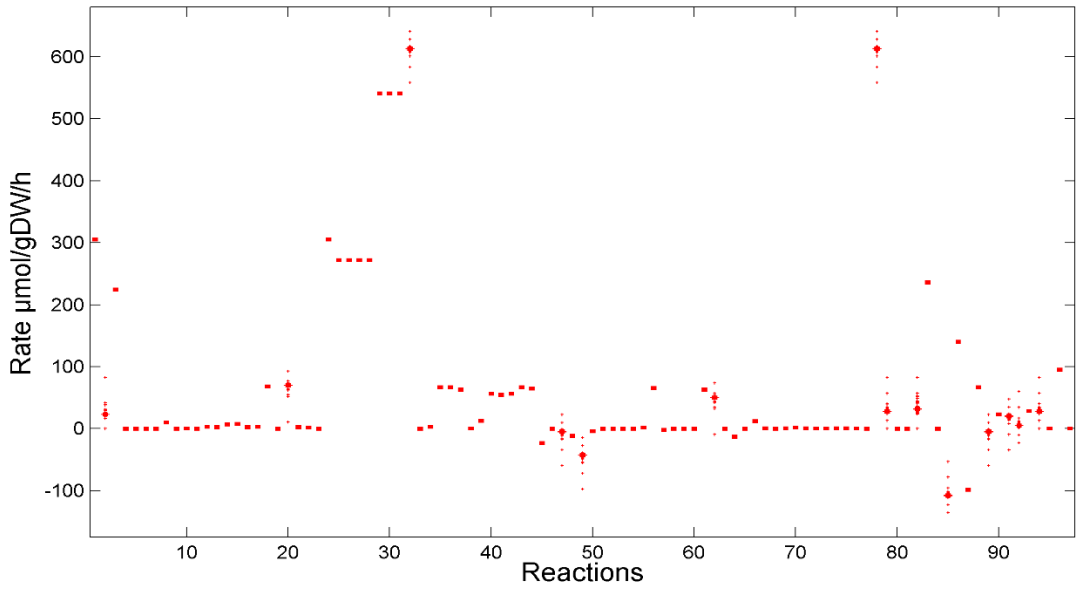

***Figure 3.10**: Rate values predicted by Flux variability analysis for FBA results with assumed biological objective constraint.*

Surprisingly only little variations can be seen in **Figure 3.10**. This means that many predicted rates are uniquely determined by FBA. This phenomena might be due to the constraints (glucose, oxygen and cysteine uptake) used to perform FBA. It was observed that the variability in the fluxes grows when the cysteine constraint is dropped (data not shown). The fluxes that vary most seem to be pyruvate uptake (r_2), serine uptake (r_{20}), alanine and glutamate catabolism (r_{49} and r_{47}) and most of transport reactions (r_{85}, r_{89}, $r_{90\text{-}94}$).

It has been shown that flux variability can be decomposed into three components: internal, external and growth variability and that variability in biomass growth gives rise to flexibility in the usage of different metabolites observed in vivo in [27]. In the present case, no variability in the growth rate value was observed as well as low or no variation in the other rate values. The question was whether the cell could still adapt to environments when a certain metabolite is not present.

3.4.2 Flux variability with different condition environment simulation

Given the deleted reactions, the biomass growth rate prediction is used to assess the viability of each scenario. In **Table 4.2**, values of the biomass growth rate predictions are provided for each scenario.

***Table 3.2**: Growth rate predictions for each tested FVA scenarios.*

Scenario	Growth rate (h^{-1})
No Pyr uptake	0.0188
No Gln uptake	0.0188
No Glu uptake	0.0188
No Asn uptake	0.0188
No Asp uptake	0.0188
No Cys uptake	0.0188
No Arg uptake	0.000
No His uptake	0.000
No Ile uptake	0.000
No Leu uptake	0.000
No Lys uptake	0.000
No Val uptake	0.000
No Met uptake	0.000
No Phe uptake	0.000
No Thr uptake	0.000
No Trp uptake	0.000
No Pro uptake	0.000-0.0142
No Ser uptake	0.000-0.0139
No Gly uptake	0.000-0.0136
No Tyr uptake	0.000-0.0141
No Lac release	0.0054-0.0101
No Amm release	0.0094-0.0131
No Ala release	0.000-0.00134
No Uric release	0.0016-0.0146
No Ala uptake	0.000-0.00133
No Pyr release	0.000-0.0121

No impact on the predicted biomass growth rate when some amino acids uptake rates of are set to zero can be observed. It is known that some amino acids are essential to cellular growth, while others can be integrally compensated with no overall impact on the biomass growth rate [34]. If an amino acid uptake rate is set to zero and the FVA predicts a biomass growth rate similar to the experimentally observed, it is likely that these amino acids are not essential. In theory, this also suggests that cultivation could be viable without these amino acids supplements. This would

be very beneficial since it diminishes the associated costs with industrial processes. In practice though, this may not be the case. For other amino acids the predicted biomass growth rate is zero when their uptake rates are set to zero. It is very likely that these amino acids are essential.

It can also be seen that when some reactions rates are set to zero, a prediction of partial growth rate is observed. The most interesting result is that when lactate production is set to zero. In this case, the predicted biomass growth rate by the FVA is 33-66% of the experimentally observed growth rate of 0.0188 h^{-1}. If no lactate was produced, most likely cellular respiration would take place. As mentioned above, this is a slower way to produce ATP for biomass growth [30], which may be the reason why the predicted growth rate is lower. When some amino acids uptake rates are set to zero and a partial biomass growth rate is predicted, this is most likely due to widely known interconversion amino acids [46] .

Overall, these results suggest that the non-essential amino acids are glutamine, glutamate, cysteine, arginine, aspartate and asparagine. These results are in agreement to [34]. Although cysteine is not known to be essential for most eukaryotic cells, it has been found to be essential to avian cells under specific conditions, as reported in [34]. They also suggest that the essential amino acids for CR.pIX are Arginine, Histidine, Isoleucine, Leucine, Lysine, Valine, Methionine, Phenylalanine, Threonine and tryptophan. This also is in agreement to [34]. Finally, partial growth is predicted in FVA results when Proline, Serine, Glycine and Tyrosine uptake rates are set to zero, which is due to the fact that these are considered semi-essential in avian cells [34].

Since flux variability was applied to in each mentioned scenario, the space of solutions where biomass growth rate is feasible was identified. In **Figure 3.11** the results from FVA for all the scenarios where predicted biomass growth rate is greater than zero are shown.

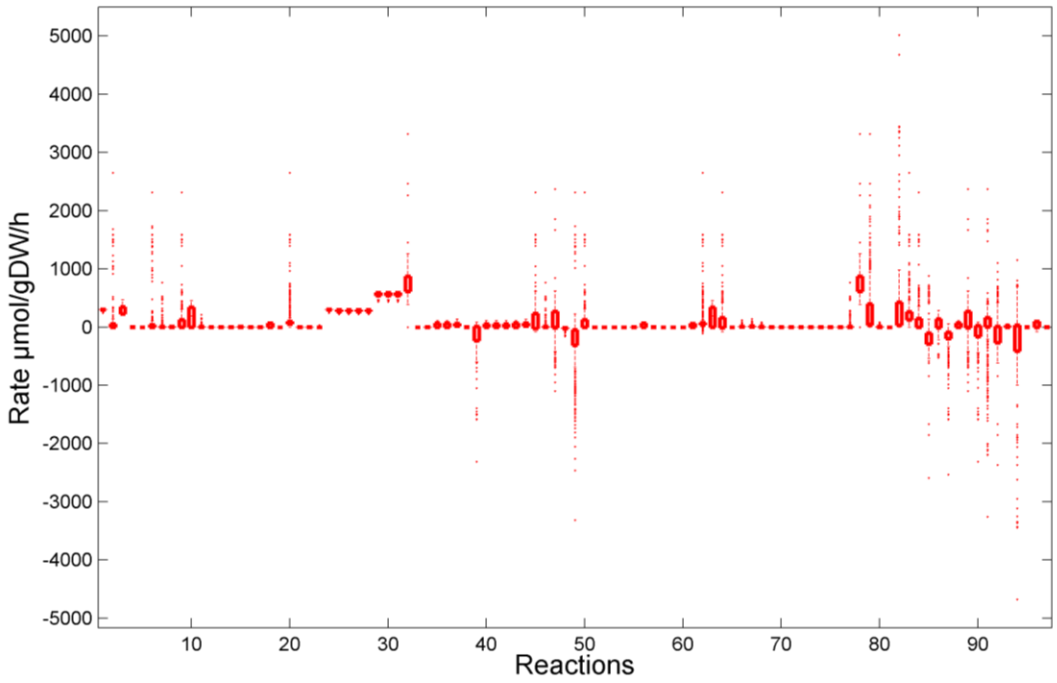

***Figure 3.11**: Rates predicted by Flux variability analysis for all the scenarios where the predicted biomass growth rate was greater than zero.*

It can be seen that most fluxes have a wide range of variation, while others do not. Detailed analysis of this data showed that estimated uptake and catabolism rates of the essential amino acids (such as r_8, $r_{12\text{-}18}$, $r_{21\text{-}22}$ and $r_{51\text{-}56}$) have insignificant variation in all scenarios. Another group of estimated reactions rates with insignificant variation is the lipid synthesis ($r_{69\text{-}76}$). Most of TCA cycle reactions also have low variability (such as $r_{35\text{-}37}$ and $r_{41\text{-}45}$). On the other hand, the estimated uptake and catabolism rates for the non-essential amino acids (such as $r_{4\text{-}7}$, $r_{9\text{-}11}$, r_{20} and $r_{47\text{-}50}$) showed high variations. Another group of reactions with high variability are the transport reactions (r_{85}, r_{89}, $r_{90\text{-}94}$). It can be assumed that the reactions that have the least variability rates through them are likely to be the most important. It has been shown in [27] that the biomass growth rate modulates the cellular metabolic flexibility. Thus, it is very likely that, when the CR.pIX cells are growing, the most important fluxes in order to grow are related to the essential amino acids uptake and catabolism, the lipid synthesis and ATP production via TCA.

3.4.1 The glutamine free medium flux variability

In this study using FVA, it was predicted that the avian cells have no need for glutamine supplement while growing. This is because the predicted biomass growth rate when glutamine uptake rate was set to zero was close to the experimentally observed. Previous studies have shown that CR.pIX is indeed capable of growth in a glutamine free medium after few passages [1]. A detailed analysis of this scenario is made, in order to better understand this phenomenon. In the following Figure, the flux distribution when glutamine uptake rate was set to zero is presented.

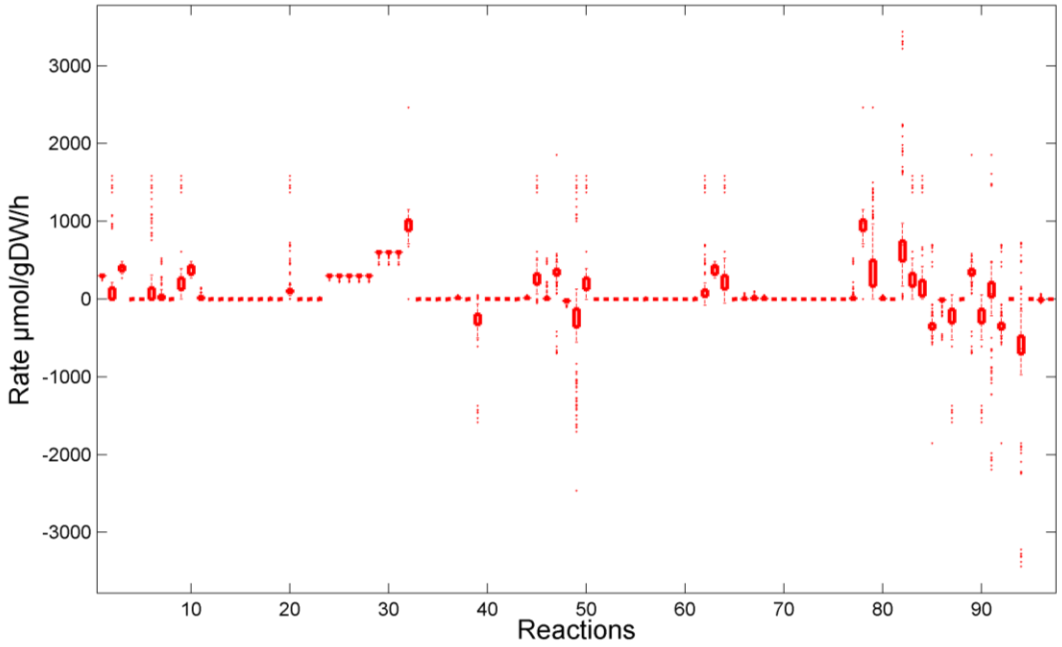

***Figure 3.12**: Rates predicted by Flux variability analysis when glutamine uptake rate is set to zero.*

These results show that when glutamine uptake is set to zero some reactions rates have a certain degree of variability. This FVA result for glutamine has a typical variability observed in all other scenarios. This suggests a certain metabolic flexibility in order to compensate for depletion.

3.5 Principal Component Analysis

Principal Component Analysis (PCA) is statistical procedure that uses orthogonal transformation to convert a set of possible correlated data into a linearly uncorrelated subspace. This transformation is essential to analyze high dimensional or large experimental data sets and identify correlations [28]. PCA has been shown to highlight the most active reactions through metabolic e genomic network [28]. Furthermore, the PCA scores plotting has been shown to be capable to cluster similar metabolic states together [28].

The PCA was performed in MATLAB using the N-way toolbox, described in [47].

3.5.1 Number of components

PCA captures the variation in a data set and reduces the noise or similarly variations, which can be found in data. PCA transforms the input data into a latent space, which dimension needs adapted to capture the variance contained in the input. It can be used to find interpretable steady state metabolic states in a data set of flux measurements [26]. This analysis highlights the small number of fluxes whose variability accounts for almost all the other observed in flux variability analysis [26]. This means that a small number of reactions determines the cellular metabolic state. A reaction that has been found to have the most impact on flux variability is the biomass growth rate [26], [27].

In **Figure 3.13** the captured variance of the FVA data set by PCA relative to the number of number of principal components is presented.

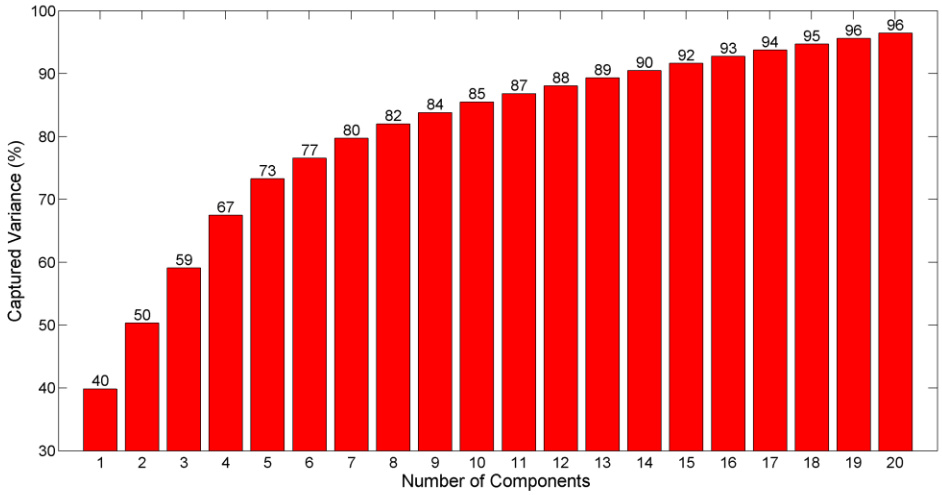

***Figure 3.13**: Captured variance in the FVA data by PCA vs the number of components.*

It can be seen that almost half of the variance in the FVA data set is captured by one component, while the rest of variance would need much more components. The idea is to retain sufficient variance in order to explain the original data set, but with as less number of components as possible [48]. This is important because noise can be included if too much components are used. In this case, the chosen number of principal components was based on the level of variance captured by PCA.

In **Figure 3.14** the comparison between experimental and PCA estimations is presented.

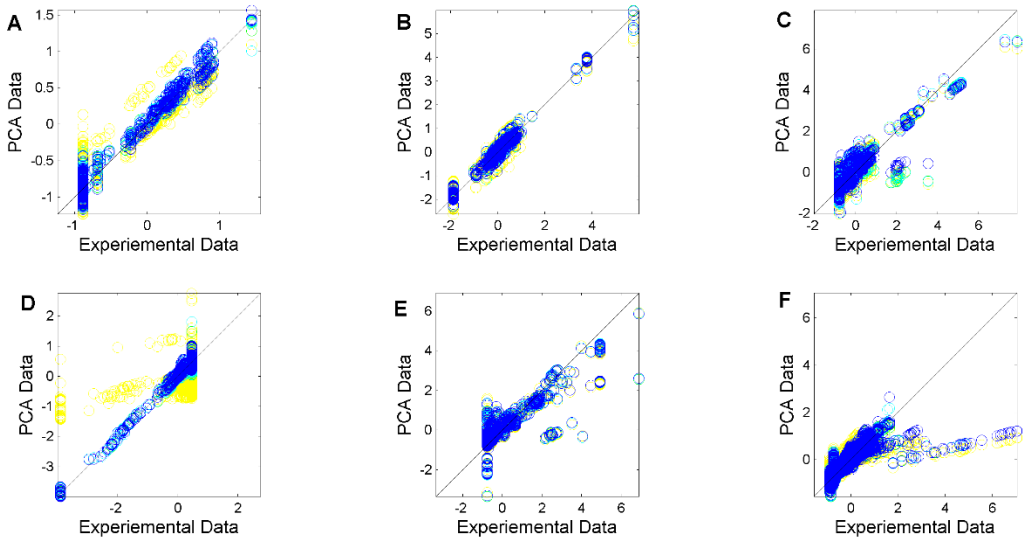

***Figure 3.14**: Comparison between PCA and scaled FVA data. A: Biomass data. B: Lactate date. C: Glucose data. D: Ammonia data. E: Alanine Data. F: Essential amino acids data. Number of components: Four ( ), five ( ), six ( ), seven ( ).*

The results presented in **Figure 3.14** clearly show that the chosen number of principal components should be five as there is no significant difference between five, six or seven components. With this number of component, the captured variance is roughly 73% of the original FVA data set. The remaining 27% seem not to be essential and they might describe the rather fixed relation between the not varying compounds.

Further analysis of the components was performed. In **Figure 3.15**, a matrix mapping of the PCA loadings from each of the five principal components, loadings, is presented.

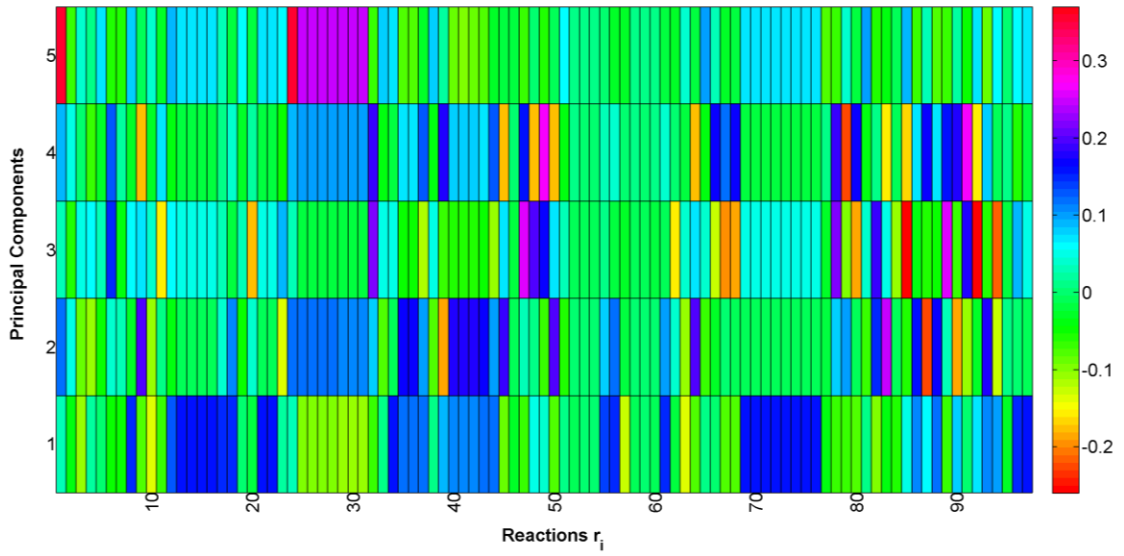

***Figure 3.15**: Principal components, loadings, and contribution to explain each reaction.*

As mentioned above, all the variability obtained using FVA analysis is likely to be explained by the control of variability of a small set of reactions. It has been found that the loading values, which varies between -1 and 1, correlates to the amount of explained variability by a given principal component. For the first principal components or latent variable (LV_1), the most relevant reactions include uptake of essential amino acid (r_8, r_{12-18}, r_{21-22}) and metabolism and lipid synthesis (r_{69-76}) and biomass formation (r_{97}). As discussed above, these reactions are the most import for biomass growth. This component is then likely to explain biomass growth rates variability. The second latent variable (LV_2) is connected to glucose uptake (r_1), part of TCA (r_{35-37} r_{39-43}) and part of Anaplerosis (r_{34-45}) cycle and arginine uptake (r_8). This component is likely to explain TCA cycling variability. The third component (LV_3) is associated to MTHF & uric acid synthesis (r_{66-68}) and lactate formation (r_{32}) and excretion (r_{78}). This component is likely to explain both lactate and MTHF & uric acid synthesis variability. The fourth component (PC_4) is related to Ala uptake (r_6), metabolism (r_{49}) and release (r_{79}) as well as other substrates catabolism that originates precursors for Ala formation (r_{47}, r_{48} and r_{91}). This latent variable is then likely to explain Ala variability. Finally, for the fifth component (PC_5) it corresponds to glucose uptake (r_1) and most part of glycolysis (r_{24-31}) meaning that most likely this component explains the variations observed in glycolysis.

3.5.2 Metabolic states

Since PCA scores plotting has been shown to cluster similar metabolic states together [28], the next step is the analysis of the scores generated from the FVA data set by PCA. The loadings matrix was used to calculate the scores for the extracellular fluxes (substrate uptake and product formation) and the intracellular flux distribution, which had been estimated with metabolic flux analysis (MFA). These data are designated MFA scores in the following. In the following Figure, FVA scores and MFA scores are presented for the first three components.

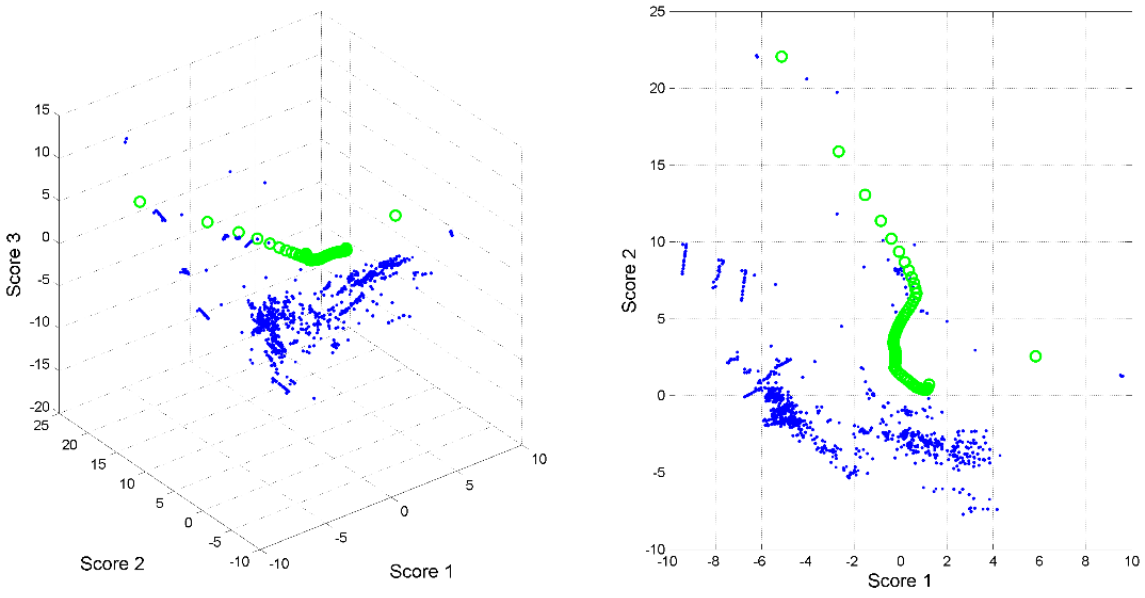

***Figure 3.16**: Score plot of PCA scores from the FVA data. A: 3D score plot. B: 2D score plot. FVA scores (■), MFA scores (○).*

It can be seen in **Figure 3.16** that there exist at least two distinct clusters and other small ones. These clusters are likely to represent different metabolic states [28].

Using the PCA coefficients generated from the FVA data set, it is possible to calculate the score for each optimal flux distribution predicted by FBA for each scenario where a reaction elimination was simulated and the predicted biomass growth rate was greater than zero. In the **Figure 3.17** these optimal flux distribution are presented along the cluster from FVA data set.

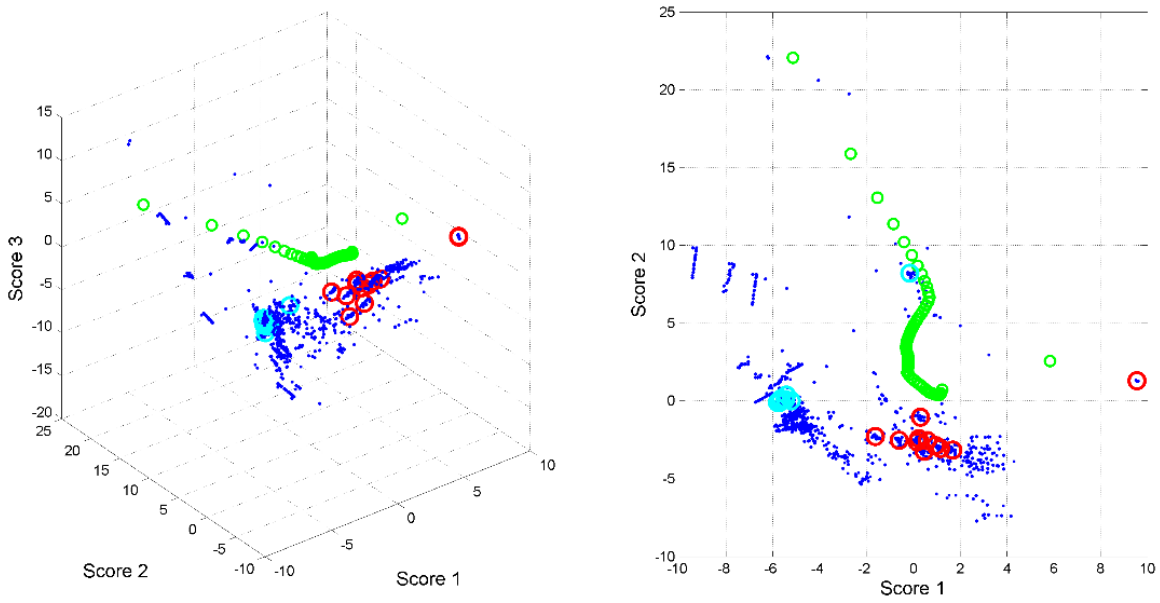

***Figure 3.17**: Plot of PCA scores from the FVA data and scores for each FBA optimal for each compound deletion simulation. A: 3D score plot. B: 2D score plot. FVA scores (■), MFA scores (○), FBA with no predicted growth rate scores (○), FBA with predicted growth rate scores (○).*

It can clearly be differentiated between the cluster clustering from FBA optimal flux distributions where no biomass growth rate was predicted (cyan) and the cases where biomass growth was predicted (red). This can be expected since growth and no growth means, theoretically, different cells metabolic states.

In **Figure 3.18** the score from optimal flux distribution predicted by FBA with the assumed CR.pIX biological objective is added.

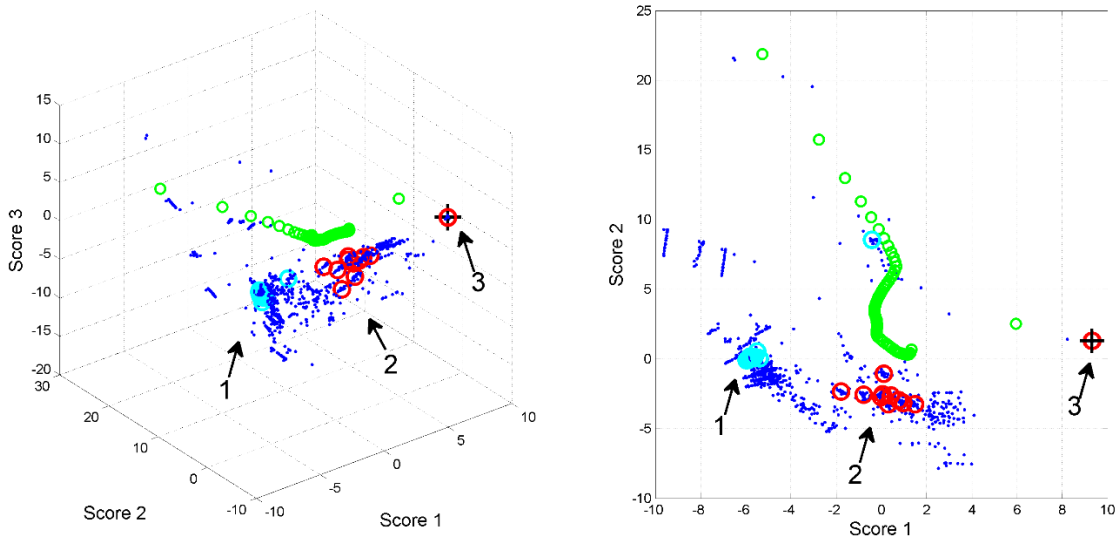

***Figure 3.18**: Plot of PCA scores from the FVA data and scores for each FBA optimal for each compound deletion simulation and FBA optimal for the cellular objective. A: 3D score plot. B: 2D score plot. FVA scores (■), MFA scores (○), FBA with no predicted growth rate scores (○), FBA with predicted growth rate scores (○). FBA optimal flux distribution with assumed CR.pIX biological objective scores (+).*

Essentially three cluster can be observed in **Figure 3.18**. The first cluster contains the FVA scenario where a compound elimination, results in a predicted biomass growth rate of zero. In the second cluster, the FVA scenario where a compound elimination results in a predicted biomass growth rate lower than the optimal biomass growth rate predicted by FBA with the assumed CR.pIX biological objective can be found. Finally, the third cluster contains the FVA scenario where a compound elimination results in a predicted biomass growth rate equal to the optimal biomass growth rate predicted by FBA with the assumed CR.pIX biological objective.

Considering the optimal flux distribution predicted by FBA that best described the CR.pIX biological objective as the normal metabolic state, it becomes clear that the elimination of some compounds in the media theoretically will have no on cell's metabolic state, which implies no impact on biomass growth rate. These reactions include the non-essential amino acids uptakes, previously discussed, and the pyruvate uptake. This means that most likely there is no need for a supplement of non-essential amino acids or pyruvate when cultivating the CR.pIX. The same experimentally observed growth rate could be achieved in a simpler media.

3.5.3 The glutamine free medium

In the PCA score plot it can be seen that the glutamine free medium metabolic state is close to the normal avian cell metabolic state, once again suggesting that the avian cell has no need for glutamine supplement while growing, which has been confirmed experimentally [5].

As for other predictions, such as growth without non-essential amino acids or pyruvate, experiments should be performed where each of these compounds is eliminated from the medium composition. The biomass growth rate of each culture should be compared to an experimental batch where the eliminated compounds were present in the medium. Furthermore, other methods such as MFA, FBA and PCA should be performed with the new experimental data. This will allow to further characterize and understand the metabolic changes.

4

4 Conclusion

In this study, the impact of eliminating compounds from the extracellular environment on the intracellular flux distributions and physiological state were studied for an avian cell line based on a metabolic model of the central metabolism [1].

At first, Metabolic Flux Analysis was used in which the central metabolic model proposed by Lohr et al [1] and was shown to be consistent in describing the flux distributions during exponential growth phase of the avian cell line cultivated in a stirred batch reactor, tough a decrease in the model consistency over time was observed. It was discussed that the pseudo-steady-state assumption might be violated, which most likely is due to cellular agglomerations or imperfect mixing. In addition, the network used might not fully describe the cells metabolic behavior. The sensitivity of the overall system consistency to changes in the standard deviations of the flux values was studied, highlighting the need for precise and careful quantification of certain compounds (e.g.: asparagine, alanine and serine) in order to improve/validate the modelling results. In addition, the quantification of additional compounds could improve the model consistency further, as well as the modelling predictability based on the central metabolic model.

In a second stage, Flux Balance Analysis was applied to elucidate the possible objective of the cell. The biological objective that resulted in the best fit of the predictions to the experimental data was the minimization of oxidative phosphorylation, suggesting a cellular respiration shutdown mechanism in order to maximize the ATP production rates. This resulted (and is known to result [30], [38], [49]) in high lactate production, which in fact was observed experimentally.

In the next step, Flux Variability Analysis (FVA) was used to obtain the variations in the optimal flux distributions of FBA, where the identified objective was employed as a constraint. Several eliminations of one compound at a time from the media (equivalent to the deletion of one reaction at a time) were simulated using this method. The Flux Variability Analysis results are in agreement with the findings of essential and non-essential amino acids for avian cells [34]. In addition, it was also found that the reactions with the least flux variability through them were essential amino acids uptake and catabolism, lipid synthesis and most of the TCA cycle reactions. This might indicate that the tight control of these reactions is of uttermost importance for CR.pIX exponential growth.

Finally, Principal Component Analysis (PCA) was applied to the data gathered in the previous steps in order to assess potential changes in the physiological states. Three metabolic states were found, which correspond to zero, partial and maximum biomass growth. It was found that the maximum biomass growth is similar to the assumed normal metabolic state. Elimination of non-essential amino acids or pyruvate from the media showed no impact on the cell's normal metabolic state. These results suggest that CR.pIX might grow and reach high concentration, as previously observed, without non-essential amino acids or pyruvate supplement, but not without the essential ones.

Three experiments should be performed where it is tested: 1) whether biomass can grow without the nonessential amino-acids; 2) whether biomass can grow without pyruvate; and 3) an experiment to test whether the minimization of the oxidative phosphorylation can truly be assumed to be the cell's objective, e.g. by studying the impact of reduced oxygen availability.

5 Future Work

In this thesis experimental data of only one batch were used. First and foremost more experiments should be performed, if possible the three suggested ones, which would allow to validate the model and modeling hypothesis. It should be evaluated whether it is possible to quantify those compounds whose fluxes' standard deviations have the most impact on the model consistency (i.e.: asparagine, alanine and serine) more accurately. Further, the quantification of more metabolites, e.g. intracellular metabolites, also has the potential to improve the modeling results. It has been shown that accurate oxygen determination also can improve the modeling methods results, wherefore quantification of the oxygen fluxes would also be beneficial. Alternatively, methods such as C_{13} labels for intracellular compounds quantification would be a leap forward in the improvement of the model and its consistency for the estimation of other intracellular fluxes.

Dynamical modeling could be applied in order to understand the regulation of the transitions between the physiological states better.

Finally, while more studies, like medium optimization are highly needed for this new avian cell line, also studies of the downstream process should be considered to further access the viability of virus production at an industrial scale.

6

6 Bibliography

[1] V. Lohr, O. Hädicke, Y. Genzel, I. Jordan, H. Büntemeyer, S. Klamt, and U. Reichl, "The avian cell line AGE1.CR.pIX characterized by metabolic flux analysis.," *BMC Biotechnol.*, vol. 14, p. 72, 2014.

[2] V. Lohr, a. Rath, Y. Genzel, I. Jordan, V. Sandig, and U. Reichl, "New avian suspension cell lines provide production of influenza virus and MVA in serum-free media: Studies on growth, metabolism and virus propagation," *Vaccine*, vol. 27, no. 36, pp. 4975–4982, 2009.

[3] I. Jordan, A. Vos, S. Beilfuss, A. Neubert, S. Breul, and V. Sandig, "An avian cell line designed for production of highly attenuated viruses.," *Vaccine*, vol. 27, no. 5, pp. 748–56, Jan. 2009.

[4] V. Sandig and I. Jordan, "Immortalized avian cell lines for virus production." WO 2005/042728 A2, 2005.

[5] V. Lohr, Y. Genzel, I. Jordan, D. Katinger, S. Mahr, V. Sandig, and U. Reichl, "Live attenuated influenza viruses produced in a suspension process with avian AGE1.CR.pIX cells," *BMC Biotechnol.*, vol. 12, no. 1, p. 79, 2012.

[6] Y. Genzel, T. Vogel, J. Buck, I. Behrendt, D. V. Ramirez, G. Schiedner, I. Jordan, and U. Reichl, "High cell density cultivations by alternating tangential flow (ATF) perfusion for influenza A virus production using suspension cells," *Vaccine*, vol. 32, pp. 2770–2781, 2014.

[7] G. Stephanopoulos, A. A. Aristidou, and J. Nielsen, *Metabolic engineering*. Academic Press, 1998.

[8] S. S. Ozturk and B. O. Palsson, "Chemical decomposition of glutamine in cell culture media: effect of media type, pH, and serum concentration.," *Biotechnol. Prog.*, vol. 6, no. 2, pp. 121–128, 1990.

[9] M. Schneider, I. W. Marison, and U. Von Stockar, “The importance of ammonia in mammalian cell culture,” *J. Biotechnol.*, vol. 46, no. 3, pp. 161–185, 1996.

[10] A. Müller and A. Bockmayr, “Thermodynamic Constraints for Metabolic Networks,” p. 83, 2012.

[11] B. a. Boghigian, G. Seth, R. Kiss, and B. a. Pfeifer, “Metabolic flux analysis and pharmaceutical production,” *Metab. Eng.*, vol. 12, no. 2, pp. 81–95, 2010.

[12] D. Machado, R. S. Costa, M. Rocha, E. C. Ferreira, B. Tidor, and I. Rocha, “Modeling formalisms in Systems Biology,” *AMB Express*, vol. 1, no. 1, p. 45, 2011.

[13] B. J. Stewart, A. Navid, K. W. Turteltaub, and G. Bench, “YEAST DYNAMIC METABOLIC FLUX MEASUREMENT IN NUTRIENT-RICH MEDIA BY HPLC AND ACCELERATOR MASS SPECTROMETRY,” *Anal Chem.*, vol. 82, no. 23, pp. 9812–9817, 2010.

[14] G. Stephanopoulos and D. E. Stafford, “Metabolic engineering: A new frontier of chemical reaction engineering,” *Chem. Eng. Sci.*, vol. 57, no. 14, pp. 2595–2602, 2002.

[15] G. Stephanopoulos and a J. Sinskey, “Metabolic engineering--methodologies and future prospects.,” *Trends Biotechnol.*, vol. 11, no. 9, pp. 392–396, 1993.

[16] S. Klamt, S. Schuster, and E. D. Gilles, “Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria,” *Biotechnol. Bioeng.*, vol. 77, no. 7, pp. 734–751, 2002.

[17] J. D. Orth, I. Thiele, and B. Ø. Palsson, “What is flux balance analysis?,” *Nat Biotechnol*, vol. 28, no. 3, pp. 245–248, 2010.

[18] J. M. Lee, E. P. Gianchandani, and J. a. Papin, “Flux balance analysis in the era of metabolomics,” *Brief. Bioinform.*, vol. 7, no. 2, pp. 140–150, 2006.

[19] E. P. Gianchandani, A. K. Chavali, and J. a Papin, “The application of flux balance analysis in systems biology,” *WIREs Syst. Biol. Med.*, vol. 2, no. 3, pp. 372–382, 2009.

[20] J. P. Mazat and B. Beauvoit, *Modelling metabolic networks-the theories of metabolism*, 1st ed., vol. 67. Elsevier Ltd., 2013.

[21] C. E. García Sánchez, C. A. Vargas García, and R. G. Torres Sáez, “Predictive potential of flux balance analysis of Saccharomyces cerevisiae using as optimization function combinations of cell compartmental objectives,” *PLoS One*, vol. 7, no. 8, 2012.

[22] J. a. Papin, N. D. Price, S. J. Wiback, D. a. Fell, and B. O. Palsson, “Metabolic pathways in the post-genome era,” *Trends Biochem. Sci.*, vol. 28, no. 5, pp. 250–258, 2003.

[23] B. Papp, R. a Notebaart, and C. Pál, “Systems-biology approaches for predicting genomic evolution.,” *Nat. Rev. Genet.*, vol. 12, no. 9, pp. 591–602, 2011.

[24] S. Gudmundsson and I. Thiele, “Computationally efficient flux variability analysis.,” *BMC Bioinformatics*, vol. 11, no. 1, p. 489, 2010.

[25] B. Teusink, A. Wiersma, D. Molenaar, C. Francke, W. M. De Vos, R. J. Siezen, and E. J. Smid, “Analysis of growth of Lactobacillus plantarum WCFS1 on a complex medium

using a genome-scale metabolic model," *J. Biol. Chem.*, vol. 281, no. 52, pp. 40041–40048, 2006.

[26] C. L. Barrett, M. J. Herrgard, and B. Palsson, "Decomposing complex reaction networks using random sampling, principal component analysis and basis rotation.," *BMC Syst. Biol.*, vol. 3, p. 30, 2009.

[27] M. San Román, H. Cancela, and L. Acerenza, "Source and regulation of flux variability in Escherichia coli.," *BMC Syst. Biol.*, vol. 8, p. 67, 2014.

[28] B. Sariyar, S. Perk, U. Akman, and A. Hortaçsu, "Monte Carlo sampling and principal component analysis of flux distributions yield topological and modular information on metabolic networks," *J. Theor. Biol.*, vol. 242, no. 2, pp. 389–400, 2006.

[29] G. Stephanopoulos, "Metabolic fluxes and metabolic engineering.," *Metab. Eng.*, vol. 1, no. 1, pp. 1–11, 1999.

[30] C. B. Vander Heiden, M. G., Cantley, L.C., Thompson, "Understanding the Warburg Effect: the metabolic requirements of cells proliferation.," *Science (80-.).*, vol. 324, no. 5930, pp. 1029–1033, 2009.

[31] N. Slavov, B. a. Budnik, D. Schwab, E. M. Airoldi, and A. van Oudenaarden, "Constant Growth Rate Can Be Supported by Decreasing Energy Flux and Increasing Aerobic Glycolysis," *Cell Rep.*, vol. 7, no. 3, pp. 705–714, 2014.

[32] I. Amelio, F. Cutruzzolá, A. Antonov, M. Agostini, and G. Melino, "Serine and glycine metabolism in cancer," *Trends Biochem. Sci.*, vol. 39, no. 4, pp. 191–198, 2014.

[33] E. J. Chenette, "Glycine fuels cancer cells," *Nat. Cell Biol.*, vol. 14, no. 7, pp. 658–658, 2012.

[34] W. R. Dawson, *Avian physiology.*, vol. 37. 1975.

[35] J. S. Britton and B. a Edgar, "Environmental control of the cell cycle in Drosophila: nutrition activates mitotic and endoreplicative cells by distinct mechanisms.," *Development*, vol. 125, no. 11, pp. 2149–2158, 1998.

[36] Y. Sidorenko, A. Wahl, M. Dauner, Y. Genzel, and U. Reichl, "Comparison of metabolic flux distributions for MDCK cell growth in glutamine- and pyruvate-containing media," *Biotechnol. Prog.*, vol. 24, no. 2, pp. 311–320, 2008.

[37] T. Wahlström and M. A. Henriksson, "Impact of MYC in regulation of tumor cell metabolism.," *Biochim. Biophys. Acta*, 2014.

[38] C. T. Goudar, J. M. Piret, and K. B. Konstantinov, "Estimating Cell Specific Oxygen Uptake and Carbon Dioxide Production Rates for Mammalian Cells in Perfusion Culture," *Biotechnol. Prog.*, vol. 27, no. 05, pp. 1347–1357, 2011.

[39] K. Raman and N. Chandra, "Flux balance analysis of biological systems: Applications and challenges," *Brief. Bioinform.*, vol. 10, no. 4, pp. 435–449, 2009.

[40] A. Bordbar, J. M. Monk, Z. a King, and B. O. Palsson, "Constraint-based models predict metabolic and associated cellular functions.," *Nat. Rev. Genet.*, vol. 15, no. 2, pp. 107–20, 2014.

[41] J. R. Neely and H. E. Morgan, “RELATIONSHIP BETWEEN CARBOHYDRATE AND LIPID METABOLISM AND THE ENERGY BALANCE OF HEART MUSCLE,” *Annu. Rev. Physiol.*, vol. 36, pp. 413–459, 1974.

[42] J. Schellenberger, R. Que, R. M. T. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke, and B. Ø. Palsson, “Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0.,” *Nat. Protoc.*, vol. 6, no. 9, pp. 1290–1307, 2011.

[43] T. Liu, W. Zou, L. Liu, and J. Chen, “A constraint-based model of Scheffersomyces stipitis for improved ethanol production.,” *Biotechnol. Biofuels*, vol. 5, no. 1, p. 72, 2012.

[44] W. R. Harcombe, N. F. Delaney, N. Leiby, N. Klitgord, and C. J. Marx, “The Ability of Flux Balance Analysis to Predict Evolution of Central Metabolism Scales with the Initial Distance to the Optimum,” *PLoS Comput. Biol.*, vol. 9, no. 6, 2013.

[45] S. Y. Lee, J. M. Park, and T. Y. Kim, *Application of metabolic flux analysis in metabolic engineering*, 1st ed., vol. 498. Elsevier Inc., 2011.

[46] P. K. Macnicol, “Synthesis and Interconversion of Amino Acids in Developing Cotyledons of Pea (Pisum sativum L.).,” *Plant Physiol.*, vol. 60, no. 3, pp. 344–348, 1977.

[47] C. a. Andersson and R. Bro, “The N-way Toolbox for MATLAB,” *Chemom. Intell. Lab. Syst.*, vol. 52, no. 1, pp. 1–4, 2000.

[48] S. Valle, W. Li, and S. J. Qin, “Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods †,” *Ind. Eng. Chem. Res.*, vol. 38, no. 11, pp. 4389–4401, 1999.

[49] J. Turner, L.-E. Quek, D. Titmarsh, J. O. Krömer, L.-P. Kao, L. Nielsen, E. Wolvetang, and J. Cooper-White, “Metabolic Profiling and Flux Analysis of MEL-2 Human Embryonic Stem Cells during Exponential Growth at Physiological and Atmospheric Oxygen Concentrations,” *PLoS One*, vol. 9, no. 11, p. e112757, 2014.

7

7 Appendix

***Table 7.1**: Table with the corresponding known fluxes applied during MFA.*

Names	Reaction ID	Number
Glc	r2	1
Pyr	r2	2
Gln	r4	3
Glu	r5	4
Asp	r7	5
Arg	r8	6
Asn	r9	7
Cys	r10	8
Gly	r11	9
His	r12	10
Ile	r13	11
Leu	r14	12
Lys	r15	13
Val	r16	14
Met	r17	15
Phe	r18	16
Pro	r19	17
Ser	r20	18
Thr	r21	19
Trp	r22	20
Tyr	r23	21
Lac	r78	22
Nh3	r83	23
Ala	r79	24
Uric Acid	r80	25
Ala	r6	26
ANPL Pyr	r46	27
Pyr	r84	28
Biomass	r97	29

***Table 7.2**: Intracellular metabolites included on the metabolic model.*

Glycolytic Metabolites	G6P	F6P	FBP	DHAP
	GAP	PG	PEP	PDH
TCA Metabolites	OAA	Cit	Fum	Mal
	SCoA	AKG		
Lipids	CH	PC	PE	PI
	DPG	PGL	SM	PS
Other Metabolites	R5P	PC	GDH	GS
	NADPH	ATP	NADH	FADH2
	FADH	THF	MTHF	

8

8 Annex

***Table 8.1**: Concentrations and analytic methods used for each compound measurements. Adapted from Lohr et al in* [1].

Measured concentrations	**Standard deviation**	**Device**
Cell	2.5 %[a]	ViCELL XR
Glucose	0.39 mM[b]	BioProfile 100plus (Nova Biomedicals)
Lactate	0.30 mM[b]	
Ammonium	4.5 %[a]	
Pyruvate	2.1 %[a]	HPLC (DX-320, Dionex)
Alanine	8.6 %[a]/5 %[c]	HPLC (ICS-5000, Dionex)/ RP-HPLC (Kontron D450)
Arginine	3.2 %[a]/5 %[c]	
Asparagine	0.44 mM[b]/5 %[c]	
Aspartate	0.55 mM[b]/5 %[c]	
Cysteine	0.09 mM[b]/5 %[c]	
Glutamate	0.03 mM[b]/5 %[c]	
Glutamine	12.8 %[a]/5 %[c]	
Glycine	2.7 %[a]/5 %[c]	
Histidine	0.54 mM[b]/5 %[c]	
Isoleucine	0.64 mM[b]/5 %[c]	
Leucine	0.56 mM[b]/5 %[c]	
Methionine	0.33 mM[b]/5 %[c]	
Phenylalanine	0.13 mM[b]/5 %[c]	
Threonine	0.11 mM[b]/5 %[c]	
Tryptophan	0.22 mM[b]/5 %[c]	
Tyrosine	0.32 mM[b]/5 %[c]	
Valine	0.82 mM[b]/5 %[c]	

[a]relative standard deviations of the method were taken for those parameters that have shown an inhomogeneity of variances.

[b]absolute standard deviations of the method were taken for those parameters that have shown homogeneous variances

[c]extracellular amino acid concentrations for metabolic flux analysis were measured with a derivatization method having a measurement error of 5 %.

***Table 8.2**: Reactions included on the metabolic model, adapted from Lohr et al in* [1].

	Uptake rates
r1, Glc	$Glc \rightarrow Glc_{cyt}$
r2, Pyr	$Pyr + 0.33\ ATP_{cyt} \rightarrow Pyr_{cyt}$
r3, O2	$O_2 \rightarrow O_{2,cyt}$
r4, Gln	$Gln + 0.33\ ATP_{cyt} \rightarrow Gln_{cyt}$
r5, Glu	$Glu + ATP_{cyt} \rightarrow Glu_{cyt}$
r6, Ala	$Ala + 0.33\ ATP_{cyt} \rightarrow Ala_{cyt}$
r7, Asp	$Asp + ATP_{cyt} \rightarrow Asp_{cyt}$
r8, Arg	$0.33\ ATP_{cyt} \rightarrow Arg_{cyt}$
r9, Asn	$0.33\ ATP_{cyt} \rightarrow Asn_{cyt}$
r10, CYS	$0.33\ ATP_{cyt} \rightarrow Cys_{cyt}$
r11, Gly	$Gly + 0.33\ ATP_{cyt} \rightarrow Gly_{cyt}$
r12, His	$His + 0.33\ ATP_{cyt} \rightarrow His_{cyt}$
r13, Ile	$Ile + 0.33\ ATPcyt \rightarrow Ile_{cyt}$
r14, Leu	$Leu + 0.33\ ATP_{cyt} \rightarrow Leu_{cyt}$
r15, Lys	$Lys + 0.33\ ATP_{cyt} \rightarrow Lys_{cyt}$
r16, Val	$Val + 0.33\ ATP_{cyt} \rightarrow Val_{cyt}$
r17, Met	$Met + 0.33\ ATP_{cyt} \rightarrow Met_{cyt}$
r18, Phe	$Phe + 0.33\ ATP_{cyt} \rightarrow Phe_{cyt}$
r19, Pro	$Pro + 0.33\ ATP_{cyt} \rightarrow Pro_{cyt}$
r20, Ser	$Ser + 0.33\ ATP_{cyt} \rightarrow Ser_{cyt}$
r21, Thr	$Thr + 0.33\ ATP_{cyt} \rightarrow Thr_{cyt}$
r22, Trp	$Trp + 0.33\ ATP_{cyt} \rightarrow Trp_{cyt}$
r23, Tyr	$Tyr + 0.33\ ATP_{cyt} \rightarrow Tyr_{cyt}$

Table 8.2: Reactions included on the metabolic model, adapted from Lohr et al in [1] **(Continued)**.

	Glycolysis
r24, G6P	$ATP_{cyt} + Glc_{cyt} \rightarrow G6P_{cyt}$
r25, F6P	$G6P_{cyt} \leftrightarrow F6P_{cyt}$
r26, FBP	$F6P_{cyt} + ATP_{cyt} \leftrightarrow FBP_{cyt}$
r27, DHAP	$FBP_{cyt} \leftrightarrow GAP_{cyt} + DHAP_{cyt}$
r28, GAP	$DHAP_{cyt} \leftrightarrow GAP_{cyt}$
r29, PG	$GAP_{cyt} \leftrightarrow NADH_{cyt} + ATP_{cyt} + PG_{cyt}$
r30, PEP	$PG_{cyt} \leftrightarrow PEP_{cyt}$
r31, PEP_Pyr	$PEP_{cyt} \rightarrow Pyr_{cyt} + ATP_{cyt}$
r32, Pyr _Lac	$Pyr_{cyt} + NADH_{cyt} \leftrightarrow Lac_{cyt}$
r33, PDH	$Pyr_{mit} + CoA_{mit} \rightarrow CO_{2,mit} + ACoA_{mit} + NADH_{mit}$
	Pentose phosphate pathway
r34, R5P	$G6P_{cyt} \rightarrow R5P_{cyt} + CO_{2,cyt} + 2\ NADPH_{cyt}$
	TCA cycle
r35, OAA	$Mal_{cyt} \leftrightarrow NADH_{cyt} + OAA_{cyt}$
r36, Cit	$OAA_{cyt} + ACoA_{cyt} \leftrightarrow Cit_{cyt} + CoA_{cyt}$
r37, Fum_Mal	$Fum_{cyt} \leftrightarrow Mal_{cyt}$
r38, Cit_{mito}	$ACoA_{mit} + OAA_{mit} \rightarrow Cit_{mit} + CoA_{mit}$
r39, OAA_{mito}	$Mal_{mit} \leftrightarrow OAA_{mit} + NADH_{mit}$
r40, Fum_Mal_{mito}	$Fum_{mit} \leftrightarrow Mal_{mit}$
r41, $SCoA_{mito}$	$aKG_{mit} + CoA_{mit} \rightarrow SCoA_{mit} + CO_{2,mit} + NADH_{mit}$
r42, Fum_{mito}	$SCoA_{mit} \leftrightarrow Fum_{mit} + ATP_{mit} + FADH2mit + CoA_{mit}$
r43, aKG_{mito}	$Cit_{mit} \leftrightarrow aKG_{mit} + CO_{2,mit} + NADH_{mit}$
	Anaplerosis
r44, Ana_PyrI	$Mal_{cyt} \rightarrow Pyr_{cyt} + CO_{2,cyt} + NADPH_{cyt}$
r45, Ana_PyrII	$Mal_{mit} \leftrightarrow CO_{2,mit} + Pyr_{mit} + NADPH_{mit}$
r46, PC	$CO_{2,mit} + Pyr_{mit} + ATP_{mit} \rightarrow OAA_{mit}$

Table 8.2: Reactions included on the metabolic model, adapted from Lohr et al in [1] **(Continued)**.

	Amino acid catabolism
r47, GDH	$Glu_{mit} \leftrightarrow aKG_{mit} + Amm_{mit} + NADPH_{mit}$
r48, GS	$Gln_{cyt} \leftrightarrow ATP_{cyt} + Glu_{cyt} + Amm_{cyt}$
r49, Ala_{cat}	$Ala_{cyt} + aKG_{cyt} \leftrightarrow Pyr_{cyt} + Glu_{cyt}$
r50, Asn_Asp	$Asn_{cyt} \leftrightarrow Asp_{cyt} + Amm_{cyt}$
r51, His_{cat}	$His_{cyt} \rightarrow CO_{2,cyt} + Glu_{cyt} + 2\ Amm_{cyt}$
r52, Ile_{cat}	$Ile_{cyt} + ATP_{mit} + aKG_{cyt} + 2\ CoA_{mit} \rightarrow SCoA_{mit} + ACoA_{mit} + NADH_{mit} + FADH2_{mit} + Glu_{cyt}$
r53, Leu_{cat}	$Leu_{cyt} + 2\ ATP_{mit} + aKG_{cyt} + 3\ CoA_{mit} \rightarrow 3\ ACoA_{mit} + NADH_{mit} + FADH2_{mit} + Glu_{cyt}$
r54, Lys_{cat}	$Lys_{cyt} + NADPH_{cyt} + 2\ aKG_{cyt} + 2\ CoA_{mit} \rightarrow 2\ CO_{2,mit} + 2\ ACoA_{mit} + 2\ NADH_{cyt} + 2\ NADH_{mit} + FADH2_{mit} + 2\ Glu_{cyt}$
r55, Met_{cat}	$Met_{cyt} + Ser_{cyt} + 3\ ATP_{cyt} + ATP_{mit} + CoA_{mit} \rightarrow SCoA_{mit} + Cys_{cyt} + NADH_{mit} + CO_{2,cyt} + Amm_{cyt}$
r56, Phe_{cat}	$O_{2,cyt} + Phe_{cyt} + NADPH_{cyt} \rightarrow Tyr_{cyt}$
r57, Pro_{cat}	$Pro_{cyt} \leftrightarrow 2\ NADH_{cyt} + Glu_{cyt}$
r58, Thr_{cat}	$Thr_{cyt} \rightarrow CO_{2,mit} + Amm_{mit} + NADH_{cyt} + Pyr_{mit} + NADH_{mit} + FADH2_{mit}$
r59, Trp_{cat}	$3\ O_{2,cyt} + Trp_{cyt} + NADPH_{cyt} + 2\ CoA_{mit} \rightarrow Ala_{cyt} + 2\ CO_{2,mit} + 2\ ACoA_{mit} + NADH_{cyt} + 2\ NADH_{mit} + 2\ CO_{2,cyt} + FADH2_{mit} + Amm_{cyt}$
r60, Val_{cat}	$Val_{cyt} + ATP_{mit} + aKG_{cyt} + CoA_{mit} \rightarrow SCoA_{mit} + 3\ NADH_{mit} + CO_{2,cyt} + FADH2_{mit} + Glu_{cyt}$
r61, Tyr_{cat}	$2\ O_{2,cyt} + Tyr_{cyt} + aKG_{cyt} + 2\ CoA_{cyt} \rightarrow Fum_{cyt} + CO_{2,cyt} + 2\ ACoA_{cyt} + Glu_{cyt}$
r62, Ser_{cat}	$Ser_{cyt} \leftrightarrow Pyr_{cyt} + Amm_{cyt}$
r63, Cys_{cat}	$O_{2,cyt} + Cys_{cyt} + aKG_{cyt} \rightarrow Pyr_{cyt} + Glu_{cyt}$
r64, Asp_{cat}	$aKG_{mit} + Asp_{cyt} \leftrightarrow OAA_{mit} + Glu_{cyt}$
r65, Arg_{cat}	$Arg_{cyt} + aKG_{cyt} \rightarrow NADH_{cyt} + 2\ Gln_{cyt} + Urea_{cyt}$

Table 8.2: Reactions included on the metabolic model, adapted from Lohr et al in [1] **(Continued)**.

	MTHF & uric acid synthesis
r66, MTHF_I	$Ser_{cyt} + THF_{cyt} \rightarrow Gly_{cyt} + MTHF_{cyt}$
r67, MTHF_II	$Gly_{cyt} + THF_{cyt} \rightarrow NADH_{cyt} + CO_{2,cyt} + Amm_{cyt} + MTHF_{cyt}$
r68, UricAcid	$Asp_{cyt} + 2\ Gln_{cyt} + Gly_{cyt} + 7\ ATP_{cyt} + CO_{2,cyt} + 2\ MTHF_{cyt} \rightarrow Fum_{cyt} + 2\ Glu_{cyt} + 2\ THF_{cyt} + UricAcid_{cyt}$
	Lipid synthesis
r69, CH_{Lip}	$11\ O_{2,cyt} + 18\ ATP_{cyt} + 18\ ACoA_{cyt} + 27\ NADPH_{cyt} \rightarrow 9\ CO_{2,cyt} + CH + 18\ CoA_{cyt}$
r70, PC_{Lip}	$GAP_{cyt} + 2\ NADHcyt + Ser_{cyt} + 27.6\ ATP_{cyt} + 17.6\ ACoA_{cyt} + 31.2\ NADPH_{cyt} + 4\ MTHF_{cyt} \rightarrow PC + 17.6\ CoA_{cyt} + 4\ THF_{cyt}$
r71, PE_{Lip}	$GAP_{cyt} + 2\ NADH_{cyt} + Ser_{cyt} + 18.6\ ATP_{cyt} + 17.6\ ACoA_{cyt} + 31.2\ NADPH_{cyt} + MTHF_{cyt} \rightarrow PE + 17.6\ CoA_{cyt} + THF_{cyt}$
r72, PS_{Lip}	$GAP_{cyt} + 2\ NADH_{cyt} + Ser_{cyt} + 18.6\ ATP_{cyt} + 17.6\ ACoA_{cyt} + 31.2\ NADPH_{cyt} + 2\ MTHF_{cyt} \rightarrow PS + 17.6\ CoA_{cyt} + 2\ THF_{cyt}$
r73, PGL_{Lip}	$2\ GAP_{cyt} + 4\ NADH_{cyt} + 17.6\ ATP_{cyt} + 17.6\ ACoA_{cyt} + 31.2\ NADPH_{cyt} \rightarrow PGL + 17.6\ CoA_{cyt}$
r74, PI_{Lip}	$G6P_{cyt} + GAP_{cyt} + 2\ NADH_{cyt} + 17.6\ ATP_{cyt} + 17.6\ ACoA_{cyt} + 31.2\ NADPH_{cyt} \rightarrow PI + 17.6\ CoA_{cyt}$
r75, SM_{Lip}	$2\ NADH_{cyt} + 2\ Ser_{cyt} + 27.8\ ATP_{cyt} + 16.8\ ACoA_{cyt} + 29.6\ NADPH_{cyt} + 3\ MTHF_{cyt} \rightarrow SM + 16.8\ CoA_{cyt} + 3\ THF_{cyt}$
r76, DPG_{Lip}	$3\ GAP_{cyt} + 6\ NADH_{cyt} + 35.2\ ATP_{cyt} + 35.2\ ACoA_{cyt} + 62.4\ NADPH_{cyt} \rightarrow 35.2\ CoA_{cyt} + DPG$
	Release rates
r77, ATP_{main}	$ATP_{cyt} \rightarrow$
r78, Lac_{out}	$Lac_{cyt} \rightarrow$
r79, Ala_{out}	$Ala_{cyt} \rightarrow$
r80, $UricAcid_{out}$	$UricAcid_{cyt} \rightarrow$

Table 8.2: Reactions included on the metabolic model, adapted from Lohr et al in [1] **(Continued)**.

	Release rates
r81, $Urea_{out}$	$Urea_{cyt} \rightarrow$
r82, amm_{out}	$Amm_{cyt} \rightarrow$
r83, $CO_{2,\ out}$	$CO2_{cyt} \rightarrow$
r84, Pyr_{out}	$Pyr_{cyt} \rightarrow$

	Transport reactions, oxidative phosphorylation
r85, $NADH_{cyt,\ trans}$	$NADH_{cyt} \leftrightarrow NADH_{mit}$
r86, ATP_{trans}	$ATP_{mit} \leftrightarrow ATP_{cyt}$
r87, $CO_{2,\ trans}$	$CO_{2,cyt} \leftrightarrow CO_{2,mit}$
r88, Mal_{trans}	$Mal_{mit} + Cit_{cyt} \leftrightarrow Mal_{cyt} + Cit_{mit}$
r89, Glu_{trans}	$Glu_{cyt} \leftrightarrow Glu_{mit}$
r90, Pyr_{trans}	$Pyr_{cyt} \leftrightarrow Pyr_{mit}$
r91, aKG_{trans}	$aKG_{mit} \leftrightarrow aKG_{cyt}$
r92, Amm_{trans}	$Amm_{cyt} \leftrightarrow Amm_{mit}$
r93, $FADH_{ox}$	$O_{2,cyt} + 2\ FADH_{2,mit} \rightarrow 3\ ATP_{mit}$
r94, $NADH_{mit,\ trans}$	$NADH_{mit} \leftrightarrow NADPH_{mit}$
r95, $NADH_{ox}$	$O_{2,cyt} + 2\ NADH_{mit} \rightarrow 5\ ATP_{mit}$
r96, $NADH_{cyt,\ trans}$	$NADH_{cyt} \leftrightarrow NADPH_{cyt}$

	Synthesis of macromolecules and biomass
r97, μ	0.552 proteins[a] + 0.263 carbohydrates[b] + 0.131 lipids[c] + 0.023 DNA[d] + 0.031 RNA[e] → biomass

[a]**Proteins** [1g]=955.79 Asp_{cyt} + 1344.29 Ala_{cyt} + 543.16 Gln_{cyt} + 817.99 Glu_{cyt} + 1024.72 Arg_{cyt} + 362.10 Asn_{cyt} + 19.61 Cys_{cyt} + 913.16 Gly_{cyt} + 271.58 His_{cyt} + 240.54 Ile_{cyt} + 588.41 Leu_{cyt} + 724.21 Lys_{cyt} + 191.97 Val_{cyt} + 114.57 Met_{cyt} + 201.91 Phe_{cyt} + 205.97 Pro_{cyt} + 247.44 Ser_{cyt} + 202.07 Thr_{cyt} + 21.17 Trp_{cyt} + 61.87 Tyr_{cyt} + 24046.3 ATP_{cyt}

[b]**Lipids** [1g]= 181 CH + 661.4 PC + 250.3 PE + 90.9 PI + 24.9 PS + 12.6 PGL + 81.4 SM + 26.8 DPG

[c]**DNA** [1g]=3009 $R5P_{cyt}$ + 3912 Asp_{cyt} + 5717 Gln_{cyt} + 2106 $NADH_{cyt}$ + 1505 Gly_{cyt} + 22569 ATP_{cyt} + 903 $NADH_{mit}$ + 2407 Mal_{cyt} + 903 $NADPH_{cyt}$ + 5717 Glu_{cyt} + 5417 $MTHF_{cyt}$ + 5417 THF_{cyt}

[d]**RNA** [1g]= 3020 $R5P_{cyt}$ + 3606 Asp_{cyt} + 6316 Gln_{cyt} + 293 $O_{2,Cyt}$ + 2435 $NADH_{cyt}$ + 1477 Gly_{cyt} + 22614 ATP_{cyt} + 586 $NADH_{mit}$ + 2069 Mal_{cyt} + 2954 $NADPH_{cyt}$ + 6316 Glu_{cyt} + 4431 $MTHF_{cyt}$ + 4431 THF_{cyt}

[e]**Carbohydrates** [1g] = 6172.8 $G6P_{cyt}$ + 21605 ATP_{cyt}