

Investigations on the Application of Complex Cell Models in the Simulation of Bioprocesses

Von der Naturwissenschaftlichen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktorin der Naturwissenschaften

Dr. rer. nat.

genehmigte Dissertation

von

M. Sc. Jin Meng

geboren am 22.04.1979 in Tianjin, China

2012

Referent: Prof. Dr.-Ing. Karl-Heinz Bellgardt

Korreferent: Prof. Dr. rer. nat. Thomas Scheper

Tag der Promotion: 20. 03. 2012

Erklärung

Hierdurch erkläre ich, dass die vorliegende Dissertation selbstständig verfasst und alle benutzten Hilfsmittel sowie evtl. zur Hilfeleistung herangezogene Institutionen vollständig angegeben wurden.

Die Dissertation wurde nicht schon als Diplom- oder ähnliche Prüfungsarbeit verwendet.

Hannover, März 2012

Jin Meng

Acknowledgments

I would like to express my warmest gratitude to all of my colleagues and the people who helped me throughout my PhD study. Special thanks to:

My supervisor, Prof. Dr. Karl-Heinz Bellgardt, who kindly provided me the opportunity to make this dissertation possible, and more importantly introduced me into this “metabolic engineering” field that is filled with challenges and excitement. I thank him for his constant guidance, encouragement, trust and patience all these years. He has been an incredible mentor and friend to me. Without his constant help and support, I cannot imagine I could have finished my research for this dissertation.

I am grateful to Prof. Dr. Thomas Scheper for his support on my research as well as being my co-referee.

I would like to thank Prof. Dr. Jürgen Alves for kindly accept my invitation to be my third examiner.

I would also like to give my appreciation to Dr. Frank Stahl for the inspiration and support of my work.

I would like to thank Dr. Ivo Havlik, Dr. Michael Dors and Martina Weiß for technique assistance and helping me to solve computer problems.

I would like to thank all the colleagues at TCI, Angelika Behnsen, Cornelia Alic, Dr. Johanna G. Walter, Dr. Öznur Kökpınar, Dr. Patrick Lindner, Dr. Ran Chen, Dr. Yanhong Li, Yangxi Zhao, Dr. Friederike Sander, Patrick Jonczyk, Anna Glyk, etc., for providing a warm and comfortable working atmosphere.

I appreciate all the help I received from TCI and Leibniz Hanover University which are too numerous to mention here.

Finally, my heartfelt thank to my parents and my husband for their endless love, encouragement and support. This work could not be done without them.

Abstract

Advances in metabolic engineering in recent years have provided thorough information of cellular metabolism that enables mathematical modeling and *in silico* simulation of metabolic networks. Among various analytical methods, Elementary Pathway Analysis is one of the efficient tools for supporting simulation and optimization of bioprocesses. This thesis is focused on exploring the possibility of using the metabolic pathway analysis in simulation of bioprocesses and characterization of metabolic networks. A series of case studies is performed on metabolic models of *Saccharomyces cerevisiae*.

Firstly, a metabolic model of *Saccharomyces cerevisiae* is established based on available data from literature and inspected with the help of metabolic flux analysis.

Secondly, a new metabolic pathway analysis based method is proposed to estimate and predict the flux coefficients of elementary flux modes for a metabolic model. The method determines an optimal combination of elementary pathways by maximizing the biomass production. This results in a predictive metabolic model that can be used to simulate various bioprocesses. The estimated reaction rates using the new method show good agreements with the experimental data. The method is also used to analyze the metabolic flux distribution on different pathways and its changing trend under varying growth condition. The simulation results provide meaningful information about the metabolic characteristics of the organism and helpful guidance for the design of bioprocess systems.

Finally, visualization software of metabolic networks and pathways is developed, which intuitively demonstrates their structure and composition, and facilitates the comparison the elementary pathways under different circumstances.

In conclusion, the presented method and software provide efficient approaches to analyze, simulate and visualize the metabolic reaction systems. The simulation framework established in this study could also provide beneficial information for designing future experiments or industrial productions to accelerate the development process.

Keywords: metabolic model, flux coefficients of elementary flux modes, visualization of metabolic network

Zusammenfassung

Fortschritte bei der Untersuchung des zellulären Stoffwechsels in den letzten Jahren haben umfangreiche Informationen über den Zellstoffwechsel erbracht, die eine mathematische Modellierung und *in silico* Simulation von metabolischen Netzwerken ermöglichen. Unter den verschiedenen Analysemethoden ist *Elementary Pathway Analysis* (Analyse elementare Stoffwechselwege) ein effizientes Werkzeug für die Simulation und Optimierung von Bioprocessen. Diese Arbeit beschäftigt sich mit Untersuchungen zur Simulation der chemischen Reaktionssysteme und Charakterisierung der metabolischen Netzwerke mittels Stoffwechselweganalyse. Eine Reihe von Fallstudien wird mit den Stoffwechselmodellen von *Saccharomyces cerevisiae* durchgeführt.

Zunächst wird ein Stoffwechselmodell von *Saccharomyces cerevisiae* auf der Grundlage von vorhandenen Daten aus der Literatur aufgestellt und mit Hilfe der Stoffflussanalyse überprüft.

Zweitens wird eine neue auf der Stoffwechselweganalyse basierende Methode vorgeschlagen, um die Flusskoeffizienten von elementaren Flussmoden für ein Stoffwechselmodell zu schätzen bzw. vorherzusagen. Diese Methode bestimmt eine optimale Kombination von elementaren Wegen durch Maximierung der Biomasse-Produktion. Es ergibt ein vorhersagendes Stoffwechselmodell, das zur Simulation der verschiedenartigen Bioprozesse verwandt werden kann. Die mit der neuen Methode geschätzten Reaktionsgeschwindigkeiten zeigen gute Übereinstimmungen mit den experimentellen Daten. Die Methode wird auch eingesetzt, um die Stoffflussverteilung in verschiedenen Stoffwechselwegen und deren Variation unter verschiedenen Wachstumsbedingungen zu analysieren. Die Ergebnisse der Simulation stellen aussagekräftige Informationen über die metabolischen Eigenschaften des Organismus bereit und geben Hinweise auf die optimale Führung von Bioprocessen.

Schließlich wird eine Visualisierungssoftware für metabolische Netzwerke und Stoffwechselwege entwickelt, die intuitiv die Struktur und Zusammensetzung des metabolischen Netzwerks zeigt und den Vergleich der elementaren Wege unter verschiedenen (Wachstums-)Bedingungen erleichtert.

Zusammenfassend bieten die vorgestellten Verfahren und die Software effiziente Ansätze, um die Stoffwechselwege zu analysieren, simulieren und visualisieren.

Das in dieser Arbeit erstellte Simulations-Werkzeug könnte auch Hilfestellungen zur Planung zukünftiger Experimente oder industrieller Produktionen bieten, und dadurch den Entwicklungsprozess beschleunigen.

Stichworte: Stoffwechselmodell, Flusskoeffizienten von elementaren Flussmoden, Visualisierungssoftware von metabolischen Netzwerken

Table of Contents

| | |
|---|-----|
| Acknowledgments..... | i |
| Abstract | ii |
| Zusammenfassung..... | iii |
| List of Symbols | iv |
| Table of Contents | v |
| 1 Introduction..... | 1 |
| 2 Metabolic Model..... | 3 |
| 2.1 Theoretical Background | 3 |
| 2.1.1 Cellular Metabolism..... | 3 |
| 2.1.2 Mathematical Modeling | 7 |
| 2.1.3 Metabolic Flux Analysis | 23 |
| 2.1.4 Metabolic Pathway Analysis..... | 28 |
| 2.2 Stoichiometric Metabolic Model of <i>S. cerevisiae</i> | 33 |
| 2.2.1 Establishing Stoichiometric Metabolic Models | 34 |
| 2.2.2 Improving the Consistency of the Stoichiometric Metabolic Models..... | 46 |
| 2.3 Homeostatic Metabolic Pathway Model | 58 |
| 2.3.1 Determination of Elementary Flux Modes..... | 58 |
| 2.3.2 Elementary Flux Modes of <i>S. cerevisiae</i> | 59 |
| 2.4 Predictive Homeostatic Metabolic Model | 62 |
| 2.4.1 Algorithms for Flux Regulation Coefficients | 62 |
| 2.4.2 Growth Condition Simulations | 63 |
| 2.4.3 Flux Coefficient Changes under Oxygen Limitation | 67 |
| 2.4.4 Contributions and Limitations..... | 72 |
| 3 Visualization of Metabolic Pathways | 75 |
| 3.1 Theoretical Background | 75 |
| 3.1.1 Graphic Representation of Reactions..... | 75 |
| 3.1.2 Generating Layout..... | 76 |
| 3.1.3 Drawing and User's Interaction | 79 |
| 3.1.4 Available Software..... | 80 |
| 3.2 Materials and Methods | 82 |
| 3.2.1 Design of PathView | 82 |
| 3.2.2 Design of GraphView | 83 |

| | | |
|-------|--|-----|
| 3.2.3 | Testing Environment..... | 83 |
| 3.3 | Results and Discussion..... | 84 |
| 3.3.1 | Metabolic Pathways Visualized with PathView Software..... | 84 |
| 3.3.2 | Metabolic Pathways Visualized with GraphView Software..... | 86 |
| 4 | Conclusions..... | 96 |
| 5 | References..... | 98 |
| 6 | Appendixes | 107 |
| 6.1 | Measured Data from Literature | 107 |
| 6.2 | Stoichiometric Matrixes | 107 |
| 6.2.1 | Stoichiometric Matrix \mathbf{Y} of Model 1..... | 108 |
| 6.2.2 | Stoichiometric Matrix \mathbf{Y} of Model 2 α | 109 |
| 6.2.3 | Stoichiometric Matrix \mathbf{Y} of Model 2..... | 110 |
| 6.2.4 | Stoichiometric Matrix \mathbf{Y} of Model 3..... | 111 |
| 6.3 | Overview of the Models | 111 |
| 6.3.1 | Model 2 | 111 |
| 6.3.2 | Model 2A | 114 |
| 6.3.3 | Model 2B..... | 116 |
| 6.3.4 | Model 2C..... | 117 |
| 6.3.5 | Model 2D | 118 |
| 6.3.6 | Model 2E..... | 119 |
| 6.3.7 | Model 2F | 120 |
| 6.4 | Complete Simulation Results | 121 |
| 6.4.1 | Modeling of Biomass Components..... | 121 |
| 6.4.2 | Calibrating ATP Balance | 122 |
| 6.4.3 | Effect of Methyl-FH4 Balance..... | 123 |
| 6.4.4 | Effects of Model Simplification..... | 124 |
| 6.5 | Elementary Flux Mode of Model 3 | 125 |
| | Curriculum Vitae..... | 126 |

1 Introduction

Microorganisms and their cellular metabolism have since long been used for production of organic molecules, for example alcohols, antibiotics and amino acids. Since beginning of last century, researchers have been looking for ways to optimize the production of these bioprocesses. However, for more than half of a century, the product rate optimization was purely relying on ad hoc experiments and random mutation and selection. In 1980s, the development of genetic engineering techniques finally enabled researchers to directly modify the genome and thereby influence the expression of corresponding enzymes. These provide powerful tools to alter the metabolism of an organism [1–4]. Meantime, thanks for the recent advances in sequencing technologies, researchers can finally look at the metabolic network in a more complete picture and pinpoint almost every single metabolic reaction inside the cells [5–7]. But, to find the bottleneck steps in the swamp of metabolic reactions is not trivial. Bioinformatics tools are therefore needed to facilitate the reconstruction of cellular metabolism of different organisms based on information encoded in their genomes [7–11]. Quantitative analysis methods of pathway operations were developed to characterize cellular metabolism and explore the phenotypic capabilities of the organisms [12–15].

To define the cellular physiological state under a certain growth condition, the concept of metabolic flux is used in many quantitative analysis methods. A metabolic flux vector consists of the rates of all reactions in a metabolic network. To date most internal reaction rates can not be directly measured. Although the isotopic tracing technique such as using ^{13}C -labeled substrates [16–19] has provided some hope of more precise reaction rate measurement, or at least of adding some extra constrains, metabolic flux estimation is still largely depending on mathematical modeling under some theoretically assumptions. Depending on the operation level, these analysis tools can be grouped into two families: reaction-level analysis and pathway-level analysis. At reaction level, all reactions are treated as independent units, and their reaction rates are directly determined using metabolic flux analysis (MFA) [19–21] or flux balance analysis (MBA) [22–25]. At pathway level, reactions are organized into biologically “meaningful” pathways, these so called Elementary Modes. These modes are used as the smallest units for the further analysis [26–28]. The latter family is usually called metabolic pathway analysis (MPA). It should be pointed out that even though the

differences between these two families are distinct, these analysis methods are developed from the same mathematical basis discussed in detail later in the text.

Although the earliest MPA can be traced back to 1960s [29], it started to draw people's attention only after the introduction of several elementary pathway extraction methods, namely extreme pathway enumeration [30–32] and elementary modes enumeration [27][33–35]. With these automatic pathway extraction methods, the invariant structures of the cellular metabolism can be identified using only the stoichiometric structure and thermodynamic constraints of reactions [36–38]. This type of analysis has been successfully applied to various organisms to investigate metabolic network structure, robustness, fragility, regulation, metabolic flux vector, and rational strain design [39–49].

In the study presented in this thesis, a new MPA-based metabolic flux vector estimation method is proposed, which simulates and predicts organism's metabolic flux by finding the combination of elementary pathways that maximizes the biomass production. This is based on the hypothesis that, due to the evolutionary pressure, organism will always prioritize the biomass synthesis under various growth conditions. The proposed method results a predictive metabolic model which can be used to simulate a bioprocess under a given condition and predict the metabolic flux. This model is first validated on the simplified metabolic network of *Saccharomyces cerevisiae* while comparing with experimental data from different culture conditions. Then it is used to analyze the changes of metabolic flux distribution on different pathways under varying growth condition. At last, in-house graph visualization software was developed to present the elementary pathways in graphic views.

This thesis is organized as follows: In *Chapter 1 — Introduction*, a brief overview of this study is given. In *Chapter 2 — Metabolic Model*, some background knowledge about metabolic modeling is first introduced, and followed by the description of models used in the study; finally, a new metabolic pathway analysis method is introduced and used to establish predictive metabolic models which are verified using experimental data. In *Chapter 3 — Visualization of Metabolic Pathways*, an overview of common graph visualization method and software is first given; then a new pathway visualization software tool is presented; some preliminary results are also provided and compared with other software tools.

2 Metabolic Model

2.1 Theoretical Background

2.1.1 Cellular Metabolism

2.1.1.1 Metabolism

Metabolism is the collection of biochemical reactions that are behind the organism's growth and reproduction. It allows the cells to maintain their structures, respond to their environment and reproduce itself. To some extent, metabolism equals to life, and is the fundamental difference between a living creature and an inanimate object, the thin line between living and death. Scientifically speaking, metabolism is the biochemical process involved in the conversion of matter and energy within cells of an organism [50]. Usually the transformation of energy in the organism is bidirectional. Thus the related metabolism can be categorized as catabolism and anabolism. Catabolic processes decompose complex molecules, such as carbohydrates, into simpler molecules and produce energy and precursor molecules, while anabolic processes synthesize complex biomolecules by consuming energy and precursor molecules.

Despite the great number of the chemical reactions involved in organisms' metabolism, it is easy to identify series of chemical reactions acting as an assembled line through which a principal chemical is modified or produced. These sequences of chemical reactions are known as metabolic pathways. Metabolism of organisms normally consists of numerous distinct pathways, which are responsible for various catabolic and anabolic processes. A metabolic pathway is enabled by a sequence of enzymes that act as catalysts for each individual reaction. In some sense, the enzymes characterize the metabolism.

If an organism can be seen as a factory, then the metabolic pathways are its production lines and enzymes are the workers. These pathways determine which substances it can be fed on and what products can be expected. For example, green plants can use water, light and CO₂ to create carbohydrate, while animals can only be fed on carbohydrate created by plants. However, the similarity of the basic metabolic pathways and components is rather high even between vastly different species [51]. For

example, glycolysis, aerobic respiration and/or anaerobic fermentation, citric acid cycle and oxidative phosphorylation occur within many living organisms from bacteria to human being. These similarities in metabolic pathways demonstrate the evolutionary connection between species, and also allow the knowledge learned from studying the metabolism of simple organism such as bacteria and yeast to be deduced to higher level organism.

2.1.1.2 Metabolic Flux

Many metabolic pathways together constitute the complex metabolic network of an organism. However the importance of different pathways are different, and even within the same cell, the roles of various pathways may change under varying grow condition. To characterize a complicated metabolic network, the concept of metabolic flux is introduced to quantify the rate of conversion among the metabolites through a pathway [20][52]. The flux of a path is defined by molecule turnover through the whole pathway per unit time under steady state (also called homeostasis, which means the reaction rate can be regarded as constant) [50]. The distribution of metabolic fluxes can reveal the *in vivo* physiological features and structures of the network; it is regarded as a fingerprint of the metabolism [20][21][52]. A direct measurement of metabolic fluxes seems impossible because the reaction rates are time-dependent and the concentration of many metabolites cannot be directly determined. However, under steady state the flux of a biochemical reaction can be deemed as its rate and be calculated by means of a mathematical model based on measurable components of the biochemical interaction network, such as substrates, products, and biomass [53]. In most cases, the uptake or production rate of these substances can be indirectly measured by monitoring their concentration in the culture media. Using these reaction rates as input, the unknown flux of an internal reaction can be determined through MFA and MPA (see below).

The metabolic flux can be regulated by enzymes in the pathway according to the expression of corresponding genes, and by this, the activity of the metabolic pathway is adapted to different growth conditions. The analysis of metabolic pathway structure and flux based on metabolic homeostasis is essential for investigations of the metabolic network model.

2.1.1.3 Metabolism of *Saccharomyces cerevisiae*

Saccharomyces cerevisiae (*S. cerevisiae*), also known as Baker's yeast, Ale yeast, Brewer's yeast and Top-fermenting yeast, is one of the most common species of yeast, and has been widely used for food and alcoholic beverage production. As unicellular fungi, these 5 ~ 10 micrometers wide ovoid cells were one of the most intensively investigated eukaryotic microorganisms in the early history of biotechnology. Since the development of recombinant DNA technologies in the 1980s, it has become the most widely used eukaryotic expression system for the productions of various biological products and also for many different processes within the pharmaceutical industry [54–58]. Meanwhile, because of a number of positive characteristics of this organism, *S. cerevisiae* is an ideal eukaryotic model system in metabolic engineering, systems biology and molecular biology [59–62]. Firstly, *S. cerevisiae* can be easily cultivated because it is robust with short generation time (doubling time 1.25 ~ 2 hours [63] at 30 °C). This makes it possible to swiftly culture and maintains multiple specimen lines at low cost. Secondly, the relative compact genome and metabolic system makes the analysis of metabolic pathways and modification of the genome easier and more predictable. A relative small number of reactions reduces the computation effort and complexity. Thirdly, well-studied genome and protein sequences lay down a good foundation for fully understanding the metabolic network and regulation. With the first completely sequenced eukaryotic genome [64–66], *S. cerevisiae* is an ideal model organism for studying general cell physiology. The *S. cerevisiae* databases, i.e. the *Saccharomyces* Genome Database (SGD) [67] and the Munich Information Center for Protein Sequences (MIPS) [68], provide highly useful information for cell modeling and also for the investigation of eukaryotic cell processes at a whole-genome level. All these characteristics make *S. cerevisiae* the most popular organism for metabolic modeling, including the study presented in this thesis.

In unicellular organisms, i.e. *S. cerevisiae*, catabolism and anabolism occur in a single cell. The central catabolic pathways of *S. cerevisiae* include Embden-Meyerhof-Parnas Pathway (EMP, also known as glycolysis), Pentose Phosphate Pathway (PPP) and the Tricarboxylic Acid Cycle (TCA Cycle, also known as Citric Acid Cycle, Krebs Cycle) and produce precursor metabolites and important small molecular, such as ATP and NAD(P)H, which are used in anabolic metabolism. The anabolic processes mainly involve the synthesis of macromolecular cellular components including proteins,

nucleic acids (deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)), cellular structures, and lipids [50].

The metabolic flux changes of *S. cerevisiae* are also well studied [69–71]. As a facultative anaerobic organism *S. cerevisiae* can grow in both metabolic models, respiration and fermentation, according to different oxygen supply condition [71–73]. The aerobic growth consumes substrate glucose and produces a large amount of energy using oxygen as a final electron acceptor of the mitochondrial respiratory chain by oxidative phosphorylation. When oxygen supply is limited or absent, glucose can be metabolized through an alternate pathway to produce ethanol. Glucose is also reduced to glycerol to maintain the redox balance, which results in the formation of significantly less energy and different end products [74][75]. The mechanisms of NAD^+/NADH conversion in respiratory and fermentative metabolism are illustrated in Fig. 2.1. It should be pointed out that, when oxygen availability is restricted, *S. cerevisiae* tends to use oxygen preferentially for oxidation of assimilatory NADH because glycerol production leads to net hydrolysis of ATP and loss of carbon [60][76].

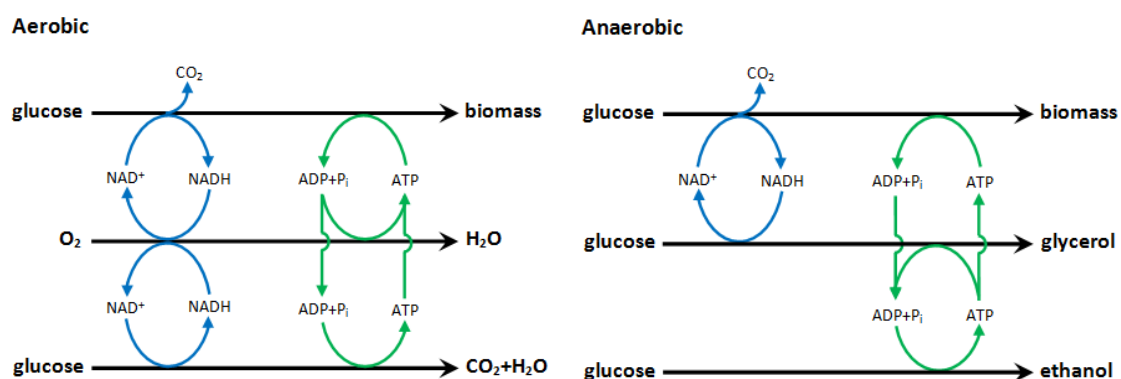


Fig. 2.1 Schematic mechanism of NAD^+/NADH conversion in respiratory (left) and fermentative (right) metabolism of *S. cerevisiae*.

Under the condition of sufficient oxygen availability, the pure respiratory metabolism of *S. cerevisiae* on glucose only exists when the supply of sugar is restricted to a certain limit. Above this limit a mixed respiratory-fermentative metabolism will occur (Crabtree effect [77]), which is partly because excess glucose represses the synthesis of respiratory enzymes [78–80]: excess of repressive sugar causes complex transcriptional regulation and limits respiratory capacity, which leads to higher alcoholic fermentation [81]. Ethanol can also be consumed as carbon source and

provide energy through oxidation in the respiration chain, which is not considered in this thesis.

2.1.2 Mathematical Modeling

In this chapter, some basic concepts of metabolic modeling are introduced. Firstly, the concept of bioprocess is explained and its impact on biological modeling is discussed. Secondly, the principle of the mass balance equation is introduced to describe the mass exchanges in bioreactors in a mathematical language. Thirdly, the homogenous model of a bioreactor is extended to a structured model where the cell mass is divided into several intracellular compounds and functional groups. Fourthly, a concrete example is given to demonstrate how to convert a metabolic network to a stoichiometric model. Finally, a more complete roadmap of mathematical modeling is summarized.

2.1.2.1 Bioprocesses and Mathematical Modeling

As mentioned above, living cells and their metabolism can be used to obtain a desired product, such as alcohol or protein. This procedure is called bioprocess and can be seen as the foundation of biotechnology [82–84]. In a bioprocess, organisms' metabolism is processed in a controlled manner to transform the given raw materials to the desired end-products. The cells are normally cultured in a relative stable artificial environment, like in bioreactors, to maximize desired output product. The typical reagents and products in the reactor can be liquids, gases and solid particles. Thus, for mathematical modeling, the bioreactor is usually divided into liquid phase, gas phase and solid phase. In most cases, because the size of cells is very small, the entire population of cells in a bioreactor is treated as a homogenous system, and the corpuscular character of cells is ignored. This makes it possible to quantify biomass or cell constituents using a simple concentration expression. However in some cases, especially when cells grow in flocks, or are attached to surfaces as biofilm, the system is evidently not a homogenous one, more complex models need to be considered to simulate the system. In this thesis, the latter situation is not covered (for more information, readers are referred to [85–89]). The bioprocess in the reactor is a dynamic system and the reactions can be at transient state or steady state. The culture in the

bioreactor can be aerobic or anaerobic, and the conversion of energy and mass under certain culture conditions is crucial for the production rates in the bioprocess.

The industrial interests of maximizing the product yield and productivity of fermentation processes have encouraged researchers to explore various ways to optimize the production. The traditional methods for raising the product yield usually focus on improving the breed conditions in the cultivation processes by trial and error or based on simple process models. With the development of gene technology, the gene-altered organisms have become an additional attracting approach [90–93]. Nevertheless, all these methods require a good understanding of the metabolism network of the organisms to eliminate numerous possible options and locate the bottleneck of a metabolic pathway. Mathematical modeling of the cellular metabolism naturally becomes a central step in the rapidly developing metabolic engineering. Through these models, researcher can not only interpret metabolic phenotypes of organisms but also predict experimental results [52][94–102].

Early mathematical models of metabolism only focused on the external fluxes of uptake and secretion and ignored intracellular reactions, the conversion of substrates to products was normally summarized in a single equation. Even though such black box models [82] have advantages (for calculation of mass balances, elemental balances and process variables like yields and productivities), the metabolic actions inside the black box have got more and more attention nowadays, because they play critical roles for improvement and regulation of cell metabolism [18][20][50]. The modern metabolic modeling normally refers to stoichiometric models of metabolism which describe the internal reaction network and the reaction stoichiometry in the system. They can integrate available biochemical information of organisms and are widely and frequently used in metabolic studies, especially for large networks. One of the first applications of a stoichiometric model for quantitative analysis of metabolic fluxes has been published by Rabkin and Blum in 1985 [103], which analyzes the structure of rat's hepatocytes metabolic model using experimental measurements under quasi-steady-state conditions. With the rapid development of genomics and bioinformatics, the increasing information of biochemical reactions of organisms offers a powerful data foundation for metabolic network modeling. People have attempted to analyze the metabolic network at genome scale [104–106] or inter-genome scale [107–109]. It can be foreseen that based on the explosion of biochemical, genomic and genetic knowledge, mathematical modeling will

be able to more accurately describe the metabolic network structure of microorganism, predict the cellular behavior under different environmental conditions and guide technical manipulation of culture conditions [50][94][98].

2.1.2.2 Reactor Model, Balance Equations of a CSTR

Based on the operation model of bioreactors, bioprocesses may be categorized as batch, fed batch and continuous (e.g. a continuous stirred-tank reactor (CSTR)). For the latter, a chemostat is commonly used for continuous culture, in which the inflow of substrate and outflow of medium (including cells) from the reactor are constant [110–112].

In an ideal CSTR, the global mass balances in the liquid phase for each component i can be described with the convective transport part (mass inflow, \dot{m}_i^I , and mass outflow, \dot{m}_i^O), the reaction part of mass, \dot{m}_i^R , and the mass flow of gas-liquid transfer, \dot{m}_i^T , as shown in the following equation:

$$\underbrace{\frac{dm_i(t)}{dt}}_{\text{accumulation}} = \underbrace{\dot{m}_i^I(t)}_{\text{inflow}} - \underbrace{\dot{m}_i^O(t)}_{\text{outflow}} + \underbrace{\dot{m}_i^R(t)}_{\text{reaction}} + \underbrace{\dot{m}_i^T(t)}_{\text{gas-liquid transfer}} \quad (2.1)$$

where $m_i(t)$ can be represented by the product of liquid volume $V_L(t)$ and component's concentration $C_i(t)$:

$$m_i(t) = V_L(t) \cdot C_i(t) \quad (2.2)$$

Therefore the mass change rate $\frac{dm_i(t)}{dt}$ is influenced by both the volume change rate $\frac{dV_L(t)}{dt}$ and the concentration change rate $\frac{dC_i(t)}{dt}$ as shown in Eq. (2.3),

$$\frac{dm_i(t)}{dt} = V_L(t) \frac{dC_i(t)}{dt} + C_i(t) \frac{dV_L(t)}{dt} \quad (2.3)$$

The mass inflow with concentration C_i^I and the mass outflow for leaving the reactor with concentration C_i can be given by the volumetric flow rates F^I and F^O respectively as

$$\dot{m}_i^I(t) = F^I(t) \cdot C_i^I(t) \quad (2.4)$$

$$\dot{m}_i^O(t) = F^O(t) \cdot C_i(t) \quad (2.5)$$

and the mass flows of reaction and gas-liquid transfer can be calculated from liquid volume V_L and the volumetric rate of biological reaction R_i or of exchange gas-liquid Q_i as

$$\dot{m}_i^R(t) = R_i(t) \cdot V_L(t) \quad (2.6)$$

$$\dot{m}_i^T(t) = Q_i(t) \cdot V_L(t) \quad (2.7)$$

Introducing Eq. (2.3) – (2.7) into the balance Eq. (2.1) and dividing by the liquid volume V_L gives the model equation of the liquid phase in terms of concentrations

$$\frac{dC_i(t)}{dt} = \frac{F^I(t)}{V_L(t)} C_i^I(t) - \frac{F^O(t)}{V_L(t)} C_i(t) + R_i(t) + Q_i(t) - \frac{C_i(t)}{V_L(t)} \frac{dV_L(t)}{dt} \quad (2.8)$$

where the change of liquid volume can be written as

$$\frac{dV_L(t)}{dt} = F^I(t) - F^O(t) \quad (2.9)$$

Using the definition of the dilution rate

$$D(t) = \frac{F^I(t)}{V_L(t)} \quad (2.10)$$

and substitution of Eq. (2.9), the balance equation Eq. (2.8) can be represented for the concentration of a compound i :

$$\frac{dC_i(t)}{dt} = D(t) \left(C_i^I(t) - C_i(t) \right) + R_i(t) + Q_i(t) \quad (2.11)$$

For non-gaseous substances the gas exchange rate can be neglected.

$$Q_i(t) = 0 \quad (2.12)$$

The volumetric reaction rate R_i , which is the rate of change of unit mass of reactant per unit volume of reactor, can be expressed by a product of cell concentration C_X and the specific rates r_i , which is the rate of change of unit mass of reactant per unit cell mass, as

$$R_i = r_i \cdot C_X \quad (2.13)$$

Therefore the Eq. (2.12) can be written as

$$\frac{dC_i(t)}{dt} = D(t) \left(C_i^I(t) - C_i(t) \right) + r_i C_X(t) \quad (2.14)$$

which can be written with vector notation by using the definitions for the vector of concentrations $\mathbf{C}(t)$ and the vector of specific rates \mathbf{r}

$$\mathbf{C}(t) = \begin{pmatrix} C_1(t) \\ C_2(t) \\ \vdots \end{pmatrix} \quad \mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \end{pmatrix} \quad (2.15)$$

as

$$\frac{d\mathbf{C}(t)}{dt} = D(t)(\mathbf{C}^I(t) - \mathbf{C}(t)) + \mathbf{r}C_X(t) \quad (2.16)$$

where \mathbf{r} is given by the dynamic metabolic model as described in the following.

2.1.2.3 Dynamic Model of Cellular Compounds, Structured Cell Model

The mathematic model discussed so far view a bio-reactor as a homogeneous biomass. There are no intracellular elements or states considered, and the model does not include inner balances. To consider internal structural elements of the cells, the structured cell model is introduced, in which cell mass is structured into several intracellular compounds and functional groups. In a simple structured model the concentrations \mathbf{C} in the liquid phase of the reactor can be separated into the concentration of the abiotic phase for extracellular compounds, \mathbf{C}_Z ,

$$\mathbf{C}_Z = \begin{pmatrix} C_S \\ C_O \\ \vdots \\ C_P \end{pmatrix} \quad (2.17)$$

and the concentration of the biotic phase for intracellular metabolites, storage material, enzymes, RNA, DNA and so on, \mathbf{C}_X ,

$$\mathbf{C}_X = \begin{pmatrix} C_{metabolite} \\ C_{enzyme} \\ \vdots \\ C_{DNA} \end{pmatrix} = (C_{Xi}) \quad (2.18)$$

and thus be written as

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_Z \\ \mathbf{C}_X \end{pmatrix}. \quad (2.19)$$

By partitioning the specific rate vector \mathbf{r} as

$$\mathbf{r} = \begin{pmatrix} \mathbf{r}_Z \\ \mathbf{r}_X \end{pmatrix} \quad (2.20)$$

the balance equation for the biotic phase becomes

$$\frac{d\mathbf{C}_X(t)}{dt} = D(t)(\mathbf{C}_X^I(t) - \mathbf{C}_X(t)) + \mathbf{r}_X C_X(t) \quad (2.21)$$

For the derivation the concentration of the biotic phase at the inflow from the reservoir can be expressed as

$$\mathbf{C}_X^I(t) = C_X^I \mathbf{c}(t) \quad (2.22)$$

and the concentration of the biotic phase in the reactor can be represented as

$$\mathbf{C}_X(t) = C_X \mathbf{c}(t) \quad (2.23)$$

where

$$\mathbf{c}(t) = c_i(t) \quad (2.24)$$

is the vector of the cell mass related intrinsic concentrations

$$c_i = \frac{C_{Xi}}{C_X} \quad (2.25)$$

Here the total cell mass C_X is the sum of all its components C_{Xi} ($\mathbf{1}^T$ is a row vector containing all ones)

$$C_X = \sum_i X_i = \mathbf{1}^T \mathbf{C}_X \quad (2.26)$$

and the sum of all intrinsic components can be expressed as

$$\sum_i c_i = \mathbf{1}^T \mathbf{c} = 1 \quad (2.27)$$

therefore

$$\frac{d(\mathbf{1}^T \mathbf{c})}{dt} = 0 \quad (2.28)$$

By substituting Eqs. (2.22) and (2.23) into Eq. (2.21) the balance for the cellular compounds can be expressed as

$$\frac{dC_X(t)\mathbf{c}(t)}{dt} = C_X(t) \frac{d\mathbf{c}(t)}{dt} + \mathbf{c}(t) \frac{dC_X(t)}{dt} = \mathbf{r}_X C_X(t) + D(t)\mathbf{c}(t)(C_X^I(t) - C_X(t)) \quad (2.29)$$

Multiplying $\mathbf{1}^T$ to each summand in Eq. (2.29) and using the conditions of Eqs. (2.27) and (2.28), the balance equation of the total cell mass can be given:

$$\frac{dC_X(t)}{dt} = \mathbf{1}^T \mathbf{r}_X C_X(t) + D(t)(C_X^I(t) - C_X(t)) \quad (2.30)$$

where the factor

$$\mathbf{1}^T \mathbf{r}_X \equiv \mu(t) \quad (2.31)$$

is identical to the specific growth rate. Using the substitution of Eqs. (2.30) and (2.31) into Eq. (2.29) the balance equation for the intrinsic concentrations can be written as

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{r}_X - \mu(t)\mathbf{c}(t) \quad (2.32)$$

2.1.2.4 Creating a Stoichiometric Metabolic Model

Since Förster et al. reconstructed the first genome-scale metabolic network of *S. cerevisiae* [113], an increasing number of studies have started to focus on analysis of genome-scale metabolic models. A stoichiometric metabolic model is usually established based on organism physiology and biochemical information, which can be obtained from database or literature. Using this information, the metabolic network under investigation can be reconstructed and represented with a stoichiometric matrix in which each column represents a reaction and each row represents a metabolite. Corresponding to the matrix, a reaction rate vector can also be constructed where each component represents the rate of the reaction in the corresponding column of the matrix. With the stoichiometric matrix and vector of reaction rates, the steady state of intracellular substrates, intracellular metabolites, biomass and the products of the organism can then be represented by simple mathematical equations. The procedure of building a stoichiometric model is explained below in more details through an example in Fig. 2.2.

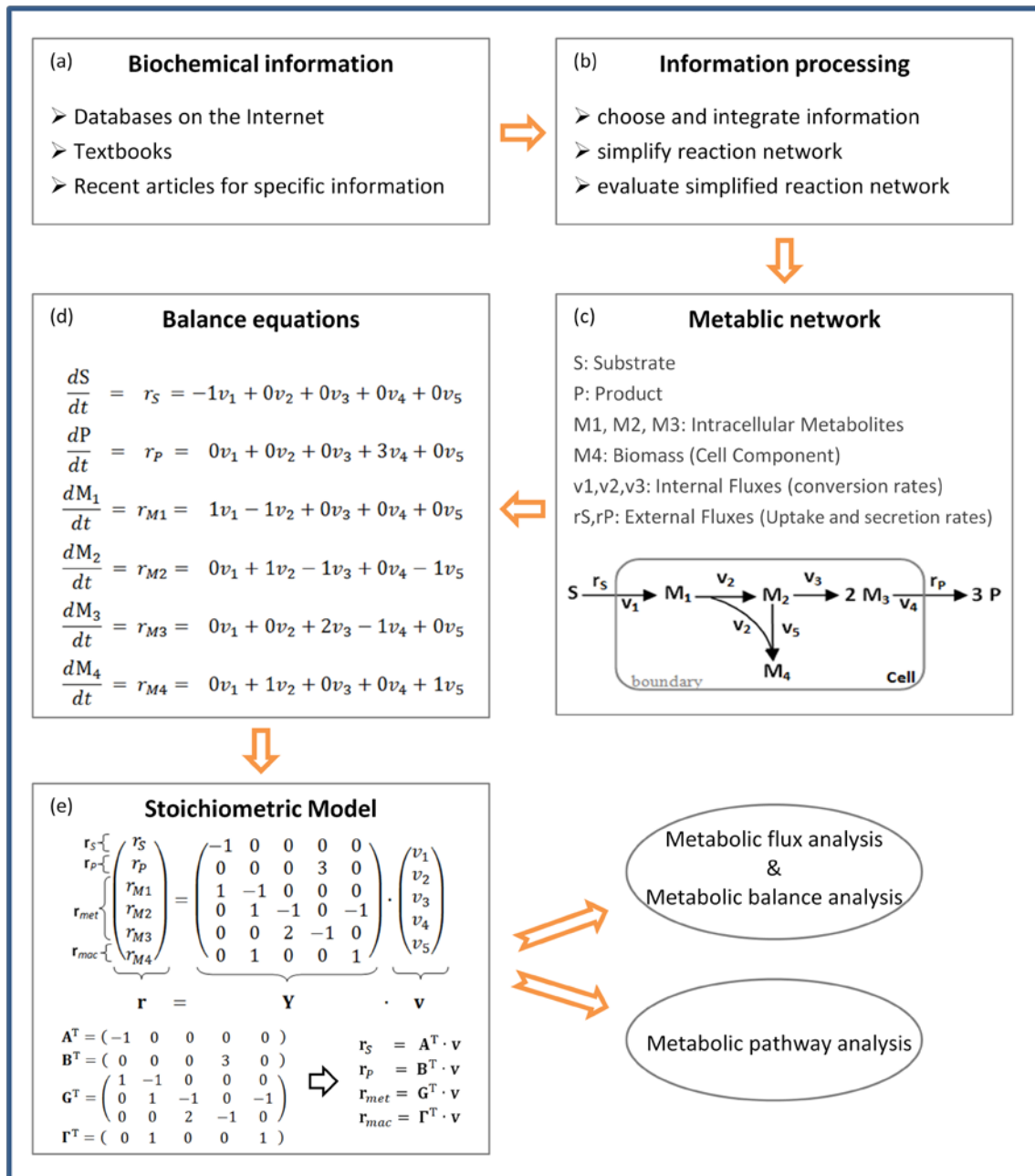
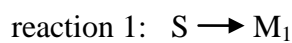
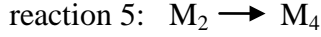


Fig. 2.2 A general scheme for establishing a stoichiometric metabolic model, illustrated by a simple example.

In this example, the small metabolic network (Fig. 2.2 (c)) consists of a single substrate S, a single product P and four metabolites: M₁ ~ M₃ are internal metabolites and M₄ is a cell component and forms a part of the biomass. In the network there are five reactions with rates v₁ ~ v₅:





The change of all the reactants can then be described by the balance equations given in Fig. 2.2(d), where r_i is the net change rate of corresponding reactant i and v_j is the rate of the reaction j . In matrix notation the net reaction rates can be written as

$$\mathbf{r} = \mathbf{Y} \cdot \mathbf{v} \quad (2.33)$$

where vector \mathbf{r} represents all net transformation rates of reactants, vector \mathbf{v} stands for all the rates of intracellular reactions, and the stoichiometric matrix \mathbf{Y} is shown in Fig. 2.2(e). Splitting the vector \mathbf{r} into \mathbf{r}_S , \mathbf{r}_P , \mathbf{r}_{met} and \mathbf{r}_{mac} , respectively, representing the vectors of the net transformation rates of substrate, product, internal metabolite and cell component, the stoichiometric matrix \mathbf{Y} can be partitioned into four parts (as shown in Fig. 2.2(e)), where the matrixes \mathbf{A}^T , \mathbf{B}^T , \mathbf{G}^T and $\mathbf{\Gamma}^T$ respectively represent the stoichiometric coefficients of substrate, product, internal metabolite and cell component. Therefore Eq. (2.33) can be expressed as

$$\mathbf{r}_S = \mathbf{A}^T \cdot \mathbf{v} \quad (2.34)$$

$$\mathbf{r}_P = \mathbf{B}^T \cdot \mathbf{v} \quad (2.35)$$

$$\mathbf{r}_{met} = \mathbf{G}^T \cdot \mathbf{v} \quad (2.36)$$

$$\mathbf{r}_{mac} = \mathbf{\Gamma}^T \cdot \mathbf{v} \quad (2.37)$$

where $\mathbf{r}_s < 0$, $\mathbf{r}_p > 0$, $\mathbf{r}_{met} > 0$ and $\mathbf{r}_{mac} > 0$, i.e. negative reaction rates for substrates, positive reaction rate for products.

During a continuous culture, the net transformation rates of cell component (\mathbf{r}_{mac}) can be regarded as at stationary status, i.e.:

$$\frac{d\mathbf{c}_{mac}}{dt} = \mathbf{r}_{mac} - \mu \cdot \mathbf{c}_{mac} = 0 \quad (2.38)$$

therefore:

$$\mathbf{r}_{mac} = \mu \cdot \mathbf{c}_{mac} \quad (2.39)$$

This equation suggests that the rate of biomass synthesis can be estimated by measuring the dilution rate if \mathbf{c}_{mac} is known. It also indicates that the \mathbf{r}_{mac} is proportional to μ .

The vector of rates \mathbf{v} can be regarded as the flux distribution in the network. Given a stoichiometric matrix, the reaction rates \mathbf{v} can in principle be determined if the

transformation rates of the consumption/production \mathbf{r}_s and \mathbf{r}_p are known, and vice versa. In some cases, the relative reaction rates, yield coefficients, are more important than the absolute values of reaction rates. The yield coefficient of a measured metabolite can be calculated using the following equation

$$Y_{ij} = \frac{v_i}{v_j} \quad (2.40)$$

where v_i is the reaction rate of the metabolite i , v_j is the reaction rates of the metabolite j . The respiratory quotient, RQ,

$$\text{RQ} = \left| \frac{r_{\text{CO}_2}}{r_{\text{O}_2}} \right| \quad (2.41)$$

is a special yield coefficient for characterizing aerobic cultivations. Here the absolute value is used (Note: $r_{\text{O}_2} < 0$). RQ is frequently used as an indicator for the type of metabolism.

It should be pointed out that, although a detailed model including every known reaction into the metabolic network can be created through annotated genome information analysis, it is almost impossible to determine a meaningful metabolic flux by such a model. This is not only because of the enormous data processing effort, but also due to many parallel pathways and futile cycles complicating the network and making it extremely hard to solve, if not impossible. Simplified stoichiometric models created with meaningful fluxes are more feasible for most metabolic network analysis cases. Using the modular approach mentioned in section 2.1.2.5, the metabolic network can be reduced by lumping some parts of it into a single reaction. In this approach the stoichiometric model can be simplified into modules and the reaction details inside each module are hidden. The reduced model shows a clear structure of overall metabolic network, if the lumping is appropriate.

There is no standard way of simplifying a metabolic network, but the simplified model must satisfy the balances of metabolites. It is extremely important to ensure the correct balancing of some small molecules, which are widely involved in the metabolic network, such as ATP and NAD(P)H. During a metabolic process, the generation and consumption of energy and proton carriers are the essential connections between the catabolism and anabolism. For example in the ATP balancing, the parameter P/O ratio

$$P/O = \frac{\text{the number of synthesized ATP molecules}}{\text{the number of consumed oxygen atoms}} \quad (2.42)$$

in the oxidative phosphorylation reaction $[O_2 + (2) \text{ NADH} \rightarrow (2 \text{ P/O}) \text{ ATP}]$ should be carefully tuned to achieve accurate simulation, since the P/O ratio can be different depending on the growth conditions [114–118]. The energy production (ATP production) is in turn coupled with the biosynthesis, because the energy generated in the catabolism must match the requirement in the biosynthesis. Although it is difficult to model the energy consumption in the same detailed way as the energy production (some forms of energy consumption are not exactly tangible), the trend must be that the ATP consumption is proportional to the growth rate. Therefore an empirical approach is usually used, which establishes a relation between the rate of consumed ATP (right part in the following equation) and the rate of produced ATP r_{ATP} of the cell:

$$r_{\text{ATP}} = Y_{\text{ATP}} \cdot \mu + m_{\text{ATP}} \quad (2.43)$$

Here the yield coefficient Y_{ATP} represents the amount of consumed ATP during the cell growth, in another word for biomass synthesis. The maintenance coefficient m_{ATP} summarizes the ATP consumption that is not coupled with the cell growth including the processes for maintenance of cell structure and membrane potentials [119]. In the above equation, the ATP production in catabolism can be estimated very accurately, given accurate measurements of substrate uptake rates and product formation rates. But the two parameters Y_{ATP} and m_{ATP} are usually estimated with empirical approach. This is due to the uncertainty of some cellular processes that vary with the growth rate. These include the decomposition and reconstruction of macromolecules, the compensation of membranes leak and growth coupled conservation.

Finally, before used for metabolic flux analysis as described in next section the sensitivity of the stoichiometric model to small changes in reaction rates should be evaluated, because a well-conditioned matrix should not be influenced by small variations of the measurements during the calculation of flux distribution. One of the most used indicators for evaluating the model sensitivity is the condition number

$$C(\mathbf{G}^T) = \|\mathbf{G}^T\| \cdot \|(\mathbf{G}^T)^{\#}\| \geq 1 \quad (2.44)$$

where $\|\cdot\|$ is a matrix norm and $\#$ denotes the pseudo-inverse of a matrix. The condition number of a model is always higher than or equal one. The lower the number is, the more the model will be of good nature. Taking into account the accuracy of

measurements with common culture techniques, the system is considered as well-conditioned if the condition number is lower than 100 [120][121].

2.1.2.5 Roadmap of Mathematical Modeling

While the principles of metabolic modeling introduced in previous sections seem simple, in real cases it is a complicated process and involves several iterative steps. And a mathematic model of metabolic network normally needs to be tailor made for each organism. In principle a mathematic model should incorporate all the available biochemical information of the studied organism such as stoichiometric information, kinetic information and dynamic regulation on the levels of enzymes and genes. Based on the experience in this work, a typical procedure of modern cellular mathematical modeling is summarized in Fig. 2.3. This roadmap can also be seen as the guideline of this thesis. The main contents of the current study are presented in the order as summarized in Fig. 2.3. The roadmap consists of seven major steps: constructing a detailed reaction network, simplifying the complex network, establishing a stoichiometric metabolic model, extracting a homeostatic metabolic pathway model, generating a predictive homeostatic metabolic model, extending it to a predictive dynamic metabolic model, and finally establishing a bioprocess model by adding the reactor model.

In step 1, a complex reaction network is constructed by listing all the known reaction of an organism. Before, this usually relied on the collection of all potentially available enzymes based on textbooks or literature. Recently, annotated genome information together with transcriptome information was directly used to pinpoint the expressed genes of available enzymes, which has become an attractive alternative way to construct a detailed metabolic model. This may easily generate a reaction network of more than 1000 metabolic reactions. Although it is possible to perform some analysis directly on these genome-scale metabolic models, many sophisticated analyses require the model to be simplified to reduce complexity.

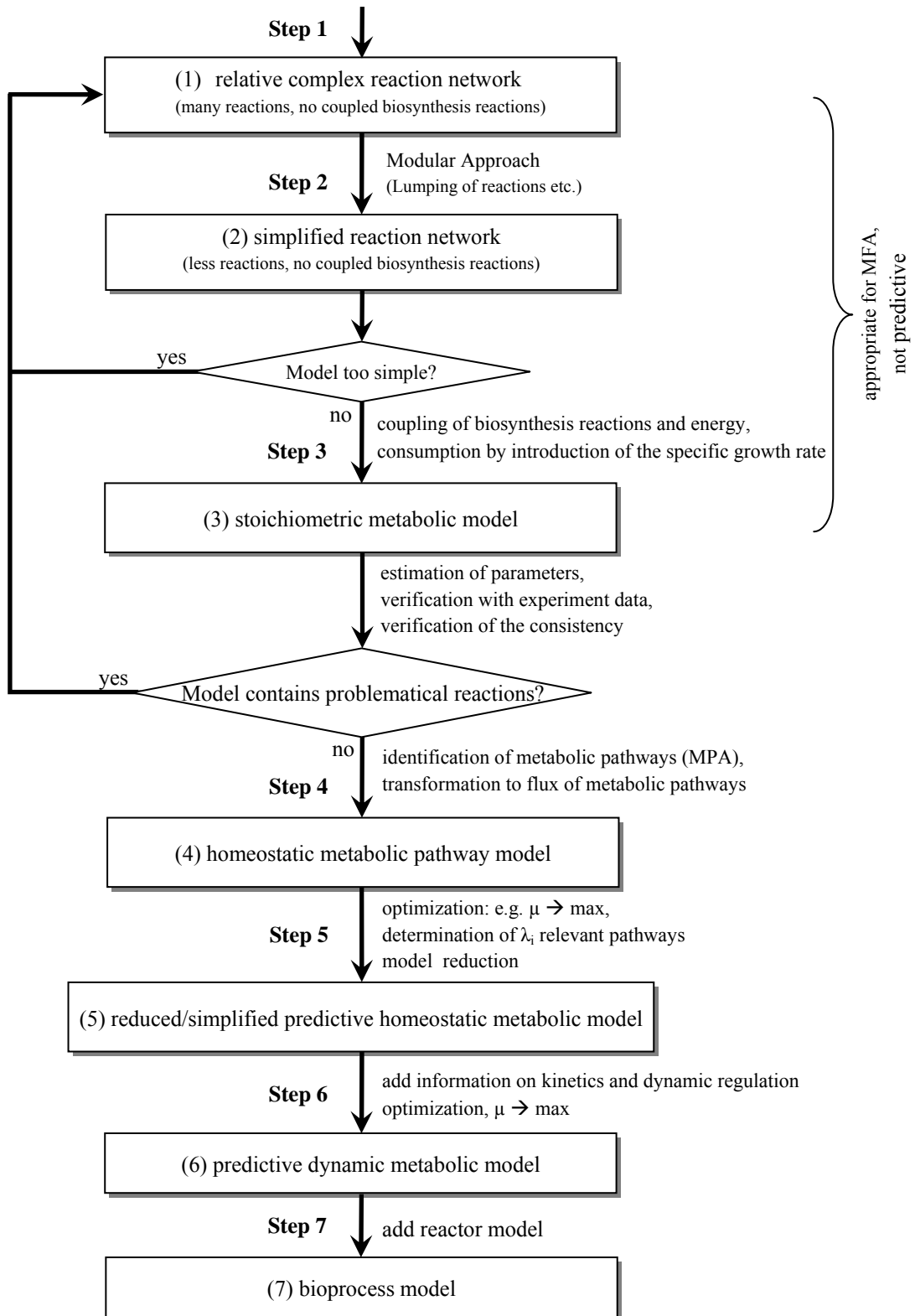


Fig. 2.3 Seven major steps for modeling: Step 1) constructing a detailed reaction network. Step 2) simplifying the complex network. Step 3) establishing a stoichiometric metabolic model. Step 4) extracting a homeostatic metabolic pathway model. Step 5) generating a predictive homeostatic metabolic model. Step 6) extending it to a predictive dynamic metabolic model. Step 7) establishing a bioprocess model by adding the reactor model. Arrows represent the operations of every step, boxes represent the resulted models and diamonds represent model verification.

In step 2, the complicated reaction network is simplified to reduce complexity. One common strategy is to lump some parts of the reaction network into a single reaction (so-called modular approach), for example the biosynthesis of the monomers, i.e. amino acids, nucleotides, fatty acids, lipopolysaccharides, carbohydrates. In ideal cases, the number of reaction should be reduced to 100 ~ 200 or less. It should be realized that the lumping procedure might result in losing detailed stoichiometric information on the overall stoichiometry of metabolism. Therefore the simplified model needs to be carefully validated so that the main characteristic of the network is not influenced, for example the network should still cover the derived range of growth condition. If not, the complex network must be revisited and the lumped reactions need to be unwrapped and re-lumped into a different pattern. The validation and model refining is usually repeated several times to balance the accuracy and simplicity.

In step 3, the simplified reaction network is converted into a stoichiometric metabolic model by coupling of all biosynthesis reactions and their energy consumption to the specific growth rate (as explained in the previous section). In this modeling step, the metabolic reactions of the cell are thought to be at a steady state. This situation can ideally be achieved in continuous culture, reducing the computation complexity and allowing researchers to measure the real reaction rate through experiments. This then enables quantitative testing of the established model. Any systematic derivation from the experimental data could mean there are errors during the simplification step. A new refining cycle then needs to be started again by revisiting the complex reaction network.

In step 4, characteristic metabolic pathways are extracted from the stoichiometric metabolic network (further explained in section 2.1.4). These pathways add more structural restrictions to the reaction network, in addition to the reaction rates. The focus of these models is to study the fluxes in pathways rather than the individual reaction rates, which means less variables. And in step 5, these characteristic metabolic pathways are used to predict the metabolic flux in the organism under a certain growth conditions (further explained in section 2.1.4). In these two steps, the metabolism of an organism is usually also thought to be at a quasi-steady state (homeostatic state). This is because the time constants of catabolic reactions are very small compared to anabolic reactions and there is no accumulation of metabolites. Also enzyme kinetics (including regulation by activation/inhibition) is not required. Although these homeostatic metabolic models can predict or describe many growth phenomena of the organism, and

cover the rapid metabolic regulation (such as in response to changes in glucose, oxygen supply) on the level of enzyme activity by using the optimal strategy (see section 2.1.4), they lack the description of dynamics of the long-term regulation by enzyme induction and repression, which can be observed for example for the pathways of respiration and for gluconeogenesis.

In step 6, by adding information of kinetics about substrate uptake and dynamic regulation (such as the induction and repression of key enzymes), predictive dynamic metabolic models (also called structured models) can be built to predict dynamic growth conditions [122]. In previous models the substrate uptake rates were assumed to be constant. In a dynamic bioprocess, however the substrate is usually limited, which requires the model to incorporate kinetics of substrate uptake. Another important fact that needs to be considered is the regulation of key enzyme by induction and repression. This dynamic regulation normally happens during microbial adaption. One example is switching the substrate in bioprocess. Because some key enzymes may be repressed by one substrate, when switching to another substrate, there is usually a lag phase until the synthesis of key enzymes is induced fully. To modeling such cellular behavior, structured models, such as the metabolic regulator approach (proposed by Bellgardt et al. [123]) or cybernetic models [53], can be used. These methods are based on the following general formulation.

As enzyme can be seen as an internal metabolite, based on Eq. (2.32) the balance equation for the intrinsic concentrations of key enzyme becomes,

$$\frac{dc_{enzyme\ i}}{dt} = r_{enzyme\ i} - \mu \cdot c_{enzyme\ i} \quad (2.45)$$

where $enzyme\ i$ is the enzyme for the i^{th} reaction and $r_{i\ max}$ is proportional to $c_{enzyme\ i}$. It is reasonable to assume the net enzyme synthesis rate r_{enzyme} follows the actual demand r_i . Therefore, the metabolic regulator model for dynamics of enzyme induction/repression can be expressed in Eq. (2.46),

$$\frac{dr_{i\ max}}{dt} = k \cdot r_i - \mu \cdot r_{i\ max} \quad (2.46)$$

and

$$r_i = \min(r_i, r_{i\ max}) \quad (2.47)$$

The constant k must be big enough to enable increase of $r_{i\ max}$ if the enzymatic reaction is rate limiting, i.e. $r_i \equiv r_{i\ min}$. In the adapted state the above model then ensures $r_{i\ max} \gtrsim r_i$.

The Result of adding the above conditions into the homeostatic model is a so called structured model that formally describes induction and repression of key enzymes of metabolic pathways. Unlike the homeostatic models in previous steps, the regulation on reaction level by inhibition/activation can seamlessly be integrated into the mathematic model without further assumptions due to the optimization approach.

In step 7, further information about the reactor model is added to construct a complete mathematic model for bioprocesses. A reactor model describes the dynamic concentration changes in the gas and liquid phase of the reactor and the mass exchanges between gas and liquid phase. The gas phase can be further divided into dispersed gas phase and headspace of the reactor. These changes are usually related to initial conditions, manipulating variables, and reactions microorganisms. The reactor models provide the concentrations of substrates, products and cells as input variables to the models of the biotic phase (results from previous steps). On the other hand, the reactor model also relies on the actual reaction rates as output of the biological models described above [124]. Besides a reactor model, in some special situation, a population model is required to describe an in-homogeneous population by adding one or more independent variables in addition to time. The additional variables are usually mass, volume, or age of a single cell. All the differential equations described in previous sections then become partial differential equations. Such models are usually very complicated as the catabolism also involves cell cycles, which implies the metabolic pathways will be different from those in previous models. Together with the other metabolic model for stoichiometry of growth (from step 5) and the regulation model for metabolic long-term regulation (from step 6), these three parts constitute a complete dynamic bioprocess model that can simulate complicated behaviors of bioprocesses. Further discussion of the regulation model, reactor model and population model is out of the scope of this thesis, interested readers are referred to [53][102][124] for more detailed discussions.

2.1.3 Metabolic Flux Analysis

2.1.3.1 Introduction

As a crucial tool in metabolic engineering, metabolic flux analysis (MFA) is used for over ten years [19][20][125]. With the help of measured experimental data, metabolic flux analysis can depict the distribution of the fluxes along the metabolic pathways using a stoichiometric model according to the mass balance of metabolites [20][125]. MFA makes it possible to evaluate the degree of engagement of each reaction in the metabolic network and obtain comprehensive knowledge of the metabolic state of an organism. Studying the flux changing under various environmental conditions could provide theoretical guidance to optimize the culture condition in the bioreactors.

MFA is based on the assumption that the bioprocess inside the chemostat can be seen as dynamic system in steady state, i.e. the production and consumption of metabolites then stay balanced. As described in section 2.1.2.3, the balance equation of the internal metabolites can be written as:

$$\frac{d\mathbf{c}_{met}}{dt} = \mathbf{r}_{met} - \mu \cdot \mathbf{c}_{met} \quad (2.48)$$

where \mathbf{c}_{met} is the vector of intrinsic concentrations for internal metabolites and \mathbf{r}_{met} their specific rate vector.

At steady state, the concentration of all metabolites is constant, so the above Eq. (2.48) is equal to zero. This model then becomes the foundation of metabolic flux analysis discussed in the rest of this chapter. It is important to notice that, at steady state in a chemostat, the specific growth rate (i.e. biomass formation rate) μ in the cell mass balance is equal to the dilution rate, i.e. $\mu \equiv D$. This implies that the specific growth rate of the cells can be directly controlled by setting a proper dilution rate. In addition, as the concentration of internal metabolites is usually very low, \mathbf{c}_{met} can be set to zero. This indicates that the net formation rate of internal metabolites is zero. The metabolic system then can be treated as a linear algebraic matrix equation and the flux balance for internal metabolites then simplifies to

$$\mathbf{r}_{met} = \mathbf{G}^T \cdot \mathbf{v} = \mathbf{0} \quad (2.49)$$

For a matrix \mathbf{G}^T with K metabolites by L reactions, if $K < L$ (this is mostly the case in metabolic models), Eq. (2.49) is not solvable for \mathbf{v} . Strictly speaking there is no

unique solution for \mathbf{v} , because, from a linear algebra point of view, there are fewer equations than unknowns. However if some reaction rates can be measured, the unknown components of vector \mathbf{v} become less, the linear equation can then be solved, depending on the number of measured rates of independent reactions J . According to linear algebra, if $J = L - K$, Eq. (2.49) is a determined system and there will be a unique solution for \mathbf{v} (given $|\mathbf{G}^T| \neq 0$); if $J > L - K$, the solution will be overdetermined, an approximate solution can be probably found; otherwise, i.e. $J < L - K$, it is an underdetermined system and there will be infinite possible solutions. In the last situation, if there are additional restrictions and bounding conditions for example several reaction rates cannot be negative, an optimal solution may still be found by using linear programming (as discussed below). In the following, it will be further explained how to solve this equation under different situations.

2.1.3.2 Linear Model

First, Eq. (2.49) is reorganized to separate the measured reaction rates \mathbf{v}_m from the other to-be-determined reaction rates \mathbf{v}_c as shown in Eq. (2.50),

$$\mathbf{G}^T \cdot \mathbf{v} = (\mathbf{G}_m^T \ ; \ \mathbf{G}_c^T) \begin{pmatrix} \mathbf{v}_m \\ \vdots \\ \mathbf{v}_c \end{pmatrix} = \mathbf{G}_m^T \mathbf{v}_m + \mathbf{G}_c^T \mathbf{v}_c = 0 \quad (2.50)$$

where \mathbf{G}_m and \mathbf{G}_c are their corresponding coefficient matrices respectively. [20] For the determined system, the matrix \mathbf{G}_c^T is a K by K square matrix, and is usually invertible (if $|\mathbf{G}_c^T| \neq 0$), thus the unknown flux \mathbf{v}_c can be easily calculated through the Eq. (2.51),

$$\mathbf{v}_c = -(\mathbf{G}_c^T)^{-1} \mathbf{G}_m^T \mathbf{v}_m \quad (2.51)$$

For the overdetermined system, the \mathbf{G}_c^T is a $K \times (L - J)$ matrix. This makes the Eq. (2.51) unsolvable, as there are more equations than unknowns ($K > L - J$). But the system can be approximately solved as least square problem by replacing the inverse of \mathbf{G}_c^T with its pseudo-inverse as given in Eq. (2.52),

$$(\mathbf{G}_c^T)^\# = (\mathbf{G}_c \mathbf{G}_c^T)^{-1} \mathbf{G}_c \quad (2.52)$$

It is worth noticing that the measured reaction rates are not necessary to be treated as known variables, instead they can remain unknown, while the measurements are added as extra equations to the balance equation as shown in Eq. (2.53),

$$\begin{pmatrix} \mathbf{v}_m \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{G}_c^T & \end{pmatrix} \mathbf{v} = \mathbf{T} \mathbf{v} \quad (2.53)$$

where \mathbf{I} is an identity matrix and its dimension is equal to the number of measured fluxes J . [20] For a determined system, \mathbf{T} is a L by L square matrix, and for overdetermined systems, \mathbf{T} is a $(J + K) \times L$ matrix with $(J + K > L)$. \mathbf{v} can be calculated by using the pseudo-inverse matrix of \mathbf{T} as shown in Eq. (2.54),

$$\mathbf{v} = \mathbf{T}^\# \mathbf{v} = (\mathbf{T}^T \mathbf{T})^{-1} \mathbf{T}^T \begin{pmatrix} \mathbf{V}_m \\ \mathbf{0} \end{pmatrix} \quad (2.54)$$

However, in \mathbf{v} new values for \mathbf{v}_m will also be calculated. This, in theory, increases the flexibility of the model to tolerate measurement errors.

Eq. (2.54) can be further improved by taking into account the variance of measuring errors. A variance-covariance-matrix \mathbf{F} ,

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_d \end{pmatrix} \quad (2.55)$$

where \mathbf{F}_m is $J \times J$ dimension submatrix for measurement error of fluxes,

\mathbf{F}_d is $K \times K$ dimension submatrix for the error of metabolite balance equations,

is used as weighting factor of the least square solution as shown in Eq. (2.56) [20][126]

$$\mathbf{v} = (\mathbf{T}^T \mathbf{F}^{-1} \mathbf{T})^{-1} \mathbf{T}^T \mathbf{F}^{-1} \begin{pmatrix} \mathbf{V}_m \\ \mathbf{0} \end{pmatrix} \quad (2.56)$$

The measurement error of fluxes is inevitable in reality, thus the latter two methods provide a more thorough foundation for the metabolic flux analysis in circumstances where a relative large number of measurements were performed.

For the underdetermined system, \mathbf{v} has infinite solutions if only the mass balance equation is given. However, in most metabolic models, several additional constraints, may apply to the system and result in a bounded solution space of all feasible fluxes. The most common examples of additional constraints are thermodynamic constraints, which allow some reactions only to proceed in the appropriate direction, i.e. they are irreversible. These constraints can be expressed by inequality as in Eq. (2.57),

$$v_i \geq 0 \quad (2.57)$$

where v_i is the reaction rate of reaction i . Yet, the bounded solution space under the condition of the Eq. (2.49) and Eq. (2.57) that describes the capacity of the metabolic network, cannot provide a unique solution of \mathbf{v} . Nevertheless, an optimum solution of the system can be searched for by using a suitable objective function, for instance, maximizing a certain product. These types of optimization problems can be solved with

the help of a linear programming method. The above process is known as flux balance analysis. It is evident, that the predictive capability of this approach relies on whether the defined objective function could adequately represent the cellular metabolism under the given growth condition [127].

It should be mentioned that flux balance analysis identifies only one optimal solution while alternative optimal solutions or suboptimal solutions can exist. In general, flux balance analysis can calculate metabolic flux vectors based on limited experimental data, but requires specification of objective functions for cellular metabolism. The more reaction rates can be measured, the more accurate the flux vector estimation will be. However, the metabolic flux vector may not be unique.

2.1.3.3 Error Correction

As mentioned above, the experimental data usually contains measurement noise (random errors), and in some cases even systematic errors. While random error can be overcome by repeating measurement multi-times, systematic errors must be identified before the values are used for MFA, as they may cause severe deviation on the estimated flux and ruin the metabolic model. One way of finding measurement errors is to check the material balances and data consistency. It is evident that in the metabolic system, mass is conserved in the overall reaction of substrates to metabolic products and biomass. In a black box model (if the whole metabolic network is lumped into one reaction: substances \rightarrow products + biomass), the elements, for example carbon, entering the system via the substrates must equal the elements flowing out of the system via products and biomass. A carbon balance can be represented by Eq. (2.58)

$$1 + \sum_{i=1}^U h_{P,i} Y_{XP_i} - \sum_{i=1}^N h_{S,i} Y_{XS_i} = 0 \quad (2.58)$$

where U and N are the numbers of the metabolic products and substrates, and $h_{P,i}$ and $h_{S,i}$ represent their carbon content (C-mol/mol). Y_{XP_i} and Y_{XS_i} , respectively, stand for the yield coefficients of the i^{th} metabolic product and substrate in relation to the specific rate of biomass formation. Eq. (2.58) can be multiplied with μ . According to the definition of the yield coefficients, the carbon balance can be rewritten in the form of specific rates as Eq. (2.59):

$$\mu + \sum_{i=1}^U h_{P,i} r_{P,i} - \sum_{i=1}^N h_{S,i} r_{S,i} = 0 \quad (2.59)$$

Using the above Eq. (2.59), the consistency of experimental data can be conveniently checked, because an inconsistency will result in unbalance between the carbon in the substrates and the carbon in the biomass and metabolic products. For example Nielsen and Villadsen [128] discovered the classical data of von Meyenburg [129] did not meet the carbon balance. Through a more elaborate data analysis, they revealed that the measurements of ethanol were inaccurate, which might be caused by evaporation due to intense aeration of the bioreactor.

While looking at the magnitude of the residual can suggest whether a large error(s) is present in the measurements, it cannot indicate where the error comes from. A procedure of measurement elimination is suggested in [20], to quickly determine the probable source of a systematic error. This procedure is carried out by eliminating one measurement at a time from the given set of data and using the rest to compute a new residual. As it is an overdetermined system, the Eq. (2.54) can still be solved. If the elemental residual is reduced significantly when eliminating a certain measurement, it strongly suggests a presence of an unignorable systematic error in the measurement that was eliminated.

Besides the material balances, a certain interdependence may exist between two rates due to some fundamental biochemical and physiological relationships. Some of these linear correlations are very useful to correlate growth data, especially those obtained from steady state continuous cultures. One example is the energy and electrons balance that involves balancing ATP, NADH and NADPH production and consumption. In a simple aerobic process without metabolite formation, the reaction uptake rate of glucose and oxygen and the production rate of CO₂ can be expressed as:

$$r_{glc} = Y_{xs}^{true} \mu + m_s \quad (2.60)$$

$$r_{O_2} = Y_{xo}^{true} \mu + m_o \quad (2.61)$$

$$r_{CO_2} = Y_{xc}^{true} \mu + m_c \quad (2.62)$$

In the above equations, the coefficient Y is the true yield rate of each element for biomass production and m is the coefficient that represents the element consumed or produced for maintenance purposes. In a bioprocess, these three equations are not

independent from each other because the energy and electrons balance which means the required ATP and NADPH for biomass synthesis have to be supplied by the catabolic pathways, and the NADH formed in the biosynthetic reactions has to be reoxidized by transfer of electrons to oxygen via the electron transport chain. Careful investigations reveal there is a linear correlation between these three reaction rates that can be summarized as follow:

$$r_{glc} = (a + 1.261)\mu + b \quad (2.63)$$

$$r_{O_2} = (a + 0.229)\mu + b \quad (2.64)$$

$$r_{CO_2} = (a + 0.261)\mu + b \quad (2.65)$$

where

$$a = \frac{Y - 0.458 P/O}{0.667 + 2 P/O} \quad (2.66)$$

$$b = \frac{m_{ATP}}{0.667 + 2 P/O} \quad (2.67)$$

The derivation steps can be found in [20] chapter 3.4. Note: Eqs. (2.63) – (2.67) are based on the metabolic model for *P. chrysogenum* as presented by Nielsen [130]. Here they serve only as an example to demonstrate the linear relationship between reaction rates. For different models the coefficients may be different but there will be similar linear relationships.

The above linear equations can be used not only to correlate experimental data, but also to evaluate the key parameters Y_{ATP} and the P/O ratio. The latter is usually done by estimating the true yield and maintenance coefficients of Eqs. (2.60) – (2.62) (by linearly regressing experimental data for the glucose, oxygen, and carbon dioxide rates against the specific growth rate), and then substitute the values of a and b into Eqs. (2.66) and (2.67). More examples of error correction using energy and electrons balance can be found in [20] chapter 3.4.

2.1.4 Metabolic Pathway Analysis

2.1.4.1 Introduction

In contrast to the metabolic flux analysis discussed above, metabolic pathway analysis (MPA) does not focus on providing a unique flux vector of the metabolic

model under a certain steady state, but identifies the topology and properties of the cellular metabolism networks. The fundamental step of MPA is to extract biologically “meaningful” pathways from an intricate metabolic network. These so-called elementary flux modes/pathways represent a set of all possible routes that involve the minimum number of enzyme to achieve steady states and can be used as smallest units for the further calculation, such as identifying the structure of a metabolic network or determining the metabolic flux vector of the specific cellular physiological state under a given growth condition [33][131–134]. For MPA, the calculation is only based on the stoichiometric Eq. (2.49) and thermodynamic constraints (Eq. (2.57)) without requiring additional measurement of any flux rates or imposing any objective function for cellular metabolism. With the development of the related disciplines, the studies of metabolic pathways has gradually become a very important tool for study of metabolic network structure, robustness, regulation, metabolic flux vector, and rational strain design [45][133–136].

2.1.4.2 Determine Elementary Flux Modes

Finding simplest biologically meaningful pathways from a metabolic network is an essential step for MPA. Although the literal meaning sounds rational, its mathematical definition is non-trivial. First of all, the elementary pathway should satisfy the stoichiometric Eq. (2.49) and inequality constraints (Eq. (2.57)) i.e. some reactions are irreversible. However, as discussed in previous sections, the number of all possible solutions is usually infinite. Thus, additional constraints such as non-decomposability and systematic independence are required to select a finite set of simplest solutions. Different definitions of “simplest” result in different applications of these additional constraints, thereby leads to different techniques for metabolic pathway analysis, for example elementary mode analysis [27] and extreme pathway analysis [28].

In elementary mode analysis after Schuster and Hilgetag [27], a flux \mathbf{v}^* is called elementary mode if and only if \mathbf{v}^* fulfills the simplicity condition, i.e. there exists no couple of vector \mathbf{v}' and \mathbf{v}'' with the following properties:

- (i) \mathbf{v}^* is a non-negative linear combination of \mathbf{v}' and \mathbf{v}''
- (ii) \mathbf{v}' , \mathbf{v}'' are the solution to Eq. (2.49) and inequality (Eq. (2.57))
- (iii) both \mathbf{v}' and \mathbf{v}'' contain at least the same number of zero elements as \mathbf{v}^* and at least one of them contains more zero elements than \mathbf{v}^* .

- (iv) for all indices i corresponding to boundary reactions, the components are not of opposite sign.

In another words, an elementary flux mode is the solution for a possible route with maximum number of zeros. The more zeros, the less number of reactions are involved in the route. As a cellular system usually consists of many possible routes, several elementary modes satisfying the above conditions will be found.

Mathematically speaking, the solutions of Eq. (2.49) and inequality (Eq. (2.57)) constitute a convex polyhedral cone [137]. And the elementary flux modes of an irreversible metabolic network (all reactions are irreversible) are the edges (generating vectors) of the cone [27]. This converts the elementary mode searching problem to enumerating extreme rays of polyhedral cones, which has been used as the foundation of many computer algorithms to extract the complete set of elementary modes.

Extreme pathway analysis utilizes the same criteria as elementary mode analysis, but in addition, it imposes a systematic independence constraint, which requires an extreme pathways cannot be expressed as a non-trivial non-negative linear combination of the other pathways [28]. Both extreme pathway analysis and elementary mode analysis will extract the same set of pathways if all the reactions are irreversible, i.e. the cone is located in the positive orthant of n -dimension space (n is the number of reactions). However, if there are reversible reactions, extreme pathway analysis will generate a smaller number of extreme pathways than the number of elementary modes, while all extreme pathways will also be included in the elementary modes. [26] This follows, because reversible reactions extend the convex cone to the negative orthants, in another words, the new convex is created by merging two half-cones in the positive and negative orthants. The problem comes from the shared edges between the two half-cones, as they are now located in the middle of faces of the new cone, or even inside the cone if more than one reversible reactions are related with them. By definition, extreme pathway analysis will discard these pathways as extreme pathways, however for elementary mode analysis they will still be kept in the results. Fig. 2.4 gives a symbolic geometric illustration of difference of flux analysis, elementary modes and extreme pathways. Let the flux cone (see Fig. 2.4A) represents the null space of Eq. (2.49), i.e. all possible pathways that satisfy the stoichiometry balance, then the five vectors on the edges of the cone represent five extreme pathways (as the circles shown in Fig. 2.4A and C). However elementary modes (see circles and squares in Fig. 2.4A

and C) are not guaranteed to lie on the face and edges of the cone, sometimes they can be inside the cone (as the squares shown in Fig. 2.4A and C). Metabolic flux analysis normally results in a single vector that lies anywhere in the cone (e.g. the green star shown in Fig. 2.4B). Theoretically, extreme pathway analysis is mathematically more sound as it defines the minimum sets of pathways to span the convex cone space, however, in practice, elementary modes could make further analysis much easier, as it is easier to interpret a single elementary pathway rather than a linear combination of two or several extreme pathways. [33]

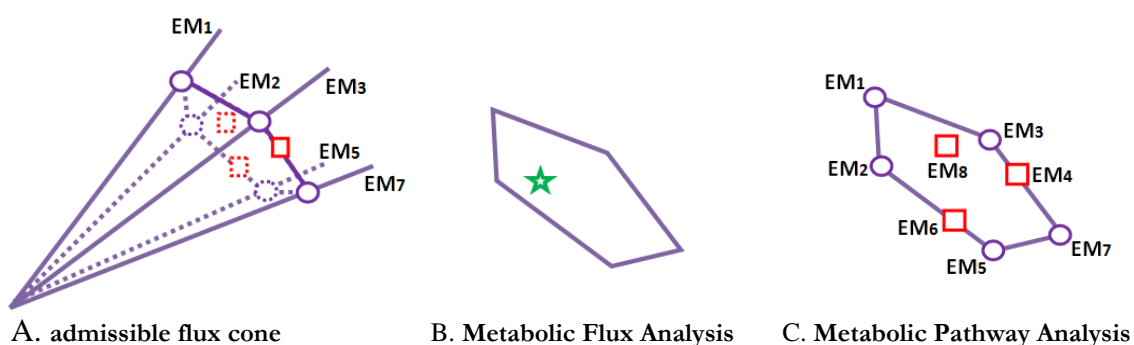


Fig. 2.4 Geometric interpretation of flux cone. Green star: a flux from MFA; purple circles and red squares: elementary modes; purple circles: extreme pathways.

Algorithms and Software

Although the definition of an elementary mode seems intuitive, finding all the EMs in a metabolic system is not straightforward. Several methods based on extreme ray enumeration algorithms from computational geometry have been proposed. Among them, the double description method is the most used method in many available software. A detailed introduction of this algorithm can be found in [138]. In most applications, the performance and memory requirements are both critical aspects. A nullspace initial matrix proposed by Wagner [139] was proved to be a successful strategy to simplify the algorithm and improve performance. Subsequently, binary vectors for flux value storing and rank computations for testing elementarity were proposed by Gagneur and Klamt [140] and Klamt et al. [141], which boosted the performance even further. More recent reviews of variants of the double description method can be found in [141][142].

Based on various EM enumeration algorithms, several available tools for extracting EMs are summarized in Tab. 2.1.

Tab. 2.1 EM algorithms and software.

| Algorithm | Software | | | Ref. | Remark |
|---------------------|---------------------------------------|--------------|-----------------------|-------|---|
| | Name | Environment | Version & Year | | |
| Original approach | METATOOL | C | 1999 | [143] | |
| | GEPASI & COPASI | C | 2006 | [144] | |
| | FluxAnalyzer | Matlab | 2003 | [26] | |
| Null space approach | SNA (MATHEMATICA) | Mathematica | 2006 | [145] | SNA in Mathematica |
| | METATOOL >=5 | C, Matlab | 2006 | [146] | |
| | YANA& YANAsquare | Java | 2005 | [147] | based on METATOOL; user-friendly interface |
| Binary approach | FluxAnalyzer 5.1 / CellNetAnalyzer | Matlab | 2004 / 2007 | [140] | optionally use METATOOL & efmtool for em |
| Bit pattern tree | efmtool | Java, Matlab | 2008/2009 (v4.7.1) | [142] | |

In this study, only software based on Matlab are tested. During the tests, METATOOL and Efmtool generate different sets of EMs even on the same model. While the results from Efmtool seem most reasonable, METATOOL usually generates more EMs than Efmtool.

2.1.4.3 Applications of Metabolic Pathway Analysis

Firstly, the elementary mode analysis can be used to identify the most efficient pathways from the given substrate (S) to the product (P). It is straightforward to calculate the relative molar yields of the product with an elementary mode i (EM_i) with the equation $Y_{EM_i} = r_p/r_s$ where r_s and r_p is the reaction rate of related reactions within the elementary mode that consume S or produce P. As any steady state can be expressed by the non-negative linear combination of elementary modes, the highest Y_{EM_i} then represents the highest molar yields potential of the underlying metabolic network. One among the first real-biological-system applications was reported in [12] by Liao et al., where elementary mode analysis was used to optimize of the production of 3-deoxy-D-arabino-heptulosonate-7-phosphatate (DAHP) in *E. coli*. Liao et al. first identified the most efficient DAHP producing pathway through elementary mode analysis, then optimized the key reactions by over-expressing the corresponding enzymes. Eventually, the achieved molar yield of DAHP is very close to the theoretically predicted value.

Secondly, the elementary modes can be used to compute the metabolic flux of a steady state under a specific grow condition. As, by definition, any flux \mathbf{v} can be expressed as a weighted sum of all elementary modes, a steady state can be seen as an overlap of several extreme/elementary pathways. To determine the weighting factors,

additional constraints have to be introduced. Similar to the flux balance method discussed in a previous section, the problem can again be seen as an optimization problem where the goal is to achieve optimized biomass production, energy consumption or difference between measured and predicted fluxes. One example is published by Carlson and Srienc [39], who identified four most efficient glucose-to-energy pathways in *E. coli* by studying the weighting factors of elementary modes under varying oxygen levels. In their study, the optimization goal is both biomass-producing efficiency and energy-producing efficiency, as they argued “evolutionary pressures under carbon-limited growth conditions likely select organisms that utilize highly efficient pathways”. The experimental results confirmed that the computed flux vectors under different oxygen limitation agree well with the measured parameters.

Thirdly, metabolic pathway analysis allows better understanding of the metabolic network properties. As the huge number of chemical reactions collapses to individual elementary pathways, it is easier to understand the metabolic network at inter- and intra-pathway level. Within each elementary mode, the effect of one or several knock off genes can be predicted, thereby characterizing the robustness and fragility of the metabolic networks [148][149]. On the other hand, at inter-pathway level, the functionality and efficiency of different pathways can be compared and the key reactions that are involved in several pathways can be identified [150–152]. More important, by comparing the flux distribution in different pathways under different growth conditions, the regulation pattern of the cellular metabolism can be investigated [153–155]. At an even higher level, comparing the elementary pathways between different species will help researchers to find the genetic footprint of evolution and predict the best living condition of organisms or, even further, design new species using gene-altering techniques. With the knowledge of all unique pathways existing in a metabolic network researchers can eliminate inefficient pathways, and force the new strains to process substrate only through efficient pathways [47][48], or introduce foreign pathways by insert new enzymes genes to the host and boost the production rate [48].

2.2 Stoichiometric Metabolic Model of *S. cerevisiae*

Although the main goal of this study is to construct metabolic pathway models of *S. cerevisiae*, many efforts were made to prepare reliable stoichiometric metabolic

models, as it is the foundation of further pathway analysis. This chapter describes methods that are corresponding to step 1 ~ 3 in the modeling roadmap (see Fig. 2.3 in section 2.1.2.5).

2.2.1 Establishing Stoichiometric Metabolic Models

All the investigations/case-studies presented in this thesis were based on metabolic networks of *S. cerevisiae*. A simplified stoichiometric model (Model 1 with 40 metabolites and 35 intracellular reactions) were established based on data from the internet databases KEGG [156] and literature [157–160], and simplified using the modular approach described in literature [50]. The metabolic network structure of Model 1 is demonstrated in Fig. 2.5.

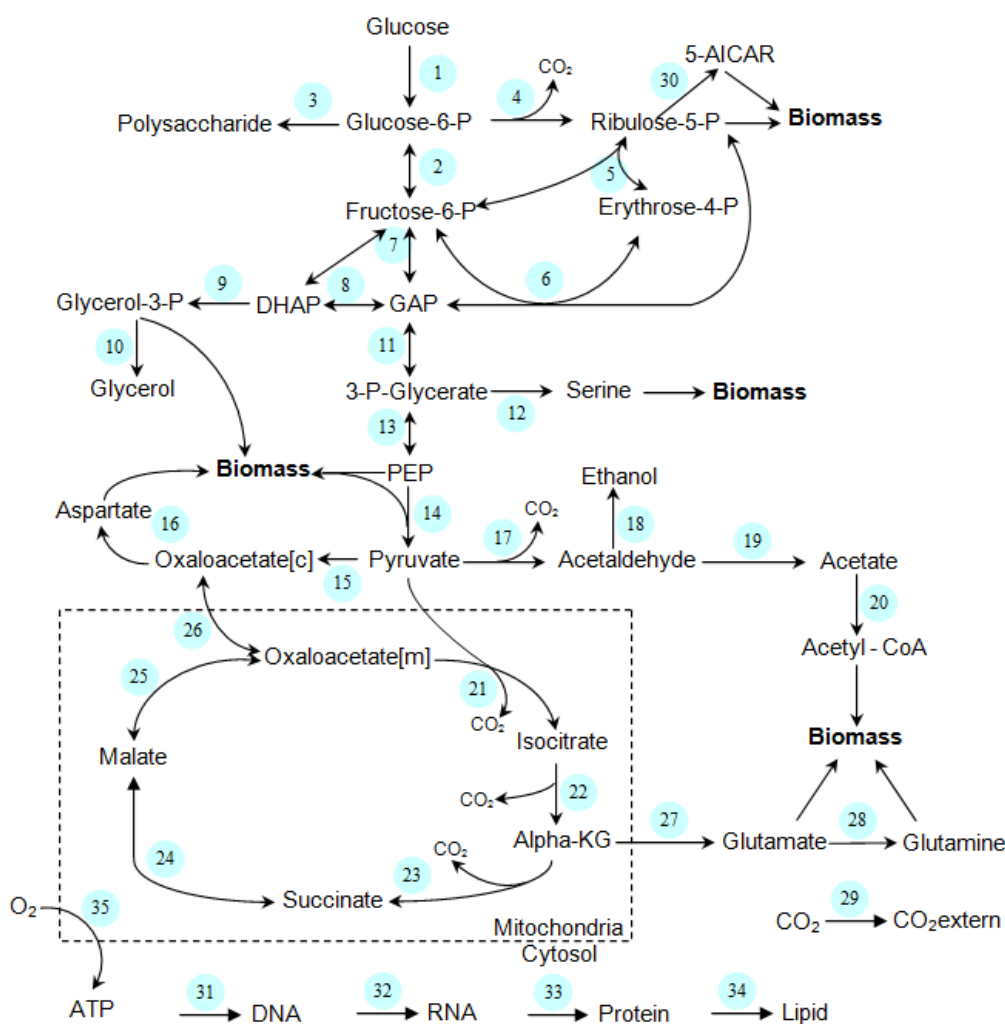


Fig. 2.5 Schematic illustration of metabolic Network for simplified metabolic model (Model 1) of *S. cerevisiae*.

2.2.1.1 From Reaction Network to Initial Stoichiometric Model

The simplified network was converted into an initial stoichiometric model based on C-mol stoichiometry. The metabolic network of Model 1 is constituted by thirty-five reactions, which are given in Tab. 2.2. All the stoichiometric coefficients in this network are on C-mol basis. In this stoichiometric model, H^+ and H_2O are not included because they cannot be measured by instruments and their balances are not important for the remaining reactions of the model. For simplification the cell compartments are not considered, i.e. no differentiation is made between NADH in cytosol and NADH in mitochondria. For the pair of cofactors namely ATP/ADP, $NAD^+/NADH$ and $NADP^+/NADPH$, only one of each pair (ATP, NADH and NADPH) is considered in the model, since their roles in reactions are opposite and exchange rates can be easily calculated to substitute one with the other. The proton carriers $FADH_2$ and Ubiquinol-6 are pooled together with NADH, while NADPH is considered separately. Ten of thirty-five reactions are considered as reversible. The reversibility of reactions is shown as “ \rightleftharpoons ” in the reactions. In Model 1, there are forty metabolites involved (listed in Tab. 2.3). The stoichiometric matrix \mathbf{Y} of Model 1 is a 40×35 matrix as shown in the appendix (section 6.2.1, Tab. 6.3). In this stoichiometric model, glucose and oxygen can be regarded as substrates; CO_2 , ethanol, glycerol and biomass are produced; however, because only the external CO_2 in the bioreactor (shown as “co2extern” in the models) is measurable, a CO_2 transport reaction is added (reaction 29). In this way, the CO_2 inside cells is treated as internal metabolite. Among all metabolites, twenty-nine metabolites are considered as internal metabolites, thereby \mathbf{G}^T is a 29×35 matrix as shown in Tab. 2.4.

Tab. 2.2 Biochemical reactions of Model 1 for *S. cerevisiae* in C-mol stoichiometry.

| No. | Designation | Biochemical Reactions |
|-----|-------------|--|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <==> f6p |
| 3 | pol | g6p + (0.1667) atp --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <==> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <==> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <==> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <==> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <==> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akg + (0.125) nadh |
| 13 | 3pg-pep | 3pg <==> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akg |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akg | iso --> (0.833333) akg + (0.1666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akg-succ | akg --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <==> mal + (0.25) nadh |
| 25 | mal-oxam | mal <==> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <==> oxam |
| 27 | akg-glu | akg + (0.2) nadph --> glu |
| 28 | glu-gln | glu + (0.2) atp --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 + (0.26087) atp --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.0842) mfh4 + (0.3313) asp + (0.2544) r5p + (0.4625) atp + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.0568) mfh4 + (0.2993) asp + (0.2400) r5p + (0.4890) atp --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (1.0396) atp + (0.1579) nadph --> prot + (0.4927) akg + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0165) mfh4 + (0.0188) 5aic + (0.0725) nadh |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.4) atp + (0.7111) nadph --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |

Tab. 2.3 Metabolites in Model 1 of *S. cerevisiae*.

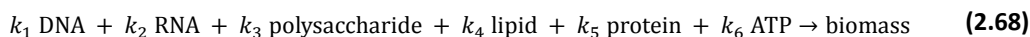
| No. | Abbreviation | Full name of metabolites |
|-----|--------------|---|
| 1 | glc | Glucose |
| 2 | atp | Adenosine triphosphate |
| 3 | g6p | Glucose 6-phosphate |
| 4 | f6p | Fructose 6-phosphate |
| 5 | pol | Polysaccharide |
| 6 | r5p | Ribose 5-phosphate |
| 7 | nadh | Nicotinamide adenine dinucleotide - reduced |
| 8 | nadph | Nicotinamide adenine dinucleotide phosphate - reduced |
| 9 | co2 | Carbon dioxide |
| 10 | co2extern | Carbon dioxide (extern) |
| 11 | e4p | Erythrose 4-phosphate |
| 12 | gap | Glyceraldehyde 3-phosphate |
| 13 | dhap | Dihydroxyacetone phosphate |
| 14 | g3p | Glycerol 3-phosphate |
| 15 | glyc | Glycerol |
| 16 | 3pg | 3-Phosphoglycerate |
| 17 | ser | Serine family |
| 18 | pep | Phosphoenolpyruvate |
| 19 | pyr | Pyruvate |
| 20 | oxac | Oxaloacetate (cytosol) |
| 21 | oxam | Oxaloacetate (mitochondrion) |
| 22 | asp | Aspartate |
| 23 | acald | Acetaldehyde |
| 24 | etoh | Ethanol |
| 25 | ace | Acetate |
| 26 | akg | alpha-Ketoglutarate |
| 27 | mal | Malate |
| 28 | glu | Glutamate |
| 29 | actcoa | Acetyl-CoA |
| 30 | gln | Glutamine |
| 31 | fum | Fumarate |
| 32 | 5aic | 5-AICAR |
| 33 | mfh4 | Methylenetetrahydrofolate (Methyl-FH4) |
| 34 | dna | Deoxyribonucleic acid |
| 35 | rna | Ribonucleic acid |
| 36 | prot | Protein |
| 37 | lip | Lipid |
| 38 | iso | Isocitrate |
| 39 | succ | Succinate |
| 40 | o2 | Oxygen |

Tab. 2.4 G^T matrix of Model 1 with 29 internal metabolites (rows) and 35 reactions (columns).

| Metabolites | Reactions | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------------|-----------|----|---------|--------|---------|---------|---------|----|---------|----|--------|--------|----|-------|-------|---------|--------|----|-----|----|----------|----------|-----|---------|-----|-------|------|------|----|---------|---------|---------|---------|---------|---------|---------|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | | |
| 1 atp | -0.1667 | 0 | -0.1667 | 0 | 0 | 0 | -0.1667 | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0.333 | -0.25 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.2 | 0 | 0 | -0.25 | 0 | -0.2 | 0 | -0.2609 | -0.4625 | -0.489 | -1.0396 | -0.4 | 2 | | |
| 2 g6p | 1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 3 f6p | 0 | 1 | 0 | 0 | 0.6 | 0.6667 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 4 r5p | 0 | 0 | 0 | 0.8333 | -1 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.2174 | -0.2544 | -0.24 | -0.0404 | 0 | 0 | | |
| 5 nadh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0 | 0.3333 | 0.125 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.142857 | 0.1 | 0.2 | 0.25 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0.154 | 0.1348 | 0.0725 | 0 | -2 | | |
| 6 nadph | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0 | -0.2 | 0 | 0 | 0.04348 | -0.0509 | 0.0568 | -0.1579 | -0.7111 | 0 | | |
| 7 co2 | 0 | 0 | 0 | 0.1667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.25 | 0 | 0.3333 | 0 | 0 | 0 | 0.142857 | 0.2 | 0.2 | 0 | 0 | 0 | 0 | 0 | -1 | -0.0435 | 0 | 0 | 0.1054 | 0 | 0 | | |
| 8 e4p | 0 | 0 | 0 | 0.4 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0623 | 0 | 0 | |
| 9 gap | 0 | 0 | 0 | 0 | 0.3333 | 0.5 | 1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0117 | 0 | 0 | |
| 10 dhap | 0 | 0 | 0 | 0 | 0 | 0.5 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 11 g3p | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0662 | 0 | |
| 12 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.375 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 13 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.375 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.1182 | -0.1012 | 0 | |
| 14 pep | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0935 | 0 | 0 | |
| 15 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.75 | 0 | -1 | 0 | 0 | 0 | -0.42857 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.338 | 0 | 0 | |
| 16 oxac | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 17 oxam | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.57143 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 18 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.1739 | -0.3313 | -0.2993 | -0.2153 | 0 | 0 | | |
| 19 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 20 ace | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 21 akc | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.625 | 0 | 0 | 0 | 0 | 0.5556 | 0 | 0 | 0 | 0 | 0 | 0.8 | -1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4927 | 0 | 0 | |
| 22 mal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 23 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.625 | 0 | 0 | 0 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0.43478 | 0.4371 | 0.5271 | -0.609 | 0 | 0 | | |
| 24 actcoa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.8326 | 0 |
| 25 gln | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -0.4348 | -0.4371 | -0.5271 | -0.2078 | 0 | 0 | | |
| 26 5aic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3913 | -0.4579 | -0.5112 | 0.0188 | 0 | 0 | | |
| 27 mfn4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0842 | -0.0568 | 0.0165 | 0 | 0 | | |
| 28 iso | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 29 succ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.85714 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |

2.2.1.2 Modeling of Biomass Components

In the above model, the biomass is represented by five components: DNA, RNA, polysaccharide, lipid and protein. To simplify calculation, Model 2 α (as shown in Fig. 2.6 and Tab. 2.5) is built from Model 1 by adding a pseudo-reaction representing the formation of biomass:



Here the stoichiometric coefficients k_n (k_1, k_2, \dots, k_5) represent the proportions of different components in the cellular structure (biomass). Eq. (2.68) can be interpreted as that DNA, RNA, polysaccharide, lipid and protein are consumed to build up the cell structure. Therefore they can be regarded as internal metabolites. The stoichiometric matrix \mathbf{G}^T of Model 2 α then become a 34×36 matrix as shown in Tab. 2.6 (The corresponding \mathbf{Y} matrix is listed in appendix Tab. 6.4). The coefficient k_6 can also be seen as Y_{ATP} , which represents the sum of ATP consumption for all biomass components, while the ATP consumptions in the individual synthetic reactions of DNA, RNA, polysaccharide etc. are removed. The coefficients k_n are determined by using the data of Nissen et al. [157] to estimate the proportion of biomass components. Y_{ATP} together with P/O value in the oxidative phosphorylation reaction are estimated using the measured O_2 , CO_2 and glucose reaction rates. It should be noticed that the proportion of different components of biomass may vary slightly under different growth conditions. This variation is ignored in the model for simplicity.

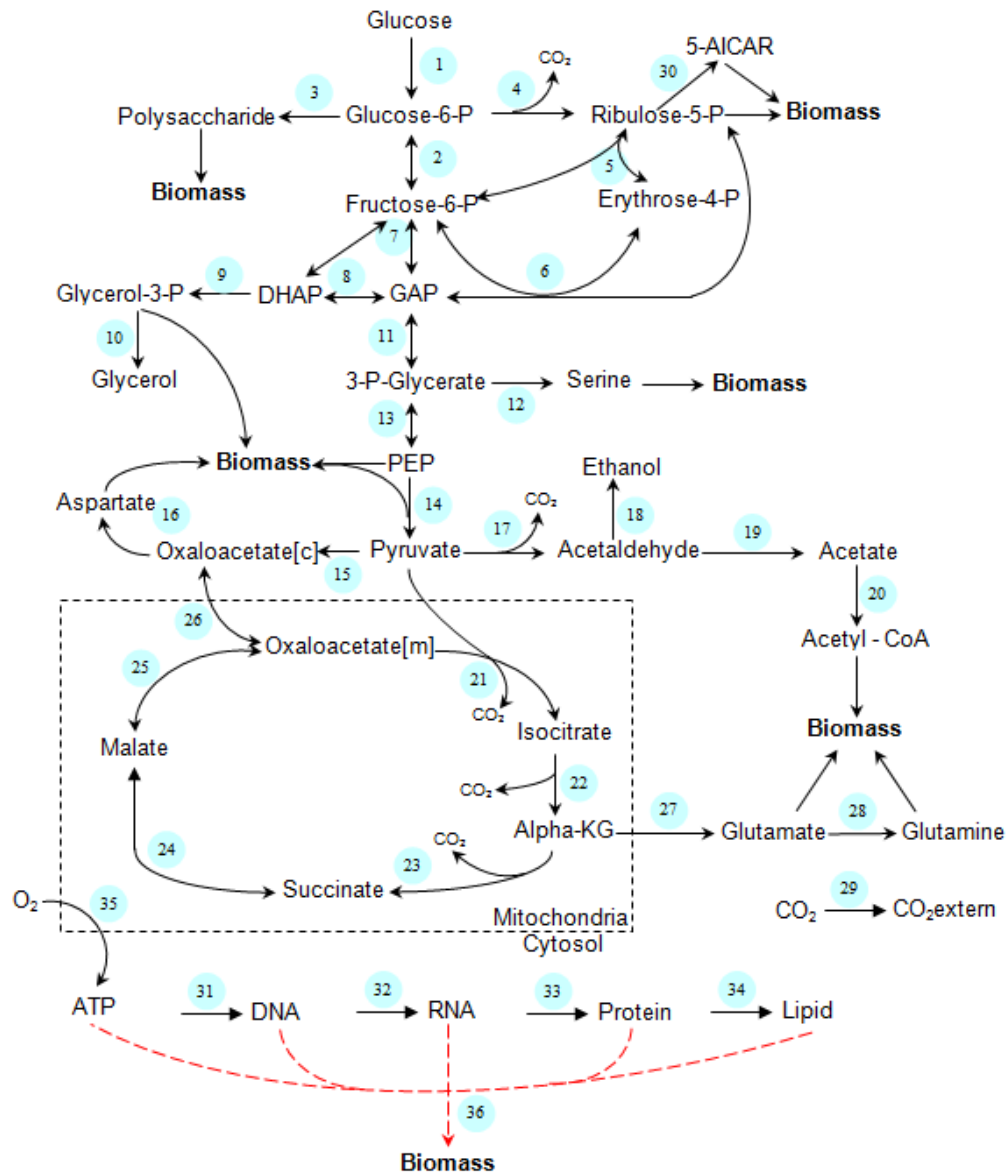


Fig. 2.6 Schematic illustration of metabolic Network for Model 2 α of *S. cerevisiae*. The red dash arrow represents the additional pseudo-reaction of biomass synthesis compared to Model 1.

Tab. 2.5 Biochemical reactions of Model 2 α for *S. cerevisiae* in C-mol stoichiometry. Compared to Model 1, Model 2 α has an additional pseudo-reaction of biomass synthesis (shown with grey background) and ATP consumption for all biomass components is summarized in the pseudo-reaction using the yield coefficient Y_{ATP} , while the ATP consumptions in the individual synthetic reactions of DNA, RNA, polysaccharide etc. were removed.

| No. | Designation | Biochemical Reactions |
|-----|-------------|---|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <==> f6p |
| 3 | pol | g6p --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <==> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <==> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <==> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <==> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <==> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akg + (0.125) nadh |
| 13 | 3pg-pep | 3pg <==> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akg |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akg | iso --> (0.833333) akg + (0.16666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akg-succ | akg --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <==> mal + (0.25) nadh |
| 25 | mal-oxam | mal <==> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <==> oxam |
| 27 | akg-glu | akg + (0.2) nadph --> glu |
| 28 | glu-gln | glu --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.0842) mfh4 + (0.3313) asp + (0.2544) r5p + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.0568) mfh4 + (0.2993) asp + (0.2400) r5p --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (0.1579) nadph --> prot + (0.4927) akg + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0165) mfh4 + (0.0188) 5aic + (0.0725) nadh |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.7111) nadph --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 36 | biomass | (0.000117103) dna + (0.001757709) rna + (0.01386991) pol + (0.000966667) lip + (0.017876333) prot + (Y_{ATP}) atp --> biomass |

Tab. 2.6 G^T matrix of Model 2 α with 34 internal metabolites (rows) and 36 reactions (columns). The parts with grey background represent the differences compared to Model 1.

| Metabolites | Reactions | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------------|-----------|----|----|--------|-----|---------|-----|----|---------|--------|--------|-------|--------|-------|-------|----|---------|--------|----|----|----------|-----|-----|----------|-------|----|------|----|---------|---------|---------|---------|---------|---------|---------|---------|----------|---------|---|----------|----------|---|----------|----------|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | | | | | | | | | | | |
| 1 atp | -0.1667 | 0 | 0 | 0 | 0 | -0.1667 | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0.333 | -0.25 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.2 | 0 | 0 | -0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2*PO | -vatp | | | | | | | | | | |
| 2 g6p | 1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | |
| 3 f6p | 0 | 1 | 0 | 0 | 0.6 | 0.6667 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | |
| 4 pol | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.01387 | | | | | | | | | | |
| 5 r5p | 0 | 0 | 0 | 0.8333 | -1 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | | |
| 6 nadh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0 | 0.3333 | 0.125 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.142857 | 0.1 | 0.2 | 0.25 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0 | 0.154 | 0.1348 | 0.0725 | 0 | -2 | 0 | | | | | | | | | | |
| 7 nadph | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0 | -0.2 | 0 | 0 | 0.04348 | -0.0509 | 0.0568 | -0.1579 | -0.7111 | 0 | 0 | 0 | | | | | | | | | | |
| 8 co2 | 0 | 0 | 0 | 0.1667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.25 | 0 | 0.3333 | 0 | 0 | 0 | 0.142857 | 0.2 | 0.2 | 0 | 0 | 0 | 0 | -1 | -0.0435 | 0 | 0 | 0 | 0.1054 | 0 | 0 | 0 | 0 | | | | | | | | | | |
| 9 e4p | 0 | 0 | 0 | 0 | 0.4 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0623 | 0 | 0 | 0 | | | | | | | | |
| 10 gap | 0 | 0 | 0 | 0 | 0 | 0.3333 | 0.5 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0117 | 0 | 0 | 0 | 0 | | | | | | | | |
| 11 dhap | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | |
| 12 g3p | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0662 | 0 | 0 | 0 | | | | | | | |
| 13 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.375 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | |
| 14 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.375 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.1182 | -0.1012 | 0 | 0 | 0 | 0 | | | | | | | |
| 15 pep | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0935 | 0 | 0 | 0 | 0 | | | | | | | |
| 16 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.75 | 0 | -1 | 0 | 0 | 0 | 0 | -0.42857 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.338 | 0 | 0 | 0 | 0 | | | | | | | |
| 17 oxac | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | |
| 18 oxam | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.57143 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | |
| 19 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.1739 | -0.3313 | -0.2993 | -0.2153 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | |
| 20 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | |
| 21 ace | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | |
| 22 akc | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.625 | 0 | 0 | 0 | 0 | 0.5556 | 0 | 0 | 0 | 0 | 0 | 0.8 | -1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4927 | 0 | 0 | 0 | 0 | | | | | | | |
| 23 mal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | |
| 24 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.625 | 0 | 0 | 0 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0.43478 | 0.4371 | 0.5271 | -0.609 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | |
| 25 actcoa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.8326 | 0 | 0 | 0 | | | | | | |
| 26 gln | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -0.4348 | -0.4371 | -0.5271 | -0.2078 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | |
| 27 5aic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3913 | -0.4579 | -0.5112 | 0.0188 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | |
| 28 mfn4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0842 | -0.0568 | 0.0165 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | |
| 29 dna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00012 | 0 | | | | | | |
| 30 rna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00176 | 0 | | | | | |
| 31 prot | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.01788 | 0 | | | |
| 32 lip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00097 | 0 | | |
| 33 iso | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 34 succ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

By introducing the pseudo-reaction of biomass synthesis, cell biomass is seen as a final “product” of the metabolic network. This makes the next maximization step of specific biomass formation rate much easier in practice, as now only one reaction rate needs to be maximized. This is theoretically reasonable because the proportion among different components of the cell structure (biomass) shows less variation under different growth conditions. To verify the equivalency of the proposed method and conventional methods, MFA were performed on Model 2 and Model 2A. Model 2 is an improved version of Model 2 α after calibrating the Y_{ATP} coefficient and removing some invalidated constrains. (Improving steps are described in the next section). Model 2A is modified from Model 2, in which the pseudo-reaction is removed and different biomass components are calculated independently (The detailed biochemical reactions and stoichiometric matrixes of Model 2 and Model 2A can be found in appendixes 6.3.1 and 6.3.2). This pair of models is used to simulate the anaerobic, glucose-limited continuous cultures using data published by Nissen et al. in [157]. Through MFA, the reaction flux \mathbf{v} computed from Model 2 given the dilution rate is almost identical to the one computed from Model 2A given several measured reaction rate. Tab. 2.8 shows one example of one state of Model 2, where by introducing the pseudo-reaction of biomass synthesis, the dilution rate D can be directly used as the reaction rate of the pseudo-reaction (see the column of given rate in Tab. 2.8). The individual components of biomass, such as nucleic acid, protein, lipid and carbohydrate, now become internal metabolites, thereby the need of determining the synthesis rates of those individual components is eliminated. Tab. 2.7 shows the minimum number of reactions needed to be measured to solve the stoichiometric Eq. (2.49) in different models. The more measurements are needed, the higher is the risk of introducing errors into the calculation. Using the biomass reaction can not only simplify the calculation but also increase the robustness of the model.

Tab. 2.7 Dimension of stoichiometric matrixes of internal metabolites (G^T) for Model 1, Model 2 α , Model 2 and Model 2A. Model 1 and Model 2A are without explicit biomass reaction, while Model 2 α and Model 2 are with pseudo-reaction of biomass synthesis.

| | Model 1 | Model 2 α | Model 2 | Model 2A |
|--|---------|------------------|---------|----------|
| Reactions | 35 | 36 | 36 | 36 |
| Internal Metabolites | 29 | 34 | 33 | 28 |
| Free columns (minimum needed measured data) | 6 | 2 | 3 | 8 |

Tab. 2.8 Flux distribution of Model 2 and Model 2A in anaerobic growth. Model 2A is Model 2 without the pseudo-reaction of biomass synthesis. The numbers with dark background are the given rates v_m . v_c represent calculated rates. The results of Model 2A are generated via flux balance analysis with non-weighted multiple objective functions or weighted single objective function. (Complete results can be found in appendixes 6.4.1.)

| No. | Reaction Designation | Given Rate C-mol/(gh) | Reaction Rates in Model 2 C-mol/(gh) | Reaction Rates in Model 2A C-mol/(gh) | |
|-------|----------------------|--------------------------|--|--|------------------------------|
| | | | | Multiple objective functions | Single objective function |
| 1 | glc | 0.0333 | 0.0333 | 0.0333 | 0.0333 |
| 10 | glyc | 0.00286 | 0.00312 | 7.2E-15 | 0.00312 |
| 18 | etoh | 0.01756 | 0.01756 | 0.00740 | 0.01756 |
| 29 | co2 | 0.0095 | 0.00931 | 0.00370 | 0.00931 |
| 35 | respiration | 0 | 0 | 0 | 0 |
| 36 | biomass/ Y_{ATP} | 0.10302 | 0.10217 | (Y_{ATP}) 1.2E-16 | (Y_{ATP}) 0.00762 |
| μ | | | 0.10217 | 0.71303 | 0.10217 |

The rate of the pseudo-reaction of biomass synthesis here serves as a convenient objective function for flux balance analysis using the underdetermined system when the number of given reaction rates is less than needed as shown in Tab. 2.7 (see section 2.1.3.2). Since it is reasonable to assume that growth and reproduction have the highest priority in organisms, maximizing the biomass synthesis rate can be added as a profit function in addition to the stoichiometry equation. In this study, flux balance analyses were carried out on Model 2 and Model 2A while using the measured glucose and oxygen uptake rate as given parameters. For Model 2, the object function was set to maximize the rate of the pseudo-reaction. For Model 2A, several object functions (multi-objective function) were used to maximize the production of each biomass component simultaneously. The difference between optimizing the biomass synthesis and optimizing individual biomass component is explained in the following. As shown in Tab. 2.8, when optimizing the biomass synthesis rate in Model 2 resulted in a relative close solution compared to the given reaction rates. The measured reaction rates and the estimated reaction rates by optimizing the biomass are very close, which are respectively 0.00286 C-mol/(gh) and 0.00312 C-mol/(gh) for glycerol, 0.0095 C-mol/(gh) and 0.00931 C-mol/(gh) for CO₂, 0.10302 C-mol/(gh) and 0.10217 C-mol/(gh) for biomass. On the other hand, optimizing all the biomass components in Model 2A resulted in irrational reaction rates, the estimated reaction rates of glycerol, CO₂ and biomass are respectively 7.2E-15 C-mol/(gh), 0.00370 C-mol/(gh) and 1.2E-16 C-mol/(gh), as shown in Tab. 2.8. These errors in latter model may be due to the unjustified weighting factor of different component in the object function, as optimizing biomass

components separately is mathematically equal to optimize a single function: 1 DNA + 1 RNA + 1 polysaccharide + 1 lipid + 1 protein + 1 ATP. If a rational object function (0.0001053927 DNA + 0.0015819381 RNA + 0.012482919 polysaccharide + 0.0008700003 lipid + 0.0160886997 protein + 0.0746 ATP) is designed to use the right weighting factor from the biomass synthesis on each biomass component, the optimizing results of Model 2A (the last column in Tab. 2.8) show the same flux distribution as Model 2. Here the weighting factor for each biomass component is identical to the coefficients (k_n) in the pseudo-reaction of biomass synthesis, so the simulated results are also identical to Model 2. Nevertheless, these results, to some extent, prove that using the pseudo-reaction of biomass synthesis does not influence the simulation accuracy and also provides a convenient way for optimizing biomass.

It is worth mentioning, the theoretical coefficients (k_n) of different component in the pseudo-reaction cannot be directly used to calculate the specific growth rate μ . A system error of 10% was found during optimization using MFA (cf. Fig. 2.7) and was deducted from the coefficients, i.e. $k_n' = k_n \times 0.9$. This system error may be due to the model simplification or deviation of parameter i.e. Y_{ATP} . However it should not affect the integrity of the established models as the proportional change of k_n only affect the rate of biomass synthesis reaction. In Tab. 2.9 shows the estimated biomass coefficients in Model 2 based on the available data of biomass component from literature [157], and the calibrated coefficients. This system error is model specific therefore, for other metabolic network this correction needs to be re-determined.

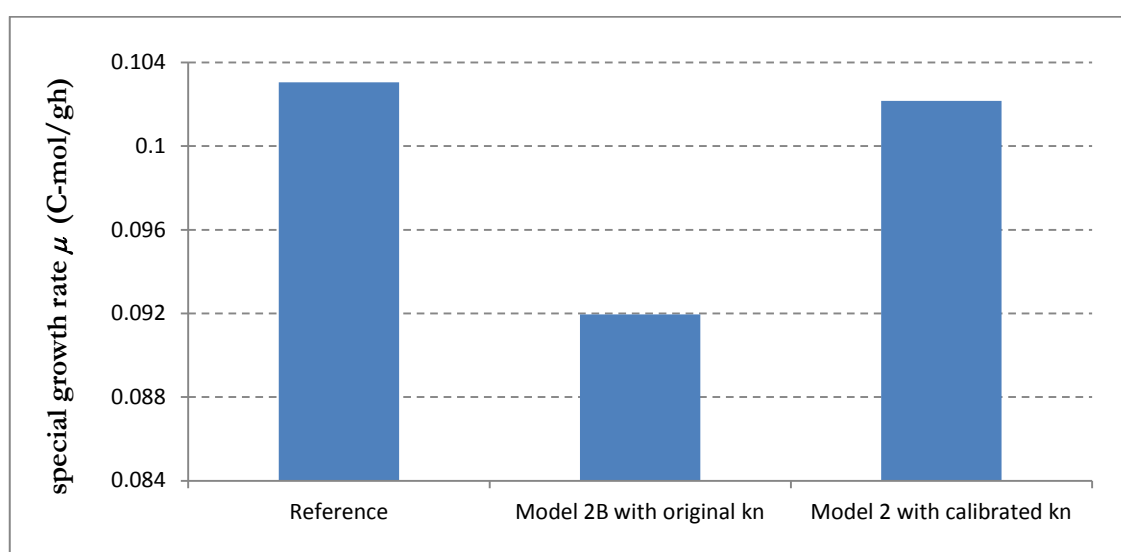


Fig. 2.7 Biomass formation rates of *S. cerevisiae* during anaerobe growth. Model 2B is a model with original k_n for biomass components and is detailed in section 6.3.3. The calculation of flux distribution in Model 2B and Model 2 is used determined system by giving uptake rates of glucose and oxygen and produced rate of ethanol.

Tab. 2.9 Calculation and calibration of the biomass components coefficients (k_n). The values of biomass share are from the literature [157] and the calibrated k_n is original $k_n \times 90\%$. Since it is difficult to determine the mol weight of the ash in biomass, k_n s are calculated without ash.

| | Mol weight g/c-mol | Biomass share g/g | original k_n * | calibrated k_n |
|---------------------------|-----------------------|----------------------|------------------|------------------|
| carbohydrate | 29.34409799 | 0.407 | 0.01386991 | 0.012482919 |
| DNA | 34.15789474 | 0.004 | 0.000117103 | 0.000105393 |
| RNA | 35.84210526 | 0.063 | 0.001757709 | 0.001581938 |
| protein and amino acid | 25.78828619 | 0.461 | 0.017876333 | 0.0160887 |
| lipid | 30 | 0.029 | 0.000966667 | 0.00087 |
| ash | | 0.05 | | |
| sum with ash | | 1.014 | | |
| sum without ash | | 0.964 | 0.034587723 | 0.03112895 |

*: The original k_n are computed from “biomass share” divided by “mol weight”.

2.2.2 Improving the Consistency of the Stoichiometric Metabolic Models

Preliminary tests suggested that the initial metabolic model (Model 2 α) created above are incomplete or contain potential errors. An improvement step was then carried out to refine the models. Measured reaction rates, which were acquired from the experiment described in [157] and evaluated with least square method, were used to find a more consistent stoichiometric matrix \mathbf{G}^T . The improvement strategy is based on the metabolic flux analysis, in which J measured reactions are selected as the input to compute the remaining measurable reaction rates. The optimization goal is to minimizing the difference between computed reaction rates and measured reaction rates from experimental data. This resulted in an improved metabolic model (Model 2) which was used for further investigations. For the convenience of the discussion, the final improved model (Model 2) is given here first (see Tab. 2.10 and section 6.3.1 in appendixes). Compared with Model 2 α , Model 2 has no methyl-FH4 balance constrain, i.e. a row vector is removed from \mathbf{G}^T , and the coefficients of ATP were also adjusted in Model 2 according to some calibration steps. The details of the data evaluation and matrix adjustment steps are presented in the flowing sub-sections. Because these two aspects (methyl-FH4 balance and ATP balance) will both contribute to the errors in Model 2 α , in the follow discussions, to demonstrate the effects of each of them, Model 2 is used as a reference while the problematic methyl-FH4 balance and ATP balance are added back respectively.

Tab. 2.10 G^T matrix of Model 2 with 33 internal metabolites (rows) and 36 reactions (columns). The parts with grey background represent the differences compared to Model 2 α and Model 2 has no methyl-FH4 balance (a row vector less than G^T of Model 2 α).

| Metabolites | Reactions | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------------|-----------|----|----|--------|-----|---------|---------|----|---------|----|--------|--------|----|--------|-------|---------|---------|--------|-----|----|----------|-----|-----|------|-----|-------|------|----|---------|---------|---------|---------|---------|---------|---------|---------|----------|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | | |
| 1 atp | -0.1667 | 0 | 0 | 0 | 0 | 0 | -0.1667 | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0.3333 | -0.25 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.2 | 0 | 0 | -0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2*PO | -yatp | |
| 2 g6p | 1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 3 f6p | 0 | 1 | 0 | 0 | 0.6 | 0.6667 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 4 pol | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.01248 | |
| 5 r5p | 0 | 0 | 0 | 0.8333 | -1 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.2174 | -0.2544 | -0.24 | -0.0404 | 0 | 0 | 0 | | |
| 6 nadh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0 | 0.3333 | 0.125 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.142857 | 0.1 | 0.2 | 0.25 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0.154 | 0.1348 | 0.0725 | 0 | -2 | 0 | | |
| 7 nadph | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0 | -0.2 | 0 | 0.04348 | -0.0509 | 0.0568 | -0.1579 | -0.7111 | 0 | 0 | 0 | | |
| 8 co2 | 0 | 0 | 0 | 0.1667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.25 | 0 | 0.3333 | 0 | 0 | 0 | 0.142857 | 0.2 | 0.2 | 0 | 0 | 0 | 0 | 0 | -1 | -0.0435 | 0 | 0 | 0.1054 | 0 | 0 | 0 | | |
| 9 e4p | 0 | 0 | 0 | 0 | 0.4 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0623 | 0 | 0 | 0 | |
| 10 gap | 0 | 0 | 0 | 0 | 0 | 0.3333 | 0.5 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0117 | 0 | 0 | 0 | 0 | |
| 11 dhap | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 12 g3p | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0662 | 0 | 0 |
| 13 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.375 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 14 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.375 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.1182 | -0.1012 | 0 | 0 |
| 15 pep | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0935 | 0 | 0 | 0 | |
| 16 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.75 | 0 | 0 | 0 | 0 | 0 | -0.42857 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.338 | 0 | 0 | 0 | |
| 17 oxac | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 18 oxam | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.57143 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 19 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.1739 | -0.3313 | -0.2993 | -0.2153 | 0 | 0 | 0 | 0 | |
| 20 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 ace | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 akg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.625 | 0 | 0 | 0 | 0 | 0.5556 | 0 | 0 | 0 | 0 | 0 | 0.8 | -1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0.4927 | 0 | 0 | 0 | |
| 23 mal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.625 | 0 | 0 | 0 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0.43478 | 0.4371 | 0.5271 | -0.609 | 0 | 0 | 0 | 0 | |
| 25 actcoa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.8326 | 0 | 0 |
| 26 gln | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -0.4348 | -0.4371 | -0.5271 | -0.2078 | 0 | 0 | 0 | 0 | |
| 27 5aic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3913 | -0.4579 | -0.5112 | 0.0188 | 0 | 0 | 0 | 0 | |
| 28 dna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | -0.00011 | |
| 29 rna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | -0.00158 | |
| 30 prot | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -0.01609 | |
| 31 lip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | -0.00087 |
| 32 iso | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 33 succ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

2.2.2.1 Experimental Data Collection and Evaluation

For calculation of the flux distribution of the stoichiometric models, three groups of measured reaction rates were used in the investigations, which were collected from different literature. For example, Nissen et al. published reaction rates in anaerobic, glucose-limited continuous cultures in [157]. Reaction rates during aerobic culture in chemostat were taken from [71] and calibrated using a least square method based on the linear system constraints by carbon and electron balances (see section 2.1.3.3). The corrected reaction rates were then used for estimation of unknown stoichiometric coefficients in the stoichiometric matrix \mathbf{G}^T as explained below. The details of these data can be found in the appendix section 6.1.

The measured reaction rates in this study include the specific uptake rate of glucose and oxygen, the specific production rate of ethanol, glycerol, CO_2 , and also protein, nucleic acid, lipids, carbohydrates.

2.2.2.2 Adjustment of the Stoichiometric Matrix \mathbf{G}^T

To detect the source of error in the stoichiometric matrix, metabolic flux analysis (MFA) on matrix \mathbf{G}^T was performed while excluding one metabolite at a time. Problematic metabolites and related reactions were then identified if the predicted reaction rates became close to the measured reaction rates. The possible errors were then further investigated and corrected via reversed prediction.

In the reversed prediction, the problematic coefficients in \mathbf{G}^T were treated as unknown variables. First, the vector \mathbf{v} is calculated by solving $\mathbf{G}_{\mathbf{k-p}}^T \cdot \mathbf{v} = 0$, where $\mathbf{G}_{\mathbf{k-p}}^T$ is a sub-matrix of \mathbf{G}^T without the rows containing the problematic coefficients. In the current study, $\mathbf{G}_{\mathbf{k-p}}^T \cdot \mathbf{v} = 0$ is an overdetermined system and can be solve with least square method. Thus \mathbf{v} is substituted into the equation $\mathbf{G}_{\mathbf{p}}^T \cdot \mathbf{v} = 0$ where $\mathbf{G}_{\mathbf{p}}^T$ are the rows of \mathbf{G}^T that contain unknown coefficients. The problematic coefficients are finally calculated by solving the second equation.

Using the model improvement strategy described above, several problematic stoichiometric coefficients were identified and corrected. These metabolites include ATP and methyl-FH₄. After removing the methyl-FH₄ balance (a row vector in \mathbf{G}^T matrix) and adjusting the ATP coefficients in Model 2 α , the improved model, Model 2

was created and used for various experiments in the rest chapters. In the following sub-sections, the refining process is further discussed.

2.2.2.2.1 Calibrating ATP Balance

When the original ATP coefficients (given in [157]) are used and the ATP balance is included in the stoichiometric equation, the calculated reaction rates tend to have great errors compared to the measured reaction rates. This is demonstrated through Model 2C (see Tab. 6.12 in section 6.3.4), which is a replicate of Model 2 where the ATPs appear in each synthesis reaction of biomass components and their coefficients are assigned to the original coefficients (given in [157]). Two experiments were designed. One used the measured glucose and oxygen uptake rates and ethanol production rates (plus CO₂ production rate in Model 2C) as known variable to predict the rest reactions rates via MFA (determined system). The other used only glucose and oxygen uptake rates to predict the rest reactions rates via flux balance analysis (underdetermined system). As shown in Tab. 2.11, if ATP balance is enforced, the predicted rates is far from accurate, glycerol production rate became negative (-0.00437 C-mol/(gh)) and biomass synthesis rates (0.37118 C-mol/(gh)) are far from the measured values (0.10302 C-mol/(gh)). For determined system, i.e. given enough measured reaction rates, it can be solved by simply removing the ATP balance and giving one more measured value. For example, in the column of Model 2C without ATP balance in Tab. 2.11, the simulated reaction rates of glycerol and biomass are closer to the given reaction rates. However for underdetermined systems using biomass maximization will lead to irrational results. For example in Model 2C when the ATP balance constraint was removed, biomass synthesis rate became unreasonable high (0.48801 C-mol/(gh)).

Tab. 2.11 Flux distribution of Model 2C in anaerobic growth. Model 2C is Model 2 without energy summary ($Y_{ATP} \cdot \mu$) in the pseudo-reaction of biomass synthesis. The numbers with dark background are the given rates v_m . (Complete results can be found in appendixes 6.4.2.)

| Reaction | | Given Rate C-mol/(gh) | Reaction Rates in Model 2C C-mol/(gh) | | | |
|----------|-------------|--------------------------|--|-----------------|---------------------|-----------------|
| No. | Designation | | with ATP balance | | without ATP balance | |
| | | | Determined | Underdetermined | Determined | Underdetermined |
| 1 | glc | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 |
| 10 | glyc | 0.00286 | -0.00437 | 0.00613 | 0.00312 | 0.01494 |
| 18 | etoh | 0.01756 | 0.01756 | 0.01309 | 0.01756 | 2.54E-16 |
| 29 | co2 | 0.0095 | 0.00810 | 0.00759 | 0.00931 | 0.00256 |
| 35 | respiration | 0 | 0 | 0 | 0 | 0 |
| 36 | Biomass | 0.10302 | 0.37118 | 0.20041 | 0.10217 | 0.48801 |

To overcome this problem, the ATP coefficients were calibrated using the measured reaction rates. A possible reason that causes the ATP coefficient error can be the simplifications process (Step 2 in Fig. 2.3) that caused overestimate/underestimate of ATP production/consumption, especially in those reactions where biomass components are produced. Because there are many reactions involved in each synthesis pathway of those biomass components, proteins for example, it is very difficult, if not impossible to locate the error. To solve this problem, the ATP consumption in those reactions is moved into the pseudo biomass synthesis reaction (represented by Y_{ATP}) and its coefficient is optimized via MFA. As discussed in section 2.2.1.2, since the proportion of biomass components is assumed to be constant, it is evident that the ATP consumption to build one unit of biomass should also be constant. The benefit of lumping all ATP consumption into one reaction is that it simplifies the coefficient calibration process as now there is only one parameter Y_{ATP} instead of several stoichiometric coefficients in different reactions. As it is not necessary to know how much ATP is consumed in each individual component production, the simplification is reasonable when the biomass composition remains constant.

It should be pointed out that in reality the P/O ratio in reaction 35 of Model 2 and reaction 16 of Model 3 varies under different growth conditions. This is because in the simplified models the series of reaction involved in the oxidative phosphorylation is replaced by a single linear equation, in which the P/O ratio is used to represent the number of synthesized ATP by consumed oxygen for each pair of electrons. Under aerobic conditions (see Tab. 6.2 in section 6.1), the estimated value of these two parameters are $Y_{ATP} = 0.1168$ mol/g and $P/O = 1.394$. For anaerobic conditions (see Tab. 6.1 in section 6.1) the optimized Y_{ATP} was determined as 0.0746 mol/g, which is used for further calculation in metabolic flux analysis and pathway analysis. Although the Y_{ATP} and P/O ratio may vary under different specific growth rates [161], in this study, the changes under the same growth condition (either anaerobic or aerobic growth), were ignored.

The consumed ATP for cell maintenance (m_{ATP}) is also ignored in these models, because the proportion of m_{ATP} is very small compared to $Y_{ATP} \cdot \mu$ if μ is not too low and the difference by including this extra reaction is negligible. Tab. 6.13 in section 6.3.5 shows the reactions of Model 2D including the ATP-maintenance reaction. Compared

with the measured reaction rates, including this additional reaction does not increase the correlation between the simulated results and measurement.

2.2.2.2.2 Effect of Methyl-FH₄ Balance

Methyl-FH₄ was also found causing errors in the simulated results. Tab. 2.12 demonstrates the simulated flux distributions in Model 2E (where the balance of methyl-FH₄ was added back to Model 2) under anaerobic growth condition. Three methods, determined system MFA, overdetermined system MFA and overdetermined system MFA with **F**-matrix were tested to estimate the metabolic flux by giving proper number of measured rates as input. Due to the methyl-FH₄ restriction, all the reaction rates related to biomass synthesis tend to zero when solved as a determined system given enough measured reaction rates (e.g. the result for determined system of Model 2E in Tab. 2.12). Further investigation by using an overdetermined system where all measured rates were given revealed the estimated flux is deflected most by the reactions involving methyl-FH₄. The value of defect for methyl-FH₄ (see Fig. 2.8) is calculated from the dot product of the row vector of **G**^T corresponding to methyl-FH₄ and the estimated flux vector **v**. This value means the net residual (positive or negative) of methyl-FH₄ which should be zero in ideal cases. It is worth mentioning that a defect of 10⁻⁵ is not ignorable because the flux distribution rates are normally in the range of 10⁻² ~ 10⁻⁵. The problem with methyl-FH₄ balance is again confirmed by the flux determination with the advanced method using **F**-Matrix (Eq. (2.55) on page 25) for the same overdetermined system (see Tab. 2.12). The calculation with **F**-Matrix reduces the effect of the deviation of **v**_m and fulfills the balances of internal metabolites as much as possible. When using **F**-Matrix, the defect of methyl-FH₄ balance is adjusted to 3.64E-11 C-mol/(gh) (see Fig. 2.8), but all the biomass synthesis related fluxes tend to zero (2.21E-07 C-mol/(gh) in Tab. 2.12) again. The sensitivity analysis of this model (see explanation in section 2.2.2.4) reveals the condition number of **G**^T matrix for Model 2E is also extremely high, which suggests the Model 2E with methyl-FH₄ balance is unstable and very sensitive to variations of measurement. All these results implied that there is problem with the methyl-FH₄ balance.

The reason of methyl-FH₄'s unbalancing maybe due to the fact that it appears in several highly integrated reactions (reaction 31 ~ 33 in Model 2) in the simplified model. Error may occur during the simplification procedure (Step 2 in Fig. 2.3) and it is difficult to locate it. Removing this metabolite is less harmful than removing the ATP

restrictions because those three reactions (reaction 31 ~ 33) have been highly constrained by many other involved metabolites. Therefore, the balance of methyl-FH₄ is ignored in Model 2.

Tab. 2.12 Flux distribution Model 2E (with methyl-FH₄ balance) in anaerobic growth. The numbers with dark background are the given rates v_m . For the overdetermined system of Model 2E two calculation methods were performed: one with only matrix G^T (overdetermined (G^T)) and the other also using matrix F (overdetermined (F)), which are covered in Eq. (2.52) on page 24 and Eq. (2.56) on page 25 respectively. (Complete results can be found in appendixes 6.4.3.)

| Reaction | | Given Rate C-mol/(gh) | Rate in Model 2E (with methyl-FH ₄ balance) C-mol/(gh) | | |
|----------|-------------|--------------------------|--|--------------------------|--------------------|
| No. | Designation | | determined | overdetermined (G^T) | overdetermined (F) |
| 1 | glc | 0.0333 | 0.0333 | 0.0333 | 0.03345 |
| 3 | pol | | 0 | 0.00128 | 2.75E-09 |
| 10 | glyc | 0.00286 | 0.01846 | 0.00312 | 0.00898 |
| 18 | etoh | 0.01756 | 0.00784 | 0.01756 | 0.01692 |
| 29 | co2 | 0.0095 | 0.00700 | 0.00931 | 0.00755 |
| 31 | dna | | 0 | 1.22E-05 | 2.63E-11 |
| 32 | rna | | 0 | 0.00016 | 3.51E-10 |
| 33 | protein | | 0 | 0.00164 | 3.55E-09 |
| 34 | lipid | | 0 | 8.89E-05 | 1.92E-10 |
| 35 | respiration | 0 | 0 | 0 | -0.00240 |
| 36 | biomass | 0.10302 | 0 | 0.10217 | 2.21E-07 |

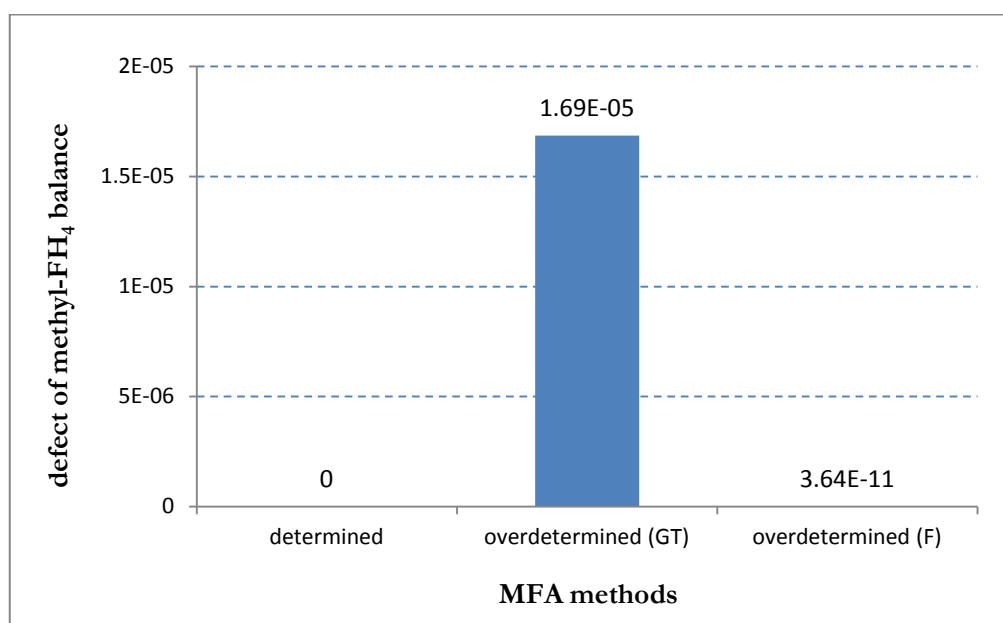


Fig. 2.8 Distribution defect of methyl-FH₄ balance using different MFA methods for Model 2E. The first defect is for determined system; the second and third defects are determined as overdetermined system and calculated respectively by simple method (using only matrix G^T) and advanced method (with matrix F).

2.2.2.3 Effects of Model Simplification

Modeling with the modular approach usually means several reactions are integrated to a single one in order to simplify the system and make it solvable and easily manipulatable. In this study, a more simplified model, Model 3, was also tested as an attempt to further simplify calculations. Model 3 is a highly simplified version and has more abstractly combined reactions than Model 2 (see Fig. 2.9). Several metabolic stages of Model 2 such as EMP, TCA Cycle and PPP are replaced with integrated single reactions in Model 3 (cf. Tab. 2.13 and Tab. 6.7 in section 6.3.1) and NADPH is replaced with NADH. The dimension of its \mathbf{G}^T matrix is almost half size of Model 2 (\mathbf{G}^T of Model 3 is a 14×18 matrix, see Tab. 2.15). Although Model 3 could provide useful information in some simulation or prediction tasks, the accuracy is compromised due to simplification. By comparing flux distribution of Model 2 and Model 3 in Tab. 2.16, the effect of reaction combination can be partly shown. In Tab. 2.16, the reaction rates under three different growth conditions were determined as underdetermined system by providing the uptake rates of substrate and oxygen (from [71]). The underdetermined systems were solved by maximizing the specific growth rate using linear programming optimization on Matlab. The same parameters of Y_{ATP} and P/O were used for both models.

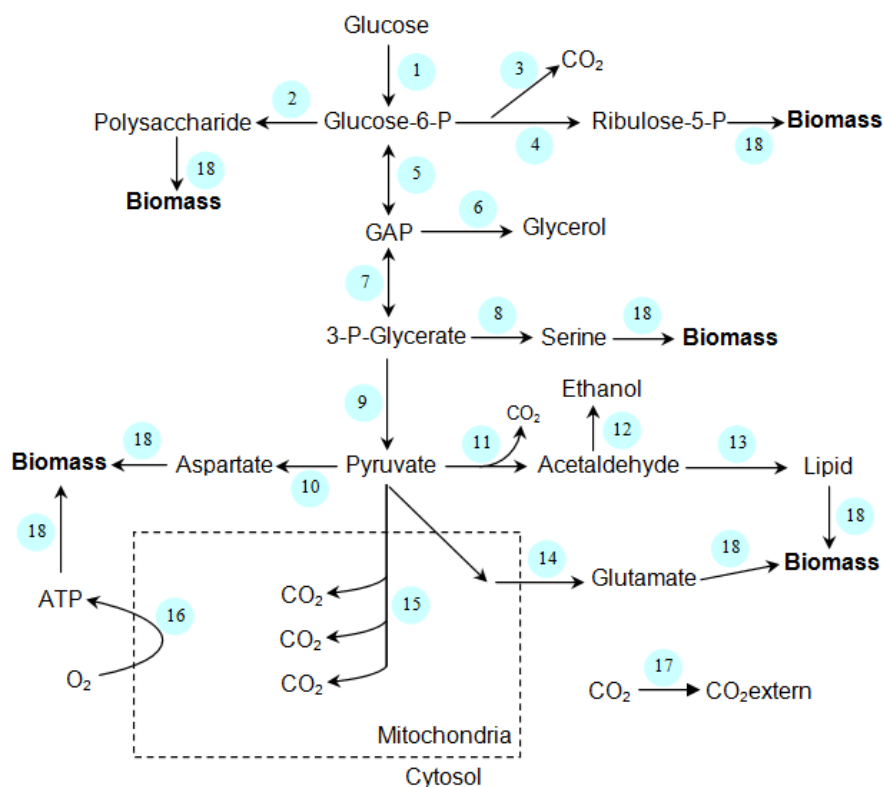


Fig. 2.9 Network of highly simplified metabolic model (Model 3) of *S. cerevisiae*.

Tab. 2.13 Biochemical reactions of Model 3 for *S. cerevisiae* in C-mol stoichiometry. The abbreviations of metabolites are listed in Tab. 2.14.

| No. | Designation | Biochemical Reactions |
|-----|-------------|--|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | pol | g6p --> pol |
| 3 | g6p-co2 | g6p --> co2 |
| 4 | g6p-r5p | g6p + (0.1667) atp --> r5p |
| 5 | g6p-gap | g6p + (0.1667) atp <=> gap |
| 6 | glyc | gap + (0.3333) nadh --> glyc |
| 7 | gap-3pg | gap <=> 3pg + (0.3333) nadh + (0.3333) atp |
| 8 | ser | 3pg --> ser |
| 9 | 3pg-pyr | 3pg --> pyr + (0.3333) atp |
| 10 | asp | (0.857) pyr + (0.143) co2 + (0.214) atp --> asp |
| 11 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 12 | etoh | acald + (0.5) nadh --> etoh |
| 13 | lip | acald --> lip |
| 14 | glu | (1.2) pyr + (0.2) atp --> glu + (0.4) nadh + (0.2) co2 |
| 15 | pyr-co2 | pyr --> co2 + (1.6667) nadh + (0.3333) atp |
| 16 | respiration | (2) nadh + o2 --> (2 P/O) atp |
| 17 | co2 | co2 --> co2extern |
| 18 | biomass | (0.0083) pol + (0.00531) r5p + (0.00145) ser + (0.00743) asp + (0.00115) lip + (0.00538) glu + (Y _{ATP}) atp --> biomass |

Tab. 2.14 Metabolites in Model 3 of *S. cerevisiae*.

| No. | Abbreviation | Full name of metabolites |
|-----|--------------|---|
| 1 | glc | Glucose |
| 2 | o2 | Oxygen |
| 3 | co2 | Carbon dioxide |
| 4 | co2extern | Carbon dioxide (extern) |
| 5 | etoh | Ethanol |
| 6 | glyc | Glycerol |
| 7 | pol | Polysaccharide |
| 8 | r5p | Ribose 5-phosphate |
| 9 | ser | Serine family |
| 10 | asp | Aspartate (Aspartate- and Pyruvat- and aromatic Family) |
| 11 | lip | Lipid |
| 12 | glu | Glutamate |
| 13 | g6p | Glucose 6-phosphate |
| 14 | atp | Adenosine triphosphate |
| 15 | gap | Glyceraldehyde 3-phosphate |
| 16 | nadh | Nicotinamide adenine dinucleotide - reduced |
| 17 | 3pg | 3-Phosphoglycerate |
| 18 | pyr | Pyruvate |
| 19 | acald | Acetaldehyde |
| 20 | biomass | Biomass |

Tab. 2.15 G^T matrix of Model 3 with 14 internal metabolites (rows) and 18 reactions (columns).

| Metabolites | Reactions | | | | | | | | | | | | | | | | | |
|-------------|-----------|----|----|---------|---------|---------|--------|----|--------|--------|--------|------|----|------|--------|------|----|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 1 co2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | -0.143 | 0.3333 | 0 | 0 | 0.2 | 1 | 0 | -1 | 0 |
| 2 pol | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0083 |
| 3 r5p | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00531 |
| 4 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00145 |
| 5 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00743 |
| 6 lip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | -0.00115 |
| 7 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -0.00538 |
| 8 g6p | 1 | -1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 atp | -0.1667 | 0 | 0 | -0.1667 | -0.1667 | 0 | 0.3333 | 0 | 0.3333 | -0.214 | 0 | 0 | 0 | -0.2 | 0.3333 | 2*PO | 0 | -Yatp |
| 10 gap | 0 | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 nadh | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0.3333 | 0 | 0 | 0 | 0 | -0.5 | 0 | 0.4 | 1.6667 | -2 | 0 | 0 |
| 12 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.857 | -1 | 0 | 0 | -1.2 | -1 | 0 | 0 | 0 |
| 14 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 |

Because some reactions of Model 2 are not handled in Model 3, Tab. 2.16 compares the simulation results from Model 2 and Model 3 using MFA when given one or two measured reaction rates as known variable. Under three different growth conditions the flux distribution of two models are very close, however some noticeable small deviations can be found between them. In Model 3 the fluxes for biomass synthesis show higher values than those in Model 2 (The real biomass synthesis rate should be 0.1, 0.24 and 0.34 C-mol/(gh) for anaerobic, oxidative and aerobic fermentative growth respectively). The reason may be the extensive simplification process of biomass synthesis system left out some detail constraints. The ignored NADPH in Model 3 might also be a possible reason.

Tab. 2.16 Comparison of flux distribution in Model 2 and Model 3 under different growth conditions. The numbers with dark background are the given rates v_m . All the rates are determined by linear programming using underdetermined system. The used measured data can be found in Tab. 6.1 and Tab. 6.2 in section 6.1. The columns of “difference” are absolute values of differences between two models.

| Reaction | Rate in Anaerobic Growth C-mol/(gh) | | | Rate in Oxidative Growth C-mol/(gh) | | | Rate in Aerobic Fermentative Growth C-mol/(gh) | | |
|-----------|--|---------|------------|--|----------|------------|---|----------|------------|
| | Model 2 | Model 3 | Difference | Model 2 | Model 3 | Difference | Model 2 | Model 3 | Difference |
| glc | 0.0333 | 0.0333 | | 0.01679 | 0.01679 | | 0.05299 | 0.05299 | |
| pol | 0.00128 | 0.00092 | 0.00036 | 0.00312 | 0.00230 | 0.00082 | 0.00319 | 0.00216 | 0.00103 |
| glyc | 0.00312 | 0.00248 | 0.00064 | 2.51E-11 | 1.56E-15 | 2.51E-11 | 2.96E-16 | 7.40E-16 | 4.44E-16 |
| ser | 0.00059 | 0.00016 | 0.00043 | 0.00145 | 0.00040 | 0.00105 | 0.00148 | 0.00038 | 0.00110 |
| asp | 0.00097 | 0.00082 | 0.00015 | 0.00238 | 0.00206 | 0.00032 | 0.00242 | 0.00193 | 0.00049 |
| pyr-acald | 0.02644 | 0.02774 | 0.00130 | 0.00027 | 0.00048 | 0.00021 | 0.03990 | 0.04165 | 0.00175 |
| etoh | 0.01756 | 0.01837 | 0.00081 | 7.87E-16 | 1.11E-15 | 3.23E-16 | 0.02642 | 0.02747 | 0.00105 |
| lip | 8.89E-05 | 0.00013 | 3.83E-05 | 0.00022 | 0.00032 | 0.00010 | 0.00022 | 0.00030 | 0.00008 |
| glut | 0.00049 | 0.00059 | 0.00010 | 0.00121 | 0.00149 | 0.00028 | 0.00123 | 0.00140 | 0.00017 |
| o2 | 0 | 0 | | 0.00865 | 0.00963 | 0.00098 | 0.00507 | 0.00507 | |
| co2 | 0.00931 | 0.00925 | 6.67E-05 | 0.00869 | 0.00876 | 7.00E-05 | 0.01832 | 0.01798 | 0.00034 |
| biomass | 0.10217 | 0.11057 | 0.00840 | 0.25021 | 0.27673 | 0.02652 | 0.25517 | 0.25986 | 0.00469 |

The phosphorylated form of NADH, coenzyme NADPH is used in anabolism, e.g. in the reactions for lipid and nucleic acid synthesis, as reducing agent to provide electrons and protons, while NADH is gained from TCA Cycle and glycolysis in catabolism and is oxidized in the respiratory chain to produce ATP. Despite of this difference, both of them participate in metabolism as reducing force. Sometimes, for simplifying models the NADPH can be considered as NADH like in Model 3. To determine how the lumping of NADPH into NADH affects the metabolic models, Model 2F, in which NADPH is substituted by NADH, was used to test the simulation accuracy. Tab. 2.17 compares the estimated metabolic flux in Model 2F and Model 2. If a determined system is used, the result still agrees with the Model 2, but the result from underdetermined system (with less measured data) has a little deviation in the reactions of pentose phosphate pathway (see rates of reaction 4 ~ 6 for underdetermined system of Model 2F in Tab. 2.17) and the synthesis of lipid and biomass (see rates of reaction 34 and 36 for underdetermined system of Model 2F in Tab. 2.17). These results imply that, even though the lumping of NADH and NADPH is practically convenient and computational equivalent in determined systems, it influences the integrity of the model and its predictive capability due to less internal restriction.

Tab. 2.17 Flux distribution of Model 2 and Model 2F in anaerobic growth. In Model 2F, NADPH is substituted with NADH. The numbers with dark background are the given rates v_m . (Complete results can be found in appendixes 6.4.4.)

| No. | Reaction Designation | Given Rate C-mol/(gh) | Rate in Model 2 C-mol/(gh) | | Rate in Model 2F C-mol/(gh) | |
|-----|----------------------|--------------------------|-------------------------------|-----------------|--------------------------------|-----------------|
| | | | determined | Underdetermined | determined | underdetermined |
| 1 | glc | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 |
| 3 | pol | | 0.00127537 | 0.00127536 | 0.00127538 | 0.00146662 |
| 4 | g6p-r5p | | 0.00200394 | 0.00200393 | 0.00200337 | 5.17E-12 |
| 5 | r5p-e4p | | 0.0011056 | 0.0011056 | 0.00110529 | -8.72E-06 |
| 6 | e4p-gap | | 0.0007647 | 0.0007647 | 0.00076442 | -0.00027285 |
| 10 | glyc | 0.00286 | 0.00312477 | 0.00312479 | 0.00312474 | 0.0012893 |
| 18 | etoh | 0.01756 | 0.01755576 | 0.01755574 | 0.01755576 | 0.01865136 |
| 29 | co2 | 0.0095 | 0.00931336 | 0.00931336 | 0.00931336 | 0.00955744 |
| 31 | dna | | 1.0768E-05 | 1.0768E-05 | 1.0768E-05 | 1.24E-05 |
| 32 | rna | | 0.00016162 | 0.00016162 | 0.00016163 | 0.00018586 |
| 33 | protein | | 0.00164376 | 0.00164376 | 0.00164378 | 0.00189027 |
| 34 | lipid | | 8.8887E-05 | 8.8887E-05 | 8.8888E-05 | 0.00010222 |
| 35 | respiration | 0 | 0 | 0 | 0 | 0 |
| 36 | biomass | 0.10302 | 0.10216887 | 0.1021687 | 0.10216976 | 0.1174903 |

2.2.2.4 Sensitivity of Models

As an index of sensitivity evaluation for models, the condition number (the ratio of the largest to the smallest eigenvalue) of \mathbf{G}^T matrices was calculated on Matlab for all the studied models. A condition number larger than 100 indicates the model is ill-conditioned and requires rechecking of all the above steps. Fig. 2.10 shows the comparison of the condition numbers of \mathbf{G}^T matrix for aforesaid models. The differences of the derivative models are summarized in Tab. 2.18. Except for Model 2 α and Model 2E, the matrix \mathbf{G}^T of all other models have reasonable condition numbers lower than 100 (note: the lower the number is, the more stable the model will be), which can be considered as insensitive systems against the variations of measurements. The increased condition number of Model 2 α and Model 2E maybe due to the unjustified methyl-FH₄ balance. As discussed in section 2.2.2.2, the coefficients of methyl-FH₄ are questionable and do not fit the model well. The non-appropriate methyl-FH₄ balance made the model unstable and very sensitive to the errors of measured data. Although Model 1 also contains the methyl-FH₄ balance, it does not enforce the fixed linear relationship among biomass components related to the pseudo-reaction of biomass synthesis, thus the sensitivity measurement is not elevated (note: it does not mean Model 1 is accurate).

The condition number of the \mathbf{G}^T matrix is a very convenient index for evaluating and adjusting the models. Combined with proper model tuning strategy, it can be helpful to locate the key elements and transform an ill-conditioned model into an eligible model.

Tab. 2.18 Overview of the investigated models.

| Model name | Differences among Models | Condition number |
|------------------|---|------------------|
| Model 1 | an initial simplified stoichiometric model | 75 |
| Model 2 α | Model 1 with an additional pseudo-reaction of biomass synthesis | 16635 |
| Model 2 | Model 2 α with adjusted coefficients of biomass components and ATP, Model 2 α without methyl-FH ₄ balance | 21 |
| Model 2A | Model 2 without the pseudo-reaction of biomass synthesis | 17 |
| Model 2B | Model 2 with original k_n in the pseudo-reaction of biomass synthesis | 21 |
| Model 2C | Model 2 without energy summary ($Y_{ATP} \cdot \mu$) in the pseudo-reaction of biomass synthesis | 81 |
| Model 2D | Model 2 with m_{ATP} reaction | 22 |
| Model 2E | Model 2 with methyl-FH ₄ balance | 18492 |
| Model 2F | NADPH in Model 2 substituted with NADH | 20 |
| Model 3 | further simplified version of Model 2 | 7 |

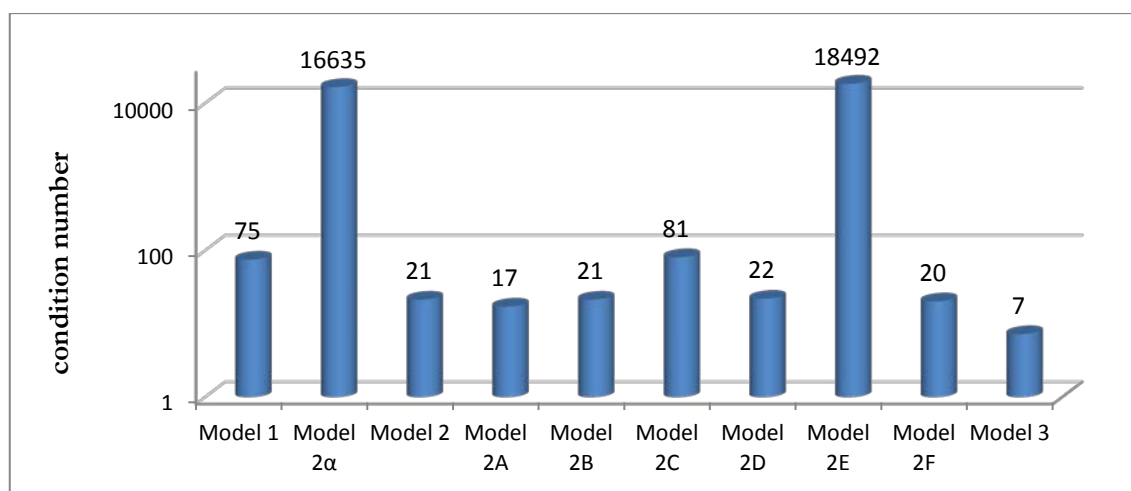


Fig. 2.10 Condition number of G^T matrix for the investigated models. For better visual perception the y-axis is drawn on a logarithmic scale of the condition number.

2.3 Homeostatic Metabolic Pathway Model

MFA can be used to analyze a steady state of the organism under a certain growth condition, however, it is difficult to directly observe the flux changes under different growth conditions, because there are so many reactions. With metabolic pathway analysis, since reactions are organized into pathway, it is easier to observe the alteration of the flux distribution in different pathways, while facilitates understanding the behavior of *S. cerevisiae* under different growth condition.

In this thesis, the studies of metabolic pathway analysis focused on exploring the variation of fluxes on different pathways using the elementary (flux) modes (EMs). As explained in section 2.1.4.2, the EMs includes all basic metabolically meaningful pathways of a model. The metabolic flux coefficients of the elementary modes can be determined by optimizing the specific formation rate of biomass using linear programming. The analysis and simulation was based on the established models of *S. cerevisiae*, which are described above. In the following part of this section, it will be explained how the EMs of the models are generated.

2.3.1 Determination of Elementary Flux Modes

To calculate the elementary flux modes, software Efmtool (version: 4.7.1) [162] was used, which is based on double description method algorithm. The coefficients of components in each pathway were normalized in relation to the first reaction (glucose uptake rate).

During some preliminary tests, it was noticed that the number of extracted elementary flux modes might vary if different P/O value and Y_{ATP} were given. A series of elementary modes extractions were performed by varying both P/O value and Y_{ATP} in a reasonable range.

2.3.2 Elementary Flux Modes of *S. cerevisiae*

Tab. 2.19 lists all the elementary modes of stoichiometric model \mathbf{G}^T for Model 2 with $P/O = 1.65$ and $Y_{\text{ATP}} = 0.0746$ mol/g, this pair of parameters is selected according to the validation step described in section 2.1.3.3, but also considering that the number of elementary flux modes should be maximized (discussed below). The calculated elementary flux modes with Efmtool are normalized according to the glucose uptake rate. This implies that all the elementary modes are assumed to have the same substrate uptake rate, therefore their biomass production/yeilding rate can be easily expressed by the normalized biomass synthesis reaction rate. The elementary modes are classified in three groups representing three different growth conditions. The first four elementary modes ($\text{EM}_1 \sim \text{EM}_4$) correspond to anaerobic growth, the five elementary modes in the middle ($\text{EM}_5 \sim \text{EM}_9$) correspond to oxidative growth and the last four elementary modes ($\text{EM}_{10} \sim \text{EM}_{13}$) correspond to aerobic fermentation. EM_1 , EM_5 , EM_{10} are three major pathways with high biomass yield and represent three growth states respectively. The remaining elementary modes are derived pathways with differences in one or two reactions.

The above classification of different elementary modes is based on the observation of relevant reaction rates. If the rate of the oxidative phosphorylation reaction is zero, then the elementary mode is regarded as anaerobic metabolic pathway. The remaining modes are divided into oxidative pathways and aerobic fermentative pathways according to the production rate of ethanol. In each group one main elementary mode can be identified. In the anaerobic metabolic group and the aerobic fermentative group, the main elementary modes are identified by finding the pathways where the reaction rates of TCA Cycle (reaction 23 ~ 25) are zero; in the oxidative group, the main elementary mode is identified by finding the pathway which has no Glycerol production. It is easy to notice these main elementary modes are all the classic pathways found in textbooks.

Tab. 2.19* Elementary flux mode of Model 2 with $P/O = 1.65$ and $Y_{ATP} = 0.0746$ mol/g.

| No. | Reaction Designation | for anaerobic growth | | | | for oxidative growth | | | | | for aerobic fermentative growth | | | |
|-----|----------------------|----------------------|-----------------|-----------------|-----------------|----------------------|-----------------|-----------------|-----------------|-----------------|---------------------------------|------------------|------------------|------------------|
| | | EM ₁ | EM ₂ | EM ₃ | EM ₄ | EM ₅ | EM ₆ | EM ₇ | EM ₈ | EM ₉ | EM ₁₀ | EM ₁₁ | EM ₁₂ | EM ₁₃ |
| 1 | Glc | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | g6p-f6p | 0.90155286 | 0.97250698 | 0.98229149 | 0.98254404 | 0.538176315 | 0.828478884 | 0.861324947 | 0.925555161 | 0.926930989 | 0.850255281 | 0.88082249 | 0.949818818 | 0.950972695 |
| 3 | Pol | 0.03828749 | 0.02003757 | 0.01752092 | 0.01745596 | 0.238989413 | 0.115088346 | 0.101069654 | 0.073656214 | 0.073069011 | 0.058237848 | 0.086859384 | 0.049649592 | 0.049027305 |
| 4 | g6p-r5p | 0.06015965 | 0.00745546 | 0.00018759 | 0 | 0.222834272 | 0.05643277 | 0.037605399 | 0.000788625 | 0 | 0.091506872 | 0.032318126 | 0.00053159 | 0 |
| 5 | r5p-e4p | 0.03319097 | 0.00402233 | 0 | -0.00010382 | 0.122363014 | 0.030663899 | 0.020288673 | 0 | -0.00043459 | 0.050485689 | 0.01743611 | 0 | -0.0002916 |
| 6 | e4p-gap | 0.02295693 | 0 | -0.0031657 | -0.00324745 | 0.066956252 | 0.00680569 | 0 | -0.01330847 | -0.01359354 | 0.034919039 | 0 | -0.00897087 | -0.00912089 |
| 7 | f6p-gap | 0.93677282 | 0.97492037 | 0.98018089 | 0.98031667 | 0.656233857 | 0.851414576 | 0.87349815 | 0.916682403 | 0.917607421 | 0.903827217 | 0.891284156 | 0.94383794 | 0.944716837 |
| 8 | dhap-gap | 0.37440086 | 0.17398632 | 0.14634932 | 0.14563597 | 0.327014273 | 0 | -0.03699978 | -0.10935256 | -0.11090237 | 0.451644909 | 0 | 0 | 0 |
| 9 | dhap-g3p | 0.09398555 | 0.31347386 | 0.34374112 | 0.34452236 | 0.001102656 | 0.425707288 | 0.473748856 | 0.567693759 | 0.569706083 | 0.000268699 | 0.445642078 | 0.47191897 | 0.472358418 |
| 10 | Glyc | 0.0938089 | 0.31338141 | 0.34366028 | 0.34444182 | 0 | 0.425176291 | 0.473282538 | 0.567353922 | 0.569368956 | 0 | 0.445241324 | 0.471689896 | 0.472132215 |
| 11 | gap-3pg | 0.85101618 | 0.66174867 | 0.63564883 | 0.63497516 | 0.681051592 | 0.429711114 | 0.401273387 | 0.345663639 | 0.344472463 | 0.916075239 | 0.446951885 | 0.469677677 | 0.470057738 |
| 12 | Ser | 0.01778608 | 0.00930826 | 0.00813917 | 0.008109 | 0.111020203 | 0.05346317 | 0.046950923 | 0.034216277 | 0.033943497 | 0.027053825 | 0.040349681 | 0.023064234 | 0.022775157 |
| 13 | 3pg-pep | 0.8443464 | 0.65825807 | 0.63259665 | 0.63193429 | 0.639419016 | 0.409662425 | 0.383666791 | 0.332832535 | 0.331743652 | 0.905930055 | 0.431820755 | 0.461028589 | 0.461517054 |
| 14 | pep-pyr | 0.83973244 | 0.65584338 | 0.63048523 | 0.6298307 | 0.610618844 | 0.395793342 | 0.371487074 | 0.323956361 | 0.322938241 | 0.898911919 | 0.421353492 | 0.455045409 | 0.455608864 |
| 15 | pyr-oxac | 0.02573443 | 0.01346839 | 0.01177691 | 0.01173325 | 0.160636122 | 0.077357183 | 0.06793465 | 0.049508964 | 0.049114281 | 0.039143802 | 0.058383122 | 0.033372607 | 0.032954337 |
| 16 | Asp | 0.02911709 | 0.01523828 | 0.01332441 | 0.01327501 | 0.18174803 | 0.087523041 | 0.076862026 | 0.056014497 | 0.055567938 | 0.04428905 | 0.066055319 | 0.037757804 | 0.037284564 |
| 17 | pyr-acald | 0.79415625 | 0.58393824 | 0.55494935 | 0.5542011 | 0.020801175 | 0.010017066 | 0.008796907 | 0.006410894 | 0.006359785 | 0.829587471 | 0.109657119 | 0.240996899 | 0.243193392 |
| 18 | etoh | 0.52724222 | 0.38814888 | 0.36896802 | 0.36847294 | 0 | 0 | 0 | 0 | 0 | 0.549706525 | 0.068068101 | 0.157791553 | 0.159292065 |
| 19 | Ace | 0.00222176 | 0.00116275 | 0.00101671 | 0.00101294 | 0.013868143 | 0.006678378 | 0.005864898 | 0.004274143 | 0.004240069 | 0.003379442 | 0.0050403 | 0.00288108 | 0.00284497 |
| 20 | actcoa | 0.00222176 | 0.00116275 | 0.00101671 | 0.00101294 | 0.013868143 | 0.006678378 | 0.005864898 | 0.004274143 | 0.004240069 | 0.003379442 | 0.0050403 | 0.00288108 | 0.00284497 |
| 21 | pyr-iso | 0.02239084 | 0.12384165 | 0.13783164 | 0.13819274 | 0.852203207 | 0.64778688 | 0.62465835 | 0.579430695 | 0.578461907 | 0.034057974 | 0.536832149 | 0.39057801 | 0.388132093 |
| 22 | iso-akg | 0.01919209 | 0.10614963 | 0.11814101 | 0.11845052 | 0.730457457 | 0.555244046 | 0.535419658 | 0.496653226 | 0.495822839 | 0.029192452 | 0.460140308 | 0.334780036 | 0.332683542 |
| 23 | akg-succ | 0 | 0.08008793 | 0.09113199 | 0.09141705 | 0.508883977 | 0.414628642 | 0.403964194 | 0.383109951 | 0.382663247 | 0 | 0.347167322 | 0.258243693 | 0.256756556 |
| 24 | succ-mal | 0 | 0.06407034 | 0.07290559 | 0.07313364 | 0.407107181 | 0.331702914 | 0.323171355 | 0.30648796 | 0.306130598 | 0 | 0.277733857 | 0.206594954 | 0.205405245 |
| 25 | mal-oxam | 0 | 0.06407034 | 0.07290559 | 0.07313364 | 0.407107181 | 0.331702914 | 0.323171355 | 0.30648796 | 0.306130598 | 0 | 0.277733857 | 0.206594954 | 0.205405245 |
| 26 | oxac-oxam | 0.0127948 | 0.00669649 | 0.00585554 | 0.00583384 | 0.079867297 | 0.038461943 | 0.033777166 | 0.024616121 | 0.024419889 | 0.019461748 | 0.029028138 | 0.016593038 | 0.016385077 |
| 27 | akg-glu | 0.06760048 | 0.03537837 | 0.03093497 | 0.03082028 | 0.421960202 | 0.203200222 | 0.178448791 | 0.130047564 | 0.129010797 | 0.102824865 | 0.153359109 | 0.087661422 | 0.086562711 |
| 28 | glu-gln | 0.01484284 | 0.00776792 | 0.0067923 | 0.00676712 | 0.092648562 | 0.044616076 | 0.039181477 | 0.028554161 | 0.028326522 | 0.022576953 | 0.033672609 | 0.019247561 | 0.01900632 |
| 29 | co2 | 0.27969687 | 0.2465263 | 0.24195211 | 0.24183404 | 0.380469498 | 0.276480255 | 0.264714471 | 0.241706575 | 0.24121374 | 0.299324499 | 0.261524885 | 0.241811578 | 0.241481898 |
| 30 | r5p-5aic | 0.00434626 | 0.00227459 | 0.00198891 | 0.00198154 | 0.027129225 | 0.013064418 | 0.011473066 | 0.00836119 | 0.008294533 | 0.006610953 | 0.009859967 | 0.005636044 | 0.005565405 |
| 31 | Dna | 0.00032326 | 0.00016918 | 0.00014793 | 0.00014738 | 0.002017776 | 0.000971686 | 0.000853326 | 0.000621876 | 0.000616918 | 0.000491699 | 0.00073335 | 0.000419189 | 0.000413935 |
| 32 | Rna | 0.00485211 | 0.00253933 | 0.00222039 | 0.00221216 | 0.030286703 | 0.014584941 | 0.012808377 | 0.009334321 | 0.009259906 | 0.007380379 | 0.011007535 | 0.006292004 | 0.006213143 |
| 33 | protein | 0.04934711 | 0.02582556 | 0.02258196 | 0.02249824 | 0.308023225 | 0.148332443 | 0.130264351 | 0.094932341 | 0.094175521 | 0.075060269 | 0.111949342 | 0.063991234 | 0.063189195 |
| 34 | Lipid | 0.00266846 | 0.00139652 | 0.00122112 | 0.0012166 | 0.01665643 | 0.008021112 | 0.007044076 | 0.005133489 | 0.005092564 | 0.004058902 | 0.006053687 | 0.003460341 | 0.003416971 |
| 35 | respiration | 0 | 0 | 0 | 0 | 0.377587305 | 0.204237788 | 0.184624289 | 0.146270248 | 0.145448695 | 0.023788688 | 0.152247474 | 0.083716834 | 0.082570745 |
| 36 | biomass | 3.06719063 | 1.60519875 | 1.40359125 | 1.39838748 | 19.1453147 | 9.219666163 | 8.096636215 | 5.900560106 | 5.853519587 | 4.665402993 | 6.958259051 | 3.977402383 | 3.927551295 |

*: All the values are normalized with glucose uptake rates.

Beside those main pathways, there are also some alternative pathways in each group. This is usually because there exist circles in a metabolic network, and some subsidiary pathways can go around the circle from another side. For example in Tab. 2.19, under anaerobic growth, EM_1 goes through the reaction 4 ~ 6 of pentose phosphate pathway (PPP) (see Fig. 2.11A), while $EM_2 \sim EM_4$ include one reaction less respectively in PPP as shown in Fig. 2.11B ~ D.

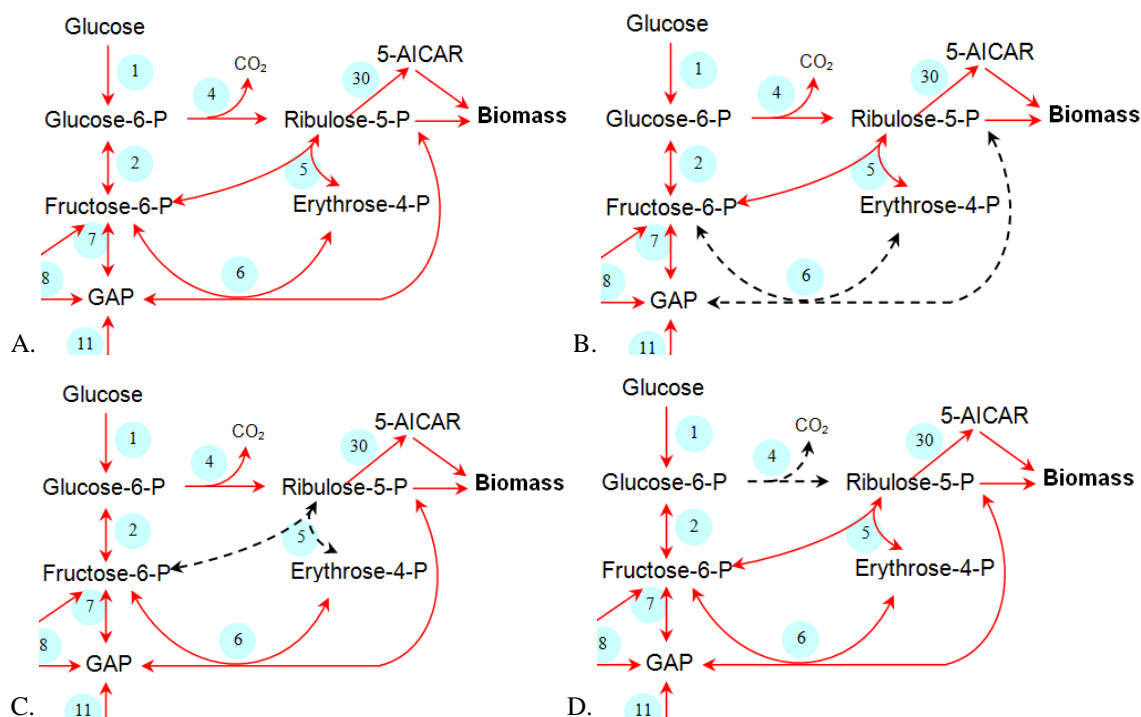


Fig. 2.11 Schematic illustration of active pathways in PPP of elementary flux mode of Model 2. A, B, C and D show partial pathways of EM_1 , EM_2 , EM_3 and EM_4 respectively. Red solid arrows represent active reactions, and black dash arrows represent inactive reaction in a pathway.

Another finding in this study is that when the stoichiometric matrix \mathbf{G}^T is changed by the alteration of P/O ratio and Y_{ATP} value, the number of elementary modes of \mathbf{G}^T will also change. The number of flux pathways (i.e. the column number of EM matrix) normally decreases while decreasing P/O ratio and keeping the Y_{ATP} constant (see Tab. 2.20). The disappearing elementary modes are usually the subsidiary pathways and involve the reactions of pentose phosphate pathway and production of glycerol. A similar phenomenon happens when altering Y_{ATP} and keeping P/O ratio constant. The number of elementary modes decreases with increasing Y_{ATP} value (see Tab. 2.20). The disappearing modes are secondary elementary modes in the oxidative growth group and aerobic fermentation group. This can again be explained by the increasing ATP consumption for biomass synthesis, which expels the consumption in other reactions.

From a geometrical point of view, it can be explained as the rotation of the null space of \mathbf{G}^T causing the edge of the convex cone to change, because the convex cone is the intersection of the null space and the inequalities in Eq. (2.57).

Tab. 2.20 The number of elementary modes of Model 2 under variation of P/O ratio and Y_{ATP} value. The circled item represents the condition that was used in the following simulation tests. This pair of P/O ratio and Y_{ATP} value is the most close to the validated Model 2 while the model still contains thirteen elementary modes.

| Y_{ATP} | PO | | | | | | |
|-----------|----|----|------|------|------|------|----|
| | 3 | 2 | 1.65 | 1.39 | 1.21 | 1.06 | 1 |
| 0.05 | 13 | 13 | 13 | 13 | 13 | 13 | 13 |
| 0.0746 | 13 | 13 | 13 | 12 | 12 | 9 | 9 |
| 0.088 | 13 | 13 | 12 | 12 | 9 | 9 | 9 |
| 0.117 | 13 | 12 | 9 | 9 | 9 | 8 | 8 |
| 0.14 | 13 | 9 | 9 | 8 | 8 | 8 | 8 |
| 0.2 | 9 | 8 | 8 | 8 | 8 | 8 | 8 |
| 0.3 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

2.4 Predictive Homeostatic Metabolic Model

In this chapter, a biomass-maximizing metabolic pathway analysis algorithm was tested to extend the pathway model from previous step to a predictive homeostatic metabolic model. It was first validated using the experimental data set, and then utilized to simulate the flux changes of *S. cerevisiae* under various growth conditions.

2.4.1 Algorithms for Flux Regulation Coefficients

The non-negative linear combinations of the EMs provide all possible states of the metabolic network. An arbitrary flux of the network can be expressed as in Eq. (2.69)

$$\mathbf{v} = \sum_{i=1}^k \lambda_i \mathbf{E}_i = \mathbf{E} \cdot \boldsymbol{\lambda}, \quad \lambda_i \geq 0 \quad (2.69)$$

where \mathbf{v} is a n -dimensional flux vector. The linear combinations of k EMs in the EM matrix \mathbf{E} (including k vector \mathbf{E}_i) are regulated by non-negative coefficients (λ_i). The vector $\boldsymbol{\lambda}$ represents all the flux coefficients λ_i . For a normalized matrix \mathbf{E} , the sum of these coefficients should equal one, since λ_i represents the relative contribution of the flux modes.

$$\sum_{i=1}^k \lambda_i = 1 \quad (2.70)$$

Given the measured/expected reaction rate of certain substrate(s) r_s and oxygen r_o (optional) and a predefined EM matrix \mathbf{E} , the flux coefficients λ can be determined by maximizing the biomass production rate (μ), using a linear programming method, as illustrated in Eq. (2.71)

$$\mu_{max} = \max \left\{ \mathbf{c}_\mu \lambda \mid \lambda_i \geq 0 \cap \sum_{i=1}^k \lambda_i = 1 \cap \mathbf{c}_s \lambda = r_s \cap \mathbf{c}_o \lambda = r_o \right\} \quad (2.71)$$

where \mathbf{c}_μ is the row vector of matrix \mathbf{E} corresponding to biomass synthesis and $c_{\mu i}$ is the reaction rate of biomass synthesis in \mathbf{E}_i . \mathbf{c}_s and \mathbf{c}_o are row vectors representing the reaction rates that involve substrate and oxygen respectively in matrix \mathbf{E} .

With this approach, the changes of metabolic flux distribution on different pathway under different growth conditions can be determined with very few measurable parameters. In the following section the estimated flux distribution will be compared with the experimental data and the results of metabolic flux analysis.

2.4.2 Growth Condition Simulations

In this study, the growth of *S. cerevisiae* at steady state in chemostat was simulated using the biomass-maximizing metabolic pathway analysis algorithm described in the previous section. This is driven by the hypothesis that a cell, no matter under what condition, will always prioritize the growth, i.e. biomass synthesis. To validate this hypothesis, a series of case studies were performed on Model 2 to simulate the flux distribution on different pathway under various growth conditions.

The principle of these studies is using the uptake rates of glucose and oxygen from literature data [71][157] and the above optimization algorithm to determine the flux coefficients of elementary modes of the yeast under aerobic or anaerobic growth.

Using the reaction rates taken from [71] (see Tab. 6.2 in section 6.1) the oxidative growth and aerobic fermentative growth of *S. cerevisiae* were simulated with various dilution rates. Two oxidative growth simulations were conducted at $D = 0.1$ and 0.24 h^{-1} , where the glucose uptake rates were set to 0.00636 and $0.01679 \text{ C-mol/(gh)}$ respectively. Five aerobic fermentative growth simulations under $D = 0.26, 0.28, 0.3, 0.34, 0.4 \text{ h}^{-1}$

were conducted by setting the glucose uptake rates to 0.02223, 0.02893, 0.03663, 0.05299, 0.07775 C-mol/(gh) respectively and the oxygen uptake rates to 0.00757, 0.00694, 0.00632, 0.00507, 0.0032 C-mol/(gh) respectively. All the simulations were performed on the elementary modes of Model 2 with $P/O = 1.65$ and $Y_{\text{ATP}} = 0.0746$ mol/g (see Tab. 2.19 on page 60). With the determined flux coefficients of each elementary mode, the reaction rates of each reaction can be computed. In all the simulations, the computed reaction rates (Ethanol, Glycerol and CO_2) were then compared with the unused reaction rates from [71] (see Tab. 6.2 in appendixes 6.1) to validate the biomass maximization method.

It is known that the growth of *S. cerevisiae* is purely oxidative and only limited by the substrate uptake rate when the glucose uptake rate is under a certain threshold and oxygen supply is high. Therefore, the flux coefficients of elementary modes would be determined by normalizing the substrate glucose uptake rate. When the substrate uptake rate is higher than the threshold, the growth of *S. cerevisiae* turns into a mixture of respiratory fermentation (Crabtree effect) as introduced in section 2.1.1.3. At this stage of growth the oxygen uptake rate could be used besides the glucose uptake rate to determine the flux coefficients of elementary modes.

Fig. 2.12 schematically illustrates the distribution of the flux coefficients λ from the elementary flux mode analysis during the aerobic growth. It can be seen that, when the dilution rate D in the Chemostat is between 0.1 h^{-1} and 0.24 h^{-1} , the glucose uptake rates do not reach the threshold and the organism is in a purely oxidative growth, i.e. only the main oxidative pathway EM_5 is active (note: the flux coefficient λ_5 is close to one and others are near zero). When the glucose uptake rate goes above the threshold and the dilution rate is within 0.26 h^{-1} to 0.4 h^{-1} , the growth turns into aerobic fermentation. The contribution of EM_5 decreases, while the main pathway for aerobic fermentation EM_{10} increases, (cf. λ_5 and λ_{10} in Fig. 2.12). These flux-changing trends coalesce closely with the known behavior of *S. cerevisiae*.

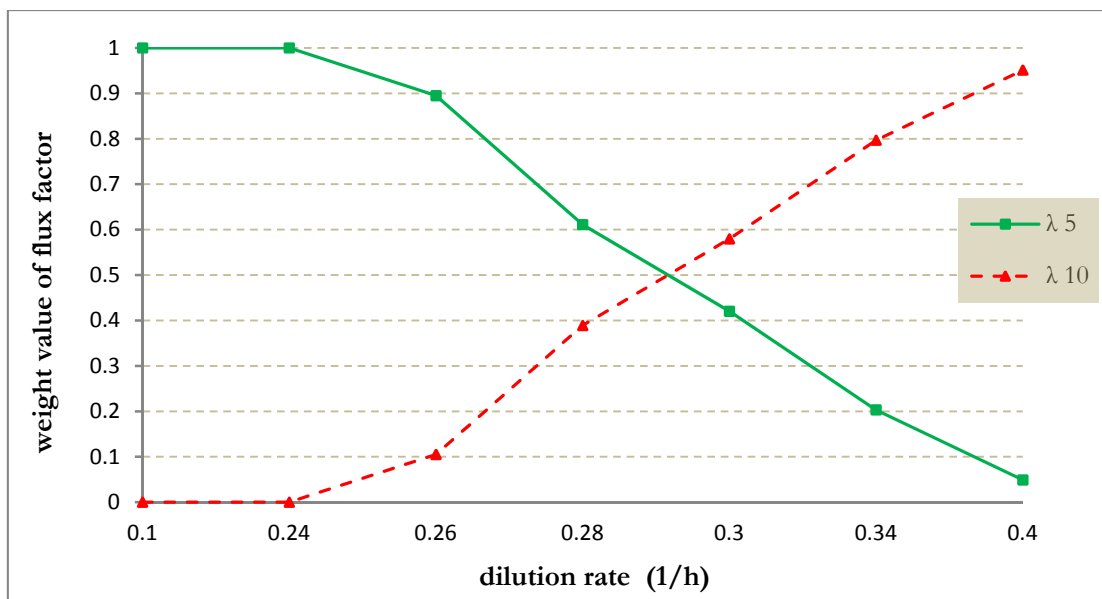


Fig. 2.12 Estimated flux coefficients for simulated aerobic chemostat growth of *S. cerevisiae* at steady state referring corrected reaction rates taken from [71]. Only two EMs (EM₅, and EM₁₀) are active during the growth, while the flux coefficients of the other eleven EMs are near zero, which are not shown in this figure.

Fig. 2.13 compares some key reaction rates of glucose, ethanol, glycerol, oxygen and carbon dioxide changing over the tested range of growth rates. Note that, except for the uptake rate of glucose and oxygen, the remaining measured reaction rates were not fixed / preset in the simulation. The experimental data from literature [71] are shown as points for comparison with the estimated reaction rates (lines). The overall agreement between the estimated reaction rates and measured reaction rates are rather good despite of some small deviations. In the purely oxidative growth state, ($D = 0.1 \text{ h}^{-1} \sim 0.24 \text{ h}^{-1}$) the consumption rate of oxygen increases when the uptake rate of glucose increases, and has almost the same value as the carbon dioxide production rate, since the respiration quotient RQ should be around one. The pathways for production of ethanol and glycerol are not active in the pure oxidative growth. For higher dilution rates (i.e. when D is between 0.26 h^{-1} and 0.4 h^{-1}), with increasing glucose uptake, the energy demand for the growth can no longer be fulfilled by the respiratory chain, in turn causing increasing aerobic fermentative metabolism with higher carbon dioxide and ethanol production. Note that the estimated ethanol production rates are higher than the measured data, which agrees with the Nielsen and Villadsen's finding that the ethanol rates are underestimated (see section 2.1.3.3). From the perspective of optimizing biomass formation, the ideal growth with sufficient oxygen should not produce glycerol as suggested in the simulation result described above. However, small amounts of glycerol

are measured during experiments in [71]. One possible explanation is that glycerol is produced to compensate the reduced force and achieve redox balance. As discussed previously, suppressing the enzymes involved in the glycerol production might be helpful to increase the production of biomass.

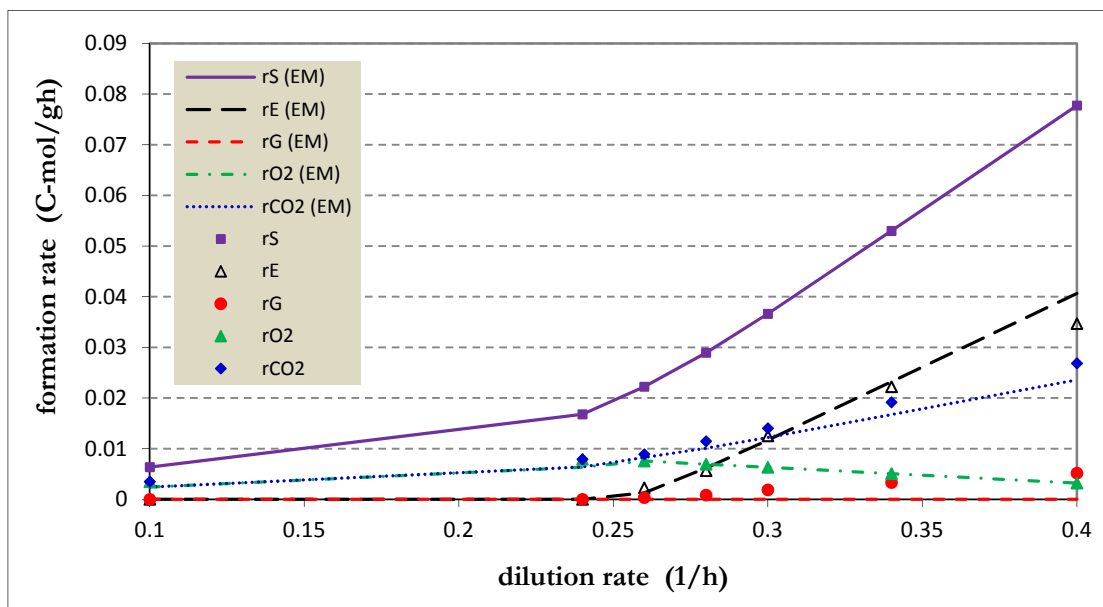


Fig. 2.13 Growth simulation of *S. cerevisiae* at steady state in aerobic chemostat culture according corrected reaction rates taken from [71]. $rS(EM)$, $rE(EM)$, $rG(EM)$, $rO_2(EM)$ and $rCO_2(EM)$ stand respectively for the calculated formation rates of glucose, ethanol, glycerol, oxygen and carbon dioxide with EMs and they are represented as lines, while rS , rE , rG , rO_2 and rCO_2 are corrected reaction rates taken from [71].

In the anaerobic growth, the simulations focused on the metabolism flux at three dilution rates D ($D = 0.1, 0.2$ and 0.3 h^{-1}). The glucose uptake rates under these three dilution rates were set to 0.0333, 0.0667 and 0.1042 C-mol/(gh) respectively according to experimental data of Nissen et al. [157] (see Tab. 6.1 in section 6.1). The oxygen uptake rates were all set to zero for the strict anaerobic growth of Nissen's experiments. The flux of Model 2 constantly goes through the EM_1 no matter at what dilution rate (see Tab. 2.21). The relation between the estimated reaction rates and the dilution rates therefore is linear. This also agrees well with the measured reaction rates as listed in Tab. 2.22.

Tab. 2.21 Estimated flux coefficients for simulated anaerobic growth of *S. cerevisiae* at stationary state referring experiment data from Nissen's study [157].

| D (h ⁻¹) | 0.1 | 0.2 | 0.3 |
|----------------------|----------|----------|----------|
| λ_1 | 1 | 1 | 1 |
| λ_2 | 2.17E-12 | 9.03E-10 | 3.44E-10 |
| λ_3 | 1.59E-12 | 2.26E-09 | 1.51E-09 |
| λ_4 | 1.58E-12 | 2.28E-09 | 1.54E-09 |
| λ_5 | 4.74E-14 | 7.86E-11 | 3.13E-10 |
| λ_6 | 1.16E-14 | 1.15E-10 | 3.69E-10 |
| λ_7 | 3.61E-15 | 1.05E-10 | 2.96E-10 |
| λ_8 | 2.89E-15 | 1.68E-11 | 6.79E-12 |
| λ_9 | 3.22E-15 | 1.33E-11 | 6.19E-13 |
| λ_{10} | 8.65E-16 | 8.89E-09 | 2.15E-08 |
| λ_{11} | 7.05E-15 | 8.41E-11 | 2.36E-10 |
| λ_{12} | 7.11E-14 | 1.68E-13 | 7.65E-10 |
| λ_{13} | 7.28E-14 | 3.24E-12 | 8.01E-10 |

Tab. 2.22 Comparison of the determined flux distributions by elementary flux mode analysis with the experiment data from [157].

| D 1/h | Difference of Rates (C-mol/(gh)) | | | | | | | |
|----------|----------------------------------|---------|-------------|---------|-------------------|---------|-------------|---------|
| | r(glycerol) | | r(ethanol) | | r(carbon dioxide) | | r(biomass) | |
| | Exper. data | EM rate | Exper. data | EM rate | Exper. data | EM rate | Exper. data | EM rate |
| 0.1 | 0.00286 | 0.00312 | 0.01756 | 0.01756 | 0.0095 | 0.00931 | 0.10302 | 0.10214 |
| 0.2 | 0.00574 | 0.00626 | 0.03315 | 0.03517 | 0.01814 | 0.01866 | 0.20634 | 0.20458 |
| 0.3 | 0.00896 | 0.00977 | 0.05179 | 0.05494 | 0.02834 | 0.02914 | 0.32235 | 0.3196 |

2.4.3 Flux Coefficient Changes under Oxygen Limitation

To study the changes of flux coefficients under oxygen limitation, the glucose uptake rate was fixed as 0.05299 C-mol/(gh), while varying the oxygen reaction rate. The biomass maximization scheme as described above was applied to find the flux coefficients. Eventually, the determined flux coefficients were plotted over the oxygen reaction rate to demonstrate the flux changes on every pathway. Fig. 2.14A shows the alteration of the flux distributions on different pathways under different oxygen supply conditions (by varying the oxygen consuming rate: reaction 35 of Model 2). λ_i is corresponding to the coefficient for the i^{th} column of the EM matrix. The changes of λ affirm the main pathways are the characteristic elementary modes (i.e. λ_1 for anaerobic EM₁, λ_5 for oxidative EM₅ and λ_{10} for aerobic fermentative EM₁₀) while the contributions of remaining elementary modes can be ignored, since their λ values are all near zero. The predicted flux coefficients clearly depict the changes of the flux

distributions under varying oxygen limitations. The first main EM for the flux of anaerobic growth is the only active pathway (EM_i is considered non-active if $\lambda_i < 10^8$) when there is no oxygen supply. As the oxygen uptake increases, the metabolic pathways change from anaerobic fermentation via aerobic fermentative stage, to the oxidative growth pathway, (note: the corresponding changes of λ_1 , λ_5 and λ_{10} of those pathways). This pathway shifting can directly be seen by comparing the pathway distribution graphs of the anaerobic fermentation (Fig. 2.15 on page 70) and oxidative respiration (Fig. 2.16 on page 71) states.

With the determined flux coefficients of elementary modes, the individual reaction rates of \mathbf{v} in Eq. (2.69) can be calculated. Fig. 2.14B illustrates the variation of some key reactions with the increasing oxygen supply under the hypothesis of constant substrate glucose uptake rate. The y-axis indicates the yield ($r_i/r_s \times 100\%$), where r_s is the fixed glucose uptake rate. As the oxygen supply increases, the production rates of biomass (red) and carbon dioxide (purple) increase; ethanol (green) and glycerol (yellow) decrease. At 0.0201 C-mol/(gh) the oxygen uptake rate reaches saturation and the organism can be seen at fully oxidative growth. The specific growth rate μ also increases with oxygen supply while the respiration quotient RQ decreases (see Fig. 2.14C).

In reality, the glucose uptake rates will change when yeast is cultured under different oxygen supply conditions, so will the distribution of the flux on different pathways. Using relative reaction rates to glucose (as shown in Fig. 2.14B) intuitively shows the relations and tendencies of the key reactions.

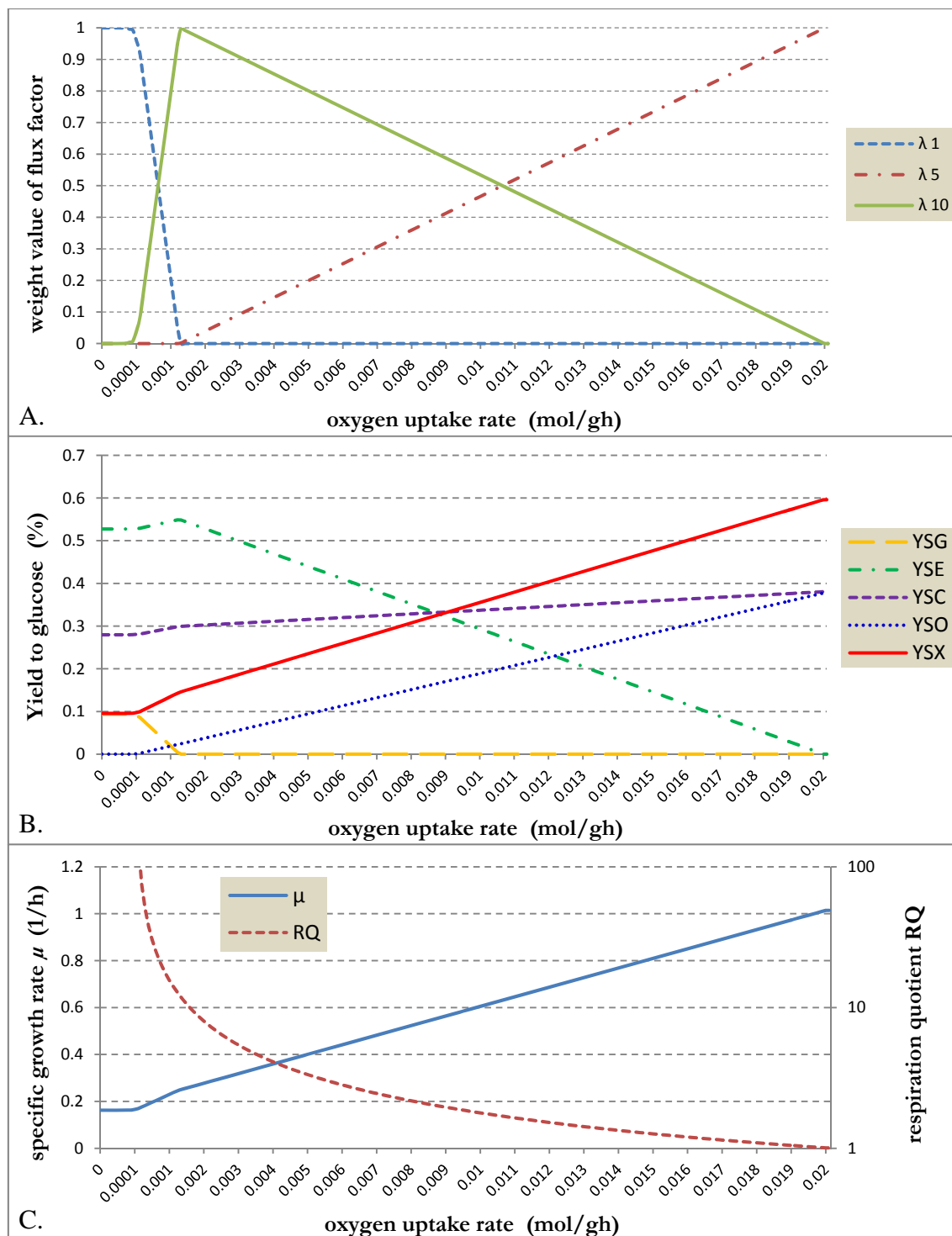


Fig. 2.14 Simulation of *S. cerevisiae* growth under different oxygen supply conditions (the same glucose uptake rate is assumed). X-axis in all three figures represents the varying oxygen uptake rate (note: from 0 to 0.0001 logarithmic scale is used, after 0.0001 linear scale is used), the maximum oxygen uptake rate is 0.0201 mol/gh). **A. Predicted flux coefficients of three main EMs.** λ_1 (blue) corresponding to flux coefficients of EM₁ stands for anaerobic growth; λ_5 (green) corresponding to flux coefficients of EM₅ stands for oxidative growth; λ_{10} (red) corresponding to flux coefficients of EM₁₀ stands for aerobic fermentative growth, which can be seen as transitional phase from anaerobic to oxidative growth. The rest flux coefficients are near zero and not shown in this figure. **B. Predicted yield to substrate ($r_i/r_s \times 100\%$).** YSG, YSE, YSC, YSO and YSX represent respectively the yield of glycerol, ethanol, carbon dioxide, oxygen and biomass to substrate glucose. **C. Predicted Specific growth rate (μ) and respiration quotient (RQ).** For better visual perception the right y-axis is drawn on a logarithmic scale of RQ. When the oxygen uptake rate approaches zero, RQ tends to explode and is not covered here.

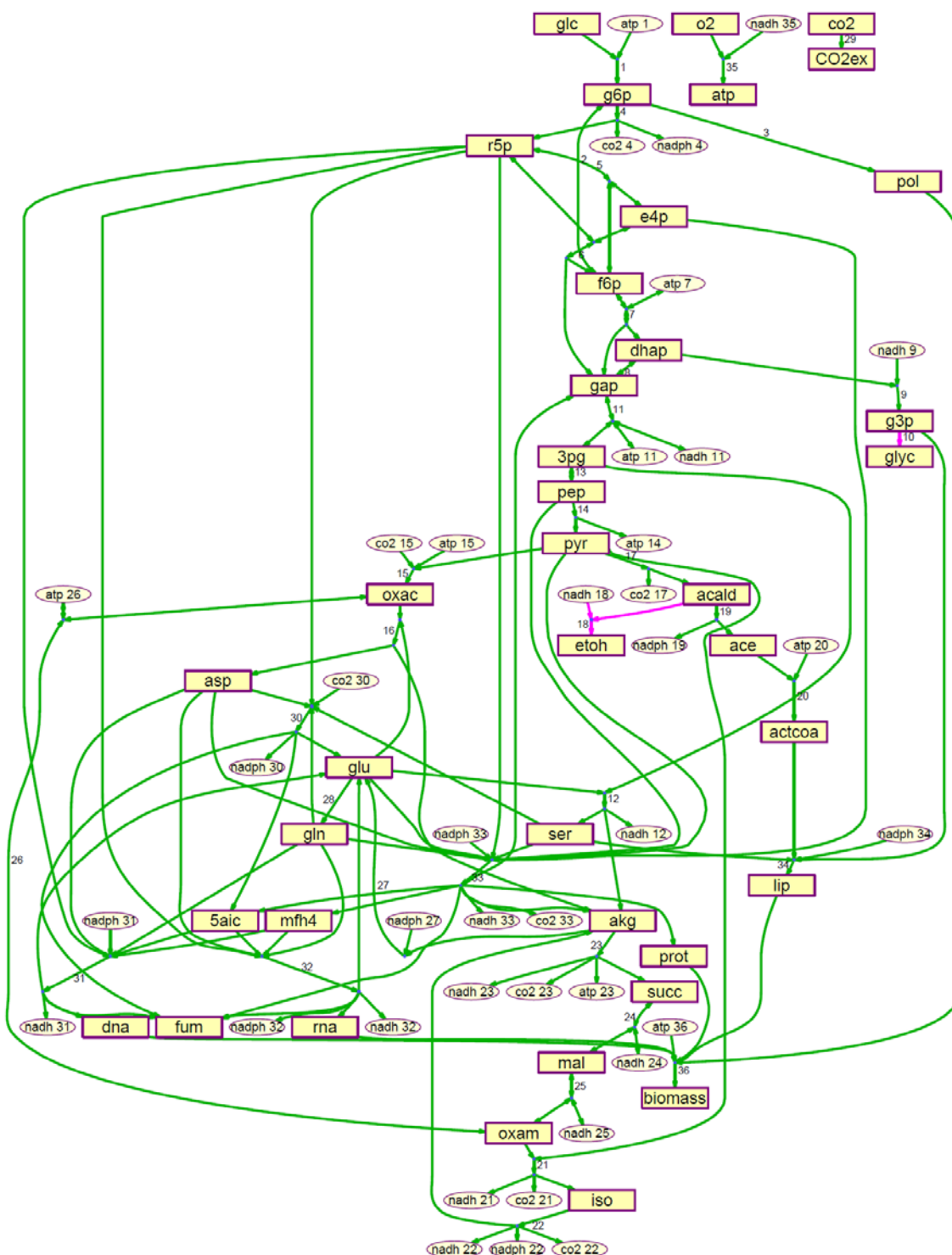


Fig. 2.16 Schematic illustration of the metabolic flux for oxidative growth of *S. cerevisiae*. The green fluxes are active metabolic pathways, while the purple fluxes are non-active pathways. This graph is created using GraphView which will be introduced in Chapter 3.

2.4.4 Contributions and Limitations

In this study, a new metabolic pathway analysis based on biomass-maximizing is proposed. The new method was tested using measured data of reaction rates, and used to simulate the metabolic changes of yeast under various growth conditions. This method can be conveniently used to predict the optimal condition for bioprocess, or find the missing pathway or bottle neck enzyme and guide the genetic engineering to improve the production rates.

Compared with other existing metabolic analysis methods introduced in section 2.1.4.3, this approach has several advantages:

Firstly, the calculation is relative simpler than most other pathway analysis methods. For example, Carlson and Srienc [39] set the optimization goal to maximize both biomass-production and energy-generation, as they argued “the energy generating pathways do not involve synthesis of biomass” and “energy can be produced independently from the biomass pathway reactions”. And in the thermodynamics method, Wlaschin et al. [49] proposed solving the pathway coefficients in classical thermodynamics as an open system in a near equilibrium steady state [163][164]. In these systems, researchers pay a lot attention to the energy efficiency. Although it is reasonable as “evolutionary pressures under carbon-limited growth conditions likely select organisms that utilize highly efficient pathways”. Mathematically, the energy-generation is positively related to the relative biomass synthesis rate (to the substrate uptake rates). Therefore, maximizing the biomass could approximately indicate the energy efficiency through this implicit relation.

Secondly, the new method results a predictive model for simulating the metabolic flux under any condition and provide an intuitive delineation of the flux shifting trend under varying simulation conditions. Unlike MFA, the MPA can not only analyze one state of cell metabolites, but also simulate the flux changing during bioprocesses and predict the behavior of cell in the bioreactor. Another strength of metabolic pathway analysis is to present the flux as a combination of several pathways rather than many individual reactions. This makes it easier to comprehend the changes between different growth conditions. Moreover, these simulations provide useful tool for finding the optimal growth to achieve maximum yield rate of a certain product.

Thirdly, through the flux-shift plotting, the essential elementary pathways can be detected. As illustrated in Fig. 2.14A, the yeast metabolic system consists of three main essential pathways (EM_1 , EM_5 and EM_{10}). Flux gradually shifts from EM_1 via EM_{10} to EM_5 while the ratio of Oxygen/Glucose uptake rates steady increase. It is important to notice that the real number of significant/meaningful elementary pathways is far less than the computed elementary modes. This is because the mathematically existing pathways may be too inefficient to contribute to the microorganism growth, i.e. they are not biological meaningful. It should be pointed out, some authors have proposed several metabolic pathway analysis method based on minimizing/maximizing the number of elementary modes [165] or minimizing the square sum of flux coefficients [43][166]. These methods may provide some information for understanding the metabolic structure from a mathematical point of view, cautions should be taken when interpreting the individual role of each pathway.

The proposed method of maximizing biomass formation rate also suffers from some drawbacks. The main prominent limit is that the established predictive model is based on linear equation and is very sensitive to the accuracy of coefficients in the stoichiometric matrix. Therefore it requires extra matrix adjustment step as discussed in section 2.2.2.2. Another limit is the difficulty to handle complicated models. For a large metabolic network, the total number of elementary modes can reach several millions. Although theoretically the above calculation can still be carried out, the plotting and analysis step will be very difficult, if not impossible. One possible solution to such situation is to set some criteria to quickly screening the mathematically generated elementary modes and eliminating biologically ineligible or inefficient pathways. It is worth mentioning the pathways found in simplified models could suggest clues for narrowing the screening condition. A from-coarse-to-fine approach could be an attractive way of modeling a genome-scale metabolic network.

In conclusion, the proposed approach of metabolic pathway analysis is a simple and efficient method for analyzing cell metabolism physiological state under a given condition and simulating cell behavior in a dynamic process. It not only determines the character of different metabolic pathways but also is helpful for studying the structure, function of elementary pathways and their dynamic relationship in complex cellular networks. The alteration of the flux coefficients expresses the changing of the metabolic pathways under varying growth conditions, which reveals the character of cell

metabolism and provides meaningful information for further metabolic regulation. Using the mutual information of the *in silico* simulation and metabolite pool measurement, a more comprehensive understanding of the proteom status of yeast can be hopefully achieved.

3 Visualization of Metabolic Pathways

3.1 Theoretical Background

Metabolic flux analysis and pathway analysis provide convenient tools for dealing with metabolic networks from a quantitative perspective. However their matrix or vector formats are very hard to comprehend for an ordinary biologist and make it extremely difficult to compare different elementary modes by showing only two columns of numbers. Graphical representations of different metabolic pathways therefore become an ideal and important tool for further analysis. As human brain is very efficient for pattern recognition, converting matrices to graphs could facilitate depicting the topology of the network and identifying common and different parts of different pathways, especially for increasingly large and complex biological networks.

Although static manually-drawn pathway graphs have long been used as an essential tool for bioprocess analysis [167] and could work very well on small networks, producing such graph is a very time-consuming task especially for large networks, and cannot be repeated very frequently. Therefore, it is not suitable for dynamic high-throughput pathway analysis encountered in this study. The following sections primarily focus on dynamic metabolic network visualization techniques, which can automatically generate a graph layout of all the pathways with reasonable aesthetic criteria from numerical inputs. Some tools also allow the graph to be updated according to the corresponding data changes, which can be used to demonstrate the flux changes under different growth conditions.

3.1.1 Graphic Representation of Reactions

In a manually created pathway graph, a reaction is conventionally represented by putting metabolites on both ends and connecting them with lines. This pattern is very similar to a mathematical definition of graph, which consists of edges and nodes, and nodes are connected with edges. However, simply defining metabolites as nodes and reactions as edges could cause ambiguity when one reaction involves more than a pair of metabolites and lead to complicated patterns, as shown in the example of Fig. 3.1A. To avoid this complexity and handle bifurcation in reactions, the representation of a

reaction with a bipartite graph is usually used in metabolic pathway visualization software, where reactions can also be considered as nodes during planning the layout when a reaction involves more than one pair of metabolites. An edge is added between two nodes and one reaction can be represented by one or more edges drawn together. The reaction nodes themselves are usually not shown in order to give an equivalent visual effect close to the man-made patterns. An example is given in Fig. 3.1B. Through this modification the reaction matrix ($m \times n$) can then be transformed to a graph representation $G = (V, E)$, where V is a set of vertices (m metabolites + some reaction nodes) and E is a set of edges (pairs of metabolites and reactions). Taking Fig. 3.1B as an example, in this graph G the set V includes seven nodes (v_1 to v_7), where v_1 to v_5 are metabolite nodes and v_6, v_7 are reaction nodes, and in the set E there are six edges (e_1 to e_6).

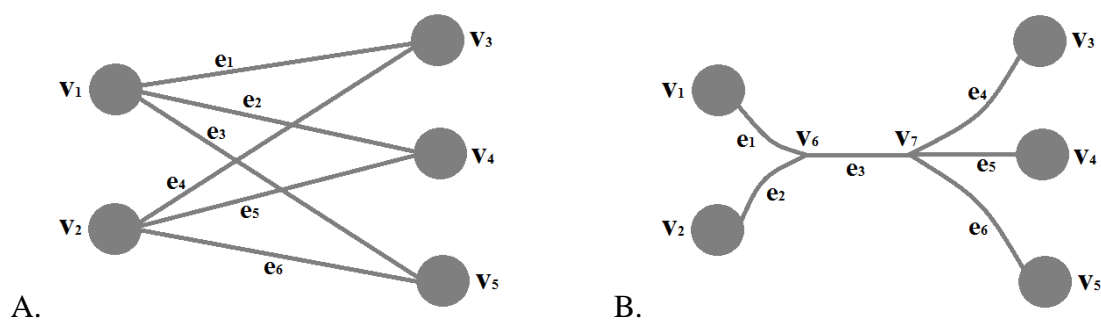


Fig. 3.1 Reaction drawing for those involve more than a pair of metabolites. A. Reactions are not considered as a node. B. Reaction nodes are hidden as a point to give a clear visual. The vertices are represented as v_i and edges e_i .

3.1.2 Generating Layout

Given the description of a graph $G = (V, E)$, the next step is to compute geometric positions for the graph elements for the final drawing step. To facilitate the human interpretation, the layout has to satisfy a set of aesthetic criteria and constraints. The most common rules include no overlap of nodes, minimal edge crossing, rational distance (closer when connect to the same reaction, further apart when not related) and maximal symmetry. Besides these aesthetic criteria, it is also attracting, although difficult, for the metabolic pathway drawing software to follow some not well-defined, conventions as can be found in some relevant biochemistry textbooks [20][50][168] (for example Fig. 3.2), because biologists are so used to the textbook representations. These requirements impose a big challenge for software developers. According to available

literature, so far there is still no automatic software that meets all these explicit and implicit rules, and network visualization remains a bottleneck of metabolic pathway analysis. In the following, some prominent efforts towards solving the problem are briefly summarized. Considering the great number of publications in this field, this summary is far from complete, but is supposed to cover the most common and basic methods in literature.

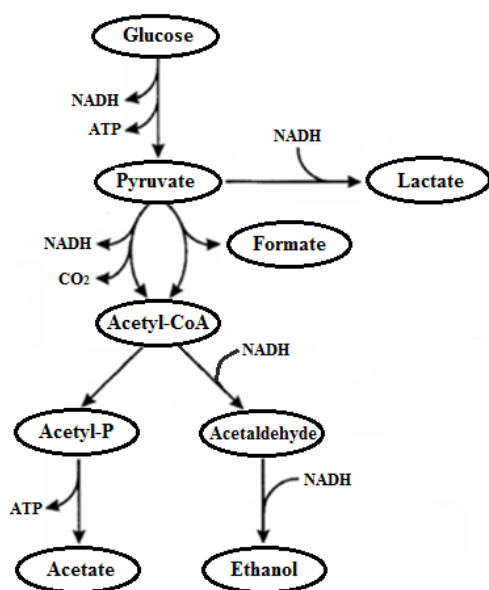


Fig. 3.2 Typical metabolic network graph in textbooks.

Hierarchy layout

A hierarchy layout draws the metabolic network as a tree shape or cascade (as shown in Fig. 3.3). This layout, if drawn properly, can emphasize the chain connection from reactants to products, therefore is the most commonly used visualization model for metabolic pathway analysis. Another reason for its popularity is because the drawing rules are straightforward and many mature drawing algorithms have been established for non-biological applications, since hierarchy structure is one the most seen structures of many data sets. Three key steps are normally involved in producing a hierarchy layout. First step is to convert graph G to a hierarchy structure form, which can be done using breadth first search (BFS) or depth first search (DFS). BFS and DFS are two graph search algorithms with different graph traversal techniques. BFS starts at a given node and travels all the nodes with priority of neighboring nodes (as the order of nodes shown in Fig. 3.3), while DFS travels the nodes with priority of branch nodes. [169–172]

The second step for producing a hierarchy layout is to minimize the crossing. This is usually solved by using the iterative permutation method such as published in [173]. The last step is locally adjustment of nodes positions. One of the most common methods is force-directed layout, which mimics a physical system by considering nodes as masses (or particles) and edges as springs (or magnetic forces). The final position is decided when the repulsion between nodes and attraction on the edges achieve an equilibrium.

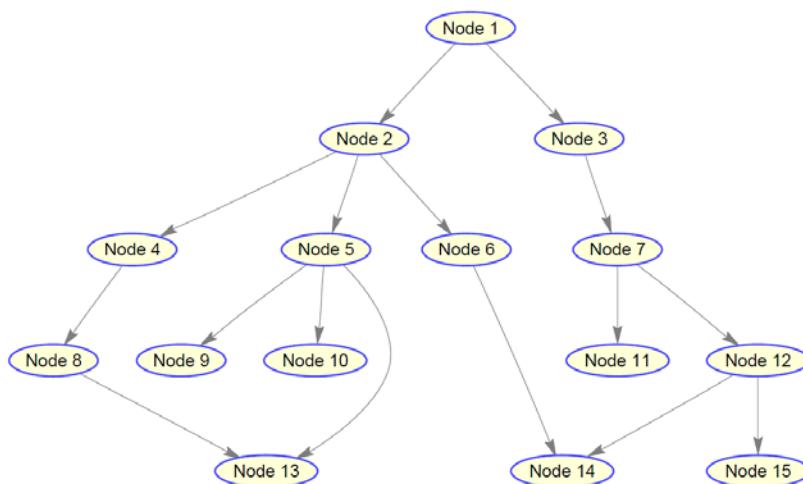


Fig. 3.3 Drawing network with hierarchy layout. The nodes were traveled with breadth first search.

Circular layout

Beside hierarchic components, in many manually drawn metabolic network figures, circular structures are also often seen, which represent a group of reactions that normally happen in a recurrent way. However the circle is defined by using some profound biochemical knowledge that cannot be explicitly expressed by rules. This raises a lot of difficulties for automatic software to identify such structures. The first attempt to detect these cyclic paths was reported in [174], where a depth first search is carried out on those strongly connected components (“subgraphs in which every two nodes are reachable from each other”). This results in the largest circle to be detected first, which usually is not the desired result, especially when side substrates, like NADH, are considered in the network. In [175], the circle detection method was replaced by the breadth first search method, which in most cases generates the smallest circles. Another contribution of this paper is to introduce a joint circle model in which two circle paths share one or more reactions on their border. These strategies lead to better visual results.

In addition to establishing the circular structure, the direct-related metabolites (e.g. substrates or products of the circle metabolites) outside of a circle need also to be positioned around this circle so that the edges crossing the circle are minimized. Therefore the circular layout is usually combined with the hierarchy layout to generate the whole metabolic network.

Grid layout

In hierarchy and circular layouts, the nodes can be placed at any distance to each other. Even when the force-directed layout algorithm is used to optimize the node positions, the inter-node distance can vary very much from a few pixels to hundreds and thousands pixels, giving the network an odd uneven looking. To overcome these drawbacks, a grid layout was proposed by Li and Kurata [176]. In a grid layout, the nodes can only be placed on a square grid and interact according to a specified cost function that is designed based on the topological structure of the network. As in the force-directed system, closely related nodes will be attracted towards each other and remotely related nodes are repulsed from each other. The size of cells in a grid limits the minimum distance between nodes and intrinsically results in an “organized” looking. Additional benefit is, as the possible positions are limited, the computation of force equilibrium is simplified. This allows more complicated restrictions like edge crossing etc. to be incorporated into optimization function [177–179].

3.1.3 Drawing and User’s Interaction

After the layout is generated, the final step of metabolic network visualization is to draw the graph on to the screen. Although it is relative easier compared to the previous step, cautions should be taken to avoid confusion. The most important consideration is to control how many details to show to the user. In large metabolic networks, showing all side reactants could easily cause distraction and confusion. On the other hand, when zooming into several closely related reactions, it is essential to provide detailed information about side reactants and even involved enzymes. Another useful feature is to highlight or distinguish different pathways in a network. For pathway analysis it will improve user perception if different color or opacity is applied on different sets of reactions.

Despite many metabolic network visualization methods exist, there is no software that can provide satisfactory results without user interventions. Therefore, to allow user's interaction during the graph layout is essential to produce easy-understandable graphic representation of metabolic pathways or networks. An advanced graph visualization software should not only allow users to reposition nodes and edges, but also be able to spontaneously adjust the surrounding nodes accordingly to reduce the needs of user input.

In conclusion, metabolic network visualization is an important tool for metabolic pathway analysis and yet a bottleneck that requires more theoretical and practical innovation. In the next section several software packages tested during this work are briefly summarized and a number of improvements will be introduced afterwards.

3.1.4 Available Software

Using the algorithms presented above and their variants, a great number of computer software was designed for metabolic network visualization. Due to different analysis focus and visualization purpose, the output from different software can be very different. Some of them were designed to generate intuitive pathway patterns of small graphs, and others were focusing on showing the complicated metabolite mesh at a large scale, in some case, even using 3D models to demonstrate the relationship between metabolites. In this section several available software suitable for this study are briefly reviewed. For a more complete review of different visualization software, the reader is referred to [180–182].

Graphviz -- Graph Visualization Software

Graphviz is open source graph visualization software. It supports a range of graph shapes and layout setting including hierarchical layout, spring model layout, circular layout, etc. The Graphviz layout program reads graph descriptions in a simple text language, and saves the graphs into a variety of output formats (PostScript, PDF, SVG, annotated text, and so on), or displays it in another interactive graph browser. A programmable (in a language inspired by EZ [3]) widget was designed to display graphs and allow the user to interactively modify them with the mouse. Some additional useful features for concrete diagrams include options for colors, fonts, node layouts, line styles, hyperlinks and custom shapes.

It should be pointed out that a tool for visualizing biochemical network “Biograph” in the Bioinformatics Toolbox of Matlab wrapped up some essential functions of Graphviz and simplified the graph description input and graph output. This makes the procedure of generating graphs much easier than in the native Graphviz environment. The drawback of Biograph is that the user is lack of control of the position of nodes and edges. Once the graph is generated, the user can only change some appearance properties, such as color and size, but not the position.

SimWiz

SimWiz is one of the pioneer software tools that attempted to represent the reaction circles on metabolic pathways. Similar to the method proposed by Becker et al. [15] SimWiz divided the pathway to cyclic and hierarchical structures. The cyclic structures are then drawn in a circle layout and the hierarchical structures are drawn in a hierarchical layout with the help of a force-directed algorithm. For small networks, SimWiz can generate nice graphics representations that are close to manually created flow graphs, however its ability to handle medium or large network is limited by its performance.

MetaViz

Metavis is a network visualization software tool focusing to present the topology of metabolic networks. Unlike other topology analysis software, it simultaneously depicts the metabolic pathway information in an intuitive way using a similar approach as SimWiz. In this software, Bourqui et al. [183] proposed a clustering step to divide the metabolic network to a largest set of independent pathways and sub-pathways based on the rules of minimizing the shared metabolites and reactions between pathways. The resulting graph clustering is drawn in a planar graph to demonstrate the topology of the network, while for each pathway or sub-pathway, metabolites and reactions are drawn using specific drawing algorithms (hierarchical and circular ones). An additional attracting feature of this software is users can provide a set of predefined pathways that have been recognized using human knowledge. These pathways will then be draw as clusters entirely.

LucidDraw

LucidDraw is a Matlab-based visual analysis tool using a grid layout algorithm. Its features include great speed and user interactivity. It usually takes LucidDraw only a

few seconds to draw a typical biological network with several hundreds of nodes and the resulting drawings are interactively modifiable. Users can also control layout styles and incorporation of other available information. However, some preliminary experiments suggest that the generated graph is not intuitive for pathway analysis. Many edge crossing may occur. Moreover the nodes sometime may overlap with edges, and cause confusion of the starting and ending points of edges. It is possible to manually adjust the result, but it is too time consuming to solve all the problems.

BioLayout 3D

BioLayout Express3D is a new tool for the visualization and analysis of networks using a 3D graph representation. Unlike other visualization tools, BioLayout 3D arranges nodes in a 3D space rather than a 2D plane. This is an interesting attempt, however, at this stage it is still difficult to find its advantage over other 2D projected methods. In the future, with addition of intuitive layout rules, this approach may be able to provide additional information that is hard to perceive in 2D graphs.

3.2 Materials and Methods

To facilitate the comparison and understanding of the flux changing under various growth conditions, in this study several different visualization methods and tools were tested.

An in-house software (PathView) based on hierarchy-layout method and breadth-first searching was first developed. Later, a more advanced visualization software — GraphView was developed to replace PathView. GraphView uses a modified version of “Biograph” program of the bioinformatics toolbox in Matlab to generate the layout.

3.2.1 Design of PathView

In this visualization software, reactions are represented as metabolite pairs. An edge is added between two metabolites when they are involved on different sides in the same reaction. Reversible reactions are drawn with bi-direction arrows. This hierarchy-layout method consists of three steps:

1. Generating a tree structure from a given substrate (usually glucose) using the breadth-first searching.

2. Reorder the node positions on horizontal level so that the metabolite that is involved in the most reactions is set in the middle and metabolites that have fewer reactions are set further outside. Meanwhile reducing crossing is another goal of this reordering operation.

3. Drawing edges using the routing method [184], which is usually used in electrical wiring, is used in the software to further reduce the edge crossing while drawing.

3.2.2 Design of GraphView

In GraphView, a reaction is represented by a bipartite graph (section 3.1.1) i.e. metabolites and reactions are all represented as nodes. Metabolites are connected to a reaction if they are all involved in that reaction. While drawing, the reaction nodes are shown as small dots and the metabolites are shown as rectangular or circular boxes.

“Biograph” as one function of the bioinformatics toolbox in Matlab, was designed for general purpose bio-network visualization. While using it for generating hierarchical layout in this application, the nodes distribution generally meets the basic aesthetic rules. However, to further emphasize on pathway illustration and comparison, several extensions are introduced.

1. Add more control of nodes size to allow user to hide the reaction nodes.
2. Add more control of edge and arrow properties, and allow the program to control how to show arrows (single, double, none) on each individual edge.
3. Add more control of weight labels on each edge, and allow to show/hide weight labels on any edge.

Other strategies to facilitate the comparison of pathways include: hiding side metabolites namely ATP, CO₂, NAD(P)H, assigning different colors to different pathways reaction groups, hiding certain paths, etc.

3.2.3 Testing Environment

All the algorithms presented in this thesis were implemented in Matlab (Version 7.6.0.324 (R2008a), The MathWorks Inc., Natick, Massachusetts, U.S.A.). The additional toolboxes and packages used in this thesis includes: optimization toolbox and bioinformatics toolbox. The *in silico* analyses and experiments were performed on a

personal computer (AMD Athlon™ 64 × dual core processor 3800+ 2.00 GHz, 1.96 GB RAM, with Windows XP 32 bit installed).

3.3 Results and Discussion

In this section, the results of metabolic pathway analysis in chapter 2 are presented using the in-house visualization software PathView and GraphView. Their advantages and limits are also further discussed.

3.3.1 Metabolic Pathways Visualized with PathView Software

Fig. 3.4 shows an example pathway graph generated from the PathView software. Starting from glucose, it searches all the reactions that take glucose as substrate and put all the products of any those reactions on the second row. The software goes on to search the direct products from all the metabolites on the second row, and in turn lists them on the third row. This iterative searching stops when all the metabolites are found. The order of metabolites on a row can be changed to minimize the crossing of reaction lines.

The advantage of the above layout method is, it can clearly demonstrate the “distance” between key metabolites, for example how many reactions are needed to convert glucose to ethanol. However, it should be pointed out that in these graphs some small molecular metabolites namely NADH, ATP, CO₂ have to be removed, otherwise most metabolites will be drawn on the level beneath those molecules. Another feature included into this software is a routing method to reduce the edge crossing. Fig. 3.4A and B compare the results of visualization with and without routing method. Apparently, with help of the routing method, the reaction 16 in Fig. 3.4A goes around other nodes and edges and avoids crossing.

A big problem with this software is that it represents reaction as pairs of metabolites, so-called paired-metabolite-method, which means one reaction that involves more than one substrate or product will introduce more than one edge on the graph. For instance, in Fig. 3.5, reaction 11 ($\text{pyr} \rightarrow (0.6667) \text{acald} + (0.3333) \text{co2}$) has been shown as two edges in the graph: $\text{pyr} \rightarrow \text{acald}$ going down and $\text{pyr} \rightarrow \text{co2}$ going up. This will cause complicated patterns, especially when the model becomes a little more complicated such as from Model 3 to Model 2. It also makes it difficult to recognize reactions.

Another problem with the graph generated from breadth first search is, the final product biomass may end up in upper layers in some pathways, as shown in Fig. 3.5, in which biomass is shown in the 4th layer of all seven layers.

Overall, PathView is relative simple and immature compared with some existing software and the GraphView software, which will be discussed below.

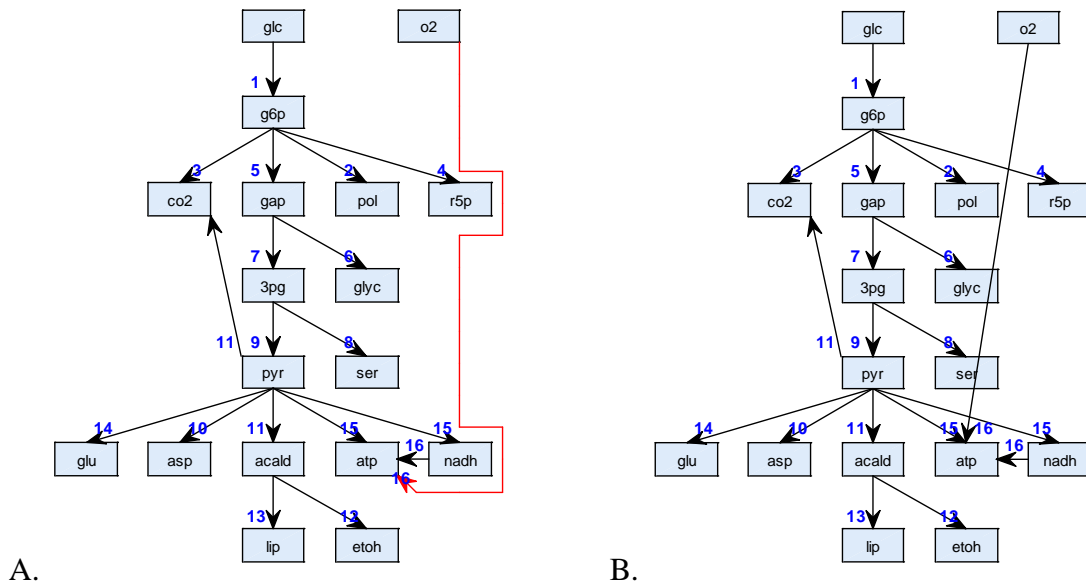


Fig. 3.4 Visualization of Model 3 using PathView software developed in this project. In Model 3 the biomass synthesis reaction is omitted. A. using routing method; B. without routing method.

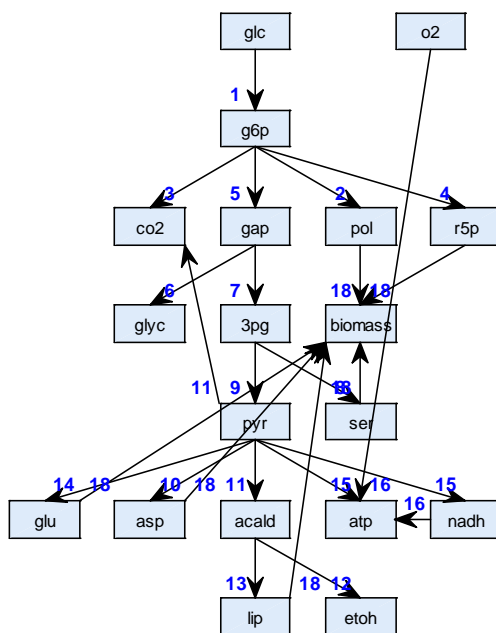


Fig. 3.5 Visualization of Model 3 using PathView software. Biomass appears on the 4th row and cause many crossing.

3.3.2 Metabolic Pathways Visualized with GraphView Software

Fig. 2.15 on page 70 shows an example pathway graph generated from the GraphView software, which is based on the “biograph” function in the Matlab bioinformatics toolbox. Unlike the PathView software, in GraphView bipartite presentation is used on reactions that involve more than two metabolites. Those reactions are recognized as nodes and drawn as small dots between substrates and products. This representation leads to a layout that is closer to the manually created metabolic pathway graph (e.g. Fig. 2.5 on page 34). However bipartite representation may cause unnecessary zigzag shape in the case where a reaction only has one substrate and one product, because the layout algorithm will often put such three nodes (two metabolites nodes and one reaction node) in a triangle shape instead of a single line. To overcome such defect, a hybrid reaction representation strategy is introduced in this study, which also uses the same paired-metabolite-method as in PathView to represent those simple reactions (one to one).

3.3.2.1 Modification of the “biograph” Function

The “biograph” function, which is based on the GraphViz package, provides three types of layout: hierarchical, radial and equilibrium. Examples of different layout patterns are shown in Fig. 3.6. It is obvious that the hierarchical layout is the best choice for metabolic pathway visualization. However, using the original “biograph” algorithm to draw metabolic pathways has several shortcomings that will violate the aesthetic criteria:

Firstly, it only provides global control of some import drawing settings, such as show/hide arrows and weight label of edges. (In this software, this label is used to show the reaction No. and/or reaction rates.)

Secondly, it does not allow drawing double arrows on an edge. To show a reversible reaction, two edges in opposite directions have to be drawn, which can easily cause confusion and increase the number of edge-crossings.

Thirdly, adding/removing edges and nodes will change the layout dramatically. As the purpose of the visualization software is to facilitate the comparison of different pathways. It is essential to keep the position of relevant metabolites in the same place while some reactions are added or removed.

Therefore several modifications are made to the original “biograph” files. These changes are summarized in Tab. 3.1 and discussed in more detail below.

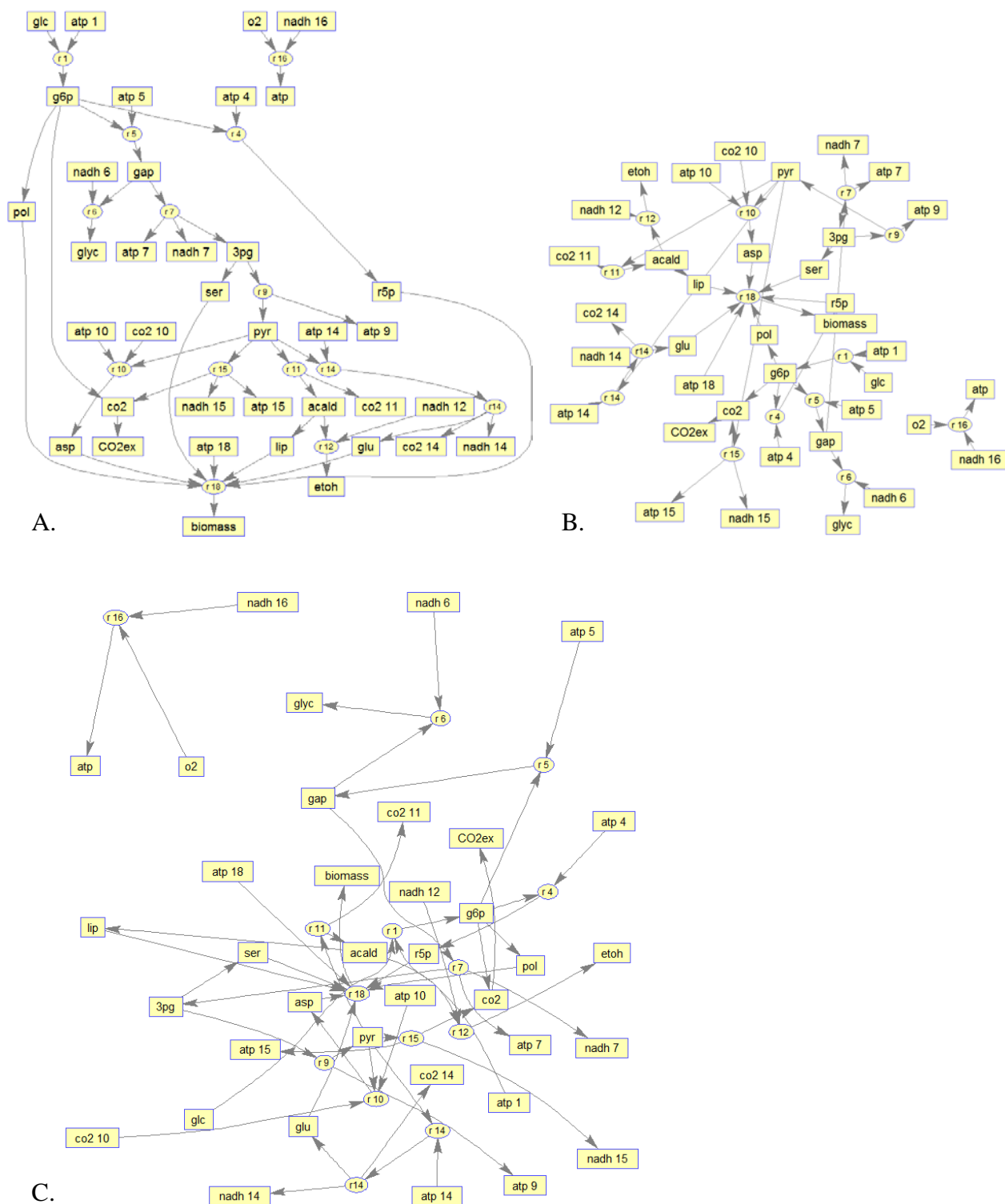
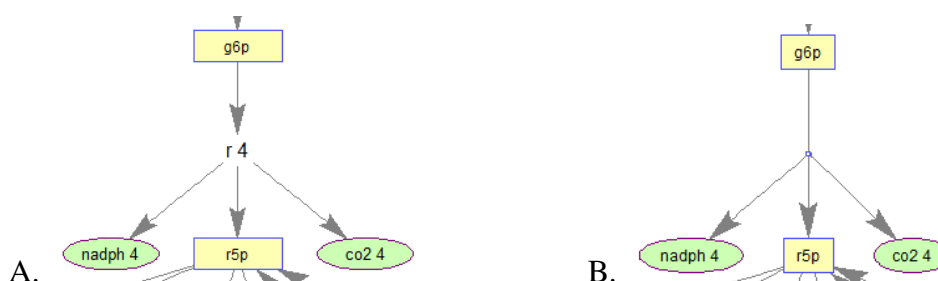


Fig. 3.6 Model 3 is visualized with three types of layout style provided by “biograph” tool. A. hierarchical layout; B. radial layout; C. equilibrium layout.

Tab. 3.1 Summary of modification to “biograph” and their functionalities.

| New functions | Modified files | Original biograph |
|--|---|--|
| Change style of single edge (line, arrow etc.) Show/hide weight label of single edge | bioinfo/@biograph/@edge/hgUpdate.m | Show/hide arrows and weight label of edges are public properties and must be changed together. |
| Show edge as double arrow, back arrow, or hide edge. | bioinfo/@biograph/@edge/hgUpdate.m bioinfo/@biograph/@edge/schema.m | There are only two styles of edge: line and front arrow. A reversible reaction will be showed by two edges in opposite directions. |
| Show selected nodes as points and the edges connected to the nodes can joint to the point. | bioinfo/@biograph/@biograph/private/createdotfile.m bioinfo/@biograph/@node/hgUpdate.m | The minimum size of node is limited and node cannot reduced to a point. |

(1) Drawing Reaction Nodes

**Fig. 3.7 Comparison of drawing reaction node.** A. the original “biograph” cannot draw reaction node as a point. B. the modified version can draw the nodes as a point and makes the graph neater.

As discussed before, in some reactions, the reaction itself is treated as a node. However the original “biograph” tool allows showing the node’s name in the middle of a node, but its minimum size cannot smaller than the character size. Even though the node can be hidden by assigning it the background color, it is not possible to fill the gap between the neighboring edges (see Fig. 3.7A). After the modification, the reaction nodes are drawn as small dots (see Fig. 3.7B) and make the graph more easily to understand.

(2) More flexible edge and arrow drawing control

In the modified version of “biograph”, more flexible arrow drawing function is introduced. First, double arrow function on an edge is added, this solves the problem with reversible reactions which can now be expressed by a single edge between two nodes (c.f. Fig. 3.8). Secondly, a new function is added to control whether to show the arrow or not on each individual edges (c.f. Fig. 3.9). This make the bipartite representation of reactions look neater. Thirdly, a function to hide edges is added, which

is later used for comparing different pathway in the same metabolic network (see more detail in section 3.3.2.3).

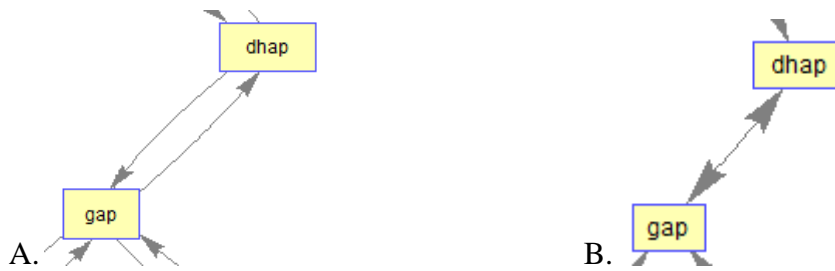


Fig. 3.8 Comparison of results for edge modified version. A. the original “biograph” does not support bidirectional arrows. When drawing a reversible reaction, double edge is used instead. B. After introducing the double arrow edges, the graph looks more succinct.

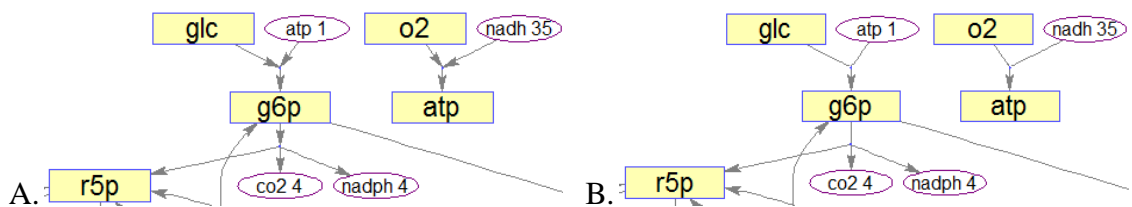


Fig. 3.9 Comparison of results for edge modified version. A. the original “biograph” only allows showing/hiding arrow being set globally. B. after hiding the arrows on some of the edge from substrates to the reaction nodes, the graph looks neater.

(3) More flexible control over weight labels

In this software the weight labels are used to show the reaction rates or reaction number. As the reactions are treated as nodes, one reaction is now divided into several segments of edges. In the original “biograph” tool, the weight label can only be set to show/hide globally, and weight labels are printed multiple times (c.f. Fig. 3.10). After adding more flexible control of the appearance of weight labels, these duplications are eliminated.

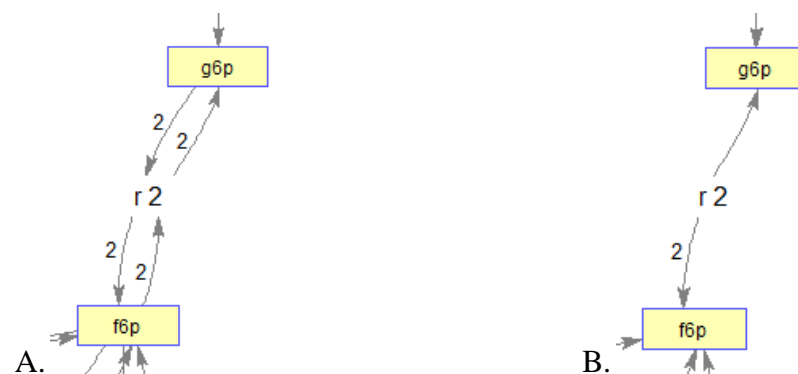


Fig. 3.10 Comparison of showing weight label A. an example shows the weight label is printed on every edges. B. the correct display of weight label after modification and for one reaction only one label is shown.

3.3.2.2 Show Metabolic Pathway with Different Level Details

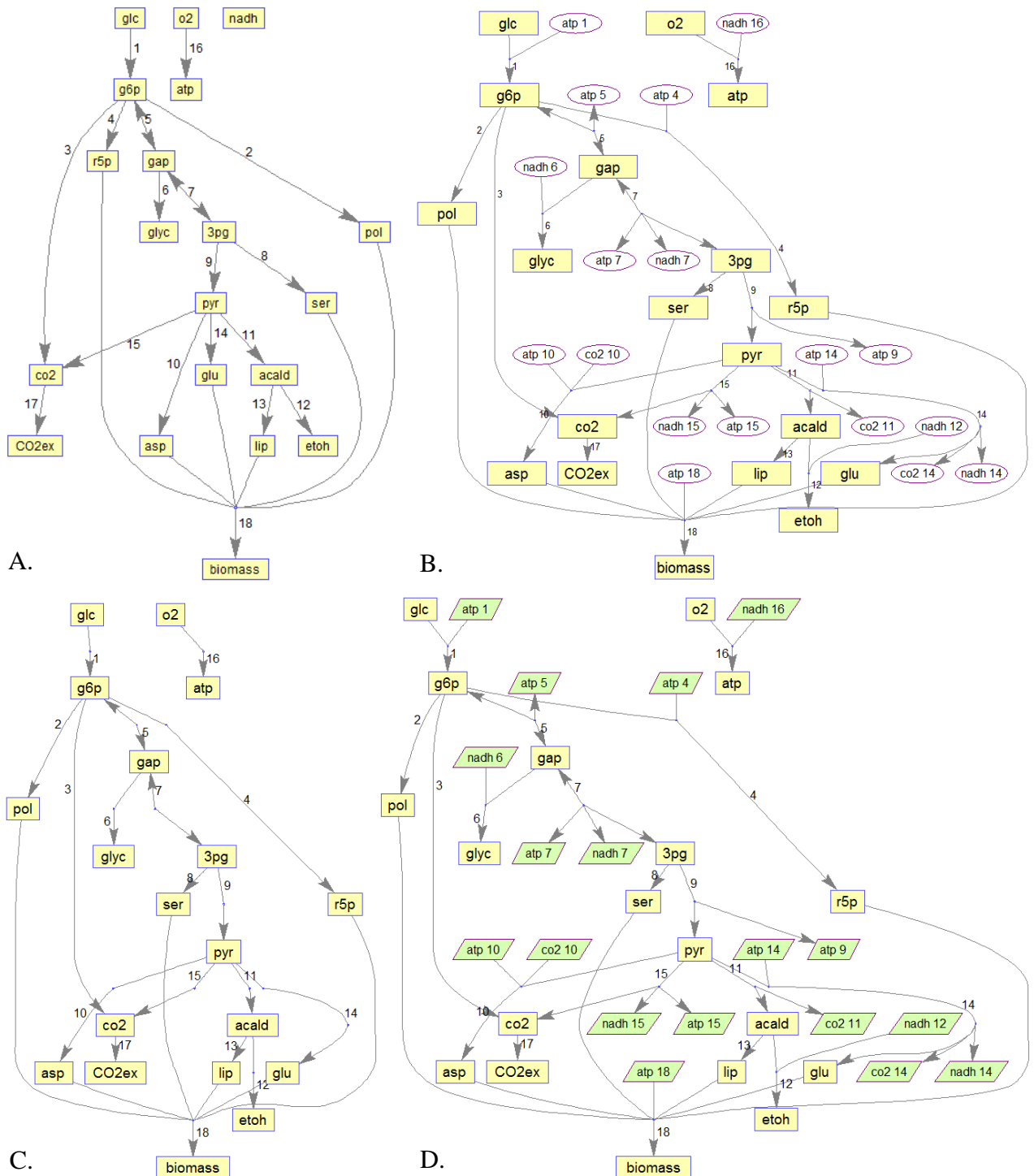


Fig. 3.11 Comparison of small molecular showing for Model 3. A. only one node for each small molecule; B. use multiple node copies for the small molecules; C. small molecules are shown in different color and shape; D. all small molecules are hidden.

In the GraphView software users are allowed to control how much detail they want to see in the graph. Besides, drawing/hiding labels, nodes and edges, a list of small molecular metabolites can be defined, for example ATP, CO₂, NADH and NADPH.

Such substances are usually involved in many reactions, if they are treated equally to other metabolites, i.e. one node for each metabolite, the graph will be distorted (c.f. Fig. 3.11A). In the GraphView software, instead these side substrate nodes can have multiple copies across the graph (Fig. 3.11B), therefore results more meaningful graph-drawing. This side substances list can be manually altered. Because the graph generating only takes a few seconds the user can interactively simplify the graph to a certain level. The side substrates are shown in different color and shape in the resulting graph (c.f. Fig. 3.11B and D), but the user can also choose to hide them all (c.f. Fig. 3.11C and D) for further simplification.

3.3.2.3 Comparing Different Pathways

One important goal of this software is to visualize the difference of several pathways in a same metabolic network.

However, the graphic representation of different pathway generated from the “biograph” function may vary dramatically, even if only one or two nodes are added/removed. The strategy to overcome this problem is to generate the layout for the whole metabolic network first (see Fig. 3.12) and then hide some reaction and metabolites that are not included in a certain pathway (c.f. Fig. 3.13). In addition, when comparing two pathways, the software can draw the different reactions in different colors and draw the overlapped part in a third color (see Fig. 3.14).

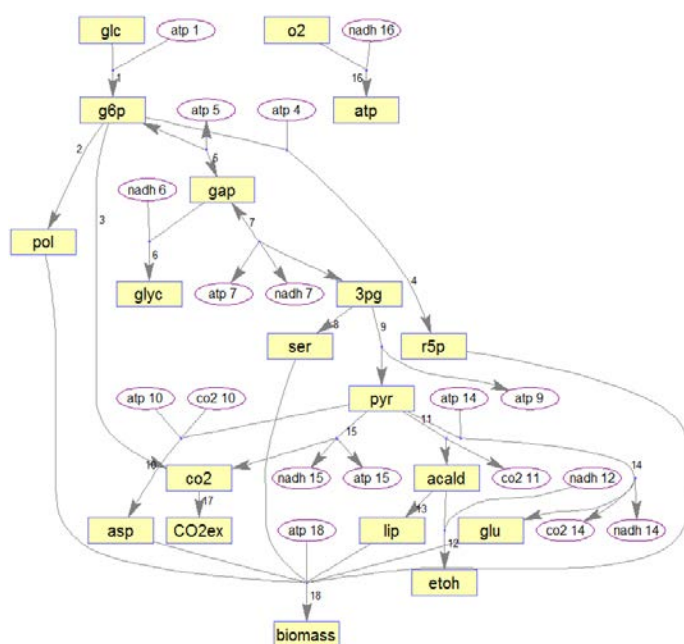


Fig. 3.12 The graph of the metabolic network of Model 3.

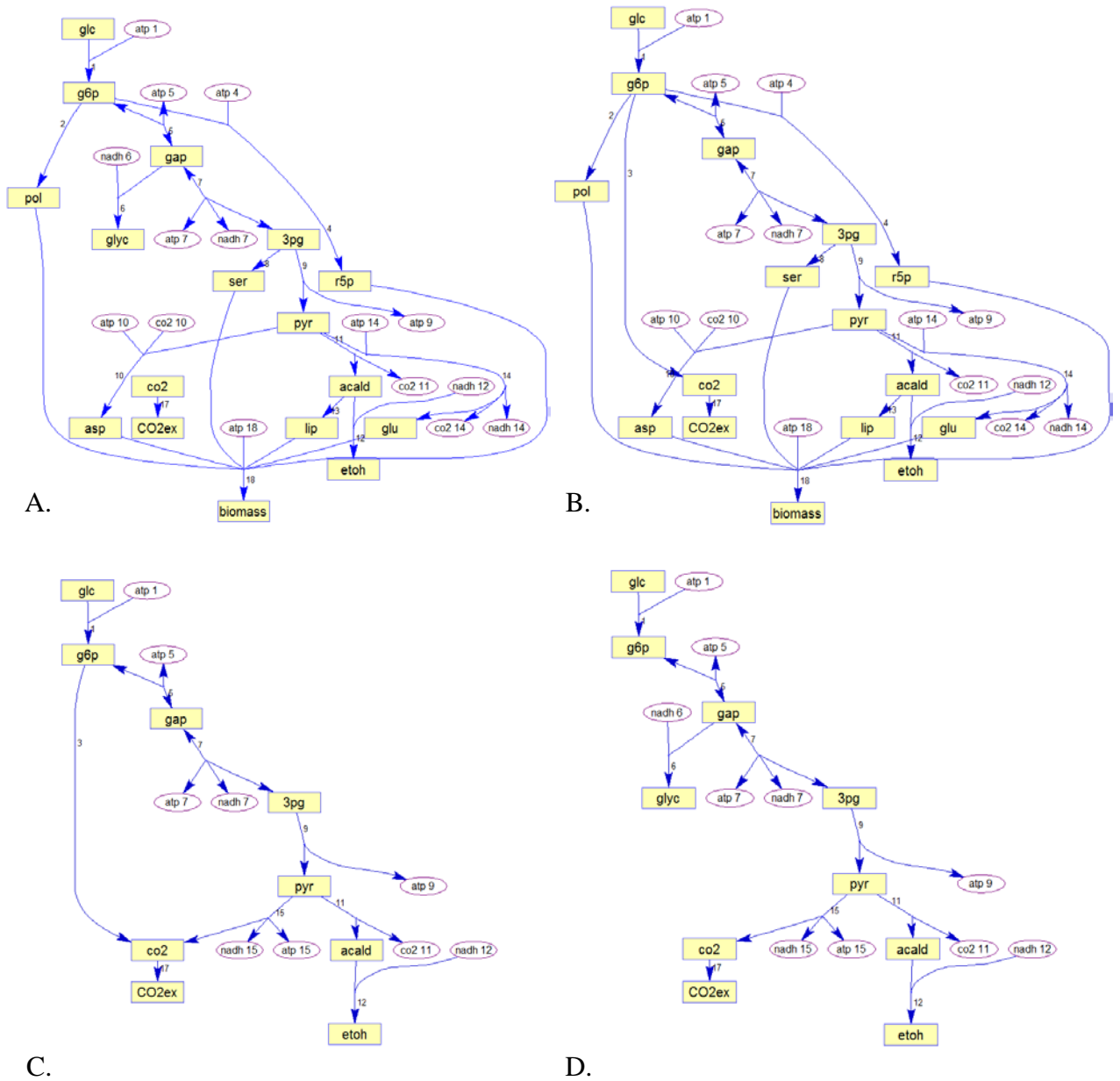
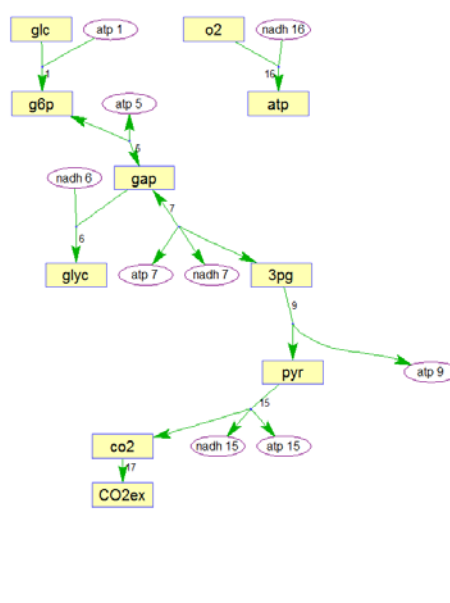
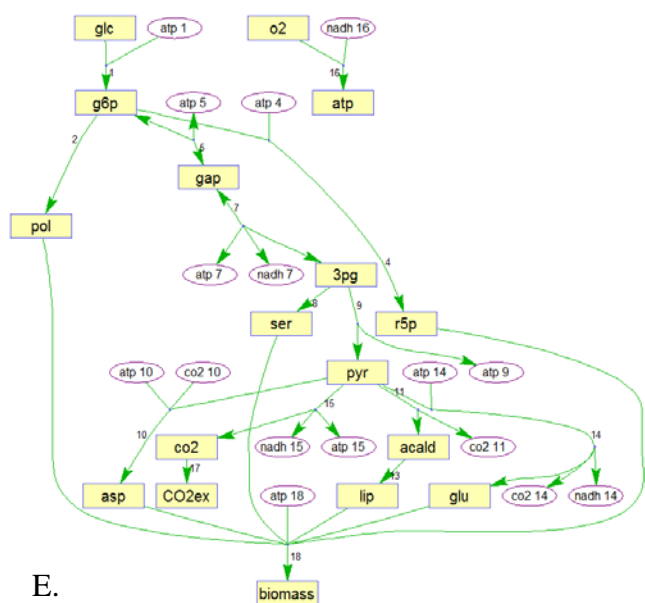
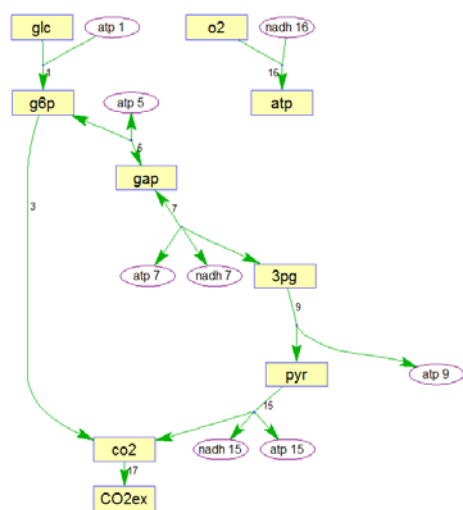


Fig. 3.13 Graphic representation of different pathway of Model 3 (elementary mode matrix is listed in Tab. 6.20 in section 6.5). A ~ D correspond to $EM_1 \sim EM_4$ which are anaerobic pathways; E ~ G correspond to $EM_5 \sim EM_7$ which are oxidative pathways (see next page); H corresponds to EM_8 which is aerobic fermentative pathway (see next page).

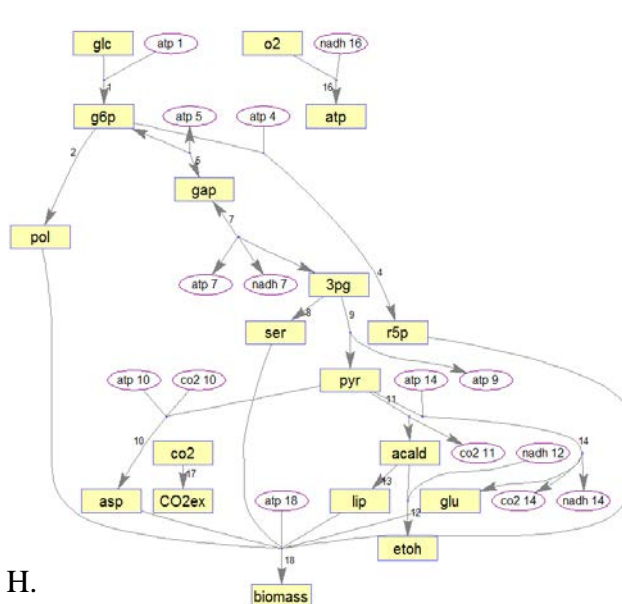


E.

F.



G.



H.

Fig. 3.13 Graphic representation of different pathway of Model 3 (elementary mode matrix is listed in Tab. 6.20 in section 6.5) (continues last page). A ~ D correspond to EM₁ ~ EM₄ which are anaerobic pathways (see last page); E ~ G correspond to EM₅ ~ EM₇ which are oxidative pathways; H corresponds to EM₈ which is aerobic fermentative pathway.

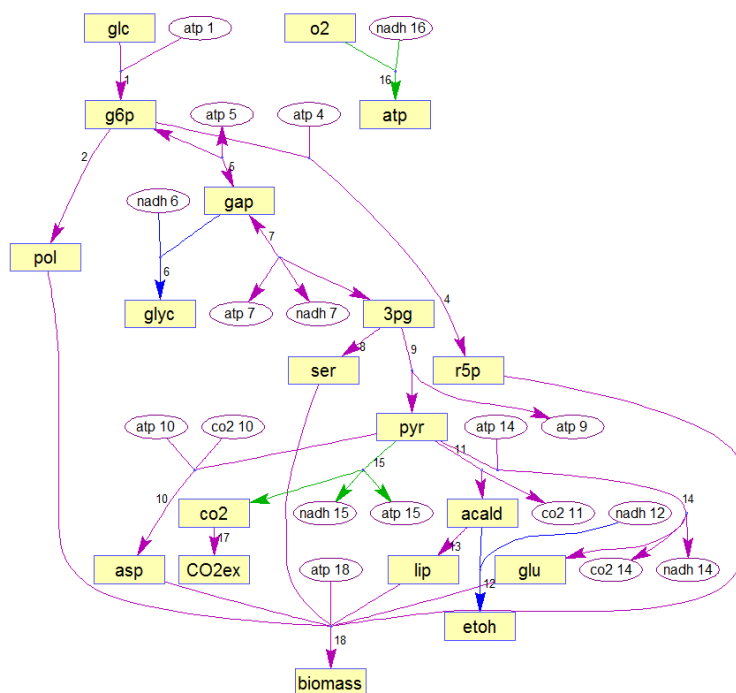


Fig. 3.14 Comparison of two pathways in Model 3 using GraphView. Overlapped parts of two pathways are shown in purple. Blue reactions belong to the anaerobic growth pathway in Fig. 3.13A, while green reactions belong to the oxidative growth pathway in Fig. 3.13E.

3.3.2.4 Comparison with Other Software

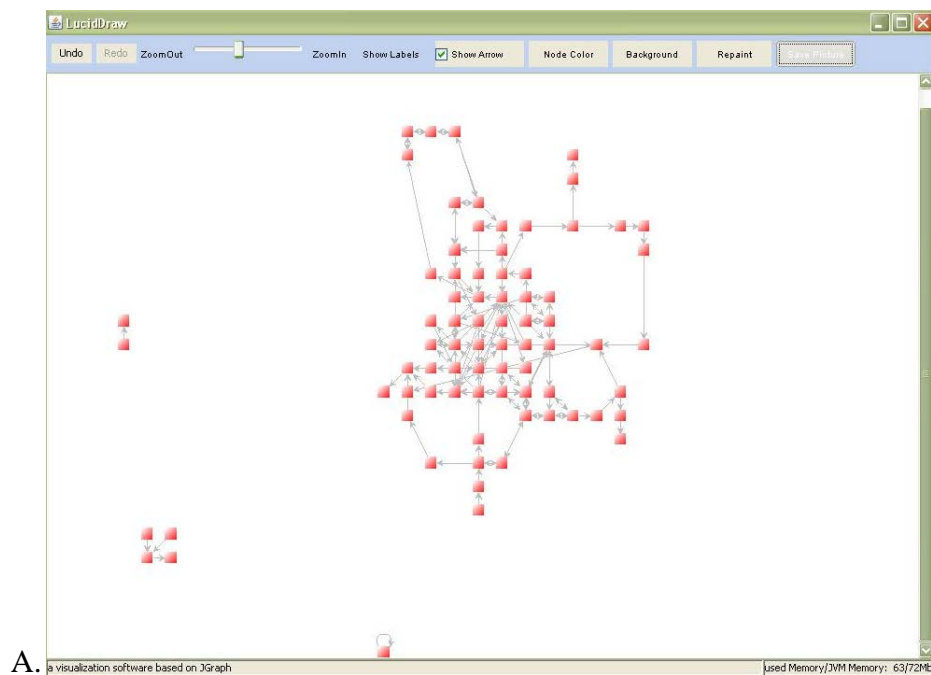
Compared with other metabolic pathway/network visualization software, GraphView generates a hierarchy layout that is similar to many manually created metabolic graphs in textbook. This allows most biologists quickly understand the graphics representations, while most other software requires special training to perceive the meaning. Fig. 3.15 compares the visualization result from GraphView and some other available software using Model 2 as an example.

Another advantage of GraphView is its ability to visualize the difference among pathways, which makes it especially suitable for studying elementary modes of small or medium size metabolic network. This function is usually unsupported in most available software.

Some limits of Graph view need also to be addressed. First of all, the ability of GraphView to handle large metabolic networks is relative limited because the complexity of a network increases sharply as the node number increase. Although it can be partly solved by dividing a big network into small sub-groups, this procedure requires many user interventions using the current software. Another limit is lacking of

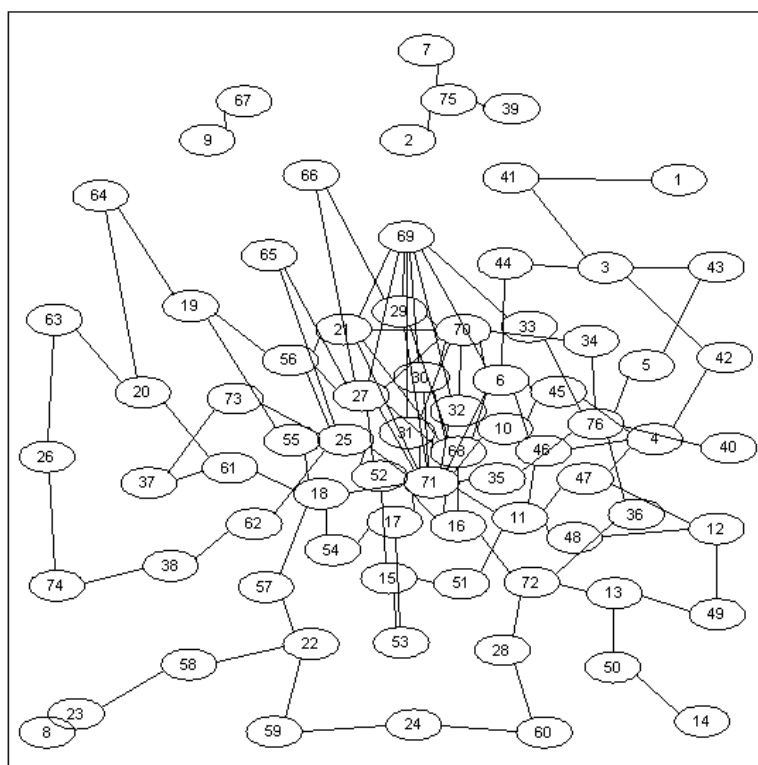
suitable representation for circular reaction chains. Becker et al. [174] have proposed several circular structure searching and drawing methods. These techniques should later be incorporated into the GraphView software to produce even more intuitive graphs.

Nevertheless, to this stage, GraphView is so far the best available software to visualize elementary modes.



A. visualization software based on JGraph

Used Memory/JVM Memory: 63/72Mb



B.

Fig. 3.15 Visualization results of Model 2 from several software. A. visualization with LucidDraw; B. visualization with Graphviz.

4 Conclusions

In this work, a framework for bioprocess modeling is established which contains a series of systematic steps to convert the common used stoichiometric models/MFA models to predictive bioprocess models. The problems of metabolic pathway visualization are also addressed. A new MPA-base flux simulation algorithm and metabolic network visualization software are proposed and tested using simplified metabolic models for *S. cerevisiae*.

Firstly, two metabolic models for *S. cerevisiae* are established using simplified reaction network from literature and problematic coefficients of stoichiometric matrix are detected and calibrated with metabolic flux analysis using experimentally measured reaction rates.

Secondly, the resulting models were then used to generate homeostatic metabolic pathway models by generating the elementary flux modes. These models are further extended to predictive metabolic pathway models by including a new MPA-based metabolic flux vector estimation method that determines an optimal combination of elementary modes by maximizing the biomass production. In the following experiments, these predictive models were validated by comparing the experimentally measured reaction rates with the simulated metabolic flux of yeast under various culture conditions. The new MPA achieved satisfactory accuracy in these preliminary tests.

The validated predictive metabolic pathway models were also used to simulate the changes of metabolic flux distribution on different pathways under varying growth condition. Three out of thirteen mathematically selected elementary modes were found to be biological “meaningful” under various growth conditions.

Compared with other existing metabolic analysis methods, the proposed approach of metabolic pathway analysis has the advantages of computational simplicity, intuitive illustration and ability of simulating cell behavior in a dynamic process. It can not only determine the character of different metabolic pathways but also is helpful for studying the structure and function of elementary modes and their dynamic relationship in complex cellular networks. The alteration of the flux coefficients expresses the changing activity of the metabolic pathways under varying growth conditions, which reveals and predicts the character of cell metabolism. These advantages make the

proposed method an attractive approach for simplifying complicated metabolic models and simulating sophisticated bioprocess. The predictive homeostatic metabolic model created with the proposed method could also provide a good foundation for further building dynamic metabolic model and bioprocess models.

Finally, in-house software, GraphView was developed to visualize the elementary modes and compare their differences. GraphView generates a hierarchy layout that is similar to many manually created metabolic graphs in textbook. This allows most biologists quickly understand the graphics representations, while most other software requires special training to perceive the meaning. These graph representations made it easier to analysis the character of the metabolic network as it generates a hierarchy layout that is similar to many manually created metabolic graph in textbook. The drawing program was also configured to facilitate the visualization of the difference among pathways, which makes it especially suitable for studying elementary modes of metabolic network and study the flux changes among different pathways in a dynamic process.

Combining all the above steps results a complete chain of software for MPA-based metabolic analysis and simulation of different organisms. Using both the quantitative information of the *in silico* simulation and the intuitive interpretation of pathway visualization, a more comprehensive understanding of the metabolome status of yeast and its dynamic changes can hopefully be achieved. The simulation framework established in this study could also provide beneficial information for future experiments or industrial productions to avoid futile attempt.

Future work includes adding information of kinetics and dynamic regulation, to make the predictive metabolic model more complete, as in the current models the enzyme suppression and induction are ignored. Furthermore a reactor model needs also to be added to construct a complete mathematic model for bioprocesses. Another task to consider in the future is to generalize both the MPA framework and pathway visualization software to handle large metabolic networks. One possible solution is to develop a from-coarse-to-fine approach where the information gathered in a simplified model could be used to screen elementary modes for biologically eligible pathways in a more complicated model. The ultimate goal is to automatically convert a genome-scale metabolic network to a manageable and perceivable bioprocess model.

5 References

- [1] N. Kleckner, J. Roth, and D. Botstein, "Genetic engineering in Vivo using translocatable drug-resistance elements: New methods in bacterial genetics," *Journal of Molecular Biology*, vol. 116, no. 1, pp. 125-159, Oct. 1977.
- [2] D. C. Cameron and I. T. Tong, "Cellular and metabolic engineering. An overview," *Applied Biochemistry and Biotechnology*, vol. 38, no. 1-2, pp. 105-140, Feb. 1993.
- [3] G. Schrimpf, *Gentechnische Methoden*, 3rd ed. Spektrum Akademischer Verlag, 2002.
- [4] R. Simon, U. Priefer, and A. Puhler, "A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria," *Nat Biotech*, vol. 1, no. 9, pp. 784-791, Nov. 1983.
- [5] B. O. Palsson, "Bioinformatics: What lies beyond bioinformatics?," *Nat Biotech*, vol. 15, no. 1, pp. 3-4, Jan. 1997.
- [6] R. C. Strohmman, "The coming Kuhnian revolution in biology," *Nat Biotech*, vol. 15, no. 3, pp. 194-200, Mar. 1997.
- [7] M. Terzer, N. D. Maynard, M. W. Covert, and J. Stelling, "Genome-scale metabolic networks," *Systems Biology and Medicine*, vol. 1, no. 3, pp. 285-297, Jun. 2009.
- [8] P. D. Karp et al., "Multidimensional annotation of the *Escherichia coli* K-12 genome," *Nucleic Acids Research*, vol. 35, no. 22, pp. 7577 -7590, Dec. 2007.
- [9] C. J. Krieger, "MetaCyc: a multiorganism database of metabolic pathways and enzymes," *Nucleic Acids Research*, vol. 32, no. 90001, p. 438D-442, Jan. 2004.
- [10] M. Kanehisa et al., "KEGG for linking genomes to life and the environment," *Nucleic Acids Research*, vol. 36, no. Database issue, pp. D480-484, Jan. 2008.
- [11] N. C. Duarte et al., "Global reconstruction of the human metabolic network based on genomic and bibliomic data," *Proceedings of the National Academy of Sciences*, vol. 104, no. 6, pp. 1777 -1782, Feb. 2007.
- [12] J. C. Liao, S. Y. Hou, and Y. P. Chao, "Pathway analysis, engineering, and physiological considerations for redirecting central metabolism," *Biotechnology and Bioengineering*, vol. 52, no. 1, pp. 129-140, Oct. 1996.
- [13] N. D. Price, J. A. Papin, and B. O. Palsson, "Determination of Redundancy and Systems Properties of the metabolic network of *Helicobacter pylori* using genome-scale extreme pathway analysis," *Genome Research*, 2002.
- [14] K. Schallau and B. H. Junker, "Simulating Plant Metabolic Pathways with Enzyme-Kinetic Models," *Plant Physiology*, vol. 152, no. 4, pp. 1763 -1771, Apr. 2010.
- [15] F. Llaneras and J. Picó, "Which metabolic pathways generate and characterize the flux space? A comparison among elementary modes, extreme pathways and minimal generators," *Journal of Biomedicine & Biotechnology*, vol. 2010, p. 753904, 2010.
- [16] K. Ugurbil, T. R. Brown, J. A. den Hollander, P. Glynn, and R. G. Shulman, "High-resolution ¹³C nuclear magnetic resonance studies of glucose metabolism in *Escherichia coli*," *Proceedings of the National Academy of Sciences*, vol. 75, no. 8, pp. 3742 -3746, 1978.
- [17] T. Szyperski, "¹³C-NMR, MS and Metabolic Flux Balancing in Biotechnology Research," *Quarterly Reviews of Biophysics*, vol. 31, no. 1, pp. 41-106, 1998.
- [18] P. Jouhten, "Metabolic modelling and ¹³C fluxanalysis," Helsinki University of Technology, Espoo, Finland, 2009.
- [19] W. Wiechert, "¹³C Metabolic Flux Analysis," *Metabolic Engineering*, vol. 3, no. 3, pp. 195-206, Jul. 2001.
- [20] G. Stephanopoulos, J. Nielsen, and G. Stephanopoulos, *Metabolic Engineering: Principles and Methodologies*. Academic Pr Inc, 1998.

- [21] J. J. Villino and G. Stephanopoulos, "Flux determination in cellular bioreaction networks: applications to lysine fermentations," 1990.
- [22] K. J. Kauffman, P. Prakash, and J. S. Edwards, "Advances in flux balance analysis," *Current Opinion in Biotechnology*, vol. 14, no. 5, pp. 491-496, Oct. 2003.
- [23] N. D. Price, J. L. Reed, and B. O. Palsson, "Genome-scale models of microbial cells: evaluating the consequences of constraints," *Nature Reviews Microbiology*, vol. 2, no. 11, pp. 886-897, 2004.
- [24] J. S. Edwards, R. Ramakrishna, C. H. Schilling, and B. O. Palsson, "Metabolic flux balance analysis," In: *Lee SY, Papoutsakis ET (eds) Metabolic engineering. Marcel Dekker, New York,* pp. 13-57, 1999.
- [25] A. Varma and B. O. Palsson, "Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use," *Nat Biotech*, vol. 12, no. 10, pp. 994-998, Oct. 1994.
- [26] S. Klamt and J. Stelling, "Two approaches for metabolic pathway analysis?," *TRENDS in Biotechnology*, vol. 21, no. 2, pp. 64-69, 2003.
- [27] S. Schuster and C. Hilgetag, "On elementary flux modes in biochemical reaction systems at steady state," *Journal of Biological Systems*, vol. 2, no. 2, pp. 165-182, 1994.
- [28] C. H. Schilling, D. Letscher, and B. Ø. Palsson, "Theory for the Systemic Definition of Metabolic Pathways and their use in Interpreting Metabolic Function from a Pathway-Oriented Perspective," *Journal of Theoretical Biology*, vol. 203, no. 3, pp. 229-248, 07 2000.
- [29] P. C. Milner, "The Possible Mechanisms of Complex Reactions Involving Consecutive Steps," *Journal of The Electrochemical Society*, vol. 111, no. 2, pp. 228-232, Feb. 1964.
- [30] M. Terzer and J. Stelling, "Parallel Extreme Ray and Pathway Computation."
- [31] D. Jevremovic, C. T. Trinh, F. Sreenc, and D. Boley, "On Algebraic Properties of Extreme Pathways in Metabolic Networks," *Journal of Computational Biology*, vol. 17, no. 2, pp. 107-119, Feb. 2010.
- [32] Y. Xi, Y.-P. P. Chen, M. Cao, W. Wang, and F. Wang, "Analysis on relationship between extreme pathways and correlated reaction sets," *BMC Bioinformatics*, vol. 10, no. 1, p. S58-S58, 2009.
- [33] C. T. Trinh, A. Wlaschin, and F. Sreenc, "Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism," *Applied Microbiology and Biotechnology*, vol. 81, no. 5, pp. 813-826, 2008.
- [34] J.-M. Schwartz and M. Kanehisa, "Quantitative elementary mode analysis of metabolic pathways: the example of yeast glycolysis," *BMC Bioinformatics*, vol. 7, no. 1, p. 186, 2006.
- [35] S. Schuster, T. Dandekar, and D. A. Fell, "Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering," *Trends in Biotechnology*, vol. 17, no. 2, pp. 53-60, 1999.
- [36] C. Reder, "Metabolic control theory: A structural approach," *Journal of Theoretical Biology*, vol. 135, no. 2, pp. 175-201, Nov. 1988.
- [37] S. Schuster, T. Pfeiffer, F. Moldenhauer, I. Koch, and T. Dandekar, "Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplasma pneumoniae*," *Bioinformatics*, vol. 18, no. 2, pp. 351 -361, Feb. 2002.
- [38] B. L. Clarke, "Stoichiometric network analysis," *Cell Biochemistry and Biophysics*, vol. 12, no. 1, pp. 237-253, 1988.
- [39] R. Carlson and F. Sreenc, "Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: Identification of reactions," *Biotechnology and Bioengineering*, vol. 85, no. 1, pp. 1-19, Jan. 2004.
- [40] R. Carlson and F. Sreenc, "Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: creation of overall flux states," *Biotechnology and Bioengineering*, vol. 86, no. 2, pp. 149-162, Apr. 2004.
- [41] S. Klamt and E. D. Gilles, "Minimal cut sets in biochemical reaction networks," *Bioinformatics*, vol. 20, no. 2, p. 226, 2004.

- [42] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles, "Metabolic network structure determines key aspects of functionality and regulation," *Nature*, vol. 420, no. 6912, pp. 190–193, 2002.
- [43] M. G. Poolman, K. V. Venkatesh, M. K. Pidcock, and D. A. Fell, "A method for the determination of flux in elementary modes, and its application to *Lactobacillus rhamnosus*," *Biotechnology and Bioengineering*, vol. 88, no. 5, pp. 601-612, Dec. 2004.
- [44] N. D. Price, J. L. Reed, J. A. Papin, I. Famili, and B. O. Palsson, "Analysis of metabolic capabilities using singular value decomposition of extreme pathway matrices," *Biophysical Journal*, vol. 84, no. 2 Pt 1, pp. 794-804, Feb. 2003.
- [45] S. Schuster, D. A. Fell, and T. Dandekar, "A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks," *Nature biotechnology*, vol. 18, no. 3, pp. 326–332, 2000.
- [46] S. J. Wiback, R. Mahadevan, and B. O. Palsson, "Using metabolic flux data to further constrain the metabolic solution space and predict internal flux patterns: the *Escherichia coli* spectrum," *Biotechnology and Bioengineering*, vol. 86, no. 3, pp. 317-331, May 2004.
- [47] C. T. Trinh, R. Carlson, A. Wlaschin, and F. Sreenc, "Design, construction and performance of the most efficient biomass producing *E. coli* bacterium," *Metabolic Engineering*, vol. 8, no. 6, pp. 628-638, Nov. 2006.
- [48] C. T. Trinh, P. Unrean, and F. Sreenc, "Minimal *Escherichia coli* Cell for the Most Efficient Production of Ethanol from Hexoses and Pentoses," *Applied and Environmental Microbiology*, vol. 74, no. 12, pp. 3634-3643, Apr. 2008.
- [49] A. P. Wlaschin, C. T. Trinh, R. Carlson, and F. Sreenc, "The fractional contributions of elementary modes to the metabolism of *Escherichia coli* and their estimation from reaction entropies," *Metabolic Engineering*, vol. 8, no. 4, pp. 338-352, Jul. 2006.
- [50] D. S. Christina, *The Metabolic Pathway Engineering Handbook: Fundamentals*, 1st ed. Crc Pr Inc, 2009.
- [51] N. R. Pace, "The universal nature of biochemistry," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 3, pp. 805-808, Jan. 2001.
- [52] G. Stephanopoulos, "Metabolic Fluxes and Metabolic Engineering," *Metabolic Engineering*, vol. 1, no. 1, pp. 1-11, Jan. 1999.
- [53] K. Schügerl and K.-H. Bellgardt, *Bioreaction Engineering: Modeling and Control*, 1st ed. Springer, 2000.
- [54] F. R. Schmidt, "Recombinant expression systems in the pharmaceutical industry," *Applied Microbiology and Biotechnology*, vol. 65, pp. 363-372, Jul. 2004.
- [55] C. Russell, J. Mawson, and P. L. Yu, "Production of recombinant products in yeasts: a review," *Australian Journal of Biotechnology*, vol. 5, no. 1, pp. 48-55, Jan. 1991.
- [56] R. G. Buckholz and M. A. G. Gleeson, "Yeast Systems for the Commercial Production of Heterologous Proteins," *Nat Biotech*, vol. 9, no. 11, pp. 1067-1072, Nov. 1991.
- [57] G. E. Grampp, A. Sambanis, and G. N. Stephanopoulos, "Use of regulated secretion in protein production from animal cells: an overview," *Advances in Biochemical Engineering/Biotechnology*, vol. 46, pp. 35-62, 1992.
- [58] D. G. Rudmann and S. K. Durham, "Utilization of Genetically Altered Animals in the Pharmaceutical Industry," *Toxicologic Pathology*, vol. 27, no. 1, pp. 111 -114, Jan. 1999.
- [59] D. Petranovic and G. N. Vemuri, "Impact of yeast systems biology on industrial biotechnology," *Journal of Biotechnology*, vol. 144, no. 3, pp. 204-211, Nov. 2009.
- [60] E. Nevoigt, "Progress in Metabolic Engineering of *Saccharomyces cerevisiae*," *Microbiol. Mol. Biol. Rev.*, vol. 72, no. 3, pp. 379-412, Sep. 2008.
- [61] O. Käppeli, "Regulation of carbon metabolism in *Saccharomyces cerevisiae* and related yeasts," *Advances in Microbial Physiology*, vol. 28, pp. 181-209, 1986.

- [62] S. Ostergaard, L. Olsson, and J. Nielsen, "Metabolic Engineering of *Saccharomyces cerevisiae*," *Microbiol. Mol. Biol. Rev.*, vol. 64, no. 1, pp. 34-50, Mar. 2000.
- [63] V. Robert, *Yeasts in food: beneficial and detrimental aspects*. Behr's Verlag, 2003.
- [64] A. Goffeau et al., "Life with 6000 Genes," *Science*, vol. 274, no. 5287, pp. 546-567, Oct. 1996.
- [65] P. Uetz et al., "A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*," *Nature*, vol. 403, no. 6770, pp. 623-627, Feb. 2000.
- [66] N. J. Krogan et al., "Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*," *Nature*, vol. 440, no. 7084, pp. 637-643, Mar. 2006.
- [67] "Saccharomyces Genome Database." [Online]. Available: <http://www.yeastgenome.org/>. [Accessed: 26-Jul-2011].
- [68] "GenRE - CYGD - The Comprehensive Yeast Genome Database." [Online]. Available: <http://mips.helmholtz-muenchen.de/genre/proj/yeast/>. [Accessed: 26-Jul-2011].
- [69] P. Duboc, I. Marison, and U. Von Stockar, "Physiology of *Saccharomyces cerevisiae* during cell cycle oscillations," *Journal of Biotechnology*, vol. 51, no. 1, pp. 57-72, Oct. 1996.
- [70] J. R. Dickinson and M. Schweizer, *Metabolism and Molecular Physiology of Saccharomyces Cerevisiae, 2nd Edition*, 2nd ed. CRC Press, 2004.
- [71] K.-H. Bellgardt, "Modellbildung des Wachstums von *Saccharomyces cerevisiae* in Rührkesselreaktoren," Ph.D. Thesis, Universität Hannover, 1984.
- [72] W. Visser, W. A. Scheffers, W. H. Batenburg-van der Vegte, and J. P. van Dijken, "Oxygen requirements of yeasts," *Applied and Environmental Microbiology*, vol. 56, no. 12, pp. 3785-3792, Dec. 1990.
- [73] K. Furukawa, E. Heinzle, and I. J. Dunn, "Influence of oxygen on the growth of *Saccharomyces cerevisiae* in continuous culture," *Biotechnology and Bioengineering*, vol. 25, no. 10, pp. 2293-2317, Oct. 1983.
- [74] B. M. Bakker et al., "Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*," *FEMS Microbiology Reviews*, vol. 25, no. 1, pp. 15-37, Jan. 2001.
- [75] M. Rigoulet et al., "Organization and regulation of the cytosolic NADH metabolism in the yeast *Saccharomyces cerevisiae*," *Molecular and Cellular Biochemistry*, vol. 256, no. 1/2, pp. 73-81, Jan. 2004.
- [76] P. Jouhten et al., "Oxygen dependence of metabolic fluxes and energy generation of *Saccharomyces cerevisiae* CEN.PK113-1A," *BMC Systems Biology*, vol. 2, pp. 60-60.
- [77] R. H. DE DEKEN, "The Crabtree Effect: A Regulatory System in Yeast," *Journal of General Microbiology*, vol. 44, no. 2, pp. 149-156, 1966.
- [78] J. H. de Winde and L. A. Grivell, "Global regulation of mitochondrial biogenesis in *Saccharomyces cerevisiae*: ABF1 and CPF1 play opposite roles in regulating expression of the QCR8 gene, which encodes subunit VIII of the mitochondrial ubiquinol-cytochrome c oxidoreductase.," *Mol. Cell. Biol.*, vol. 12, no. 6, pp. 2872-2883, Jun. 1992.
- [79] K. D. Entian, "Glucose repression: a complex regulatory system in yeast," *Microbiological Sciences*, vol. 3, no. 12, pp. 366-371, Dec. 1986.
- [80] J. M. Gancedo, "Yeast carbon catabolite repression," *Microbiology and Molecular Biology Reviews: MMBR*, vol. 62, no. 2, pp. 334-361, Jun. 1998.
- [81] G. N. Vemuri, M. A. Eiteman, J. E. McEwen, L. Olsson, and J. Nielsen, "Increasing NADH oxidation reduces overflow metabolism in *Saccharomyces cerevisiae*," *Proceedings of the National Academy of Sciences*, vol. 104, no. 7, pp. 2402-2407, 2007.
- [82] J. Nielsen, J. Villadsen, and G. Lidén, *Bioreaction Engineering Principles*, 2nd ed. Springer, 2002.
- [83] K. C.L., "Production of bioproducts through the use of transgenic animal models," *Animal Reproduction Science*, vol. 82-83, no. 0, pp. 5-12, Jul. 2004.

- [84] Technology Information, Forecasting, and Assessment Council (India); India.; Indo-US Science & Technology Forum. and S. Biswas, *Bioprocess and bioproducts: technology trends and opportunities*. New Delhi: Asiatech Publishers, 2009.
- [85] K. Schügerl, *Bioreaction Engineering: Characteristic features of bioreactors*. Wiley, 1991.
- [86] K. Schügerl, *Bioreaktionstechnik: Bioprozesse mit Mikroorganismen und Zellen*. Birkhäuser, 1997.
- [87] D. J. Ingham, *Chemical engineering dynamics: an introduction to modelling and computer simulation*. Wiley-VCH, 2000.
- [88] I. J. Dunn, *Biological reaction engineering: dynamic modelling fundamentals with simulation examples*. Wiley-VCH, 2003.
- [89] K. Muttzall, *Modellierung von Bioprozessen*. Behr, 1994.
- [90] R. D. Schmid, *Pocket Guide to Biotechnology and Genetic Engineering*, 1st ed. Wiley-VCH Verlag GmbH & Co. KGaA, 2003.
- [91] L. O. Ingram, T. Conway, D. P. Clark, G. W. Sewell, and J. F. Preston, "Genetic engineering of ethanol production in *Escherichia coli*," *Appl. Environ. Microbiol.*, vol. 53, no. 10, pp. 2420-2425, Oct. 1987.
- [92] H. Daniell, M. S. Khan, and L. Allison, "Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology," *Trends in Plant Science*, vol. 7, no. 2, pp. 84-91, Feb. 2002.
- [93] A. Wright and S. L. Morrison, "Effect of glycosylation on antibody function: implications for genetic engineering," *Trends in Biotechnology*, vol. 15, no. 1, pp. 26-32, Jan. 1997.
- [94] A. K. Gombert and J. Nielsen, "Mathematical modelling of metabolism," *Current Opinion in Biotechnology*, vol. 11, no. 2, pp. 180-186, 2000.
- [95] J. E. Bailey, "Mathematical Modeling and Analysis in Biochemical Engineering: Past Accomplishments and Future Opportunities," *Biotechnology Progress*, vol. 14, no. 1, pp. 8-20, Jan. 1998.
- [96] J. Bailey, "Toward a science of metabolic engineering," *Science*, vol. 252, no. 5013, pp. 1668 - 1675, Jun. 1991.
- [97] N. Gershenfeld, *The Nature of Mathematical Modeling*. Cambridge University Press, 1998.
- [98] "Metabolic modelling of microbes: the flux-balance approach - Edwards - 2002 - Environmental Microbiology - Wiley Online Library."
- [99] W. M. van Gulik and J. J. Heijnen, "A metabolic network stoichiometry analysis of microbial growth and product formation," *Biotechnology and Bioengineering*, vol. 48, no. 6, pp. 681-698, Dec. 1995.
- [100] C. H. Schilling, J. S. Edwards, and B. O. Palsson, "Toward Metabolic Phenomics: Analysis of Genomic Data Using Flux Balances," *Biotechnology Progress*, vol. 15, no. 3, pp. 288-295, Jun. 1999.
- [101] M. Laakso, T. Författare, and M. Laakso, "Reconstructing the Metabolic Network of a Bacterium from Its Genome," 2007.
- [102] K. Bellgardt, "Cell Models," *Biotechnology Set, Second Edition*, vol. 9, pp. 267-298, 2008.
- [103] M. Rabkin and J. J. Blum, "Quantitative analysis of intermediary metabolism in hepatocytes incubated in the presence and absence of glucagon with a substrate mixture containing glucose, ribose, fructose, alanine and acetate.," *Biochemical Journal*, vol. 225, no. 3, pp. 761-786, Feb. 1985.
- [104] M. J. Herrgård, S. S. Fong, and B. Ø. Palsson, "Identification of Genome-Scale Metabolic Network Models Using Experimentally Measured Flux Profiles," *PLoS Comput Biol*, vol. 2, no. 7, p. e72, Jul. 2006.

- [105] I. Famili, J. Förster, J. Nielsen, and B. O. Palsson, "Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network," *Proceedings of the National Academy of Sciences*, vol. 100, no. 23, pp. 13134-13139, Nov. 2003.
- [106] A. P. Burgard, E. V. Nikolaev, C. H. Schilling, and C. D. Maranas, "Flux Coupling Analysis of Genome-Scale Metabolic Network Reconstructions," *Genome Research*, vol. 14, no. 2, pp. 301-312, Feb. 2004.
- [107] J. K. Fredrickson et al., "Towards environmental systems biology of Shewanella," *Nat Rev Micro*, vol. 6, no. 8, pp. 592-603, 2008.
- [108] J. Shon, J. Y. Park, and L. Wei, "Beyond similarity-based methods to associate genes for the inference of function," *BIOSILICO*, vol. 1, no. 3, pp. 89-96, Jul. 2003.
- [109] R. E. Jinkerson, V. Subramanian, and M. C. Posewitz, "Improving biofuel production in phototrophic microorganisms with systems biology," *Biofuels*, vol. 2, pp. 125-144, Mar. 2011.
- [110] A. Novick and L. Szilard, "Description of the Chemostat," *Science*, vol. 112, no. 2920, pp. 715-716, Dec. 1950.
- [111] D. Herbert, R. Elsworth, and R. C. Telling, "The Continuous Culture of Bacteria; a Theoretical and Experimental Study," *Journal of General Microbiology*, vol. 14, no. 3, pp. 601-622, Jul. 1956.
- [112] T. W. James, "Continuous Culture of Microorganisms," *Annual Review of Microbiology*, vol. 15, pp. 27-46, Oct. 1961.
- [113] J. Förster, I. Famili, P. Fu, B. Ø. Palsson, and J. Nielsen, "Genome-Scale Reconstruction of the Saccharomyces cerevisiae Metabolic Network," *Genome Research*, vol. 13, no. 2, pp. 244-253, 2003.
- [114] M. F. Luna, C. F. Mignone, and J. L. Boiardi, "The carbon source influences the energetic efficiency of the respiratory chain of N₂-fixing Acetobacter diazotrophicus," *Applied Microbiology and Biotechnology*, vol. 54, pp. 564-569, Oct. 2000.
- [115] R. Ouhabi, M. Rigoulet, and B. Guerin, "Flux-yield dependence of oxidative phosphorylation at constant $\Delta\mu\text{H}^+$," *FEBS Letters*, vol. 254, no. 1-2, pp. 199-202, Aug. 1989.
- [116] C. Larsson, U. von Stockar, I. Marison, and L. Gustafsson, "Growth and metabolism of Saccharomyces cerevisiae in chemostat cultures under carbon-, nitrogen-, or carbon- and nitrogen-limiting conditions," *Journal of Bacteriology*, vol. 175, no. 15, pp. 4809-4816, Aug. 1993.
- [117] C. Verduyn, A. H. Stouthamer, W. A. Scheffers, and J. P. Dijken, "A theoretical evaluation of growth yields of yeasts," *Antonie van Leeuwenhoek*, vol. 59, no. 1, pp. 49-63, Jan. 1991.
- [118] E. Hernandez and M. J. Johnson, "Energy Supply and Cell Yield in Aerobically Grown Microorganisms," *Journal of Bacteriology*, vol. 94, no. 4, pp. 996-1001, Oct. 1967.
- [119] S. J. Pirt, "The Maintenance Energy of Bacteria in Growing Cultures," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 163, no. 991, pp. 224-231, Oct. 1965.
- [120] J. J. Vallino and G. Stephanopoulos, "Metabolic flux distributions in Corynebacterium glutamicum during growth and lysine overproduction," *Biotechnology and Bioengineering*, vol. 41, no. 6, pp. 633-646, Mar. 1993.
- [121] C. Varela, F. Pizarro, and E. Agosin, "Biomass Content Governs Fermentation Rate in Nitrogen-Deficient Wine Musts," *Applied and Environmental Microbiology*, vol. 70, no. 6, pp. 3392-3400, Jun. 2004.
- [122] F. M. Williams, "A model of cell growth dynamics," *Journal of Theoretical Biology*, vol. 15, no. 2, pp. 190-207, May 1967.
- [123] K.-H. Bellgardt, N. Hopf, R. Luttmann, and W.-D. Deckwer, "A new approach for structured growth models," presented at the Preprints of Fourth Int. Conf. on Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes, Cambridge, UK, Chichester (UK), 1988.
- [124] K. Bellgardt and J. Yuan, "Process Models: Optimization of Yeast Production — A Case Study," pp. 383-406.

- [125] K. Lee, F. Berthiaume, G. N. Stephanopoulos, and M. L. Yarmush, "Metabolic Flux Analysis: A Powerful Tool for Monitoring Tissue Function," *Tissue Engineering*, vol. 5, no. 4, pp. 347-368, Aug. 1999.
- [126] F. Madron, V. Veverka, and V. Vaněček, "Statistical analysis of material balance of a chemical reactor," *AIChE Journal*, vol. 23, no. 4, pp. 482-486, Jul. 1977.
- [127] R. Schuetz, L. Kuepfer, and U. Sauer, "Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*," *Molecular Systems Biology*, vol. 3, pp. 119-119.
- [128] J. Nielsen and J. Villadsen, *Bioreaction engineering principles*, 1st ed. New York: Plenum Press, 1994.
- [129] K. von Meyenburg, "Katabolit-Repression und der Sprossungszyklus von *Saccharomyces cerevisiae*," Ph.D. Thesis, ETH Zürich, 1969.
- [130] J. Nielsen, *Physiological engineering aspects of *Penicillium chrysogenum**. World Scientific, 1997.
- [131] R. Cuthbertson and R. Paton, *Computation in cellular and molecular biological systems*. World Scientific, 1996.
- [132] M. G. Poolman, K. V. Venkatesh, M. K. Pidcock, and D. A. Fell, "A method for the determination of flux in elementary modes, and its application to *Lactobacillus rhamnosus*," *Biotechnology and Bioengineering*, vol. 88, no. 5, pp. 601-612, Dec. 2004.
- [133] J. Papin, "Metabolic pathways in the post-genome era," *Trends in Biochemical Sciences*, vol. 28, no. 5, pp. 250-258, 2003.
- [134] C. H. Schilling, S. Schuster, B. O. Palsson, and R. Heinrich, "Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era," *Biotechnology Progress*, vol. 15, no. 3, pp. 296-303, Jun. 1999.
- [135] A. M. Feist, M. J. Herrgård, I. Thiele, J. L. Reed, B. O. Palsson, and P. Box, "Reconstruction of biochemical networks in microorganisms," *Nature Reviews: Microbiology*, vol. 7, no. 2, pp. 129-143, 2009.
- [136] C. Kaleta, F. Centler, and P. Dittrich, "Analyzing Molecular Reaction Networks: From Pathways to Chemical Organizations," *Molecular Biotechnology*, vol. 34, no. 2, pp. 117-124, 2006.
- [137] R. T. Rockafellar, *Convex analysis*. Princeton University Press, 1970.
- [138] K. Fukuda and A. Prodon, "Double description method revisited," in *Combinatorics and Computer Science*, 1996, pp. 91-111.
- [139] C. Wagner, "Nullspace Approach to Determine the Elementary Modes of Chemical Reaction Systems," *The Journal of Physical Chemistry B*, vol. 108, no. 7, pp. 2425-2431, Feb. 2004.
- [140] J. Gagneur and S. Klamt, "Computation of elementary modes: a unifying framework and the new binary approach," *BMC bioinformatics*, vol. 5, no. 1, p. 175, 2004.
- [141] S. Klamt, J. Gagneur, and A. von Kamp, "Algorithmic approaches for computing elementary mo... [Syst Biol (Stevenage). 2005] - PubMed result," 2005.
- [142] M. Terzer and J. Stelling, "Large-scale computation of elementary flux modes with bit pattern trees," *Bioinformatics*, vol. 24, no. 19, p. 2229, 2008.
- [143] "METATOOL: for studying metabolic networks -- Pfeiffer et al. 15 (3): 251 -- Bioinformatics."
- [144] S. Hoops et al., "COPASI - a Complex Pathway Simulator," *Bioinformatics*, vol. 22, no. 24, p. 3067, 2006.
- [145] R. Urbanczik, "SNA - a toolbox for the stoichiometric analysis of metabolic networks," *BMC bioinformatics*, vol. 7, no. 1, p. 129, 2006.
- [146] A. Kamp and S. Schuster, "Metatool 5.0: fast and flexible elementary modes analysis," *Bioinformatics*, vol. 22, no. 15, p. 1930, 2006.
- [147] R. Schwarz et al., "YANA - a software tool for analyzing flux modes, gene-expression and enzyme activities," *BMC Bioinformatics*, vol. 6, no. 1, p. 135, 2005.

- [148] T. Wilhelm, J. Behre, and S. Schuster, "Analysis of structural robustness of metabolic networks," *Systems Biology*, vol. 1, p. 114, 2004.
- [149] R. Tanaka, M. Csete, J. Doyle, and others, "Highly optimised global organisation of metabolic networks," *Systems biology*, vol. 152, no. 4, p. 179, 2005.
- [150] J. L. Galazzo and J. E. Bailey, "Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae*," *Enzyme and Microbial Technology*, vol. 12, no. 3, pp. 162-172, Mar. 1990.
- [151] M. Sonderegger, M. Jeppsson, B. Hahn-Hägerdal, and U. Sauer, "Molecular Basis for Anaerobic Growth of *Saccharomyces cerevisiae* on Xylose, Investigated by Global Gene Expression and Metabolic Flux Analysis," *Applied and Environmental Microbiology*, vol. 70, no. 4, pp. 2307 - 2317, Apr. 2004.
- [152] T. Dandekar, S. Schuster, B. Snel, M. Huynen, and P. Bork, "Pathway alignment: application to the comparative analysis of glycolytic enzymes.," *Biochemical Journal*, vol. 343, no. 1, pp. 115-124, Oct. 1999.
- [153] I. Nookaew et al., "Identification of flux regulation coefficients from elementary flux modes: A systems biology tool for analysis of metabolic networks," *Biotechnology and Bioengineering*, vol. 97, no. 6, pp. 1535-1549, Aug. 2007.
- [154] S. J. Wiback, R. Mahadevan, and B. Ø. Palsson, "Reconstructing metabolic flux vectors from extreme pathways: defining the alpha-spectrum," *Journal of Theoretical Biology*, vol. 224, no. 3, pp. 313-324, Oct. 2003.
- [155] M. W. COVERT and B. O. PALSSON, "Constraints-based models: Regulation of Gene Expression Reduces the Steady-state Solution Space," *Journal of Theoretical Biology*, vol. 221, no. 3, pp. 309-325, Apr. 2003.
- [156] "KEGG PATHWAY Database." [Online]. Available: <http://www.genome.jp/kegg/pathway.html>. [Accessed: 31-Jul-2011].
- [157] T. L. Nissen, U. Schulze, J. Nielsen, and J. Villadsen, "Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*.,," *Microbiology*, vol. 143 (Pt 1), no. 1 & 997, pp. 203-218, 1997.
- [158] I. Sarvari Horvath, C. J. Franzén, M. J. Taherzadeh, C. Niklasson, and G. Lidén, "Effects of furfural on the respiratory metabolism of *Saccharomyces cerevisiae* in glucose-limited chemostats," *Applied and environmental microbiology*, vol. 69, no. 7, p. 4076, 2003.
- [159] N. C. Duarte, M. J. Herrgård, and B. Ø. Palsson, "Reconstruction and Validation of *Saccharomyces cerevisiae* iND750, a Fully Compartmentalized Genome-Scale Metabolic Model," *Genome Research*, vol. 14, no. 7, pp. 1298 -1309, Jul. 2004.
- [160] Y. Fang, "Quantitative Analysis of Metabolic Fluxes in Anaerobic Chemostat Cultivations of Yeast Strains."
- [161] J. P. Barford and R. J. Hall, "A mathematical model for the aerobic growth of *Saccharomyces cerevisiae* with a saturated respiratory capacity," *Biotechnology and Bioengineering*, vol. 23, no. 8, pp. 1735-1762, Aug. 1981.
- [162] M. Terzer and J. Stelling, "efmtool - Elementary Flux Mode Tool," *Bioinformatics*, vol. 24, no. 19, pp. 2229-2235, Aug. 2008.
- [163] H. Qian and D. A. Beard, "Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium," *Biophysical Chemistry*, vol. 114, no. 2-3, pp. 213-220, Apr. 2005.
- [164] H. Qian, D. A. Beard, and S. Liang, "Stoichiometric network theory for nonequilibrium biochemical systems," *European Journal of Biochemistry*, vol. 270, no. 3, pp. 415-421, Feb. 2003.
- [165] Q. Wang, Y. Yang, H. Ma, and X. Zhao, "Metabolic network properties help assign weights to elementary modes to understand physiological flux distributions," *Bioinformatics*, vol. 23, no. 9, pp. 1049 -1052, May 2007.
- [166] J.-M. Schwartz, "A quadratic programming approach for decomposing steady-state metabolic flux distributions onto elementary modes," *Bioinformatics*, vol. 21, no. 2, p. ii204-ii205, Sep. 2005.

- [167] K. Laboratories, "KEGG: Kyoto Encyclopedia of Genes and Genomes." [Online]. Available: <http://www.genome.jp/kegg/>. [Accessed: 25-Apr-2011].
- [168] N. V. Torres and E. O. Voit, *Pathway Analysis and Optimization in Metabolic Engineering*, 1st ed. Cambridge University Press, 2002.
- [169] A. M. Gibbons, *Algorithmic graph theory*. Cambridge University Press, 1985.
- [170] M. C. Golumbic, *Algorithmic graph theory and perfect graphs*. Elsevier, 2004.
- [171] L. W. Beineke and R. J. Wilson, *Topics in algebraic graph theory*. Cambridge University Press, 2004.
- [172] B. Bollobás, *Graph theory: proceedings of the Conference on Graph Theory, Cambridge*. Elsevier, 1982.
- [173] K. Sugiyama, S. Tagawa, and M. Toda, "Methods for Visual Understanding of Hierarchical System Structures," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 11, no. 2, pp. 109-125, 1981.
- [174] M. Y. Becker and I. Rojas, "A graph layout algorithm for drawing metabolic pathways," *Bioinformatics (Oxford, England)*, vol. 17, no. 5, pp. 461-467, May 2001.
- [175] K. Wegner, "A new dynamical layout algorithm for complex biochemical reaction networks," *BMC Bioinformatics*, vol. 6, no. 1, p. 212, 2005.
- [176] W. Li and H. Kurata, "A grid layout algorithm for automatic drawing of biochemical networks," *Bioinformatics*, vol. 21, no. 9, pp. 2036-2042, Jan. 2005.
- [177] K. Kojima, M. Nagasaki, E. Jeong, M. Kato, and S. Miyano, "An efficient grid layout algorithm for biological networks utilizing various biological attributes," *BMC bioinformatics*, vol. 8, no. 1, p. 76, 2007.
- [178] K. Kojima, M. Nagasaki, and S. Miyano, "Fast grid layout algorithm for biological networks with sweep calculation," *Bioinformatics*, vol. 24, no. 12, pp. 1433-1441, Jun. 2008.
- [179] M. Kato, M. Nagasaki, A. Doi, and S. Miyano, "Automatic drawing of biological networks using cross cost and subcomponent data," *GENOME INFORMATICS SERIES*, vol. 16, no. 2, p. 22, 2005.
- [180] M. Suderman and M. Hallett, "Tools for visually exploring biological networks," *Bioinformatics*, vol. 23, no. 20, pp. 2651-2659, Oct. 2007.
- [181] A. Zoubarov, "Tools for visual analysis of biological networks."
- [182] P. Saraiya, C. North, and K. Duca, "Visualizing biological pathways: requirements analysis, systems evaluation and research agenda," *Information Visualization*, vol. 4, no. 3, pp. 191-205, Jun. 2005.
- [183] R. Bourqui et al., "Metabolic network visualization eliminating node redundance and preserving metabolic pathways," *BMC Systems Biology*, vol. 1, p. 29, 2007.
- [184] C. Y. Lee, "An Algorithm for Path Connections and Its Applications," *IRE Transactions on Electronic Computers*, vol. 10, no. 3, pp. 346-365, Sep. 1961.

6 Appendixes

6.1 Measured Data from Literature

Tab. 6.1 Calculated reaction rates in anaerobic, glucose-limited continuous cultures based on data from Nissen et al. [157].

| Reaction | D | Reaction Rate (C-mol/(gh)) | | |
|------------|---|----------------------------|-----------|-----------|
| | | 0.1 (1/h) | 0.2 (1/h) | 0.3 (1/h) |
| 1 glc | | 0.0333 | 0.0667 | 0.1042 |
| 10 glyc | | 0.00286 | 0.00574 | 0.00896 |
| 18 etoh | | 0.01756 | 0.03315 | 0.05179 |
| 27 co2 | | 0.0095 | 0.01814 | 0.02834 |
| 36 biomass | | 0.10302 | 0.20634 | 0.32235 |

Tab. 6.2 Corrected reaction rates in aerobic chemostat culture based on reaction rates taken from [71] with calibration of glucose and ethanol.

| Reaction | D | Rate (C-mol/(gh)) | | | | | | |
|----------------|---|-------------------|------------|------------|------------|-----------|------------|-----------|
| | | 0.1 (1/h) | 0.24 (1/h) | 0.26 (1/h) | 0.28 (1/h) | 0.3 (1/h) | 0.34 (1/h) | 0.4 (1/h) |
| 1 glc | | 0.00636 | 0.01679 | 0.02223 | 0.02893 | 0.03663 | 0.05299 | 0.07775 |
| 10 glyc | | 0 | 0 | 0.00034 | 0.00084 | 0.00187 | 0.00332 | 0.00517 |
| 18 etoh | | 0 | 0 | 0.00229 | 0.00565 | 0.01252 | 0.02223 | 0.03469 |
| 27 co2 | | 0.0035 | 0.00791 | 0.0089 | 0.01146 | 0.01403 | 0.01916 | 0.02686 |
| 35 respiration | | 0.0035 | 0.00791 | 0.00757 | 0.00694 | 0.00632 | 0.00507 | 0.0032 |

6.2 Stoichiometric Matrixes

The stoichiometric matrix is the kernel of the metabolic models. In this section the stoichiometric matrix for all metabolites (matrix **Y**) of several models will be shown. The involved metabolites are indicated with abbreviation.

6.2.1 Stoichiometric Matrix Y of Model 1

Tab. 6.3 Matrix Y of Model 1 with 40 metabolites (rows) and 35 reactions (columns).

| Metabolites | Reactions | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------------|-----------|----|--------|--------|-----|---------|---------|----|---------|----|--------|--------|----|--------|---------|---------|--------|-----|----|---------|---------|---------|------|------|----|-------|------|------|--------|--------|--------|--------|--------|--------|--------|---|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | | |
| 1 glc | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 2 atp | -0.1667 | 0 | -0.167 | 0 | 0 | 0 | -0.1667 | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0.3333 | -0.25 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.2 | 0 | 0 | -0.25 | 0 | -0.2 | 0 | -0.261 | -0.463 | -0.489 | -1.04 | -0.4 | 2 | | |
| 3 g6p | 1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 4 f6p | 0 | 1 | 0 | 0 | 0.6 | 0.6667 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 5 pol | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 6 r5p | 0 | 0 | 0 | 0.8333 | -1 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.217 | -0.254 | -0.24 | -0.04 | 0 | 0 | | |
| 7 nadh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0 | 0.3333 | 0.125 | 0 | 0 | 0 | 0 | -0.5 | 0 | 0 | 0.14286 | 0.08333 | 0.2 | 0.25 | 0.25 | 0 | 0 | 0 | 0 | 0 | 0.154 | 0.1348 | 0.0725 | 0 | 0 | -2 | | |
| 8 nadph | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0.08333 | 0 | 0 | 0 | 0 | -0.2 | 0 | 0.0435 | -0.051 | 0.0568 | -0.158 | -0.711 | 0 | 0 | | |
| 9 co2 | 0 | 0 | 0 | 0.1667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.25 | 0 | 0.3333 | 0 | 0 | 0 | 0.14286 | 0.16667 | 0.2 | 0 | 0 | 0 | 0 | 0 | -1 | -0.043 | 0 | 0 | 0.1054 | 0 | 0 | | |
| 10 co2extem | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 11 e4p | 0 | 0 | 0 | 0 | 0.4 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.062 | 0 | 0 | |
| 12 gap | 0 | 0 | 0 | 0 | 0 | 0.3333 | 0.5 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0117 | 0 | 0 | |
| 13 dhap | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 14 g3p | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.066 | 0 | |
| 15 glyc | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 16 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.375 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 17 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.375 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.118 | -0.101 | 0 | 0 | |
| 18 pep | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.094 | 0 | 0 | |
| 19 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.75 | 0 | -1 | 0 | 0 | 0 | 0 | -0.4286 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.338 | 0 | 0 | |
| 20 oxac | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 21 oxam | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.5714 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 22 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.174 | -0.331 | -0.299 | -0.215 | 0 | 0 | | |
| 23 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 24 etoh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 25 ace | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 26 akg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.625 | 0 | 0 | 0 | 0.5556 | 0 | 0 | 0 | 0 | 0 | 0.83333 | -1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4927 | 0 | 0 | 0 | |
| 27 mal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 28 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.625 | 0 | 0 | 0 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0.4348 | 0.4371 | 0.5271 | -0.609 | 0 | 0 | | |
| 29 actcoa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.833 | 0 | |
| 30 gln | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -0.435 | -0.437 | -0.527 | -0.208 | 0 | 0 | | |
| 31 fum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1739 | 0.1278 | 0.1073 | 0.0413 | 0 | 0 | | |
| 32 5aic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3913 | -0.458 | -0.511 | 0.0188 | 0 | 0 | | |
| 33 mfh4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.084 | -0.057 | 0.0165 | 0 | 0 | | |
| 34 dna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | | |
| 35 rna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | | |
| 36 prot | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | |
| 37 lip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 38 iso | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.85714 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 39 succ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 40 o2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |

6.2.2 Stoichiometric Matrix Y of Model 2 α

Tab. 6.4 Matrix Y of Model 2 α with 41 metabolites (rows) and 36 reactions (columns). Compared to Model 1, matrix Y of Model 2 α has one more metabolite (biomass) and one more reaction (pseudo-reaction of biomass formation).

| Metabolites | Reactions | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------------|-----------|----|----|--------|-----|---------|---------|---------|----|--------|-------|--------|----|--------|-------|---------|--------|--------|-----|----|---------|---------|-----|------|------|-----|------|----|--------|--------|--------|--------|--------|--------|--------|--------|---------|--------|---------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | | | | |
| 1 glc | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | |
| 2 atp | -0.1667 | 0 | 0 | 0 | 0 | 0 | -0.1667 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0.3333 | -0.25 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.2 | 0 | 0 | -0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2*PO | -Vstp | | |
| 3 g6p | 1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | |
| 4 f6p | 0 | 1 | 0 | 0 | 0.6 | 0.6667 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | |
| 5 pol | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0139 | | | |
| 6 r5p | 0 | 0 | 0 | 0.8333 | -1 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.217 | -0.254 | -0.24 | -0.04 | 0 | 0 | 0 | 0 | | | |
| 7 nadh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0 | 0.3333 | 0.125 | 0 | 0 | 0 | 0 | 0 | 0 | -0.5 | 0 | 0 | 0.14286 | 0.08333 | 0.2 | 0.25 | 0.25 | 0 | 0 | 0 | 0 | 0 | 0.154 | 0.1348 | 0.0725 | 0 | -2 | 0 | | | | |
| 8 nadph | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0.08333 | 0 | 0 | 0 | 0 | -0.2 | 0 | 0.0435 | -0.051 | 0.0568 | -0.158 | -0.711 | 0 | 0 | 0 | | | | |
| 9 co2 | 0 | 0 | 0 | 0.1667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.25 | 0 | 0.3333 | 0 | 0 | 0 | 0.14286 | 0.16667 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -0.043 | 0 | 0 | 0.1054 | 0 | 0 | | | |
| 10 co2extern | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | |
| 11 e4p | 0 | 0 | 0 | 0 | 0.4 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.062 | 0 | 0 | | | |
| 12 gap | 0 | 0 | 0 | 0 | 0 | 0.3333 | 0.5 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0117 | 0 | 0 | | | |
| 13 dhap | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 14 g3p | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.066 | 0 | 0 | | |
| 15 glyc | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 16 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.375 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 17 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.375 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.118 | -0.101 | 0 | 0 | | |
| 18 pep | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.094 | 0 | 0 | | | |
| 19 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.75 | 0 | -1 | 0 | 0 | 0 | 0 | -0.4286 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.338 | 0 | 0 | | | |
| 20 oxac | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 21 oxam | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.5714 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 22 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.174 | -0.331 | -0.299 | -0.215 | 0 | 0 | 0 | 0 | | | |
| 23 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 24 etoh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 25 ace | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 26 akg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.625 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.83333 | -1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4927 | 0 | 0 | 0 | | |
| 27 mal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 28 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.625 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0.4348 | 0.4371 | 0.5271 | -0.609 | 0 | 0 | 0 | 0 | | |
| 29 actcoa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.833 | 0 | 0 | |
| 30 gln | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -0.435 | -0.437 | -0.527 | -0.208 | 0 | 0 | 0 | 0 | | | |
| 31 fum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1739 | 0.1278 | 0.1073 | 0.0413 | 0 | 0 | |
| 32 5aic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3913 | -0.458 | -0.511 | 0.0188 | 0 | 0 |
| 33 mfn4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.084 | -0.057 | 0.0165 | 0 | 0 | 0 |
| 34 dna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -0.0001 | |
| 35 rna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -0.0018 | |
| 36 prot | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | -0.0179 | |
| 37 lip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | -0.001 |
| 38 iso | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 39 succ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 40 o2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 |
| 41 biomass | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

6.2.3 Stoichiometric Matrix Y of Model 2

Tab. 6.5 Matrix Y of Model 2 with 40 metabolites (rows) and 36 reactions (columns). Compared to Model 2 α , matrix Y of Model 2 has one less metabolite (Methyl-FH4) and the coefficients in the pseudo-reaction of biomass synthesis are adjusted.

| Metabolites | Reactions | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------------|-----------|----|----|--------|-----|---------|---------|---------|----|--------|--------|-------|--------|--------|-------|---------|---------|--------|-----|----|---------|---------|-----|------|------|-------|------|----|--------|--------|--------|--------|--------|--------|--------|--------|---------|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | | |
| 1 glc | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 2 atp | -0.1667 | 0 | 0 | 0 | 0 | 0 | -0.1667 | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0.3333 | -0.25 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0.2 | 0 | 0 | -0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2*PO | -Yatp | |
| 3 g6p | 1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 4 f6p | 0 | 1 | 0 | 0 | 0.6 | 0.6667 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 5 pol | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0125 | |
| 6 r5p | 0 | 0 | 0 | 0.8333 | -1 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.217 | -0.254 | -0.24 | -0.04 | 0 | 0 | 0 | | |
| 7 nadh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0 | 0.3333 | 0.125 | 0 | 0 | 0 | 0 | 0 | 0 | -0.5 | 0 | 0 | 0.14286 | 0.08333 | 0.2 | 0.25 | 0.25 | 0 | 0 | 0 | 0 | 0.154 | 0.1348 | 0.0725 | 0 | -2 | 0 | 0 | | |
| 8 nadph | 0 | 0 | 0 | 0.3333 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0.08333 | 0 | 0 | 0 | 0 | -0.2 | 0 | 0.0435 | -0.051 | 0.0568 | -0.158 | -0.711 | 0 | 0 | 0 | 0 | |
| 9 co2 | 0 | 0 | 0 | 0.1667 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.25 | 0 | 0.3333 | 0 | 0 | 0 | 0.14286 | 0.16667 | 0.2 | 0 | 0 | 0 | 0 | 0 | -1 | -0.043 | 0 | 0 | 0.1054 | 0 | 0 | 0 | 0 | |
| 10 co2extern | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 11 e4p | 0 | 0 | 0 | 0 | 0.4 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.062 | 0 | 0 | 0 | |
| 12 gap | 0 | 0 | 0 | 0 | 0 | 0.3333 | 0.5 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0117 | 0 | 0 | 0 | 0 |
| 13 dhap | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 g3p | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.066 | 0 | 0 |
| 15 glyc | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.375 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.118 | -0.101 | 0 | 0 |
| 18 pep | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.094 | 0 | 0 | 0 |
| 19 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.75 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.338 | 0 | 0 | 0 |
| 20 oxac | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 oxam | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.5714 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4444 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.174 | -0.331 | -0.299 | -0.215 | 0 | 0 | 0 | 0 |
| 23 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 etoh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25 ace | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 26 akg | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.625 | -1 | 0 | 0 | 0.5556 | 0 | 0 | 0 | 0 | 0 | 0.83333 | -1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0.4927 | 0 | 0 | 0 | 0 | 0 |
| 27 mal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 28 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.625 | 0 | 0 | -0.5556 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0.4348 | 0.4371 | 0.5271 | -0.609 | 0 | 0 | 0 | |
| 29 actcoa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.833 | 0 | 0 |
| 30 gln | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.435 | -0.437 | -0.527 | -0.208 | 0 | 0 | 0 | 0 | |
| 31 fum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1739 | 0.1278 | 0.1073 | 0.0413 | 0 | 0 | 0 | 0 |
| 32 5aic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3913 | -0.458 | -0.511 | 0.0188 | 0 | 0 | 0 | 0 | |
| 33 dna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | -0.0001 | |
| 34 rna | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -0.0016 | |
| 35 prot | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | -0.0161 | |
| 36 lip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -0.0009 | |
| 37 iso | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.85714 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 38 succ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 39 o2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 |
| 40 biomass | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

6.2.4 Stoichiometric Matrix Y of Model 3

Tab. 6.6 Matrix Y of Model 3 with 20 metabolites (rows) and 18 reactions (columns).

| Metabolites | Reactions | | | | | | | | | | | | | | | | | |
|-------------|-----------|----|----|---------|---------|---------|--------|----|--------|--------|--------|------|----|------|--------|------|----|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 1 glc | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 o2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 |
| 3 co2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | -0.143 | 0.3333 | 0 | 0 | 0.2 | 1 | 0 | -1 | 0 |
| 4 co2extern | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 5 etoh | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 glyc | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 pol | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.0083 |
| 8 r5p | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00531 |
| 9 ser | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00145 |
| 10 asp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.00743 |
| 11 lip | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | -0.00115 |
| 12 glu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -0.00538 |
| 13 g6p | 1 | -1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 atp | -0.1667 | 0 | 0 | -0.1667 | -0.1667 | 0 | 0.3333 | 0 | 0.3333 | -0.214 | 0 | 0 | 0 | -0.2 | 0.3333 | 2•PO | 0 | -Yatp |
| 15 gap | 0 | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 nadh | 0 | 0 | 0 | 0 | 0 | -0.3333 | 0.3333 | 0 | 0 | 0 | 0 | -0.5 | 0 | 0.4 | 1.6667 | -2 | 0 | 0 |
| 17 3pg | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 pyr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -0.857 | -1 | 0 | 0 | -1.2 | -1 | 0 | 0 | 0 |
| 19 acald | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.6667 | -1 | -1 | 0 | 0 | 0 | 0 | 0 |
| 20 biomass | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

6.3 Overview of the Models

In section 2.2 several derivative models are involved. The details of the models are listed here for reference and comparison.

6.3.1 Model 2

Model 2 is the final improved metabolic model from Model 1 and Model 2 α . Compared with Model 2 α , Model 2 has no methyl-FH4 balance constrain and the coefficients of ATP and biomass components were also adjusted.

Tab. 6.7 Biochemical reactions of Model 2 for *S. cerevisiae* in C-mol stoichiometry. The parts with grey background represent the different reactions compared to Model 2 α . The abbreviations of metabolites are listed in Tab. 6.8.

| No. | Designation | Biochemical Reactions |
|-----|-------------|---|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <==> f6p |
| 3 | pol | g6p --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <==> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <==> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <==> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <==> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <==> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akg + (0.125) nadh |
| 13 | 3pg-pep | 3pg <==> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akg |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akg | iso --> (0.833333) akg + (0.16666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akg-succ | akg --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <==> mal + (0.25) nadh |
| 25 | mal-oxam | mal <==> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <==> oxam |
| 27 | akg-glu | akg + (0.2) nadph --> glu |
| 28 | glu-gln | glu --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.3313) asp + (0.2544) r5p + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.2993) asp + (0.2400) r5p --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (0.1579) nadph --> prot + (0.4927) akg + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0188) 5aic + (0.0725) nadh |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.7111) nadph --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 36 | biomass | (0.0001053927) dna + (0.0015819381) rna + (0.012482919) pol + (0.0008700003) lip + (0.0160886997) prot + (Y _{ATP}) atp --> biomass |

Tab. 6.8 Metabolites in Model 2 of *S. cerevisiae*.

| No. | Abbreviation | Full name of metabolites |
|-----|--------------|---|
| 1 | glc | Glucose |
| 2 | atp | Adenosine triphosphate |
| 3 | g6p | Glucose 6-phosphate |
| 4 | f6p | Fructose 6-phosphate |
| 5 | pol | Polysaccharide |
| 6 | r5p | Ribose 5-phosphate |
| 7 | nadh | Nicotinamide adenine dinucleotide - reduced |
| 8 | nadph | Nicotinamide adenine dinucleotide phosphate - reduced |
| 9 | co2 | Carbon dioxide |
| 10 | co2extern | Carbon dioxide (extern) |
| 11 | e4p | Erythrose 4-phosphate |
| 12 | gap | Glyceraldehyde 3-phosphate |
| 13 | dhap | Dihydroxyacetone phosphate |
| 14 | g3p | Glycerol 3-phosphate |
| 15 | glyc | Glycerol |
| 16 | 3pg | 3-Phosphoglycerate |
| 17 | ser | Serine family |
| 18 | pep | Phosphoenolpyruvate |
| 19 | pyr | Pyruvate |
| 20 | oxac | Oxaloacetate (cytosol) |
| 21 | oxam | Oxaloacetate (mitochondrion) |
| 22 | asp | Aspartate |
| 23 | acald | Acetaldehyde |
| 24 | etoh | Ethanol |
| 25 | ace | Acetate |
| 26 | akg | alpha-Ketoglutarate |
| 27 | mal | Malate |
| 28 | glu | Glutamate |
| 29 | actcoa | Acetyl-CoA |
| 30 | gln | Glutamine |
| 31 | fum | Fumarate |
| 32 | 5aic | 5-AICAR |
| 33 | dna | Deoxyribonucleic acid |
| 34 | rna | Ribonucleic acid |
| 35 | prot | Protein |
| 36 | lip | Lipid |
| 37 | iso | Isocitrate |
| 38 | succ | Succinate |
| 39 | o2 | Oxygen |
| 40 | biomass | Biomass |

6.3.2 Model 2A

Model 2A is a model without the pseudo-reaction of biomass synthesis and the energy consumption of biomass formation is lumped together and represented with Y_{ATP} in reaction 18 (see Tab. 6.9).

Tab. 6.9 Biochemical reactions of Model 2A for *S. cerevisiae* in C-mol stoichiometry. Model 2A does not contain the lumped reaction of biomass and instead of it a reaction for Y_{ATP} is used.

| No. | Designation | Biochemical Reactions |
|-----|-------------|---|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <=> f6p |
| 3 | pol | g6p --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <=> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <=> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <=> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <=> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <=> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akg + (0.125) nadh |
| 13 | 3pg-pep | 3pg <=> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akg |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akg | iso --> (0.833333) akg + (0.16666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akg-succ | akg --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <=> mal + (0.25) nadh |
| 25 | mal-oxam | mal <=> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <=> oxam |
| 27 | akg-glu | akg + (0.2) nadph --> glu |
| 28 | glu-gln | glu --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.3313) asp + (0.2544) r5p + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.2993) asp + (0.2400) r5p --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (0.1579) nadph --> prot + (0.4927) akg + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0188) 5aic + (0.0725) nadh |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.7111) nadph --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 36 | Y_{ATP} | (Y_{ATP}) atp --> biomass |

6.3.3 Model 2B

In Model 2B the biomass coefficients k_n are original values calculated from the available data of biomass component from literature [157] (see Tab. 2.9 on page 46), while in Model 2 the calibrated k_n is used.

Tab. 6.11 Biochemical reactions of Model 2B for *S. cerevisiae* in C-mol stoichiometry. Compared to Model 2, in Model 2B the original k_n in the pseudo-reaction of biomass synthesis are used (shown with grey background).

| No. | Designation | Biochemical Reactions |
|-----|-------------|---|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <=> f6p |
| 3 | pol | g6p --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <=> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <=> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <=> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <=> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <=> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akglu + (0.125) nadh |
| 13 | 3pg-pep | 3pg <=> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akglu |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akglu | iso --> (0.833333) akglu + (0.16666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akglu-succ | akglu --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <=> mal + (0.25) nadh |
| 25 | mal-oxam | mal <=> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <=> oxam |
| 27 | akglu-glu | akglu + (0.2) nadph --> glu |
| 28 | glu-gln | glu --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.3313) asp + (0.2544) r5p + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.2993) asp + (0.2400) r5p --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (0.1579) nadph --> prot + (0.4927) akglu + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0188) 5aic + (0.0725) nadh |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.7111) nadph --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 36 | biomass | (0.000117103) dna + (0.001757709) rna + (0.01386991) pol + (0.000966667) lip + (0.017876333) prot + (Y _{ATP}) atp --> biomass |

6.3.4 Model 2C

In Model 2C the consumed ATPs for biomass synthesis appear in each synthesis reaction of biomass components and their coefficients are assigned to the original coefficients (given in [157]), while in Model 2 they are lumped together as $Y_{ATP} \cdot \mu$ in the pseudo-reaction of biomass synthesis.

Tab. 6.12 Biochemical reactions of Model 2C for *S. cerevisiae* in C-mol stoichiometry.

| No. | Designation | Biochemical Reactions |
|-----|-------------|--|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <==> f6p |
| 3 | pol | g6p + (0.1667) atp --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <==> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <==> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <==> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <==> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <==> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akg + (0.125) nadh |
| 13 | 3pg-pep | 3pg <==> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akg |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akg | iso --> (0.833333) akg + (0.16666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akg-succ | akg --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <==> mal + (0.25) nadh |
| 25 | mal-oxam | mal <==> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <==> oxam |
| 27 | akg-glu | akg + (0.2) nadph --> glu |
| 28 | glu-gln | glu + (0.2) atp --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 + (0.26087) atp --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.3313) asp + (0.2544) r5p + (0.4625) atp + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.2993) asp + (0.2400) r5p + (0.4890) atp --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (1.0396) atp + (0.1579) nadph --> prot + (0.4927) akg + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0188) 5aic + (0.0725) nadh |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.4) atp + (0.7111) nadph --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 36 | biomass | (0.0001053927) dna + (0.0015819381) rna + (0.012482919) pol + (0.0008700003) lip + (0.0160886997) prot --> biomass |

6.3.5 Model 2D

In Model 2D the consumed ATP for cell maintenance (m_{ATP}) is considered and represented as a reaction for m_{ATP} (see reaction 30 in Tab. 6.13).

Tab. 6.13 Biochemical reactions of Model 2D for *S. cerevisiae* in C-mol stoichiometry.

| No. | Designation | Biochemical Reactions |
|-----|-------------|---|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <=> f6p |
| 3 | pol | g6p --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <=> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <=> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <=> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <=> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <=> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akg + (0.125) nadh |
| 13 | 3pg-pep | 3pg <=> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akg |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akg | iso --> (0.833333) akg + (0.16666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akg-succ | akg --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <=> mal + (0.25) nadh |
| 25 | mal-oxam | mal <=> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <=> oxam |
| 27 | akg-glu | akg + (0.2) nadph --> glu |
| 28 | glu-gln | glu --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | mATP | atp --> |
| 31 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 32 | dna | (0.4579) 5aic + (0.4371) gln + (0.3313) asp + (0.2544) r5p + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 33 | rna | (0.5112) 5aic + (0.5271) gln + (0.2993) asp + (0.2400) r5p --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 34 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (0.1579) nadph --> prot + (0.4927) akg + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0188) 5aic + (0.0725) nadh |
| 35 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.7111) nadph --> lip |
| 36 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 37 | biomass | (0.000117103) dna + (0.001757709) rna + (0.01386991) pol + (0.000966667) lip + (0.017876333) prot + (Y_{ATP}) atp --> biomass |

6.3.6 Model 2E

Compared to Model 2, Model 2E is a model with methyl-FH₄ balance, i.e. Model 2E has one more metabolite (methyl-FH₄).

Tab. 6.14 Biochemical reactions of Model 2E for *S. cerevisiae* in C-mol stoichiometry. Model 2E has one more metabolite (shown with grey background) compared to Model 2.

| No. | Designation | Biochemical Reactions |
|-----|-------------|---|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <==> f6p |
| 3 | pol | g6p --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadph + (0.1667) co2 |
| 5 | r5p-e4p | r5p <==> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <==> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <==> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <==> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <==> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akG + (0.125) nadh |
| 13 | 3pg-pep | 3pg <==> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akG |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadph |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akG | iso --> (0.833333) akG + (0.16666667) co2 + (0.083333335) nadh + (0.083333335) nadph |
| 23 | akG-succ | akG --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <==> mal + (0.25) nadh |
| 25 | mal-oxam | mal <==> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <==> oxam |
| 27 | akG-glu | akG + (0.2) nadph --> glu |
| 28 | glu-gln | glu --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadph |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.0842) mfH ₄ + (0.3313) asp + (0.2544) r5p + (0.0509) nadph --> dna + (0.4371) glu + (0.1278) fum + (0.1540) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.0568) mfH ₄ + (0.2993) asp + (0.2400) r5p --> rna + (0.5271) glu + (0.1073) fum + (0.1348) nadh + (0.0568) nadph |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (0.1579) nadph --> prot + (0.4927) akG + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0165) mfH ₄ + (0.0188) 5aic + (0.0725) nadh |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.7111) nadph --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 36 | biomass | (0.0001053927) dna + (0.0015819381) rna + (0.012482919) pol + (0.0008700003) lip + (0.0160886997) prot + (Y _{ATP}) atp --> biomass |

6.3.7 Model 2F

In Model 2F, NADPH is considered as NADH and is written as NADH in the NADPH-related reactions.

Tab. 6.15 Biochemical reactions of Model 2F for *S. cerevisiae* in C-mol stoichiometry.

| No. | Designation | Biochemical Reactions |
|-----|-------------|--|
| 1 | glc | glc + (0.1667) atp --> g6p |
| 2 | g6p-f6p | g6p <=> f6p |
| 3 | pol | g6p --> pol |
| 4 | g6p-r5p | g6p --> (0.8333) r5p + (0.3333) nadh + (0.1667) co2 |
| 5 | r5p-e4p | r5p <=> (0.6) f6p + (0.4) e4p |
| 6 | e4p-gap | (0.5556) r5p + (0.4444) e4p <=> (0.6667) f6p + (0.3333) gap |
| 7 | f6p-gap | f6p + (0.1667) atp <=> (0.5) gap + (0.5) dhap |
| 8 | dhap-gap | dhap <=> gap |
| 9 | dhap-g3p | dhap + (0.3333) nadh --> g3p |
| 10 | glyc | g3p --> glyc |
| 11 | gap-3pg | gap <=> 3pg + (0.3333) atp + (0.3333) nadh |
| 12 | ser | (0.375) 3pg + (0.625) glu --> (0.375) ser + (0.625) akg + (0.125) nadh |
| 13 | 3pg-pep | 3pg <=> pep |
| 14 | pep-pyr | pep --> pyr + (0.3333) atp |
| 15 | pyr-oxac | (0.75) pyr + (0.25) co2 + (0.25) atp --> oxac |
| 16 | asp | (0.4444) oxac + (0.5556) glu --> (0.4444) asp + (0.5556) akg |
| 17 | pyr-acald | pyr --> (0.6667) acald + (0.3333) co2 |
| 18 | etoh | acald + (0.5) nadh --> etoh |
| 19 | ace | acald --> ace + (0.5) nadh |
| 20 | actcoa | ace + atp --> actcoa |
| 21 | pyr-iso | (0.42857) pyr + (0.57143) oxam --> (0.85714) iso + (0.142857) co2 + (0.142857) nadh |
| 22 | iso-akg | iso --> (0.833333) akg + (0.16666667) co2 + (0.16666667) nadh |
| 23 | akg-succ | akg --> (0.8) succ + (0.2) nadh + (0.2) co2 + (0.2) atp |
| 24 | succ-mal | succ <=> mal + (0.25) nadh |
| 25 | mal-oxam | mal <=> oxam + (0.25) nadh |
| 26 | oxac-oxam | oxac + (0.25) atp <=> oxam |
| 27 | akg-glu | akg + (0.2) nadh --> glu |
| 28 | glu-gln | glu --> gln |
| 29 | co2 | co2 --> co2extern |
| 30 | r5p-5aic | (0.21739) r5p + (0.43478) gln + (0.130435) ser + (0.173913) asp + (0.043478) co2 --> (0.3913) 5aic + (0.43478) glu + (0.173913) fum + (0.043478) nadh |
| 31 | dna | (0.4579) 5aic + (0.4371) gln + (0.3313) asp + (0.2544) r5p --> dna + (0.4371) glu + (0.1278) fum + (0.1031) nadh |
| 32 | rna | (0.5112) 5aic + (0.5271) gln + (0.2993) asp + (0.2400) r5p --> rna + (0.5271) glu + (0.1073) fum + (0.1916) nadh |
| 33 | protein | (0.0404) r5p + (0.609) glu + (0.2078) gln + (0.2153) asp + (0.338) pyr + (0.1182) ser + (0.0623) e4p + (0.0935) pep + (0.0854) nadh --> prot + (0.4927) akg + (0.0413) fum + (0.0117) gap + (0.1054) co2 + (0.0188) 5aic |
| 34 | lipid | (0.8326) actcoa + (0.0662) g3p + (0.1012) ser + (0.7111) nadh --> lip |
| 35 | respiration | o2 + (2) nadh --> (2 P/O) atp |
| 36 | biomass | (0.0001053927) dna + (0.0015819381) rna + (0.012482919) pol + (0.0008700003) lip + (0.0160886997) prot + (Y _{ATP}) atp --> biomass |

6.4 Complete Simulation Results

6.4.1 Modeling of Biomass Components

Tab. 6.16 Flux distribution of Model 2 and Model 2A in anaerobic growth. Model 2A is Model 2 without the pseudo-reaction of biomass synthesis. The numbers with dark background are the given rates v_m . v_c represent calculated rates. The results of Model 2A are generated via flux balance analysis with non-weighted multiple objective functions or weighted single objective function.

| No. | Reaction Designation | Given Rate C-mol/(gh) | Reaction Rates in Model 2 C-mol/(gh) | | | Reaction Rates in Model 2A C-mol/(gh) | |
|-----|----------------------|--------------------------|---|----------------|----------------------|--|------------------------------|
| | | | determined | | Under- determined | Underdetermined | |
| | | | μ as v_m | μ as v_c | | Multiple objective functions | Single objective function |
| 1 | glc | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 | |
| 2 | g6p-f6p | | 0.02998 | 0.03002 | 0.03002 | 0.03002 | |
| 3 | pol | | 0.00129 | 0.00128 | 0.00128 | 0.00128 | |
| 4 | g6p-r5p | | 0.00203 | 0.00200 | 0.00200 | 1.1E-06 | |
| 5 | r5p-e4p | | 0.00112 | 0.00111 | 0.00111 | 6.3E-07 | |
| 6 | e4p-gap | | 0.00078 | 0.00076 | 0.00076 | 4.3E-07 | |
| 7 | f6p-gap | | 0.03117 | 0.03120 | 0.03119 | 0.01110 | |
| 8 | dhap-gap | | 0.01258 | 0.01247 | 0.01247 | 0.00556 | |
| 9 | dhap-g3p | | 0.00300 | 0.00313 | 0.00313 | 7.7E-15 | |
| 10 | glyc | 0.00286 | 0.00300 | 0.00313 | 0.00312 | 7.2E-15 | |
| 11 | gap-3pg | | 0.02845 | 0.02834 | 0.02834 | 0.01110 | |
| 12 | ser | | 0.00060 | 0.00060 | 0.00059 | 3.0E-07 | |
| 13 | 3pg-pep | | 0.02822 | 0.02812 | 0.02812 | 0.01110 | |
| 14 | pep-pyr | | 0.02807 | 0.02796 | 0.02796 | 0.01110 | |
| 15 | pyr-oxac | | 0.00086 | 0.00086 | 0.00086 | 4.6E-07 | |
| 16 | asp | | 0.00098 | 0.00097 | 0.00097 | 4.9E-07 | |
| 17 | pyr-acald | | 0.02657 | 0.02644 | 0.02644 | 0.01110 | |
| 18 | etoh | 0.01756 | 0.01764 | 0.01756 | 0.01756 | 0.00740 | |
| 19 | ace | | 7.5E-05 | 7.4E-05 | 7.4E-05 | 5.3E-15 | |
| 20 | actcoa | | 7.5E-05 | 7.4E-05 | 7.4E-05 | 5.3E-15 | |
| 21 | pyr-iso | | 0.00069 | 0.00075 | 0.00075 | 4.3E-07 | |
| 22 | iso-akg | | 0.00059 | 0.00064 | 0.00064 | 3.7E-07 | |
| 23 | akg-succ | | -4.7E-05 | -9.3E-09 | 3.5E-11 | 1.5E-14 | |
| 24 | succ-mal | | -3.7E-05 | -7.5E-09 | 2.6E-11 | 0 | |
| 25 | mal-oxam | | -3.7E-05 | -7.5E-09 | 3.1E-11 | 0 | |
| 26 | oxac-oxam | | 0.00043 | 0.00042 | 0.00043 | 2.5E-07 | |
| 27 | akg-glu | | 0.00227 | 0.00225 | 0.00225 | 1.2E-06 | |
| 28 | glu-gln | | 0.00050 | 0.00049 | 0.00049 | 2.2E-07 | |
| 29 | co2 | 0.0095 | 0.00933 | 0.00931 | 0.00931 | 0.00370 | |
| 30 | r5p-5aic | | 0.00015 | 0.00014 | 0.00014 | 1.7E-16 | |
| 31 | dna | | 1.1E-05 | 1.1E-05 | 1.1E-05 | 1.7E-15 | |
| 32 | rna | | 0.00016 | 0.00016 | 0.00016 | 3.5E-08 | |
| 33 | protein | | 0.00166 | 0.00164 | 0.00164 | 9.5E-07 | |
| 34 | lipid | | 8.9E-05 | 8.9E-05 | 8.9E-05 | 6.4E-15 | |
| 35 | respiration | 0 | 0 | 0 | 0 | 0 | |
| 36 | biomass/ Y_{ATP} | 0.10302 | 0.10302 | 0.10217 | 0.10217 | (Y_{ATP}) 1.2E-16 | |
| | μ | | 0.10302 | 0.10217 | 0.10217 | (Y_{ATP}) 0.00762 | |

6.4.2 Calibrating ATP Balance

Tab. 6.17 Flux distribution of Model 2 and Model 2C in anaerobic growth. Model 2C is Model 2 without energy summary ($Y_{ATP} \cdot \mu$) in the pseudo-reaction of biomass synthesis. The numbers with dark background are the given rates v_m .

| No. | Reaction Designation | Given Rate C-mol/(gh) | Reaction Rates in Model 2 C-mol/(gh) | | Reaction Rates in Model 2C C-mol/(gh) | | | |
|-----|----------------------|--------------------------|---|-----------|--|-----------|---------------------|-----------|
| | | | Determined | Underdet. | with ATP balance | | without ATP balance | |
| | | | | | Determined | Underdet. | Determined | Underdet. |
| 1 | glc | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 |
| 2 | g6p-f6p | | 0.03002 | 0.03002 | 0.01996 | 0.02687 | 0.03002 | 0.01764 |
| 3 | pol | | 0.00128 | 0.00128 | 0.00463 | 0.00250 | 0.00128 | 0.00609 |
| 4 | g6p-r5p | | 0.00200 | 0.00200 | 0.00871 | 0.00393 | 0.00200 | 0.00957 |
| 5 | r5p-e4p | | 0.00111 | 0.00111 | 0.00481 | 0.00217 | 0.00111 | 0.00528 |
| 6 | e4p-gap | | 0.00076 | 0.00076 | 0.00349 | 0.00150 | 0.00076 | 0.00365 |
| 7 | f6p-gap | | 0.03119 | 0.03119 | 0.02517 | 0.02917 | 0.03119 | 0.02324 |
| 8 | dhap-gap | | 0.01247 | 0.01247 | 0.01694 | 0.00844 | 0.01247 | -0.00335 |
| 9 | dhap-g3p | | 0.00313 | 0.00313 | -0.00435 | 0.00615 | 0.00313 | 0.01497 |
| 10 | glyc | 0.00286 | 0.00312 | 0.00312 | -0.00437 | 0.00613 | 0.00312 | 0.01494 |
| 11 | gap-3pg | | 0.02834 | 0.02834 | 0.03076 | 0.02356 | 0.02834 | 0.00958 |
| 12 | ser | | 0.00059 | 0.00059 | 0.00215 | 0.00116 | 0.00059 | 0.00283 |
| 13 | 3pg-pep | | 0.02812 | 0.02812 | 0.02995 | 0.02312 | 0.02812 | 0.00852 |
| 14 | pep-pyr | | 0.02796 | 0.02796 | 0.02939 | 0.02282 | 0.02796 | 0.00778 |
| 15 | pyr-oxac | | 0.00086 | 0.00086 | 0.00311 | 0.00168 | 0.00086 | 0.00409 |
| 16 | asp | | 0.00097 | 0.00097 | 0.00352 | 0.00190 | 0.00097 | 0.00463 |
| 17 | pyr-acald | | 0.02644 | 0.02644 | 0.02674 | 0.01985 | 0.02644 | 0.00053 |
| 18 | etoh | 0.01756 | 0.01756 | 0.01756 | 0.01756 | 0.01309 | 0.01756 | 2.54E-16 |
| 19 | ace | | 7.40E-05 | 7.40E-05 | 0.00027 | 0.00015 | 7.40E-05 | 0.00035 |
| 20 | actcoa | | 7.40E-05 | 7.40E-05 | 0.00027 | 0.00015 | 7.40E-05 | 0.00035 |
| 21 | pyr-iso | | 0.00075 | 0.00075 | -0.00397 | 0.00146 | 0.00075 | 0.00356 |
| 22 | iso-akg | | 0.00064 | 0.00064 | -0.00340 | 0.00125 | 0.00064 | 0.00305 |
| 23 | akg-succ | | -9.33E-09 | 3.49E-11 | -0.00477 | 8.25E-17 | -2.13E-08 | 9.90E-17 |
| 24 | succ-mal | | -7.47E-09 | 2.64E-11 | -0.00381 | 7.28E-12 | -1.70E-08 | -2.84E-14 |
| 25 | mal-oxam | | -7.47E-09 | 3.09E-11 | -0.00381 | -5.46E-12 | -1.70E-08 | -2.84E-14 |
| 26 | oxac-oxam | | 0.00043 | 0.00043 | 0.00155 | 0.00084 | 0.00043 | 0.00204 |
| 27 | akg-glu | | 0.00225 | 0.00225 | 0.00818 | 0.00442 | 0.00225 | 0.01076 |
| 28 | glu-gln | | 0.00049 | 0.00049 | 0.00180 | 0.00097 | 0.00049 | 0.00236 |
| 29 | co2 | 0.0095 | 0.00931 | 0.00931 | 0.00810 | 0.00759 | 0.00931 | 0.00256 |
| 30 | r5p-5aic | | 0.00014 | 0.00014 | 0.00053 | 0.00028 | 0.00014 | 0.00069 |
| 31 | Dna | | 1.08E-05 | 1.08E-05 | 3.91E-05 | 2.11E-05 | 1.08E-05 | 5.14E-05 |
| 32 | Rna | | 0.00016 | 0.00016 | 0.00059 | 0.00032 | 0.00016 | 0.00077 |
| 33 | protein | | 0.00164 | 0.00164 | 0.00597 | 0.00322 | 0.00164 | 0.00785 |
| 34 | Lipid | | 8.89E-05 | 8.89E-05 | 0.00032 | 0.00017 | 8.89E-05 | 0.00042 |
| 35 | respiration | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 36 | Biomass | 0.10302 | 0.10217 | 0.10217 | 0.37118 | 0.20041 | 0.10217 | 0.48801 |

6.4.3 Effect of Methyl-FH4 Balance

Tab. 6.18 Flux distribution of Model 2 (without methyl-FH₄ balance) and Model 2E (with methyl-FH₄ balance) in anaerobic growth. The numbers with dark background are the given rates v_m . For the overdetermined system (overdet.) of Model 2E two calculation methods were performed: one with only matrix G^T (overdet. (G^T)) and the other also using matrix F (overdet. (F)), which are covered in Eq. (2.52) on page 24 and Eq. (2.56) on page 25 respectively.

| Reaction | | Given Rate C-mol/(gh) | Rate in Model 2 C-mol/(gh) | | Rate in Model 2E (with methyl-FH ₄ balance) C-mol/(gh) | | |
|----------|-------------|--------------------------|-------------------------------|-----------|--|--------------------|--------------|
| No. | Designation | | determined | underdet. | determined | overdet. (G^T) | overdet. (F) |
| 1 | Glc | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.03345 |
| 2 | g6p-f6p | | 0.03002 | 0.03002 | 0.03498 | 0.03002 | 0.03296 |
| 3 | Pol | | 0.00128 | 0.00128 | 0 | 0.00128 | 2.75E-09 |
| 4 | g6p-r5p | | 0.00200 | 0.00200 | -0.00168 | 0.00200 | 0.00049 |
| 5 | r5p-e4p | | 0.00111 | 0.00111 | -0.00093 | 0.00111 | 0.00027 |
| 6 | e4p-gap | | 0.00076 | 0.00076 | -0.00084 | 0.00076 | 0.00025 |
| 7 | f6p-gap | | 0.03119 | 0.03119 | 0.03386 | 0.03119 | 0.03329 |
| 8 | dhap-gap | | 0.01247 | 0.01247 | -0.00153 | 0.01247 | 0.00766 |
| 9 | dhap-g3p | | 0.00313 | 0.00313 | 0.01846 | 0.00313 | 0.00898 |
| 10 | Glyc | 0.00286 | 0.00312 | 0.00312 | 0.01846 | 0.00312 | 0.00898 |
| 11 | gap-3pg | | 0.02834 | 0.02834 | 0.01512 | 0.02834 | 0.02439 |
| 12 | Ser | | 0.00059 | 0.00059 | 1.24E-18 | 0.00059 | 1.28E-09 |
| 13 | 3pg-pep | | 0.02812 | 0.02812 | 0.01512 | 0.02812 | 0.02439 |
| 14 | pep-pyr | | 0.02796 | 0.02796 | 0.01512 | 0.02796 | 0.02439 |
| 15 | pyr-oxac | | 0.00086 | 0.00086 | 2.79E-08 | 0.00086 | -6.35E-09 |
| 16 | Asp | | 0.00097 | 0.00097 | 0 | 0.00097 | 2.10E-09 |
| 17 | pyr-acald | | 0.02644 | 0.02644 | 0.01176 | 0.02644 | 0.02538 |
| 18 | etoh | 0.01756 | 0.01756 | 0.01756 | 0.00784 | 0.01756 | 0.01692 |
| 19 | Ace | | 7.40E-05 | 7.40E-05 | 0 | 7.40E-05 | 1.60E-10 |
| 20 | actcoa | | 7.40E-05 | 7.40E-05 | 0 | 7.40E-05 | 1.60E-10 |
| 21 | pyr-iso | | 0.00075 | 0.00075 | 0.00783 | 0.00074 | -0.00230 |
| 22 | iso-akg | | 0.00064 | 0.00064 | 0.00671 | 0.00064 | -0.00197 |
| 23 | akg-succ | | -9.33E-09 | 3.49E-11 | 0.00560 | -1.16E-06 | -0.00164 |
| 24 | succ-mal | | -7.47E-09 | 2.64E-11 | 0.00448 | -9.28E-07 | -0.00132 |
| 25 | mal-oxam | | -7.47E-09 | 3.09E-11 | 0.00448 | -9.32E-07 | -0.00132 |
| 26 | oxac-oxam | | 0.00043 | 0.00043 | 2.79E-08 | 0.00043 | -7.28E-09 |
| 27 | akg-glu | | 0.00225 | 0.00225 | 0 | 0.00225 | 4.86E-09 |
| 28 | glu-gln | | 0.00049 | 0.00049 | 0 | 0.00050 | 1.07E-09 |
| 29 | co2 | 0.0095 | 0.00931 | 0.00931 | 0.00700 | 0.00931 | 0.00755 |
| 30 | r5p-5aic | | 0.00014 | 0.00014 | 0 | 0.00015 | 3.19E-10 |
| 31 | Dna | | 1.08E-05 | 1.08E-05 | 0 | 1.22E-05 | 2.63E-11 |
| 32 | Rna | | 0.00016 | 0.00016 | 0 | 0.00016 | 3.51E-10 |
| 33 | protein | | 0.00164 | 0.00164 | 0 | 0.00164 | 3.55E-09 |
| 34 | Lipid | | 8.89E-05 | 8.89E-05 | 0 | 8.89E-05 | 1.92E-10 |
| 35 | respiration | 0 | 0 | 0 | 0 | 0 | -0.00240 |
| 36 | Biomass | 0.10302 | 0.10217 | 0.10217 | 0 | 0.10217 | 2.21E-07 |

6.4.4 Effects of Model Simplification

Tab. 6.19 Flux distribution of Model 2 and Model 2F in anaerobic growth. In Model 2F, NADPH is substituted with NADH. The numbers with dark background are the given rates v_m . The last three columns were calculated by giving different measured data.

| Reaction | | Rate in Model 2 C-mol/(gh) | | Rate in Model 2F C-mol/(gh) | | | |
|----------|-------------|-------------------------------|-----------------|--------------------------------|-----------------------------|--------------------------------|--------------------------------|
| No. | Designation | determined | Underdetermined | determined | underdetermined System | | |
| | | | | | v_1 and v_{35} as v_m | v_1, v_{18}, v_{35} as v_m | v_1, v_{29}, v_{35} as v_m |
| 1 | Glc | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 | 0.0333 |
| 2 | g6p-f6p | 0.0300207 | 0.03002071 | 0.03002126 | 0.03183338 | 0.0319853 | 0.031977 |
| 3 | Pol | 0.00127537 | 0.00127536 | 0.00127538 | 0.00146662 | 0.0013147 | 0.001323 |
| 4 | g6p-r5p | 0.00200394 | 0.00200393 | 0.00200337 | 5.17E-12 | 1.333E-14 | 6.48E-15 |
| 5 | r5p-e4p | 0.0011056 | 0.0011056 | 0.00110529 | -8.72E-06 | -7.82E-06 | -7.87E-06 |
| 6 | e4p-gap | 0.0007647 | 0.0007647 | 0.00076442 | -0.00027285 | -0.000245 | -0.000246 |
| 7 | f6p-gap | 0.03119389 | 0.03119389 | 0.03119406 | 0.03164624 | 0.0318176 | 0.0318081 |
| 8 | dhap-gap | 0.01246629 | 0.01246627 | 0.0124664 | 0.01452705 | 0.0128656 | 0.0129571 |
| 9 | dhap-g3p | 0.00313065 | 0.00313068 | 0.00313063 | 0.00129607 | 0.0030432 | 0.002947 |
| 10 | Glyc | 0.00312477 | 0.00312479 | 0.00312474 | 0.0012893 | 0.0030371 | 0.0029409 |
| 11 | gap-3pg | 0.02833734 | 0.02833732 | 0.02833745 | 0.03028135 | 0.0287127 | 0.0287991 |
| 12 | Ser | 0.00059246 | 0.00059246 | 0.00059246 | 0.0006813 | 0.0006107 | 0.0006146 |
| 13 | 3pg-pep | 0.02811517 | 0.02811515 | 0.02811527 | 0.03002586 | 0.0284837 | 0.0285686 |
| 14 | pep-pyr | 0.02796148 | 0.02796146 | 0.02796158 | 0.02984912 | 0.0283253 | 0.0284092 |
| 15 | pyr-oxac | 0.00085722 | 0.00085722 | 0.00085723 | 0.00098577 | 0.0008836 | 0.0008893 |
| 16 | Asp | 0.0009699 | 0.0009699 | 0.00096991 | 0.00111534 | 0.0009998 | 0.0010061 |
| 17 | pyr-acald | 0.02644333 | 0.0264433 | 0.02644333 | 0.0281033 | 0.0264467 | 0.026538 |
| 18 | etoh | 0.01755576 | 0.01755574 | 0.01755576 | 0.01865136 | 0.0175558 | 0.0176161 |
| 19 | Ace | 7.4007E-05 | 7.4007E-05 | 7.4008E-05 | 8.51E-05 | 7.629E-05 | 7.68E-05 |
| 20 | actcoa | 7.4007E-05 | 7.4007E-05 | 7.4008E-05 | 8.51E-05 | 7.629E-05 | 7.68E-05 |
| 21 | pyr-iso | 0.00074583 | 0.00074584 | 0.00074605 | 0.00085769 | 0.0015005 | 0.0014651 |
| 22 | iso-akg | 0.00063928 | 0.00063929 | 0.00063947 | 0.00073516 | 0.0012861 | 0.0012558 |
| 23 | akg-succ | -9.332E-09 | 3.4895E-11 | 1.3867E-07 | 3.37E-11 | 0.0005226 | 0.0004938 |
| 24 | succ-mal | -7.466E-09 | 2.6375E-11 | 1.1093E-07 | 3.94E-11 | 0.0004181 | 0.0003951 |
| 25 | mal-oxam | -7.466E-09 | 3.0923E-11 | 1.1093E-07 | 3.17E-11 | 0.0004181 | 0.0003951 |
| 26 | oxac-oxam | 0.0004262 | 0.0004262 | 0.0004262 | 0.00049011 | 0.0004393 | 0.0004421 |
| 27 | akg-glu | 0.00225179 | 0.00225178 | 0.00225181 | 0.00258947 | 0.0023212 | 0.0023359 |
| 28 | glu-gln | 0.00049442 | 0.00049442 | 0.00049442 | 0.00056856 | 0.0005097 | 0.0005129 |
| 29 | co2 | 0.00931336 | 0.00931336 | 0.00931336 | 0.00955744 | 0.0092991 | 0.0093134 |
| 30 | r5p-5aic | 0.00014477 | 0.00014477 | 0.00014478 | 0.00016649 | 0.0001492 | 0.0001502 |
| 31 | Dna | 1.0768E-05 | 1.0768E-05 | 1.0768E-05 | 1.24E-05 | 1.11E-05 | 1.12E-05 |
| 32 | Rna | 0.00016162 | 0.00016162 | 0.00016163 | 0.00018586 | 0.0001666 | 0.0001677 |
| 33 | protein | 0.00164376 | 0.00164376 | 0.00164378 | 0.00189027 | 0.0016944 | 0.0017052 |
| 34 | Lipid | 8.8887E-05 | 8.8887E-05 | 8.8888E-05 | 0.00010222 | 9.163E-05 | 9.22E-05 |
| 35 | respiration | 0 | 0 | 0 | 0 | 0 | 0 |
| 36 | biomass | 0.10216887 | 0.1021687 | 0.10216976 | 0.1174903 | 0.105317 | 0.1059875 |

6.5 Elementary Flux Mode of Model 3

Tab. 6.20* Elementary flux mode of Model 3 with $P/O = 1.65$ and $Y_{ATP} = 0.0746$ mol/g.

| No. Designation | Anaerobic Growth | | | | Oxidative Growth | | | Aerob. F. G. |
|-----------------|------------------|-------------|-------------|-------------|------------------|-------------|-------------|--------------|
| | EM 1 | EM 2 | EM 3 | EM 4 | EM 5 | EM 6 | EM 7 | EM 8 |
| 1 glc | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 pol | 0.027559551 | 1.85081E-05 | 0 | 0 | 0.182259674 | 0 | 0 | 0.037847326 |
| 3 g6p-co2 | 0 | 0.666152344 | 0.666538865 | 0 | 0 | 0 | 0.959668054 | 0 |
| 4 g6p-r5p | 0.017631472 | 1.18407E-05 | 0 | 0 | 0.116602273 | 0 | 0 | 0.024213169 |
| 5 g6p-gap | 0.954808978 | 0.333817307 | 0.333461135 | 1 | 0.701138053 | 1 | 0.040331946 | 0.937939505 |
| 6 glyc | 0.074436337 | 0 | 0 | 0.545290897 | 0 | 0.817861189 | 0 | 0 |
| 7 gap-3pg | 0.880372641 | 0.333817307 | 0.333461135 | 0.454709103 | 0.701138053 | 0.182138811 | 0.040331946 | 0.937939505 |
| 8 ser | 0.00481462 | 3.23334E-06 | 0 | 0 | 0.031840545 | 0 | 0 | 0.006611882 |
| 9 3pg-pyr | 0.875558021 | 0.333814074 | 0.333461135 | 0.454709103 | 0.669297507 | 0.182138811 | 0.040331946 | 0.931327623 |
| 10 asp | 0.024670779 | 1.65681E-05 | 0 | 0 | 0.163155347 | 0 | 0 | 0.033880197 |
| 11 pyr-acald | 0.832978481 | 0.333785479 | 0.333452799 | 0.36382728 | 0.037877376 | 0 | 0 | 0.872853458 |
| 12 etoh | 0.551528262 | 0.222532214 | 0.222312981 | 0.242563647 | 0 | 0 | 0 | 0.576687494 |
| 13 lip | 0.003818492 | 2.56437E-06 | 0 | 0 | 0.025252846 | 0 | 0 | 0.005243907 |
| 14 glut | 0.017863902 | 1.19968E-05 | 0 | 0 | 0.118139403 | 0 | 0 | 0.024532363 |
| 15 pyr-co2 | 0 | 0 | 8.33632E-06 | 0.090881824 | 0.349828716 | 0.182138811 | 0.040331946 | 0 |
| 16 o2 | 0 | 0 | 0 | 0 | 0.432002297 | 0.045842243 | 0.040331946 | 0.017042218 |
| 17 co2 | 0.277676587 | 0.777403074 | 0.777687019 | 0.212145456 | 0.362749911 | 0.182138811 | 1 | 0.290983662 |
| 18 biomass | 3.320427802 | 0.002229886 | 0 | 0 | 21.95899686 | 0 | 0 | 4.559918803 |

*: All the values are normalized with glucose uptake rates.

Curriculum Vitae

Personal Data

Name: MENG, Jin
Gender: Female
Marital status: Married
Nationality: P. R. China
Place of birth: Tianjin
Date of birth: 22.04.1979

Education

1985 – 1991 Tianjin Jingjianglu primary school, Tianjin, China
1991 – 1994 Tianjin No. 14 middle school (junior middle school), Tianjin, China
1994 – 1997 Tianjin No. 14 middle school (highschool), Tianjin, China
1997 – 2002 Clinical Medicine education in the Tianjin Medical University (China)
Degree: B Med (equivalent to the MD degree in the U.S. system)
2004 – 2007 Life Science master program in the Leibniz Hanover University
Degree: M.Sc.
2007 – 2011 PhD study in Biotechnology at Institute for Technical Chemistry of
Leibniz Hanover University
Dissertation: Investigation on the Application of Complex Cell
Models in the Simulation of Bioprocesses

Work Experiences

2002 – 2003 Assistant lector in the Medical University of Tianjin (China)