BLACK BOX KINETIC MODELING OF GROWTH AND CITRIC ACID PRODUCTION BY *CANDIDA OLEOPHILA* ATCC20177

by
Ezgi Tanıl

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
Biotechnology

Yeditepe University
2019

BLACK BOX KINETIC MODELING OF GROWTH AND CITRIC ACID PRODUCTION BY *CANDIDA OLEOPHILA* ATCC20177

APPROVED BY:

Assist. Prof. Dr. İlkem Emrah Nikerel ..... 
(Thesis Supervisor)

Assoc. Prof. Dr. Ali Özhan Aytekin ..... 

Assoc. Prof. Dr. Kazım Yalçın Arğa ..... 

DATE OF APPROVAL: ..../..../2019

ACKNOWLEDGEMENTS

I would like to thank to my supervisor Assist. Prof. Dr. Emrah Nikerel, who supported me and encouraged me to succeed, it was a great opportunity to work with him. I am very grateful for him guidance, corrections, tolerance and teaching how to work under stress and in a short time period. I would also like to thank to members of committee, Assoc. Prof. Dr. Ali Özhan Aytekin and Assoc. Prof Dr. Kazım Yalçın Arğa for the time that they devote to read and making comments on my thesis. I also would like to thank to Assoc. Prof. Dr. Mustafa Türker, who kindly contributed to the early phase of this thesis.

I would especially thank to my colleagues Huriye Yanık, Sheyda Shakoory, Burcu Şirin, Bahtir Hyseni, Özgür Yüksel, Kerem Kaya, Ülker Alkaya, Filiz Erçelik, Gülşah Akçadağ, Gizem Banger and Erhan Özer for their help and support during preparation of this thesis and making office hours more sufferable. I would also thank to my friends Ecenur Evran, Gülsu Eskiyenentürk, Umut Aydöner and Elif Şen for tolerating my “crazy” moments and supporting me in all circumstances.

Additionally, I would like to thank to Selma Çamurdan-Altaç for supporting me and being a role model as a successful businesswoman.

Finally, I would like to thank to all of the members of my big family, from my youngest cousin to my grandmother, for their endless support, understanding and tolerance. You are the most important source of my happiness.

ABSTRACT

BLACK BOX KINETIC MODELING OF GROWTH AND CITRIC ACID PRODUCTION BY *CANDIDA OLEOPHILA* ATCC20177

Mathematical modeling allows representation of biological systems, explanation of underlying mechanisms and prediction of system behavior. As such, it is one of the key focus points for both fundamental and applied research not only to improve our understading, but also to decrease costs by reducing the necessary experiments.

Organic acids are used in several industries, such as monomers for bioplastics, food preservatives and additives, pharmaceuticals, agriculture etc. Non-petrochemical, sustainable production of these acids is of great interest which can be achieved by fermentation of cheap substrates into various organic acids. The desired form of these typically weak acids is the protonated form (HA), to decrease the overall production and downstream costs. This form can be obtained by fermentation at the pH below the pKa of the corresponding acid. In production of organic acid, key is to determine growth and acid production dynamics, as the product itself has direct and indirect intibitory effects on the host.

The aim of this thesis, is to set up and analyze a mathematical model to study the dynamics of the growth and production of citric acid by building a black-box kinetic model using published batch fermentation data. Noncompetitive inhibition on substrate uptake reflected the system dynamics better than the competitive inhibition model. Monte Carlo study is also carried to find possible correlations among parameters. Glucose saturation constants were not correlated to their respective maximum specific growth rates. Parameters related to the energetics of growth and production are also determined as growth (K_x) and non growth associated (m_{ATP}) maintenance constants. The 90 per cent confidence intervals of K_x and m_{ATP} were [0.6, 4] and [2.5, 8] respectively for published P/O ratio of 1.45. Also at dilution rates lower than 0.07 h^{-1}, the obtained model reflected well the experimentally reported changes in chemostat.

ÖZET

CANDIDA OLEOPHILA ATCC20177 BÜYÜME VE SİTRİK ASİT ÜRETİM KİNETİĞİNİN KARA-KUTU TEKNİĞİYLE MODELLENMESİ

Matematiksel modelleme biyolojik sistemlerin betimlenmesine, temel sistemlerin açıklanmasına ve sistem davranışının öngörülmesine olanak sağlar. Bu sebeple, sadece anlayışımızı geliştirme amaçlı değil, ayrıca gereken deney miktarını azaltarak maliyeti düşürmek amaçlı kullanımı ile hem temel hem de uygulamalı bilimde, ana odak noktalarından biridir

Organik asitler monomer olarak biyoplastik, koruyucu ve katkı maddeleri olarak gıda, ilaç sanayii ve ziraat vb. gibi bir çok endüstri alanında kullanılır. Bu asitlerin ucuz substratların çeşitli organik asitlere fermentasyonuyla petrokimyasal olmayan sürdürülebilir üretimi ilgi çekmektedir. Genelde zayıf asit olan bu asitlerin, tüm üretim ve üretim sonrası maliyetleri düşürmesi sebebiyle, istenen formu proton kazanmış formudur. Bu form ilgili asidin pKa'sının altında bir pH değerinde üretim yapılarak elde edilir. Organik asit üretimindeki anahtar nokta, ürünün kendisinin konak hücre üzerinde doğrudan ve dolaylı inhibisyon etkisinin olmasından dolayı, büyüme ve asit üretim dinamiklerinin belirlenmesidir.

Bu tezin amacı, yayınlanımış kesikli fermentasyon verisi kullanılarak kurulan kara-kutu kinetik model ile büyüme ve sitrik asit üretiminin araştırılmasıdır. Substrat alımındaki yarışmasız inhibisyon, sistem dinamiklerini yarışmalı inhibisyondan daha iyi yansıtmıştır. Ayrıca parametreler arası ilişkileri belirlemek için Monte Carlo çalışması yapılmıştır. Glikoz doyum katsayılarının kendi spesifik maksimum hızlarıyla bağlantısı yoktur. Büyüme ve üretim enerjetiği, büyümeyle ilişkili (K_x) ve büyümeyle ilişkili olmayan (m_{ATP}) idame parametreleri ile belirlenmiştir. K_x ve m_{ATP} için yüzde 90 güven aralığı yayınlanan 1.45 P/O değeri için sırasıyla [0.6, 4] ve [2.5, 8] olarak bulunmuştur. Ayrıca elde edilen model, 0.07 h^{-1}'den düşük seyrelme hızı için kemostat verilerini de iyi bir şekilde yansıtmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

c	Number of parameters
C_{feed}	Concentration of a compound in the feed
C_i	Concentration of compound i
C_{to}	Metabolite concentration at t_0
C_{t1}	Metabolite concentration at t_1
D	Dilution rate
E_i	i^{th} element
E_{in}	Rate of input of an element
E_{out}	Rate of output of an element
F	Feed rate
h	Hour
I	Inhibitor concentration
J	Number of intracellular reactions
k	Number of compounds involved in reaction system
K_1	Monod constant for glucose consumption
K_2	Monod constant for growth (glucose)
K_3	Monod constant for growth (nitrogen)
K_4	Monod constant for citrate production
K_i	Inhibition constant
K_p	ATP required for product formation
K_x	Growth associated maintenance
L	Liter
m_{ATP}	Non-growth associated maintenance
m_p	Maintenance coefficient associated with product
m_s	Maintenance coefficient associated with substrate
n	Number of linearly independent equations

N	Nitrogen concentration
N_{feed}	Nitrogen concentration in the feed
P	General term for product concentration / Pressure
P^{max}	Maximum product concentration that growth rate becomes zero
q	Specific rate
q_{ATP}	Specific ATP production/consumption rate
q_{ByP}	Specific by-product production rate
q_{CA}	Specific citric acid production rate
q_{CA}^{max}	Maximum specific citric acid production rate
q_{CO2}	Specific carbon dioxide production rate
q_i	Specific production / consumption rate of compound i
q_N	Specific nitrogen consumption rate
q_{O2}	Specific oxygen production rate
q_s	Specific substrate consumption rate
q_S^{max}	Maximum specific substrate consumption rate
q_p	Specific product formation rate
Q	General term for gas evolution / consumption rate
R	Gas constant
S	General term for substrate concentration
S_0	Substrate concentration in the feed
t_0	Initial time point
t_1	Time point 1
T	Absolute temperature
v_{in}	Gas inflow rate
v^{max}	Maximum specific rate
v_{out}	Gas outflow rate
V	Culture volume
$x_{c,i}$	Carbon mass fraction of ith metabolite
x^{in}	Mole fraction of O_2 / CO_2 in gas input

x^{out}	Mole fraction of O_2 / CO_2 in gas output
X	General term for biomass concentration
X_{exp}	Experimental value
X_{model}	Predicted value from model
$\bar{X}$	Average biomass concentration
Y_{ps}	Product yield per unit mass of substrate consumed
Y_{xp}	Yield of biomass on product formed
Y_{xs}	Yield of biomass on substrate consumed
α_i	Stoichiometric coefficient of the i^{th} compound
γ	Degrees of reduction value
μ	Specific growth rate
μ^{max}	Maximum specific growth rate
σ	Standard deviation
ATP	Adenosine triphosphate
ByP	By-product concentration
CA	Citric acid production
CER	Carbon dioxide evolution rate
$C_xH_yO_z$	Generalized substrate formula
$CH_\alpha O_\gamma N_\delta P_\varepsilon$	Generalized biomass formula
$C_\alpha H_\beta O_\theta N_\delta P_\varepsilon$	Generalized compound formula
CO_2%	Percent mol fraction of carbon dioxide
Crec	Carbon recovery
df	Degrees of freedom
DNA	Deoxyribonucleic acid
g/L	Grams per liter
$g/gDWh^{-1}$	Grams per gram dry weight per hour
gDW/L	Gram dry weight per liter

Glc	Glucose concentration
Glc_{feed}	Glucose concentration
L atm/molK	Liter atm per mol per Kelvin
IQR	Interquartile range
mmol/mmol X h^{-1}	Milimoles per milimole biomass per hour
MCA	Metabolic control analysis
Mw	Molecular weight
O2%	Percent mol fraction of oxygen
ODE	Ordinary differential equations
OUR	Oxygen uptake rate
pKa	Acid dissociation constant in logarithmic scale
P/O	Phosphate oxygen ratio
RSD	Relative standard deviation
RQ	Respiratory quotient
SST	Sum of squares total
TCA	Tricarboxylic acid
UB	Upper bound

1. INTRODUCTION

1.1. MATHEMATICAL MODELING

Mathematical modeling gained attention for both industrial and research purposes since these models abstract and represent the biological systems and allow the simulation of different scenarios. In particularly dynamic models for metabolism delineate the metabolic mechanisms of production or consumption of the metabolite of interest; or structural mechanisms such as transport systems, describe the dynamics of cellular processes and points to strain improvement targets, for example, by determining the most important genes in production of the desired metabolite. They can be used in optimization, scale up and scheduling of fermentation process and to predict the future behavior of the system as a response to changing environmental conditions. These models decrease the need of experiments every time, therefore decrease the production cost. Especially mathematical models are the main focus of the biotechnological applications in areas ranging from molecular genetics to industrial fermentation processes [1].

1.1.1. Types of Models

Models are classified based on different criteria. Most of them can be classified as deterministic or stochastic according to the number of species involved in the system that will be modeled. Deterministic models are used for processes that include sufficient amount of species (>1000) such as metabolic networks, however stochastic models are constructed for processes that include low number of species such as gene expression regulation [2].

Models can be further classified based on nature of the events (discrete or continuos events in time), unstructured or structured, segregated or lumped, quantitative or qualitative etc. Figure 1.1 represents the commonly used types of models in literature [3].

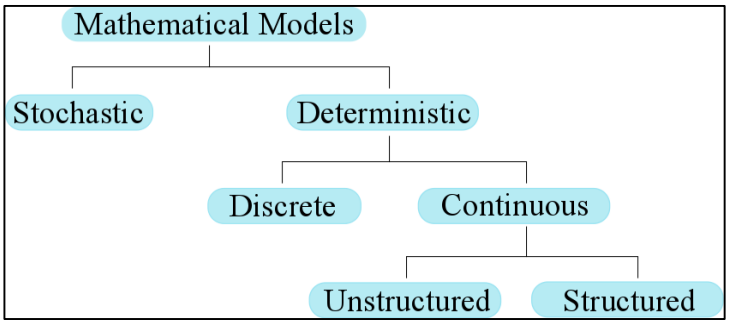

Figure 1.1. Types of mathematical models [2]

Stoichiometric models are used for determining the behavior of the cells under different environmental conditions, the differences between different mutants of an organism and predicting the effects of genes on phenotype of the organism. However, the main disadvantage of this method is, because of not having enough information about regulatory systems in the model, the predictive power of this method is limited [1]. If there is enough detailed information about certain cellular processes such as enzyme-catalyzed reactions, interactions of proteins with proteins and DNA, the dynamics of the process can be defined by using the combination of kinetics with known pathway stoichiometry. The set up of kinetic expressions to describe a metabolic process is called kinetic modeling [1].

1.1.2. Basics of Kinetic Modeling

Typical steps of mathematical modeling are summarized in Figure 1.2. First, the system boundaries and the level of complexity of the model is defined according to the aim of the study. For this purpose, the number of reactions and the corresponding stoichiometry are specified. According to the specified reactions, mass balances for each metabolite are constructed and formulated as ordinary differential equations. These mass balances will represent the time traces of metabolite concentrations over time. Next, kinetic rate expressions are defined for each reaction. Different types of kinetic expressions (e.g. Michaelis-Menten, ping pong, linlog kinetics) are considered as rate expressions and the one that satisfies the model is used. Finally, model parameters are estimated by fitting the model to available experimental data. The resulting model is used for analysis of the system e.g. to evaluate different scenarios, examining the effect of metabolite perturbations. Furthermore, by performing Metabolic Control Analysis (MCA) [4], the systemic properties (e.g.

bottlenecks of the system) can be defined. Moreover, bioprocess design involving reactor design and process optimization to improve yield and productivity can be performed. Also strain design to obtain improved strains with desired characteristics can be performed.

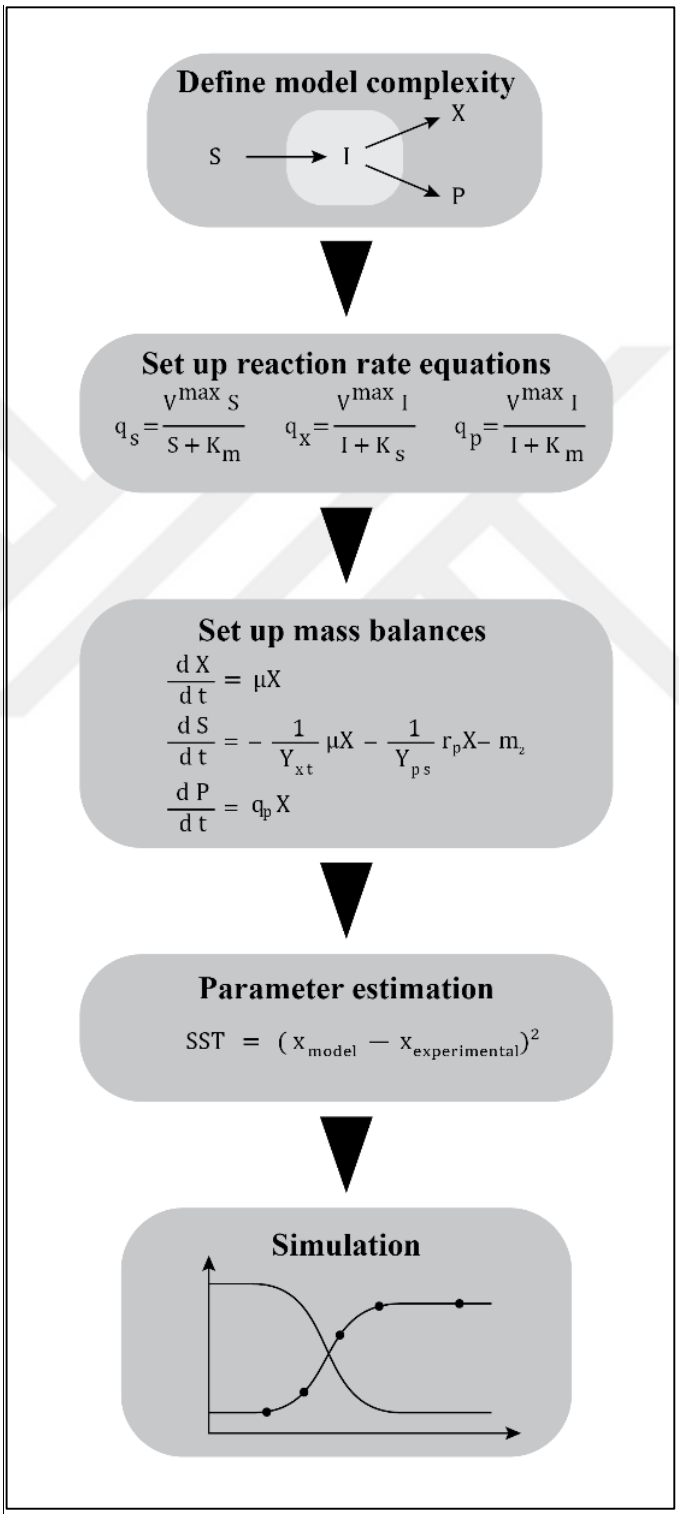

Figure 1.2. Basic steps of mathematical modeling

1.1.2.1. Definition of Model Complexity

The first step of the modeling is defining model complexity according to aim of the study. While modeling, it is important to consider that the model should be simple but not simpler than it is necessary to explain system properties. In this step, number of reactions that will be included in the model is determined. Every metabolite that is known or thought to influence the determined reactions should be included [5].

The models that are useful in bioprocess design are generally unstructured models [6]. These models ignore the diversity between cell forms or any internal structure of the cell. So these models cannot describe the system that is affected by cell composition, or the morphology of the cells [5]. In structured models, the biomass can be divided into different morphological forms and a set of intracellular reactions for defined intracellular components for each form is chosen. Basically, the model complexity increases with the increasing number of reactions and metabolites involved in the model. These complex structured models are useful for researches by providing detailed information about biological processes.

1.1.2.2. Black-box Kinetic Models

Black box kinetic models are the simplest form of unstructured models. In black box kinetic models, all cellular reactions are lumped into only one overall reaction. Experimental data is used to determine experimental yields and the yields per consumed substrate are assumed to be constant [5]. After specification of the complexity according to the aim of the study, the kinetic expressions for reactions involved in the model are described. For this purpose, reaction stoichiometry is determined. Microbial growth processes are described by determining the substrate uptake reactions for each substrate, intracellular reactions, and product excretion reactions for each product. For each reaction a reaction balance involving the conversion of substrates to cell mass and products is generated. The components of the balance are chosen according to the process, but the general form of the reaction balance for microbial growth is shown in Equation 1.1 **[6, 7]**.

$$\alpha_1\, C_x H_y O_z + \alpha_2\, O_2 + \alpha_3\, NH_3 + \alpha_4\, H_2PO_4 + \cdots \rightarrow \alpha_5\, CH_\beta O_\gamma N_\delta P_\varepsilon + \alpha_6\, CO_2 + \alpha_7\, H_2O \quad (1.1)$$

This balance shows the substrate ($C_xH_yO_z$) consumption with respiration using nitrogen (NH_3) phosphate (H_2PO_4) and potassium sources if present, to generate biomass ($CH_\alpha O_\gamma N_\delta P_\varepsilon$), carbon dioxide ($CO_2$) , and water ($H_2O$). To determine the stoichiometric coefficients, elemental mass balances are generated using each element used to form the reaction balance (C, H, O, N, P…) [7]. For each chemical species (E_i) , the elemental balances (rate of accumulation of each element) are expressed as the difference between rate of input of E_i and rate of output of E_i, is shown in Equation 1.2.

$$\frac{d(E_i)}{dt} = E_{in} - E_{out} \quad (1.2)$$

When more than one reaction occurs in a system the overall stoichiometric equation of the system can be written by a summation of each reaction. Since mass is conserved in chemical reactions, total amount of an element remains constant during process in a closed system. Also, at steady state, where the net accumulation rate is zero, each elemental balance must be equal to zero. Balances for each element are set as shown in Equation 1.3 [7, 8].

$$\sum_{i}^{k} \alpha_i \cdot E_i = 0 \quad (1.3)$$

α_i is the stoichiometric coefficient of the ith compound, E_i is the amount of element in i^{th} compound and k is the number of compounds involved in reaction system.

Using α values calculated from the elemental balances, the overall stoichiometry of the reaction can be determined. The elemental balance equation systems can be solved using linear algebra. To determine whether this system can or cannot be solved by linear algebra, one should determine the degrees of freedom. For n linearly independent equations containing c parameters, the degrees of freedom (df) is calculated using Equation 1.4.

$$df = n - c \quad (1.4)$$

The value of degrees of freedom shows the amount of information needed to solve the system. This information includes resulting experimental yields and/or theoretical

information. Generally the value of an unknown coefficient is fixed to 1 and according to the calculated experimental yields the coefficients are found. For instance, if the substrate coefficient is fixed to 1, and the stoichiometric coefficient of biomass can be calculated using biomass yield per unit mass of substrate consumed as shown in Equation 1.5. The α value in Equation 1.5 is yield coefficient expressed in mol/mol basis [7].

$$Y_{xs} = \frac{mass\ of\ biomass\ produced}{mass\ of substrate\ consumed} = \frac{\alpha \cdot Mw(biomass)}{Mw(substrate)} \tag{1.5}$$

where Y_{xs} is the biomass yield per unit mass of substrate consumed and Mw is the molecular weight of the related molecule.

1.1.2.3. Mass, Degrees of Reduction and ATP Balances

Mass, degrees of reduction and energy balances are commonly used in bioprocess engineering to determine the overall success of the fermentation and quality of the data since they provide preliminary information about the system behaviour. Therefore, these balances are used before the modeling to get a better understanding about the system and to decide how the model outputs should be [9].

Most commonly used elemental mass balance is carbon balance. Since the carbon in the substrate is one of the most important sources of growth, product formation and carbon dioxide, carbon balance and carbon recovery provides important information about the system. Using carbon content of substrate consumed in a given time per cell (q_s), biomass formed (μ) and products formed (q_p and q_{CO2}), carbon balance can be set. For a fermentation process, carbon balance for k carbon containing metabolites can be set as Equation 1.6 (without including stoichiometric coefficients).

$$\frac{dC}{dt} = \sum_{i=1}^{k} (x_{C,i}) q_i \tag{1.6}$$

Where $x_{c,i}$ is the carbon mass fraction of i^{th} metabolite and q_i is the specific consumption or production rate of i^{th} metabolite. To apply this balance, specific rates can be calculated from experimental data using Equation 1.7.

$$q = \frac{C_{t_1} - C_{t_0}}{(t_1 - t_0) \cdot \bar{X}} \tag{1.7}$$

Where q is the specific rate, C_{t0} and C_{t1} are the metabolite concentrations (g/L) in t_0 and t_1 time points, respectively and $\bar{X}$ is the average biomass concentration (gDW/L) between t_0 and t_1. Also per cent carbon recovery that defines how much of the provided carbon is used can be calculated using:

$$Crec\ (\%) = \frac{\sum Carbon\ consumed}{\sum Carbon\ provided} \cdot 100 \tag{1.8}$$

Degrees of reduction is used to define proton-electron balance of the reaction. In bioprocess perspective, degree of reduction can be defined as the number of moles of available equivalents per mol or per C-mol of each compound. Degrees of freedom of most common elements used in bioprocesses are -3, -2, 1, 4 and 5 for N,O, H, C and P [10]. If the elemental composition of the compound in interest is known degrees of reduction (γ) of that compound with a elemental composition of $C_\alpha H_\beta O_\theta N_\delta P_\varepsilon$ using Equation 1.9.

$$\gamma = 4\alpha + \beta - 2\theta - 3\delta + 5\varepsilon \tag{1.9}$$

Therefore, the degrees of reduction balance of the system containing k metabolites at a given time point can be set as shown in Equation 1.10.

$$0 = \sum_{i=}^{k} \gamma_i \cdot q_i \tag{1.10}$$

Finally to define the effects of used substrate, growth and production on ATP production, and give a better insight about energy metabolism of the organism under certain conditions, ATP balance was set. The difference between ATP producing and consuming reactions is used to express change in ATP concentration in time, similar to elemental and degrees of reduction balances [11-13]. The typical form of ATP balance is shown in Equation 1.11.

$$0 = \frac{2P}{O} \cdot q_{O_2} + \sum q_i^{ATP} + m_{ATP} + K_x \cdot \mu + K_p \cdot q_p \tag{1.11}$$

First term defines the ATP production from supplied oxygen where P/O ratio is the ATP production per unit oxygen provided and q_{O2} is the specific oxygen uptake rate. Second term defines the sum of specific ATP production rates per unit substrate consumed. Therefore the first term repsresents ATP formation from oxidative phosphorylation and second term represents ATP formation from reactions with known stoichiometry including substrate level phosphorylation. m_{ATP} is the non-growth associated maintenance coefficient that represents ATP consumed per unit biomass per unit time. Fourth term shows the consumed ATP as a growth requirement where K_x is the growth associated maintenance (ATP produced per unit biomass formed). Finally, K_pq_p is the ATP consumed for production of other metabolites of interest [13, 14]. The parameters of ATP balance (P/O, m_{ATP}, K_x, K_p) can be theoretical, obtained from previous researches of close-related organisms, or can be measured experimentally [15]. One of the estimation methods for m_{ATP} and K_x is explained by Kliphius et al. in 2011. In this method, experimental specific growth rate is plotted against specific ATP production rate. Slope of the linear trendline of the data represents K_x and y-intercept of the trendline is m_{ATP} as shown in Figure 1.3 [16].

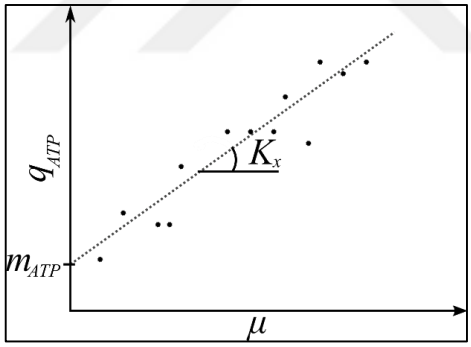

Figure 1.3. Regression line of μ and q_{ATP} (ATP production rate: represents the total ATP production, i.e. sum of first two terms that produce ATP from oxidative phosphorylation and from substrate level phosphorylation)

1.1.2.4. Reaction Rates

According to the determined reaction stoichiometries, reaction rates can be described [17]. Generally in, most basic form of black box models, all cellular reactions are lumped into only one overall reaction and the volumetric substrate utilization and the product formation

are described as a linear function of specific cell growth rate (μ). It is assumed that yield coefficients (Y_{xs}, Y_{xp}) and maintenance coefficients (m_s, m_p) are independent of substrate concentration and composition of cells. The specific cell growth rate is usually given in form of Monod kinetics [6].

$$\mu = \frac{\mu^{max} \cdot S}{S + K_s} \tag{1.12}$$

$$\frac{dS}{dt} = Y_{xs} \cdot \mu + m_s \tag{1.13}$$

$$\frac{\mathrm{dP}}{dt} = Y_{xp} \cdot \mu + m_p \tag{1.14}$$

Here, the growth rate (μ) is expressed as a function of the substrate (S) which is generally the carbon source used for growth. The substrate consumption and product formation are proportional with the growth rate. μ^{max} is the specific growth rate that shows the maximum value of the growth rate reached when all of the substrates present in the excess (nothing limits the growth). K_s is the Monod constant (aslo called as saturation constant) which is the substrate concentration when the reaction rate is the half of its maximum value. Typically, K_s is much smaller than the initial substrate concentration and reaction rate approaches to μ^{max}. In time substrate concentration decreases as biomass formed and when it is equal to K_s, reaction rate becomes half of the maximum rate [5]. In Table 1.1 most commonly used growth rate equations were listed. The model that fits the experimental data can be chosen.

Table 1.1. Various (microbial) growth models (μ_{max} : maximum specific growth rate, S: substrate concentration, K_s: Monod constant, half saturation constant, K_i: inhibition constant, P: product concentration, P_{max}: maximum product concentration at which specific growth rate becomes zero) [18]

Model	**Equation**
Monod	$\frac{\mu_{max}S}{S+K_s}$
Jerusalimsky	$\frac{\mu_{max}S}{S+K_s}\left(\frac{K_i}{K_i+P}\right)$
Dagley	$\frac{\mu_{max}S}{S+K_s}(1-K_iS)$
Levenspiel	$\frac{\mu_{max}S}{S+K_s}\left(1-\frac{P}{P_{max}}\right)^n$
Contois	$\frac{\mu_{max}S}{S+K_sX}$
Moser	$\frac{\mu_{max}S^n}{S^n+K_s}$
Haldane	$\frac{\mu_{max}S}{K_s+S+S^2/K_i}$

Although substrate uptake rate was commonly expressed based on experimental yields, production and growth rates and maintenance in unstructured models [19, 20], the rates can be expressed using different extended versions of Monod model, similar to growth models. By performing this, specific substrate uptake rate can be obtained independently.

Like substrate uptake rate, specific product formation rates can be defined independently or as a function of growth rate. Additionally, model published by Luedeking and Piret in 1959 for lactic acid production [21], can be used to define organic acid production.

$$\frac{dP}{dt} = \alpha\frac{dX}{dt} + \beta X \tag{1.15}$$

Where α and β are coefficients and X is the biomass concentration. This model can be used to define growth related production.

1.1.2.5. Mass Balances

Using these rate equations and experimental yields, mass balances for each metabolite will be set to define the time profile of that metabolite [1]. Using the laws of thermodynamics, change in concentration of a compound in a system is defined as:

$$\frac{dC}{dt} = C_{in} - C_{out} + C_{generated} - C_{consumed} \tag{1.16}$$

Where C_{in} and C_{out} are the concentrations of compound C enters to the system and leaves the system. $C_{generated}$ is the concentration of compound C produced in the system and $C_{consumed}$ is the amount of compound C consumed in the system.

Defined rate equations are used to describe the relationship between metabolite concentrations and reaction rates. For this purpose, dynamic mass balances are specified. These mass balances are used to calculate the change in concentration of metabolite or biomass in time. These expressions are set using the experimental yields or stoichiometric coefficients of the metabolite in each reaction balance. General forms of these mass balances for batch fermentations are [20]:

$$\frac{dX}{dt} = \mu X \tag{1.17}$$

$$\frac{dS}{dt} = -\frac{1}{Y_{xs}}\mu X - \frac{1}{Y_{ps}} q_p X - m_s \tag{1.18}$$

where μ is the specific growth rate, X is the biomass concentration, Y_{xs} is the biomass yield per unit mass of substrate consumed, Y_{ps} is the product yield per unit mass of substrate consumed, q_p is the specific product formation rate and m_s is the maintenance coefficient. For each metabolite involved in the model, the mass balances are generated; metabolite concentrations, reaction rates of the reactions that affect the metabolite concentration with a positive value for reactions producing the metabolite and negative sign for the reactions consuming the metabolite [2].

For chemostat, the dilution caused by feed should be included to the mass balances. So the mass balances take the form of [22]:

$$\frac{dC}{dt} = D\left(C_{feed} - C_i\right) + q_C X \tag{1.19}$$

where D is the dilution rate, C_{feed} is the concentration of compound C in feed, C_i is the concentration of compound C in the system, q_C is specific production or consumption rate of C and X is the biomass concentration. In chemostat, specific growth rate is equal to the dilution rate which is F/V. Specific reactions for change in biomass concentration and substrate concentration are:

$$\frac{dX}{dt} = -\frac{F}{V}X + \mu X \tag{1.20}$$

$$\frac{dS}{dt} = \frac{F}{V}S_o - \frac{1}{Y_{xs}}\mu X - \frac{1}{Y_{ps}}q_p X - m_s \tag{1.21}$$

where S_0 is the substrate concentration of feed, F is the feed rate and V is the culture volume.

1.1.2.6. Parameter Estimation

The final step of model construction is the estimation of parameters [5]. The equations in the model contain parameters such as maximum reaction rates, inhibition constants etc. which must be estimated to simulate the system. One of the main challenges of the kinetic modeling is estimation of the parameters. Because of the experimental noise and systematic bias that is introduced due to the in vitro measurements, it is hard to directly measure and estimate these parameters. Although numerous methods have been developed in recent years, formulating and fitting kinetic models to experimental data is still a challenge.

The parameter estimation will be performed by an optimization algorithm that fits the function to experimental data by using an objective function that minimizes total sum of squares (SST) of differences between predicted (X_{model}) and experimental (X_{exp}) values for *m* metabolites at *n* time points [23]:

$$SST = \sum_{i=1}^{i=n}\sum_{j=1}^{j=m}\left(\frac{X_{exp,ij} - X_{model,ij}}{\sigma_{ij}}\right)^2 \tag{1.22}$$

To use this method, an initial set of parameters needs to be provided and the squared residual (squared difference between experimental value of a variable and the fitted value of the same variable) is minimized iteratively until it reaches a minimum value, typically using greedy algorithms. The function is repeated iteratively using different values for each parameter at each repeat, starting from initially assigned values at the beginning. At the end, the values for each parameter that gives minimum SST value can be obtained. After finding the parameters, the fermentation process is simulated to observe results and to modify the model structure if necessary [5, 23].

1.1.3. Challenges in Modeling

Mathematical modeling has major challenges especially due to the inherent complexity and nonlinearity of the biological systems. Many different types of metabolites are present even in the simplest organism so, defining every single reaction involving each metabolite or interactions between metabolites is challenging. Additionally, the experimental information about the model parameters or theoretical information such as rate equations are either missing or incomplete. Recent advances in “omic” technologies provide high-throughput data in different levels, but the dynamics and interplay within heterogeneous experimental data is often hard to grasp. Besides the unavailability of the required information from different fields, algorithms that are necessary to solve the ODE systems or non-linear squared error problem that is used for optimization of parameters are needed [24].

1.1.4. Monte Carlo Simulations

Monte Carlo method was first proposed in 1945 as a sampling tool for statistical applications [25]. It was commonly used in sampling of random artificially generated data, estimation model parameters and optimization of complex, nonlinear objective functions. Since it is easy to use, provides randomness for computations in deterministic systems, enables improving insight about the system behavior, it has wide usage area from computational biology to modeling telecommunication systems [26]. While modeling biological systems, data where randomly generated error with a pre-defined mean and standard deviation can be

used to estimate parameters for each data set. Using the parameter distributions, one can gain more understanding about system's sensitivity on changes in parameters [27].

1.2. CITRIC ACID

As a tricarboxylic acid, citric acid ($C_6H_8O_7$) is a common metabolite of animals and plants along with citrus fruits and pineapple [28]. Citric acid has been produced industrially since the beginning of 20th century. It is one of the most exploited biochemical product and one of the few bulk chemicals produced by fermentation. Citric acid is used in the household, in the preparation of industrial products and in the food, pharmaceutical and chemical industries and as well as cleaning and complexing agent [29]. In 2010, the consumption of citric acid along with its salt, trisodium citrate has reached 1 400 000 tones with 5 per cent growth per year [30].

Pure citric acid is colorless and readily soluble in water with 210.13 g/mol molecular weight. It is safe, versatile, biodegradable and ecofriendly chemical for buffering, cleaning, wetting and dispersing [28].

Considering its three carboxylic acid functional groups, citric acid has three pKa values at pH 3.1, 4.7, and 6.4. The traces of citric acid can be found in virtually all animals and plants as being the universal intermediate product of metabolism [31].

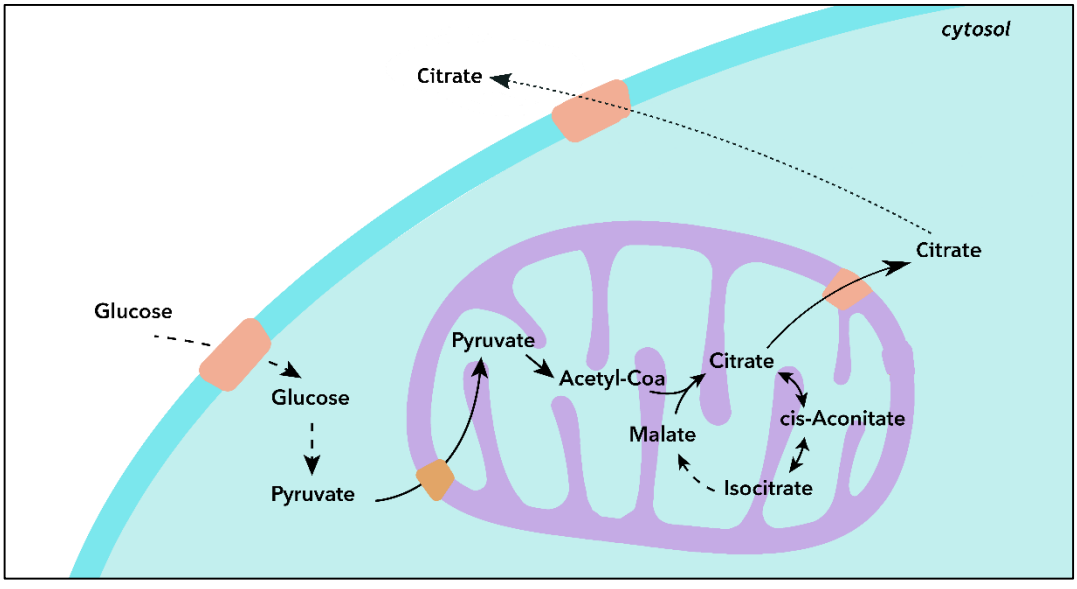

Figure 1.4. Citrate synthesis pathway

The constituents of many aerobic and anaerobic microorganisms are the products of carbohydrate metabolism with citric acid being the main intermediate. Certain microorganisms produce citric acid as an over flow product due to flawed operation of the tricarboxylic acid cycle (TCA) under specific environmental conditions. TCA is an intermediate cycle comprised with terminal steps in the conversion of carbohydrates, fats and proteins to carbon dioxide and water with accompanying release of energy for the growth. Formation of citric acid is a solely enzymatic process in biological systems. Competence of citric acid fermentation depends on the regulation of each one of the enzymes in TCA and its activity under various control mechanisms such as cofactors. Metal ions are the components of the cofactors, controlling concentration of the trace element can regulate the enzyme activity accordingly [28].

1.3. FERMENTATION

Fermentation process involves utilization of microorganisms to convert substrates to valuable products. According to area, various types of substrates can be used and products can be obtained. Most common fermentation products are cheese, soy sauce, wine, organic acids such as citric acid etc [32].

Considering the type of organism that is used in production or product that is produced, fermentation operation modes may differ. The common operation modes are batch, continuous and fed-batch operation modes [5, 10]. Batch fermentation is the most basic example. In batch fermentation systems, all of the required nutrients are supplied at the beginning of the process and inoculation is done after medium sterilization. Fermentation can be stopped according to process aim for example, when the desired final growth rate or final product concentration is reached. This operation mode is used in wine making [33] vitamin [34] and amino acid production [35].

In fed-batch mode, growth begins in a batch culture, but at a specific time point substrate or other input such as amino acids are given to the system. This can be done with or without the removal of the used media. One of the most common examples of products produced using fed-batch mode is penicillin production using *Penicillium chrysogenum* [36].

Finally, in continuous mode, exponantial growth phase is extended to maintain stable environmental conditions (steady state) by feeding the organisms with fresh media and removing used media at the same time. Different control techniques used in continuous cultures such as chemostat (substrate-limited growth), turbidostat (unlimited growth with excess nutrients) and auxostat (a variable such as pH is controlled by feed rate) etc [37].

1.4. CITRIC ACID FERMENTATION IN LITERATURE

Previously, citric acid production was performed by extraction from lemon juice. It can also be extracted from different fruits such as raspberry, tomato and pineapple. Due to the advances in microbial fermentation technologies, these procedures were replaced by microbial fermentation [38].

Different microorganisms including bacteria, yeasts and other fungi were used to produce citric acid. Some of these organisms are *Aspergillus* spp. *Bacillus licheniformis, Corynebacterium* spp.,*Penicillium janthinellum, Candida* spp., and *Yarrowia lipolytica*. The commercial production of citric acid mainly depends on *Aspergillus niger* [39].

Conventionally, *Aspergillus niger* is used to produce citric acid. However the filamentous structure of *Aspergillus* species is problematic for efficient production since it decreases oxygen transfer, causes sensitivity to hydromechanical stress and it makes the process more laborious because of increased viscosity and the presence of impurities[40]. To overcome these problems, production of citric by different yeast strains are proposed and *Candida oleophila* ATCC20177 strain shown to be a good candidate for being a citric acid producer. Its unicellular growth eliminates the disadvantages of filamentous growth and makes the organism useful for production of citric acid using different carbon sources and different fermentation types, especially for continuous processes [29, 41-43].

1.5. EFFECTS OF pH ON ORGANIC ACID FERMENTATION

As is known to all, microorganisms were evolved to survive efficiently in a range around their optimum pH values. They can be acidophiles, alkaliphiles or neutrophiles that prefer acidic, alkali or neutral pH values, however, although these ranges can be wide for yeasts,

to obtain the most efficient production in the process, pH should be close to optimum because cells should maintain their intracellular pH for better growth and production [44].

The low extracellular pH resulted from weak acids causes weak acid stress on cells. The low pH affects the membrane integrity of the cells by changing conformations of membrane proteins and lipids. This conformational changes alter the electrochemical potential of the membrane, making the membrane more permeable to ions and other small metabolites like the weak acid itself. When intracellular environment is acidified, the excess H^+ should be pumped out using ATP to avoid enzymatic activity changes. Also the pumped protons causes undissociated acid formation in extracellular environment, creating more stress. In previous researches, low extracellular pH during fermentation processes affects both fermentation rate and growth of the organism by increasing fermentation time due to delayed growth and possibly decreased production due to unnecessary ATP consumption [45].

Besides the effects of pH during fermentation, it may also affect the downstream processes to recover itaconic acid from media. Crystallization is the most basic method for itaconic acid separation [46]. On the other hand, the protonated form of organic acids are desirable because they can easily separated using methods that separates molecules based on ionic strength. In a media with a pH lower than the acid's pKa value, organic acids become protonated and have higher efficiency to ionic phases of separation methods. Therefore it will be adventageous to carry out the process in a pH lower than the pKa of the produced acid, since the most expensive and challenging step of bioprocess are separation and purification steps [47].

1.6. PRODUCT INHIBITION

Accumulation of organic acids inhibits fermentation by causing weak acid stress. Since the electrochemical potential of the cell membrane is affected from acid accumulation, the uptake of the provided substrates can also be affected. Since substrate uptake and transport through membranes are two of the possible bottlenecks of the acid fermentation processes, decreasing the acid accumulation to tolerable levels or constructing strains that have higher tolerance to low pH is crucial [48].

Although Michaelis-Menten kinetics are proposed to define enzyme kinetics, the defined inhibition types can be applied to express the kinetics at organism level in Monod models. 3 different types of inhibitions are reported, competitive, uncompetitive and noncompetitive inhibitions [49]. The models were shown in Table 1.2.

Table 1.2. Models for different inhibiton types and changes in corresponding parameters [49].

	Equation	**K_m**	**v^{max}**
Competitive	$\frac{v^{max} \cdot S}{S + K_m \cdot \left(1 + \frac{I}{K_I}\right)}$	Increases	No change
Uncompetitive	$\frac{v^{max} \cdot S}{S \cdot \left(1 + \frac{I}{K_I}\right) + K_m}$	Decreases	Decreases
Noncompetitive	$\frac{v^{max} \cdot S}{S + K_m} \cdot \left(1 + \frac{I}{K_I}\right)$	No change	Decreases

In the presence of competitive inhibition, both inhibitor (I) and substrate of the enzyme competes to bind to the active site. When compared to normal state, K_m increases but the inhibition does not affect v^{max}. In uncompetitive inhibition, the inhibitor binds to enzyme-substrate complex, which causes decrease in both v^{max} and K_m. Finally in noncompetitive inhibition, the inhibitor affects the enzyme activity by binding a certain site other than the active site, causing conformational changes in active site that prevents the binding of substrate. In this case, K_m is not affected, however v^{max} decreases [49, 50].

1.7. AIM OF THE THESIS

The main purpose of this thesis is constructing black-box kinetic model of growth and citric acid production of *Candida oleophila* ATCC20177 strain fermentation data published by Anastassiadis *et al.* [29, 41-43].

For this purpose, data was used to calculate specific production and consumption rates, carbon, degrees of reduction and ATP balances, carbon recoveries and respiratory quotients. The batch data was modeled including growth, substrate consumption, citric acid and byproduct formation using Monod kinetics. Then Monte Carlo simulations were performed to determine effects of uncertainties on model parameters and correlations between maximum specific rates and saturation constants. Next growth associated and non-growth associated maintenance parameters were determined from experimental μ and q_{ATP}. Finally this model parameters were tested on continuous data and effects of pH on final titer of the products, biomass and substrate was examined.

2. MATERIALS AND METHOD

2.1. ORGANISM AND DATA

Fermentation data of *Candida olephila* ATCC 20177 strain published by Anastassiadis *et al.* [42] containing time (h), glucose concentration (g/L), biomass concentration (gDW/L) and citrate concentration (g/L) was used to model substrate uptake, growth and citrate production kinetics. The data used for modeling was from batch fermentation, performed on 2 L working volume bioreactor and pH was controlled at 4.5. The media contains ~400 g/L glucose and 6 g/L NH_4Cl. To test the model on continuos process, data published by Anastassiadis *et al.* in 2005 [41, 43] was used. Calculations using experimental data were performed using Microsoft Excel and modeling and parameter estimation were perfomed on MATLAB.

2.2. CALCULATIONS

2.2.1. Calculation of OUR and CER using Off-gas Data

To calculate oxygen and carbon dioxide concentrations, ideal gas law was used. Following equation was used to calculate oxygen uptake rate and carbon dioxide consumption rates using off-gas data:

$$Q\left(\frac{g}{h}\right) = \frac{\%\, fraction}{100} \cdot v_{out} \cdot \frac{P}{RT} \cdot Mw \tag{2.1}$$

where v_{out} is the air flow that exits the system, P is the total pressure in the system, R is the gas constant (0.082 L atm/molK) and Mw is the molecular weight of the compound. Then, concentration of oxygen consumed or carbon dioxide produced (C) between time points t_1 and t_0 were calculated in V (L) system volume using:

$$C\left(\frac{g}{L}\right) = \frac{Q \cdot (t_1 - t_0)}{V} \tag{2.2}$$

Since N_2 is an inert gas, its number of moles remains constant in the system. It was assumed that air consists of N_2, CO_2 and O_2. Therefore, when the mole fractions of O_2 and CO_2 and gas inflow rate (v_{in}) were known, for constant temperature and pressure, gas outflow rate (v_{out}) can be calculated using:

$$v_{in}\left(1 - x_{O2}^{in} - x_{CO2}^{in}\right) = v_{out}\left(1 - x_{O2}^{out} - x_{CO2}^{out}\right) \quad (2.3)$$

Where x_{O2}^{in} x_{CO2}^{in} and x_{O2}^{out} and x_{CO2}^{out} are the mole fractions of O_2 and CO_2 in gas input and output respectively. Since the difference between concentrations of each compound in inlet and outlet gases is equal to the consumed or produced amount of gas in the time period, the oxygen uptake rates (OUR) and carbon dioxide evolution rates (CER) for each time point can be obtained from:

$$OUR = \frac{P}{RTV}\left(v_{in}x_{O2}^{in} - v_{out}x_{O2}^{out}\right) \quad (2.4)$$

$$CER = \frac{P}{RTV}\left(v_{in}x_{CO2}^{in} - v_{out}x_{CO2}^{out}\right) \quad (2.5)$$

2.2.2. Calculation of Specific Rates

Specific substrate consumption (q_s: g/gDWh^{-1}), growth (μ: h^{-1}) and citrate production rates (q_p: g/gDWh^{-1}) were calculated from experimental data. For glucose (Glc), biomass (X) and citric acid (CA) concentrations (g/gDW h^{-1}) at time t_1 and t_2 specific rates were calculated using:

$$q_s = \left(\frac{Glc_{t_1} - Glc_{t_2}}{t_2 - t_1}\right)\left(\frac{1}{\frac{X_{t_1} + X_{t_2}}{2}}\right) \quad (2.6)$$

$$\mu = \left(\frac{X_{t_1} - X_{t_2}}{t_2 - t_1}\right)\left(\frac{1}{\frac{X_{t_1} + X_{t_2}}{2}}\right) \quad (2.7)$$

$$q_p = \left(\frac{CA_{t_1} - CA_{t_2}}{t_2 - t_1}\right)\left(\frac{1}{\frac{X_{t_1} + X_{t_2}}{2}}\right) \quad (2.8)$$

2.2.3. Balances of Carbon and Degrees of Reduction

When there was no additional products are formed, the amount of carbon present in the system was provided by glucose and distributed among biomass, itaconic acid and carbon dioxide. Considering specific substrate uptake, growth, citrate formation and carbon dioxide production rates carbon balance was set.

$$\left(\frac{72}{180}q_s - \frac{12}{20.6}\mu - \frac{72}{192}q_p - \frac{12}{44}q_{CO2}\right) = 0 \quad (2.9)$$

Additionally, carbon recovery for each time point was calculated using:

$$Crec(\%) = \frac{\sum C_{consumed}}{\sum C_{supplied}} \quad (2.10)$$

Similarly, electrons provided by oxygen were assumed to be shared between glucose, biomass and products. Since known degrees of reduction for glucose, O_2, biomass and products are 24, -4, 5.27 and 18, respectively, the degrees of reduction balance was set as:

$$\frac{24}{180}q_s - \frac{4}{32}q_{O_2} - \frac{4.96}{20.26}\mu - \frac{18}{192}q_{CA} = 0 \quad (2.11)$$

2.2.4. Calculation of CER from Growth, Production and Respiration

The growth, citrate and byproduct formation and respiration stoichiometries were defined considering published biomass composition $CH_{1.86}\ O_{0.33}\ N_{0.14}$ [29]. Molar CO_2 yields on biomass, products and oxygen were used to determine the CO_2 production rates from each reaction using:

$$CER_x = \frac{\alpha_{CO2}}{\alpha_x} q_x \quad (2.12)$$

Where α_x and q_x are stoichiometric coefficient of biomass, product or oxygen and their specific rates and α_{CO2} is the stoichiometric coefficient of CO_2 in production or consumption reaction of species x. The ratios were:

$$CER^{\mu} = 0.59\mu \quad (2.13)$$

$$CER^{P} = q_P \quad (2.14)$$

$$CER^{O_2} = q_{O_2} \quad (2.15)$$

By calculating CER and OUR, respiratory quotients (RQ) for each time point were calculated using molar CER and OUR.

$$RQ = \frac{CER}{OUR} \quad (2.16)$$

2.2.5. Calculation of K_x and m_{ATP} from Specific Rates

Growth associated and non-growth associated maintenance coefficients, K_x (mmol ATP/mmol X) and m_{ATP} (mmol ATP/mmol X h^{-1}), can be obtained μ vs $2q_s + 2P/Oq_{O2}$ plot. The P/O ratio, obtained from published P/O ratio for oxidation of mitochondrial NADH of *Candida utilis* was 1.45. However, since these values are only estimations, the P/O ratio was changed in range between 1 and 2.5 chosen based on published results for yeasts in literature. Linear functions were fitted to data considering which obtained from experimental measurements using MATLAB [51]. The slopes of the trendline represent K_x and the y-intersect of these trendline represent m_{ATP}.

2.3. KINETIC MODELING

2.3.1. Model Scheme & Reaction Rates

The black box system that was modeled was shown in Figure 2.1. In this system, glucose oxygen and nitrogen enters the system and it is used for growth and citrate and by-product production. These by-products were considered as citrate equivalents. Also, carbon dioxide is produced in growth and product formation reactions.

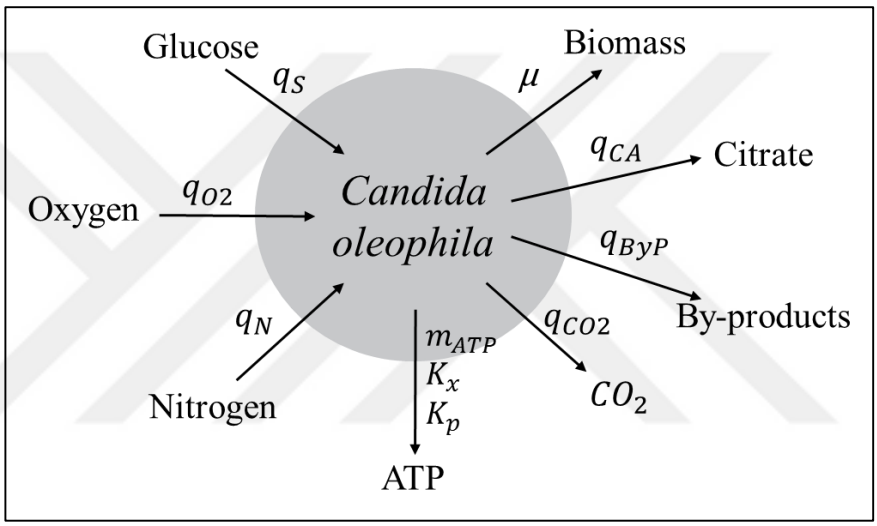

Figure 2.1. Modeled black-box system

In this system, q_S is the specific substrate uptake rate (g/gDWh^{-1}), μ is the specific growth rate (h^{-1}) and q_{CA} is the specific citrate production rate (g/gDWh^{-1}), q_{ByP} is the specific by-product (citrate equivalents) production rate (g/gDWh^{-1}), q_{CO2} is the specific CO_2 production rate (g/gDWh^{-1}), q_{O2} is the specific oxygen uptake rate (g/gDWh^{-1}) and q_N is the specific nitrogen consumption rate (g/gDWh^{-1}). To express the kinetics of these reactions, Monod model was used. The Monod kinetics with noncompetitive and competitive product inhibition were used for specific substrate uptake rate separately. Also Monod kinetics was used to represent citrate production rate, where q_S^{max} and q_{CA}^{max} are the maximum reaction rates for substrate uptake and citrate production, respectively, K_1 and K_4 are the saturation constants and K_i is the inhibition constant for substrate consumption. Glc is the glucose concentration, CA is the citrate concentration and ByP is the byproduct concentration. Since

it was known that nitrogen concentration highly influences growth, specific growth rate was expressed with Monod model with two substrates, intracellular intermediate and nitrogen using Equation 2.13 where μ^{max} is the maximum specific growth rate, K_2 and K_3 are the saturation constants, and N is the nitrogen concentration.

$$q_s = \frac{q_S^{max} \cdot Glc}{K_1 + Glc}\left(\frac{1}{1 + \frac{CA + ByP}{K_i}}\right) \tag{2.17}$$

$$q_s = \frac{q_S^{max} \cdot Glc}{K_1 \cdot \ 1 + \frac{CA + ByP}{K_i} + Glc} \tag{2.18}$$

$$\mu = \ \mu^{max} \cdot \frac{Glc}{Glc + \ K_2} \cdot \frac{N}{N + \ K_3} \tag{2.19}$$

$$q_p = \frac{q_p^{max} \cdot Glc}{K_4 + Glc} \tag{2.20}$$

Specific nitrogen consumption rate, q_N (g/gDWh^{-1}) was expressed using consumed nitrogen per g biomass from experimental data, which is 0.03. Therefore q_N becomes:

$$q_N = 0.03\mu \tag{2.21}$$

It was assumed that citrate-equivalent by-products/ citrate percent was 20 per cent based on experimental date. So q_{ByP} becomes:

$$q_{ByP} = 0.2\mu \tag{2.22}$$

To express specific oxygen consumption rate, q_{O2} (g/gDWh^{-1}), degrees of reduction balance was used by assuming there was no other product formation occurs in the system. Using same assumption, carbon dioxide evolution rate was set as carbon balance, since the remaining carbon of glucose was assumed to end up forming carbon dioxide.

$$q_{O2} = \left(-\frac{24}{180}\ q_s + \ \frac{4.96}{20.26}\ \mu + \ \frac{18}{192}\ q_p\right)\left(\frac{32}{4}\right) \tag{2.23}$$

$$q_{CO2} = \left(\frac{44}{12}\right)\left(\frac{72}{180}q_s - \frac{12}{20.26}\,\mu - \frac{72}{192}\,q_p\right) \quad (2.24)$$

2.3.2. Mass Balances

After setting the rate expressions, ordinary differential equations as mass balances were set for all species included to model (glucose (Glc), biomass (X), citrate (CA), by-products (ByP), oxygen (O_2), carbon dioxide (CO_2), nitrogen (N)), according to the known stoichiometry of the system. The final mass balance equations are:

$$\frac{dGlc}{dt} = -q_s X \quad (2.25)$$

$$\frac{dX}{dt} = \mu X \quad (2.26)$$

$$\frac{dCA}{dt} = q_{CA} X \quad (2.27)$$

$$\frac{dByP}{dt} = q_{ByP} X \quad (2.28)$$

$$\frac{dN}{dt} = -q_N X \quad (2.29)$$

$$\frac{dCO_2}{dt} = q_{CO_2} X \quad (2.30)$$

$$\frac{dO_2}{dt} = q_{O_2} X \quad (2.31)$$

2.3.3. Chemostat Model

The model parameters were tested at chemostat data to determine the effects of pH, dilution rate and air saturation on final titers of substrate, biomass and products. The ratio of byproducts on citrate was decreased to 10 per cent according to the experimental data. The updated mass balances are:

$$\frac{dGlc}{dt} = D(Glc_{feed} - Glc) - q_s X \quad (2.32)$$

$$\frac{dX}{dt} = -D \cdot X + \mu X \quad (2.33)$$

$$\frac{dCA}{dt} = -D \cdot CA + q_{CA} X \quad (2.34)$$

$$\frac{dByP}{dt} = -D \cdot ByP + q_{ByP} X \quad (2.35)$$

$$\frac{dN}{dt} = D \cdot (N_{feed} - N) - q_N X \quad (2.36)$$

$$\frac{dCO_2}{dt} = Net\ CO_2 + q_{CO_2} X \quad (2.37)$$

$$\frac{dO_2}{dt} = Net\ O_2 - q_{O_2} X \quad (2.38)$$

Where D is the dilution rate (h^{-1}), Glc_{feed} and N_{feed} are the glucose and nitrogen concentrations in the feed.

2.3.4. Parameter Estimation

Eight parameters (q_S^{max}, K_1, K_i , μ^{max}, K_2, K_3, q_P^{max} and K_4) were present in the rate equations. To estimate them, regression method was used by minimizing the objective function that represents sum of squares total, sum of the differences between the experimental results and model output. For 7 species, objective function was set. Using a MATLAB algorithm that minimizes the sum of squares total (SST) objective function for experimental concentrations ($X_{exp,ij}$) and predicted concentrations ($X_{model,ij}$) was used to estimate parameters.

2.4. MONTE CARLO SIMULATIONS

To perform Monte Carlo simulations, 100 random data sets were created by adding normally distributed noise with a mean of zero and 10 per cent relative standard deviation. Using the

parameters estimated from original data as initial value, the same parameter estimation procedure was re-applied. First and third quartile (middle numbers between lowest and highest values of dataset and median) values were determined for each parameter and interquartile range (IQR) was calculated by subtracting 1st quartile (Q_1) from 3rd quartile (Q_3). Using these interquartile ranges, upper and lower bounds for each parameter were defined. Since parameter values cannot be negative, lower bounds were fixed to zero. The upper bounds (UB) were calculated using [52]:

$$UB = Q_3 + 1.5 \cdot IQR \tag{2.39}$$

The 16 points that remain outside of these bounds were defined as outliers and removed from the system. The histograms of remaining points were plotted and a normal distribution function was fitted to each of them. Corresponding mean, standard deviation and relative standard deviations (standard deviation/mean) were reported.

2.5. ESTIMATION FROM REDUCED PARAMETER SPACE

Since Monte Carlo simulations revealed the correlations between parameters, new parameter estimation process was performed by defining dependent parameters in terms of independent parameters. The correlation functions between parameters were:

$$K_i = 7.675(q_S^{max})^{-1.2} \tag{2.40}$$

$$K_3 = 6.8203\,\mu^{max} - 0.7008 \tag{2.41}$$

$$K_4 = 2707.3\,q_{CA}^{max} - 130.04 \tag{2.42}$$

By fixing the related K values as a function of maximum specific rates, more precise parameters can be obtained for identifyable parameters.

3. RESULTS

3.1. STOICHIOMETRY OF THE SYSTEM

The batch fermentation data was used to calculate the stoichiometry of the system. The growth, acid and by-product production and respiration stoichiometries were defined.

$$C_6H_{12}O_6 + 0.14\ NH_4^+ \xrightarrow{\mu} 4.72\ CH_{1.86}O_{0.33}N_{0.14} + 1.28\ CO_2 + 1.89\ H_2O \quad (3.1)$$

$$C_6H_{12}O_6 + 2.25\ O_2 \xrightarrow{q_P} 0.83\ C_6H_8O_7 + CO_2 + 2.67\ H_2O \quad (3.2)$$

$$C_6H_{12}O_6 + 6O_2 \xrightarrow{q_{O_2}} 6CO_2 + 6H_2O \quad (3.3)$$

Equation 3.1 represents growth stoichiometry, Equation 3.2 represents citrate production stoichiometry and Equation 3.3 represents respiration.

3.2. RQ, CARBON RECOVERY AND CER CALCULATIONS

To determine the presence of by-products and carbon dioxide production, percent carbon recoveries were calculated for each time point. The carbon recoveries were shown in Figure 3.1. Since there were gaps in carbon balance for all time points, by-product formation was included to the system by assuming 20 per cent of the produced citrate was produced as well, like citrate-equivalent by-products. Also when produced CO_2 was included, 100 per cent carbon recovery was obtained.

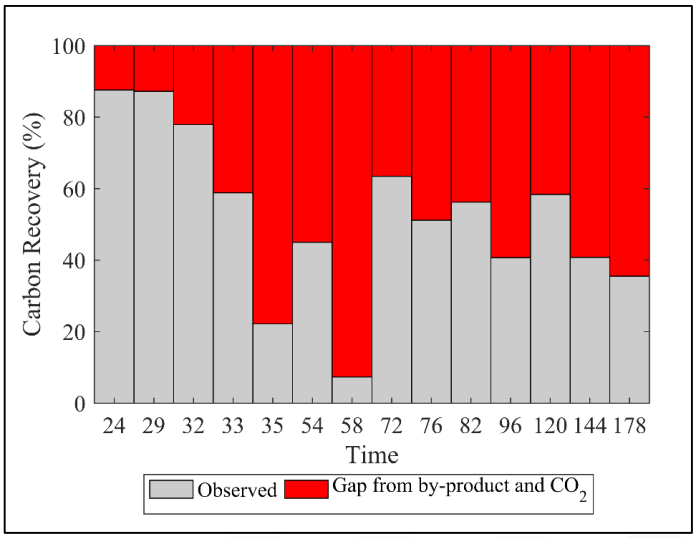

Figure 3.1 Carbon recovery calculated at each time point (per cent)

Next, to analyze the validity of assumptions, CER obtained from carbon balance and CER obtained from growth, acid and by-product formations and respiration were compared. The comparison of 2 calculated CER's were shown in Figure 3.2.a. Almost perfect fit with 0.99 R^2 was obtained. Also, CER the contribution of different reactions to CER was shown in Figure 3.2.b. In first 33 hours of fermentation, the carbon dioxide produced from product formation has the less contribution. In time, carbon dioxide evolution from growth decreases and after 76 hours, CER was only shared between product formation and respiration. The highest CER from product formation and growth were observed at 58th and 35th hours respectively.

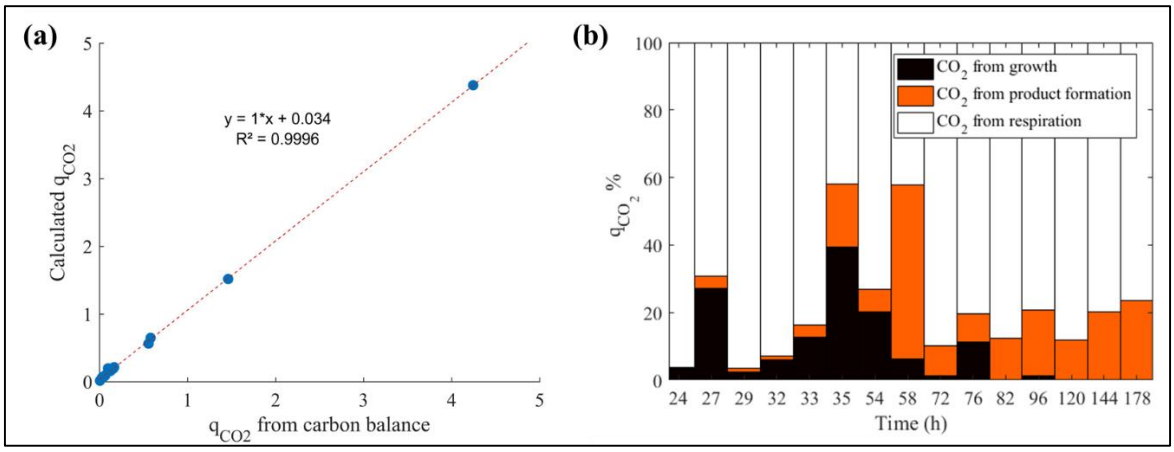

Figure 3.2. (a) Comparison of carbon dioxide evolution from carbon balance and sum of other reactions (b) Contribution of different reactions on carbon dioxide evolution rate

Additionally, respiratory quotients were calculated for each time point, the resulting plot was shown in Figure 3.3.

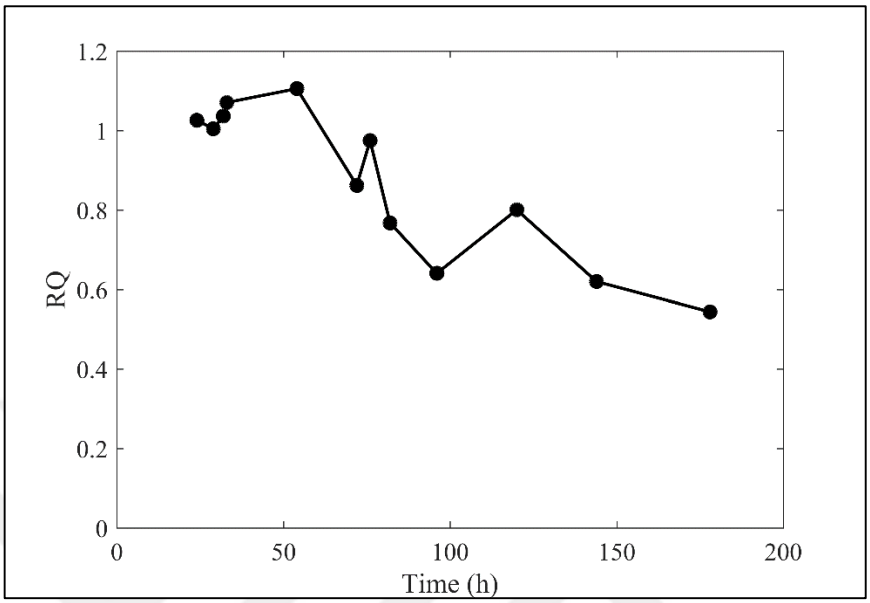

Figure 3.3. Respiratory quotients at each time point

3.3. CALCULATION OF MAINTENANCE PARAMETERS

Growth associated maintenance (K_x: mmol ATP/mmol X) and non-growth associated maintenance (m_{ATP}: mmol ATP/mmol X h) were obtained using OUR from degrees of reduction and specific growth rates. The specific growth rate was plotted against the sum of ATP produced from substrate level phosphorylation and oxidative phosphorylation considering consumed substrate and oxygen. The results were shown in Figure 3.4. The slope of trendlines represent K_x and the y-intercept of the trendlines represent m_{ATP}. Equation of the trendline shows that K_x is 2.31 mmol ATP/mmol X and m_{ATP} is 5.24 mmol ATP/mmol X h-1. 95 per cent interval of K_x was [1.04 3.58] and m_{ATP} was [3.2 7.28]. Also the maintenance parameters were estimated for P/O 1, 1.5, 2 and 2.5. The results were shown in Figure 3.4. As can be seen from the figure, K_x increases from 1.78 to 3.55 in response to increasing P/O. Also, m_{ATP} increases from 4.9 to 8.38 in response to increasing P/O.

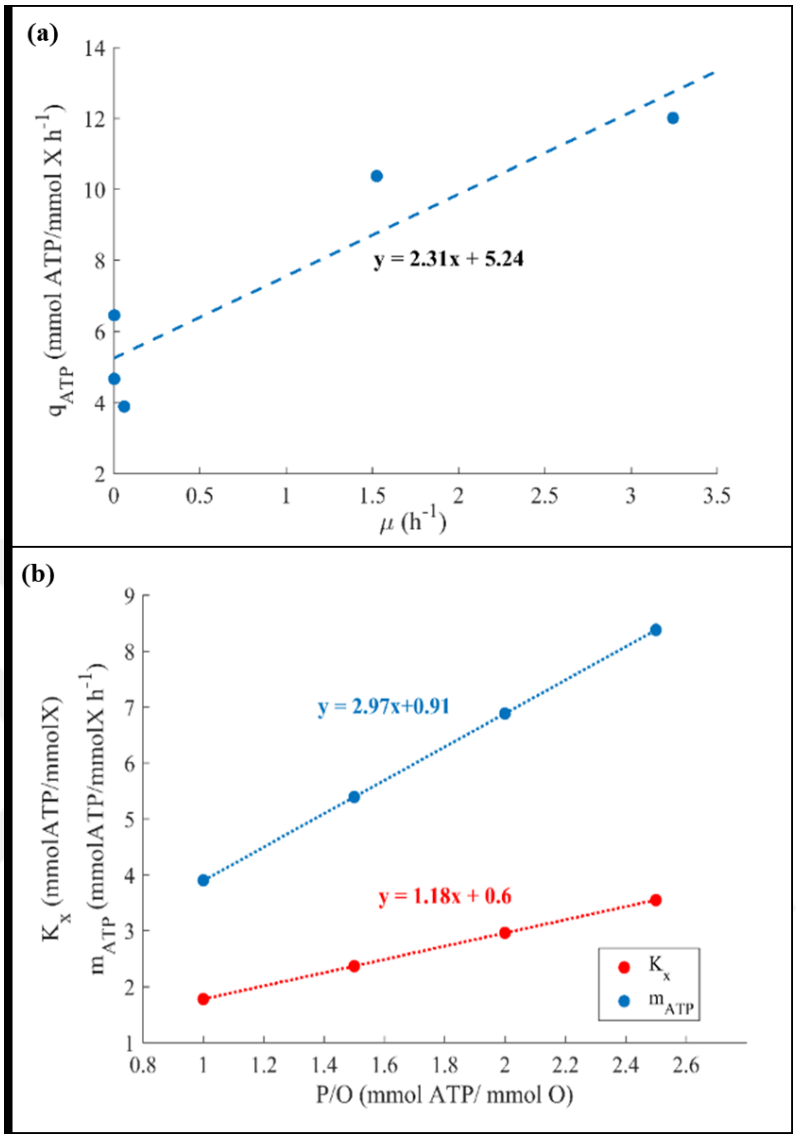

Figure 3.4. Trendline of specific growth rate vs ATP production rate for a fixed P/O ratio, (a) using published vale and (b) for various P/O ratio

3.4. SIMULATION RESULTS

Batch fermentation data was used to model citrate production and growth of *Candida oleophila.* Simulation results were shown in Figure 3.5. Model's total normalized resudial was 0.12. Growth stops when nitrogen was depleted and all carbon provided from glucose was used for citrate and by-product formation. The estimated parameters are presented in Table 3.1.

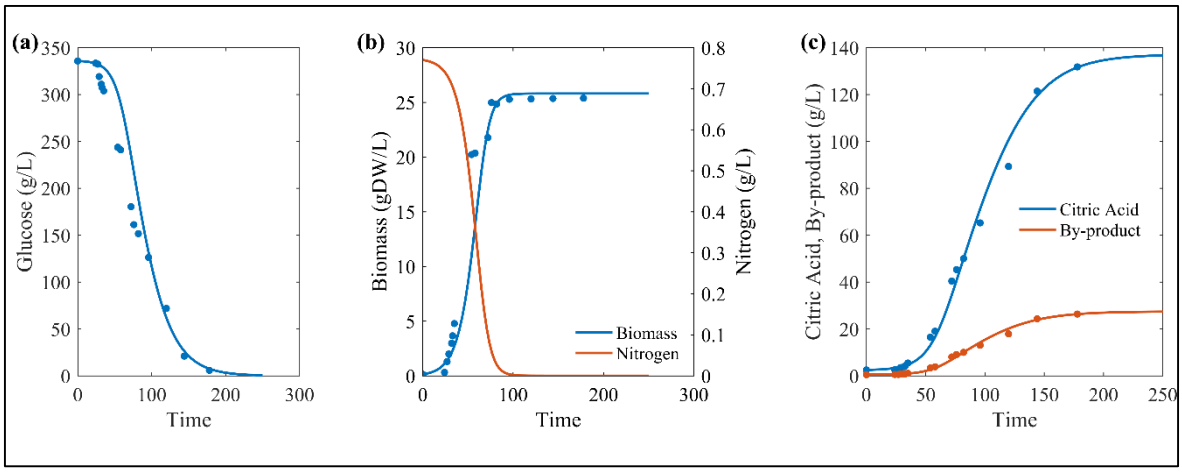

Figure 3.5. Time profiles of (a) glucose, (b) biomass, nitrogen, (c) citric acid and by-products

Table 3.1. Estimated parameters of specific glucose consumption, growth and product formation rates

Parameters	Values
q_S^{max} (g/gDWh)	0.483
K_1 (g/L)	0.020
K_i (g/L)	17.593
μ^{max} (h^{-1})	0.259
K_2 (g/L)	22.744
K_3 (g/L)	0.986
q_{CA}^{max} (g/gDWh)	0.066
K_4 (g/L)	52.917

The specific reaction rates were shown in Figure 3.6.a. q_S starts to decrease after 40th hour. Correspondingly, μ, q_{CA} and q_{ByP} decreases, however decrease in q_{CA} and q_{ByP} were slower. q_{O2} and q_{CO2} increased until they reached their peak at approximately 75th hour. Also volumetric rates were shown inn Figure 3.6.b. Growth rate reaches to maximum between 50th and 60th hour. On the other hand, all other rates reach to their maximum between approximately 60-80th hours.

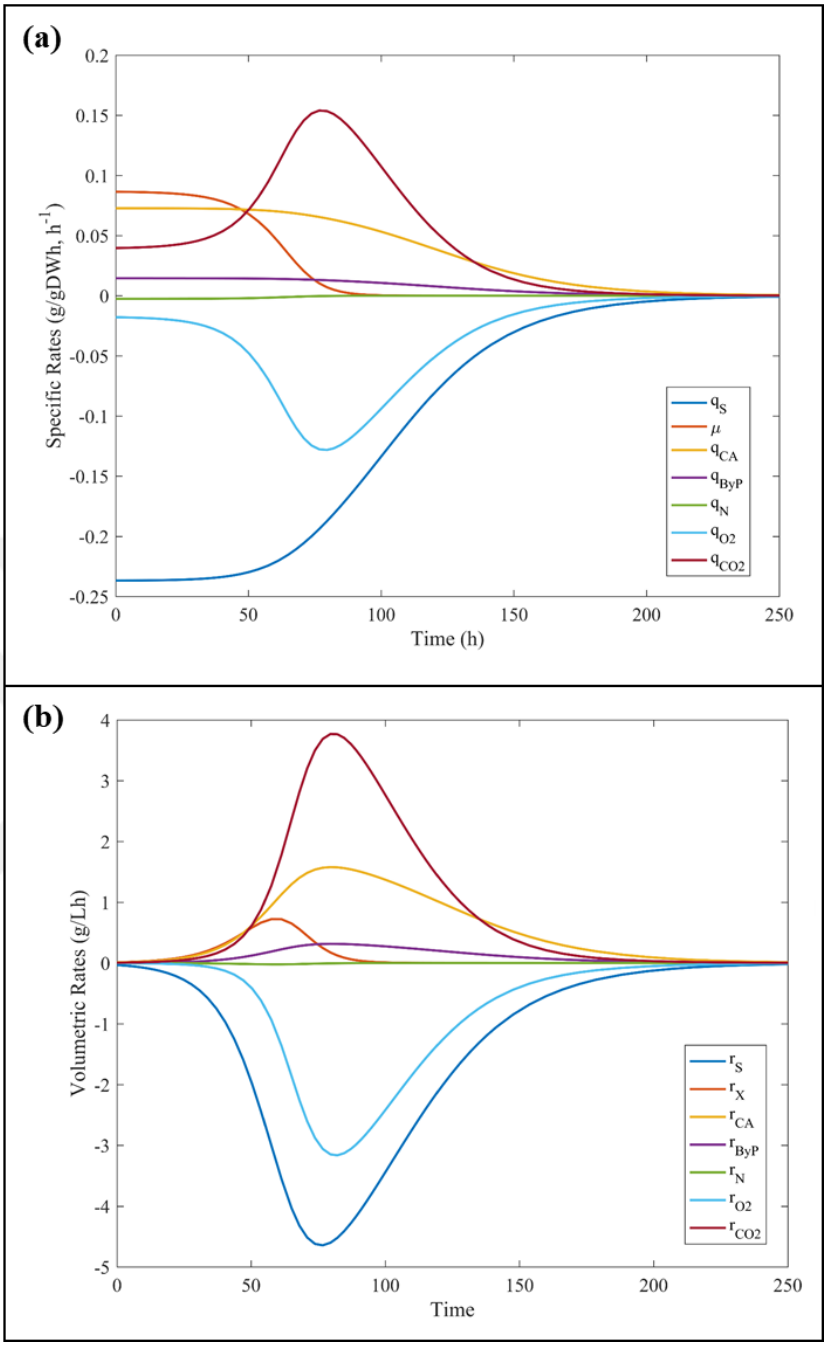

Figure 3.6 Time profiles of (a) specific and (b) volumetric reaction rates

3.5. SIMULATION OF CONTINUOUS MODEL

The estimated parameters for growth, citric acid production and substrate uptake with noncompetitive inhibition were used to simulate the continuous model. Comparison of

steady state concentrations at dilution rates 0.074, 0.063, 0.056, 0.045 and 0.04 published by Anastassiadis *et al.* [42] with simulation results were shown in Figure 3.7.

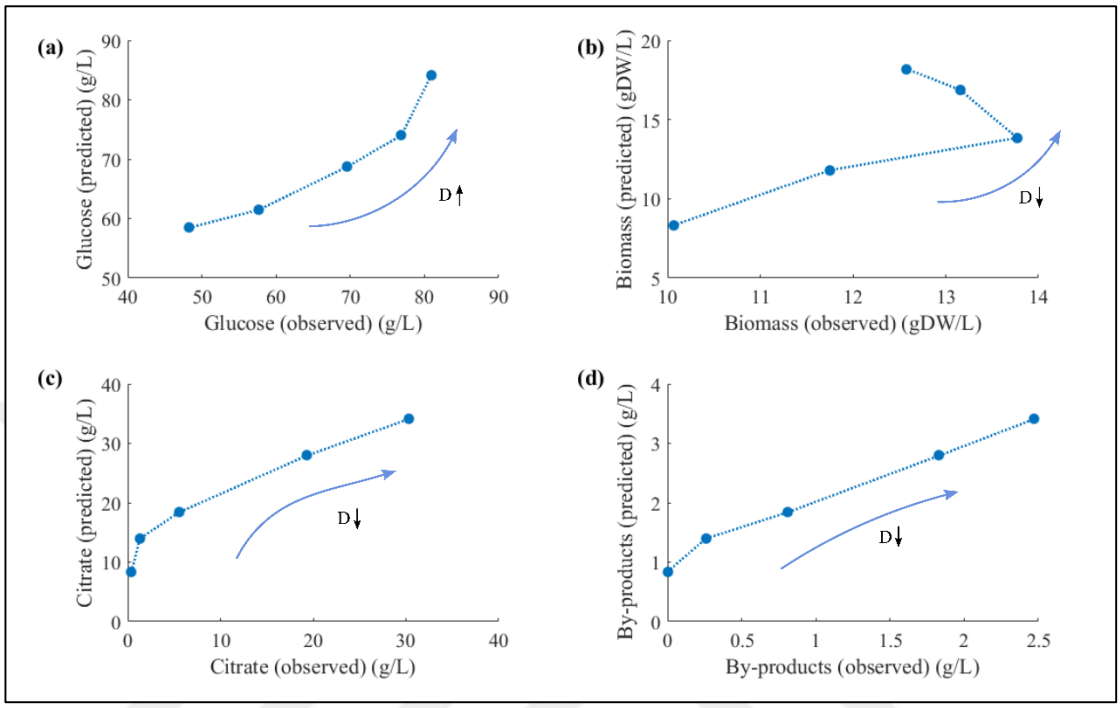

Figure 3.7. Comparison of predicted and observed steady state concentrations of (a) glucose, (b) biomass, (c) citrate and (d) by-products

As shown in Figure 3.7, less glucose was consumed when dilution rate was increased. Accordingly, less biomass, citrate and by-product were produced. The difference between observed and predicted concentrations were lowest for glucose, however, the predicted results almost (at some dilution rates) more than 1.5 times of the observed results for biomass, citrate and by-products.

3.6. MONTE CARLO SIMULATION RESULTS

Random datasets were created to determine the probability distributions of each parameter and considering the fitted distribution functions, mean, standard deviation and relative standard deviation for each parameter was expressed. The parameter distributions were shown in Figure 3.8. Although the paramaters should have positive values, it was assumed that the parameters follow a normal distribution. In noncompetitive inhibition model, highest

relative standard deviation (RSD) were observed at K_1 and K_2 (~133 per cent) and lowest relative standard deviations were observed at q_{CA}^{max}. In competitive inhibition model, highest RSD was observed at K_i (91 per cent) and lowest RSD was observed at q_{CA}^{max} In general, relative standard deviations of maximum specific rates were much lower than relative standard deviations of saturation constants.

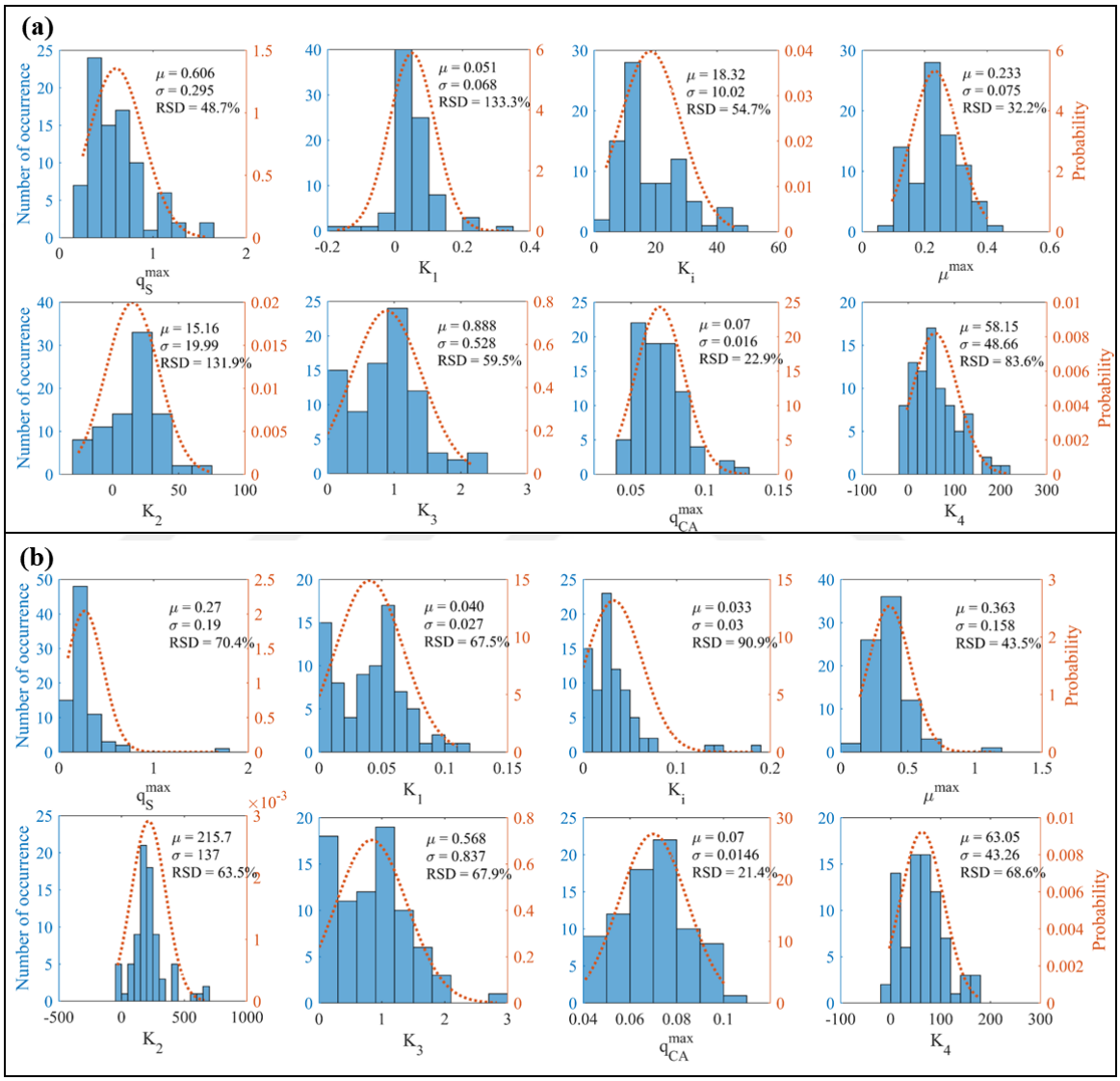

Figure 3.8. Parameter distributions obtained from Monte Carlo simulations for (a) Noncompetitve inhibition and (b) competitive inhibition

Generally RSD's were lower in noncompetitive inhibition for maximum specific rates, however RSD's of saturation and inhibition constants were lower for competitive inhibition.

To determine the existence of correlations between maximum specific rates and their respective saturation and inhibition constants, parameters plotted against each other. The resulting plot was shown in Figure 3.9.

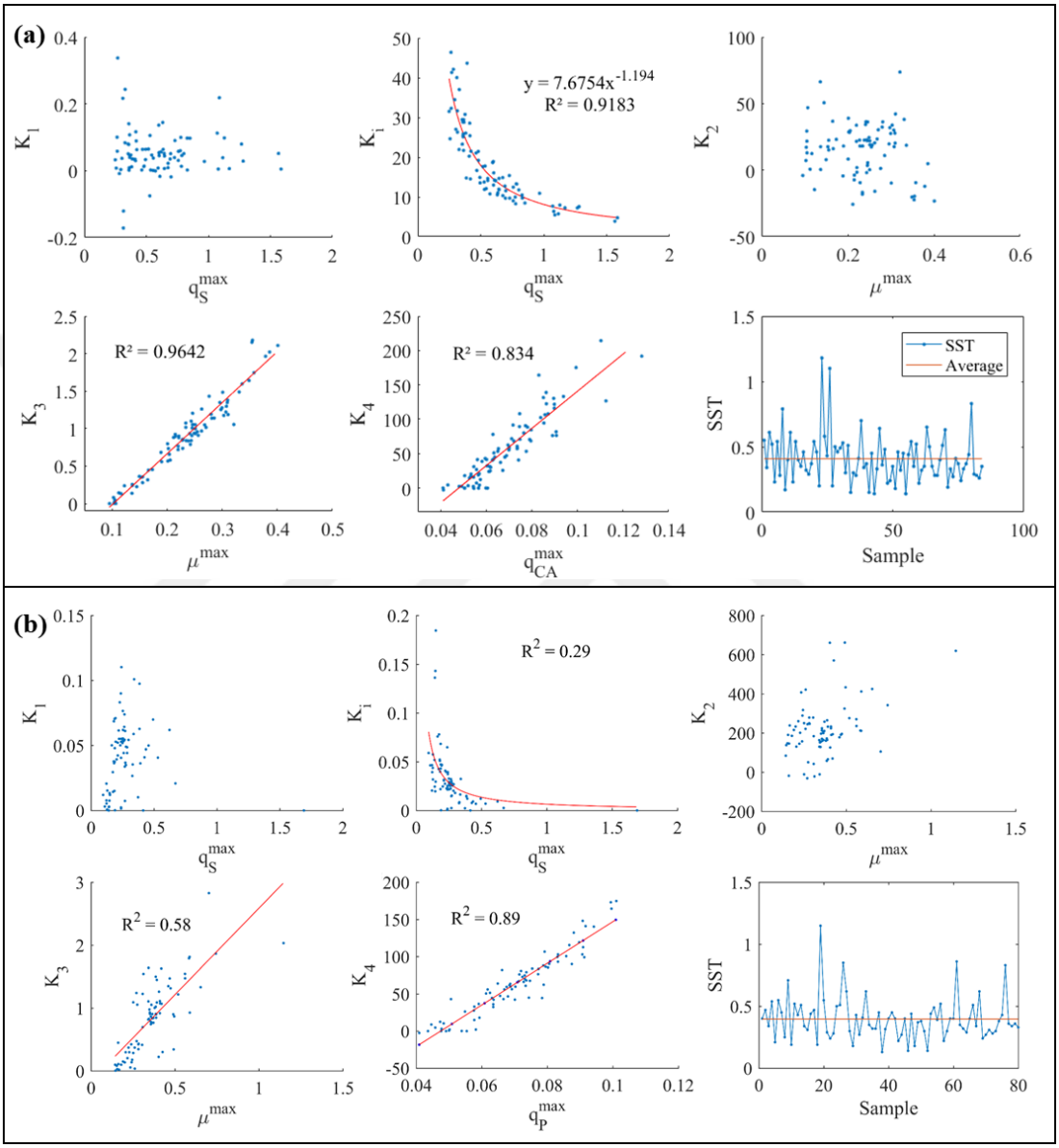

Figure 3.9. Relationship between maximum specific rates with their corresponding saturation constants for (a) noncompetitive inhibition and (b) competitive inhibition

As shown in Figure 3.9 there was no apparent correlation between q_S^{max} and K_1 and μ^{max} and K_2. The relationship between q_S^{max} and K_i can be explained by power law. μ^{max} and K_3 and q_{CA}^{max} and K_4 has a positive linear correlation with R^2 being 0.96 and 0.83. To define the goodness of fit of each sample, SST results after parameter estimation for each sample was

represented. The SST values change in range between 0.2 and 1.2 with an average of approximately 0.45. Although the SST values of the results of two inhibition models were close and RSD's of some parameters were lower in competitive inhibition samples, the R^2 for correlations were much lower for competitive inhibition.

3.7. RE-ESTIMATION OF IDENTIFIABLE PARAMETERS USING REDUCED PARAMETER SPACE

K_i, K_3 and K_4 were found to be correlated to qSmax, mumax and qPmax respectively. Therefore, the saturation constants were fixed using trendline equations. The re-estimated parameters were shown in Figure 3.10. Re-estimated K_i, and K_3 were lower than initially estimated parameters, but K_4 was higher. K_1, and μ^{max} also decreased. The difference in q_S^{max} and q_P^{max} were not significant.

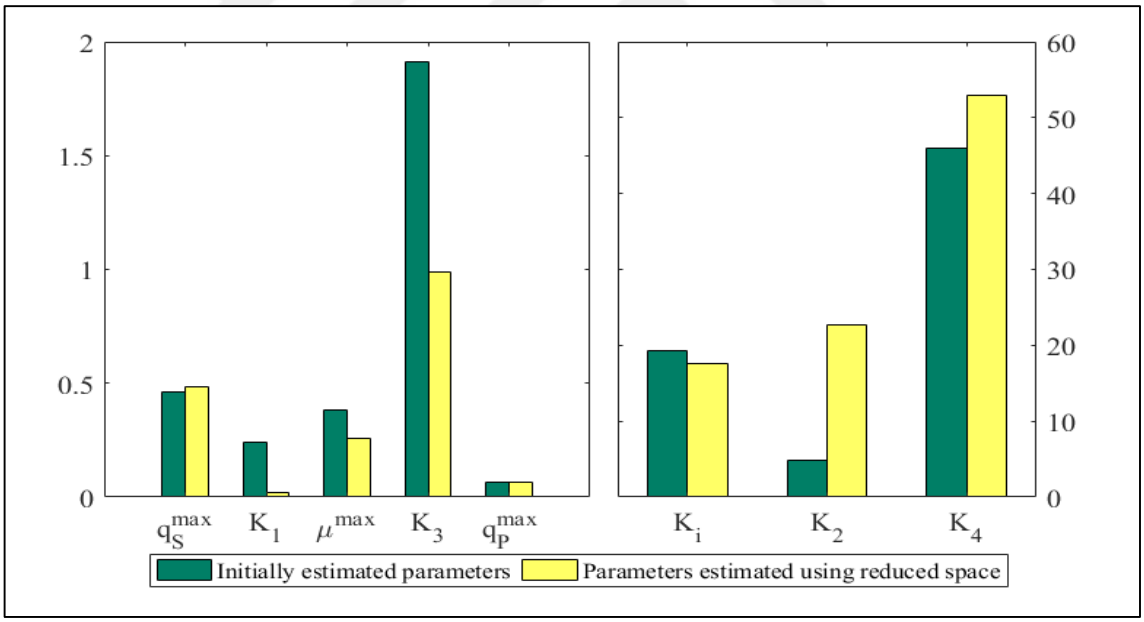

Figure 3.10. Re-estimated parameters using reduced parameter space

4. DISCUSSION

This thesis aim to develop a black-box kinetic model of growth and citric acid production by *Candida oleophila* ATCC20177 using data published by Anastassiadis *et al.* [42, 43].

For this purpose, material and ATP balances were set by assuming all carbon is transferred from glucose to citrate, biomass, by-products and carbon dioxide, and all electron equivalents are shared between glucose, oxygen, biomass, citrate and by-products. However, as shown in Figure 3.1 there is a gap between carbon recoveries calculated using specific rates from experimental data and calculated from degrees of reduction balance. The difference between balanced and calculated results may indicate the production of other metabolites, such as other organic acids and metabolites. Therefore, it was assumed that 20 per cent of the measured citrate was additionally produced as citrate equivalent by-products. Also the use of average biomass concentration while calculating the specific rates from experimental data results in underestimated rates since there was not enough information about the biomass concentration in each time point. This is a common problem, because these linear equations do not represent actual specific rates due to the nonlinear nature of the biological systems [53]. When CER calculated from carbon balance and from growth, production and respiration were compared, a trendline with R^2 1 was obtained, making the by-product assumption valid, shown in Figure 3.2. Although the q_{CA}/CER ratio was 1.2, the ratio used in model was 1 because not all of the by-products were citrate equivalents. Also results showed the validity of the assumption. By calculating RQ, preliminary information about the cell metabolism was obtained. RQ is around 1 for highly oxidized substrates such as glucose and proteins and lower for less oxidized substrates such as fatty acids [54]. Initially, the respiration using glucose was predominant with RQ around 1. At the end of the fermentation RQ decreases to 0.5 indicating the change towards fat metabolism possibly due to stress response as shown in Figure 3.3.

Another important step of calculations was determining the maintenance parameters. First, ATP parameters were calculated for published *Candida utilis* P/O [51]. The difficulty of P/O ratio measurement and prediction is well known and typically they are published as range. Since P/O ratio for *Candida oleophila* was not measured, an acceptable range based on published yeast P/O ratios in the literature was used (Figure 3.4) [55, 56]. Then the change

in m_{ATP} and K_X in response to changing P/O was examined. It was observed that effects of P/O change on K_X was much lower than on m_{ATP}. Regression diagnostics methods were employed to ensure the model's validity by eliminating the effects of steady-state assumptions done while performing ATP balance. Since these balances were done using pseudo-steady state assumptions, the highly dynamic system due to exponential growth phase at the beginning cannot be represented correctly [57]. Also, the lack of fit increases at last hours. By selecting appropriate data, a valid trenline with 2.31 mmolATP/mmolX K_x and 5.24 mmolATP/mmolXh with 90 per cent confidence intervals of (0.6, 4) and (2.5, 8) was obtained.

In addition to experimental calculations, a black-box kinetic model that represents growth and citrate production was obtained. The SST value indicates the validity of the model. The growth depends on nitrogen availability and growth continues until nitrogen is depleted as shown in Figure 3.5 [58]. In this case, since the growth rate was expressed in terms of nitrogen in addition to glucose, growth stopped after nitrogen depletion as expected. Almost all volumetric rates reached their peak around 70-100 hours, represents the end of exponential growth phase. However, the volumetric citrate and by-product production peaks were not as sharp as the other rates because growth stops after nitrogen depletion and all remaining glucose was used for citrate and by-product production.

The obtained model was updated to test the published chemostat data. The citrate/by-product ratio was decreased to 10 per cent because the ratio was reported for chemostat. The apparent μ^{max} calculated to be 0.069. The published dilution rates change 0.27 an 0.039. When dilution rate is higher than μ^{max} either the dilution rate should be equal to the μ^{max} or biomass concentration should be zero, indicating wash out. That means the cells exit the system at a higher rate than the rate they are produced [59]. So the chemostat data between 0.07 and 0.039 were examined. According to the results shown in Figure 3.7, although predicted and observed glucose levels were close to each other, higher growth and product formation was observed around dilution rates 0.06 and 0.05. At lowest dilution rate, the predicted and observed values were close, but more detailed analysis and data was required. Since only steady state concentrations were published for chemostat, direct parameter estimation from chemostat data was not possible.

Two different inhibition models were examined for substrate consumption. Since weak acid stress alters the electrochemical potential of the membrane, it affects substrate uptake into

cell. After the parameter estimation, Monte Carlo simulations were performed to determine identifiable parameters with given data. The RSD of the parameters were high due to black-box assumptions, high degrees of freedom introduced by limited number of measurements. Although the lack of fit after the parameter estimation of both models were not significantly different, the correlations between parameters were observed better at noncompetitive inhibition model. It also makes sense when weak acid inhibition mechanism was considered because the presence of weak acids affect substrate uptake indirectly by altering the permeability of the cell membrane and activities of intracellular enzymes. However, it should be noted that Monte Carlo simulation results cannot be proof. Also, the noncompetitive model fitted to chemostat data better.

To determine the effects of inhibition on parameters more explicitly, parameter space was reduced by relating the correlated parameters with respect to their trendline equations. As shown in Figure 3.10, a positive linear relationship was observed at μ^{max} and K_3 and $q_P{}^{max}$ and K_4. Also $q_S{}^{max}$ and K_i was related with power law. Substrate saturation constants for growth and substrate uptake rates have no correlation with respective maximum specific rate constants and their relative standard deviations were higher. The model with re-estimated parameters was fitted sligthly better to the model.

5. CONCLUSION

The aim of this study was to model growth and citrate production of *Candida oleophila*. The batch fermentation model with two different product inhibition mechanisms used for susbtrate uptake and growth on both glucose and nitrogen was constructed and the model reflected the dynamics of the system with a low lack of fit. Then the batch fermentation model was updated for chemostat data. Noncompetitive inhibition model reflected chemostat dynamics better than competitive inhibition model. Although the biomass and product yields were higher at dilution rates close to apparent maximum specific growth rate, at low dilution rates, model also fitted to chemostat data.

The Monte Carlo simulations were performed to determine which parameters can be identified from given data. The uncertainties introduced by data availability and the precision of the parameter estimation process caused high relative standard deviation results, especially for saturation constants for glucose. However, the correlations between parameters were clear, especially for noncompetitive inhibition model. Both model fits and Monte Carlo simulations support the indirect effect of weak acid stress on substrate uptake.

The carbon balance and RQ calculations were performed to give more insight to dynamics of the system. The results of the assumption of lower CER/q_p ratio showed that not all by-products that produce carbon dioxide were citric acid equivalents, other organic acids composed of less amount of carbon can possibly be produced.

In conclusion, a model that represents growth and citrate production well was obtained. The results indicate that batch fermentation data can be used to obtain preliminary information about the organic acid production, even for the chemostat cultures. Further studies to construct a more reliable chemostat model with time-dependent data is required. Also effects of pH on citrate production can be analyzed to have a better insight about the weak acid stress on *Candida oleophila*.

REFERENCES

1. Gombert AK, Nielsen J. Mathematical modelling of metabolism. *Current Opinion in Biotechnology.* 2000;11(2):180-186.

2. Nikerel I. *Managing complexity of cellular systems: Theoretical tools for dynamic modeling of metabolic reaction networks.* Netherlands: Delft University of Technology; 2009.

3. Szallasi Z, Stelling J, Periwal V. *System modeling in cellular biology: from concepts to nuts and bolts.* Massachusetts: the MIT Press; 2006.

4. Fell DA. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochemical Journal.* 1992;286(Pt 2):313-330

5. Nielsen J, Villadsen J, Liden G. *Bioreaction engineering principles.* New York: Kluwer Academic/Plenum Publishers; 2003.

6. Nielsen J, Villadsen J. Modelling of microbial kinetics. *Chemical Engineering Science*. 1992;47(17-18):4225-4270.

7. Bérangère F, Laurent P, Agnès P, Claude-Gilles D. Methodology for bioprocess analysis: mass balances, yields and stoichiometries. *Stoichiometry and Research-The Importance of Quantity in Biomedicine*: InTech; 2012.

8. Van der Heijden R, Heijnen J, Hellinga C, Romein B, Luyben KCA. Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. *Biotechnology and Bioengineering.* 1994;43(1):3-10.

9. Erickson L, Minkevich I, Eroshin V. Application of mass and energy balance regularities in fermentation. *Biotechnology and Bioengineering.* 2000;67(6):748-774.

10. Doran PM. *Bioprocess engineering principles*: Oxford: Elsevier; 1995.

11. Lagunas R. Energy metabolism of Saccharomyces cerevisiae discrepancy between ATP balance and known metabolic functions. *Biochimica et Biophysica Acta.* 1976;440(3):661-674.

12. Sauer U, Bailey JE. Estimation of P-to-O ratio in Bacillus subtilis and its influence on maximum riboflavin yield. *Biotechnology and Bioengineering.* 1999;64(6):750-754.

13. VanGulik W, Antoniewicz M, DeLaat W, Vinke J, Heijnen J. Energetics of growth and penicillin production in a high-producing strain of Penicillium chrysogenum. *Biotechnology and Bioengineering*. 2001;72(2):185-193.

14. Van Gulik W, Heijnen J. A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnology and Bioengineering.* 1995;48(6):681-698.

15. Taymaz-Nikerel H, Borujeni AE, Verheijen PJ, Heijnen JJ, van Gulik WM. Genome-derived minimal metabolic models for Escherichia coli MG1655 with estimated in vivo respiratory ATP stoichiometry. *Biotechnology and Bioengineering.* 2010;107(2):369-381.

16. Kliphuis AM, Klok AJ, Martens DE, Lamers PP, Janssen M, Wijffels RH. Metabolic modeling of Chlamydomonas reinhardtii: energy requirements for photoautotrophic growth and maintenance. *Journal of Applied Phycology.* 2012;24(2):253-266.

17. Rizzi M, Baltes M, Theobald U, Reuss M. In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: II. Mathematical model. *Biotechnology and Bioengineering.* 1997;55(4):592-608.

18. Thilakavathi M, Basak T, Panda T. Modeling of enzyme production kinetics. *Applied Microbiology and Biotechnology.* 2007;73(5):991-1007.

19. Fu W, Mathews A. Lactic acid production from lactose by Lactobacillus plantarum: kinetic model and effects of pH, substrate, and oxygen. *Biochemical Engineering Journal*. 1999;3(3):163-170.

20. Song H, Jang SH, Park JM, Lee SY. Modeling of batch fermentation kinetics for succinic acid production by Mannheimia succiniciproducens. *Biochemical Engineering Journal*. 2008;40(1):107-115.

21. Luedeking R, Piret EL. A kinetic study of the lactic acid fermentation. Batch process at controlled pH. *Journal of Biochemical and Microbiological Technology and Engineering.* 1959;1(4):393-412.

22. Xu B, Jahic M, Enfors SO. Modeling of Overflow Metabolism in Batch and Fed-Batch Cultures of Escherichiacoli. *Biotechnology Progress.* 1999;15(1):81-90.

23. Kittrell J, Hunter WG, Watson C. Nonlinear least squares analysis of catalytic rate models. *AIChE Journal.* 1965;11(6):1051-1057.

24. Achcar F, Kerkhoven EJ, Bakker BM, Barrett MP, Breitling R. Dynamic modelling under uncertainty: the case of Trypanosoma brucei energy metabolism. *PLoS Computational Biology*. 2012;8(1):e1002352.

25. Metropolis N. Monte Carlo Method. *From cardinals to chaos: Reflection on the life and legacy of Stanislaw Ulam.*Book News, 1989:125.

26. Kroese DP, Brereton T, Taimre T, Botev ZI. Why the Monte Carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics.* 2014;6(6):386-392.

27. Manly BF. *Randomization, bootstrap and Monte Carlo methods in biology:*Florida: Chapman and Hall/CRC; 2006.

28. Angumeenal A, Venkappayya D. An overview of citric acid production. *LWT-Food Science and Technology*. 2013;50(2):367-370.

29. Anastassiadis S, Aivasidis A, Wandrey C. Citric acid production by Candida strains under intracellular nitrogen limitation. *Applied Microbiology and Biotechnology.* 2002;60(1-2):81-87.

30. Kamzolova SV, Fatykhova AR, Dedyukhina EG, Anastassiadis SG, Golovchenko NP, Morgunov IG. Citric acid production by yeast grown on glycerol-containing waste from biodiesel industry. *Food Technology and Biotechnology.* 2011;49(1):65-74.

31. Papagianni M. Advances in citric acid fermentation by Aspergillus niger: biochemical aspects, membrane transport and modeling. *Biotechnology Advances.* 2007;25(3):244-263.

32. Robinson RK. *Encyclopedia of food microbiology* California: Academic press; 2014.

33. Kavanagh K. Fungal fermentation systems and products. *Fungi: Biology and Applications*. Wiley, 2005.

34. LeBlanc J, Laiño JE, del Valle MJ, Vannini Vv, van Sinderen D, Taranto MP, et al. B-Group vitamin production by lactic acid bacteria–current knowledge and potential applications. *Journal of Applied Microbiology.* 2011;111(6):1297-1309.

35. Leuchtenberger W, Huthmacher K, Drauz K. Biotechnological production of amino acids and derivatives: current status and prospects. *Applied Microbiology and Biotechnology.* 2005;69(1):1-8.

36. Christensen LH, Nielsen J, Villadsen J. Monitoring of substrates and products during fed-batch penicillin fermentations on complex media. *Analytica Chimica Acta.* 1991;249(1):123-136.

37. Larsson G, Enfors SO, Pham H. The pH-auxostat as a tool for studying microbial dynamics in continuous fermentation. *Biotechnology and Bioengineering.* 1990;36(3):224-232.

38. Ciriminna R, Meneguzzo F, Delisi R, Pagliaro M. Citric acid: emerging applications of key biotechnology industrial product. *Chemistry Central Journal.* 2017;11(1):22-31

39. Vandenberghe L, Rodrigues C, de Carvalho J, Medeiros A, Soccol C. Production and Application of Citric Acid. *Current Developments in Biotechnology and Bioengineering:* Elsevier; 2017. p. 557-575.

40. Klement T, Milker S, Jäger G, Grande PM, de María PD, Büchs J. Biomass pretreatment affects Ustilago maydis in producing itaconic acid. *Microbial Cell Factories.* 2012;11(1):43-56.

41. Anastassiadis S, Rehm H-J. Continuous citric acid secretion by a high specific pH dependent active transport system in yeast Candida oleophila ATCC 20177. *Electronic Journal of Biotechnology*. 2005;8(2):26-42.

42. Anastassiadis S, Rehm H-J. Citric acid production from glucose by yeast Candida oleophila ATCC 20177 under batch, continuous and repeated batch cultivation. *Electronic Journal of Biotechnology.* 2006;9(1):26-39.

43. Anastassiadis S, Wandrey C, Rehm H-J. Continuous citric acid fermentation by Candida oleophila under nitrogen limitation at constant C/N ratio. *World Journal of Microbiology and Biotechnology*. 2005;21(5):695-705.

44. Rafi MM, Hanumanthu M, Rizwana S, Venkateswarlu K, Rao DM. Effect of different physico-chemical parameters on fermentative production of itaconic acid by Ustilago maydis. *Journal of Microbiology and Biotechnology Research.* 2017;2(5):794-800.

45. Rafi MM, Hanumanthu M, Rao DM, Venkateswarlu K, Zeraib A, Ramdani M, et al. Production of itaconic acid by Ustilago maydis from agro wastes in solid state fermentation. *Journal of BioScience & Biotechnology.* 2014;3(2):163-168

46. Magalhães AI, de Carvalho JC, Medina JDC, Soccol CR. Downstream process development in biotechnological itaconic acid manufacturing. *Applied Microbiology and Biotechnology.* 2017;101(1):1-12.

47. Oliveira FS, Araújo JM, Ferreira R, Rebelo LPN, Marrucho IM. Extraction of L-lactic, L-malic, and succinic acids using phosphonium-based ionic liquids. *Separation and Purification Technology.* 2012;85:137-146.

48. Sauer M, Porro D, Mattanovich D, Branduardi P. Microbial production of organic acids: expanding the markets. *Trends in Biotechnology.* 2008;26(2):100-108.

49. Cornish-Bowden A, Cornish-Bowden A. *Fundamentals of enzyme kinetics:* Wiley-Blackwell Weinheim, Germany; 2012.

50. Cornish-Bowden A. One hundred years of Michaelis–Menten kinetics. *Perspectives in Science*. 2015;4:3-9.

51. Verduyn C, Stouthamer AH, Scheffers WA, van Dijken JP. A theoretical evaluation of growth yields of yeasts. *Antonie Van Leeuwenhoek.* 1991;59(1):49-63.

52. Hyndman RJ, Fan Y. Sample quantiles in statistical packages. *The American Statistician.* 1996;50(4):361-365.

53. Holmes CC, Held L. Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis.* 2006;1(1):145-168.

54. Ellis AC, Hyatt TC, Hunter GR, Gower BA. Respiratory quotient predicts fat mass gain in premenopausal women. *Obesity.* 2010;18(12):2255-2259.

55. Hinkle PC. P/O ratios of mitochondrial oxidative phosphorylation. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. 2005;1706(1-2):1-11.

56. Vanrolleghem PA, de Jong-Gubbels P, van Gulik WM, Pronk JT, van Dijken JP, Heijnen S. Validation of a metabolic network for Saccharomyces cerevisiae using mixed substrate studies. *Biotechnology Progress.* 1996;12(4):434-448.

57. Velleman PF, Welsch RE. Efficient computing of regression diagnostics. *The American Statistician.* 1981;35(4):234-242.

58. Maassen N, Panakova M, Wierckx N, Geiser E, Zimmermann M, Bölker M, et al. Influence of carbon and nitrogen concentration on itaconic acid production by the smut fungus Ustilago maydis. *Engineering in Life Sciences.* 2014;14(2):129-134.

59. Ziv N, Brandt NJ, Gresham D. The use of chemostats in microbial systems biology. *Journal of Visualized Experiments: JoVE.* 2013;(80).