HYBRID OF ANT COLONY OPTIMIZATION AND FLUX VARIABILITY FOR IMPROVING METABOLITES PRODUCTION

AMIRA HUSNA BINTI AZHAR

.

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Philosophy

Faculty of Computing
Universiti Teknologi Malaysia

JANUARY 2017

All praises to Allah the Almighty for
the strength and His blessing in completing this thesis.

Specially dedicated to
my beloved parents
my precious siblings

ACKNOWLEDGEMENT

"In the name of Allah, the most Gracious and the most Merciful"

In preparing this thesis, I was in contact with my supervisors, course mates, best friend, seniors and family members. They have contributed towards my understanding and thoughts. First and foremost, I would like to express my appreciation and gratitude to my supervisor for his encouragement, guidance, and valuable advices. His guides and suggestions are really helping for me to complete this research.

Besides that, I would like to thanks my course mates and seniors for their knowledge sharing while doing this research. All their sharing, teaching, and comments help me a lot to complete this research. Last but not least, sincere thanks to my family and best friend for their continuous supports, cares and encouragements all the time.

ABSTRACT

Metabolic engineering has been successfully used for the production of a variety of useful compounds such as L-phenylalanine and biohydrogen that received high demand on food, pharmaceutical, fossil fuels, and energy industries. Reaction deletion is one of the strategies of *in silico* metabolic engineering that can alter the metabolism of microbial cells with the objective to get the desired phenotypes. However, due to the size and complexity of metabolic networks, it is difficult to determine the near-optimal set of reactions to be knocked out. The complexity of the metabolic network is also caused by the presence of competing pathway that may interrupt the high production of a desireable metabolite. Consequently, this factor leads to low Biomass-Product Coupled Yield (BPCY), production rate and growth rate. Other than that, inefficiency of existing algorithms in modelling high growth rate and production rate is another problem that should be handled and solved. Therefore, this research proposed a hybrid algorithm comprising Ant Colony Optimization and Flux Variability Analysis (ACOFVA) to identify the best reaction combination to be knocked out to improve the production of desired metabolites in microorganisms. Based on the experimental results, ACOFVA shows an increase in terms of BPCY and production rate of L-Phenylalanine in Yeast and biohydrogen in Cyanobacteria, while maintaining the optimal growth rate for the target organism. Besides, suggested reactions to be knocked out for improving the production yield of L-Phenylalanine and biohydrogen have been identified and validated through the biological database. The algorithm also shows a good performance with better production rate and BPCY of L-Phenylalanine and biohydrogen than existing results.

ABSTRAK

Kejuruteraan metabolik telah berjaya digunakan untuk pengeluaran pelbagai sebatian berguna seperti *L-Phenylalanine* dan *biohydrogen* yang menerima permintaan yang tinggi dalam industri makanan, farmaseutikal, bahan api fosil, dan tenaga. Teknik penyingkiran reaksi adalah salah satu strategi di dalam kejuruteraan metabolik yang boleh mengubah metabolisme sel-sel mikrob dengan objektif untuk mendapatkan fenotip yang dikehendaki. Walau bagaimanapun, disebabkan oleh saiz dan kerumitan rangkaian metabolik, ia adalah sukar untuk menentukan set reaksi yang hampir optimum untuk disingkirkan. Kerumitan rangkaian metabolik ini disebabkan oleh kehadiran reaksi yang boleh mengganggu pengeluaran tinggi sesuatu metabolit yang diingini. Faktor ini menyebabkan nilai hasil bersama biojisim-produk, kadar pengeluaran dan kadar pertumbuhan menjadi rendah. Selain itu, masalah ketidakecekapan algoritma sedia ada dalam memodelkan kadar pertumbuhan dan pengeluaran yang tinggi juga perlu ditangani dan diatasi. Oleh itu, kajian ini mencadangkan satu algoritma hibrid iaitu Pengoptimuman Koloni Semut dan Analisis Fluks Kepelbagaian untuk mengenal pasti reaksi dan gen untuk disingkirkan bagi meningkatkan pengeluaran metabolit dikehendaki dalam mikroorganisma. Berdasarkan keputusan eksperimen, ACOFVA menunjukkan peningkatan dari segi BPCY dan kadar pengeluaran *L-Phenylalanine* dalam Yis dan *biohydrogen* dalam Cyanobacteria, di samping mengekalkan kadar pertumbuhan organisma yang optimum. Selain itu, reaksi dan gen yang dicadangkan untuk disingkir bagi meningkatkan hasil pengeluaran *L-Phenylalanine* dan *biohydrogen* juga telah dikenal pasti dan disahkan melalui pangkalan data biologi. Algoritma ini juga menunjukkan prestasi yang baik dengan kadar pengeluaran dan BPCY *L-Phenylalanine* dan *biohydrogen* yang lebih tinggi berbanding keputusan yang sedia ada.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	TITLE	**i**
	DECLARATION	**ii**
	DEDICATION	**iii**
	ACKNOWLEDGEMENT	**iv**
	ABSTRACT	**v**
	ABSTRAK	**vi**
	TABLE OF CONTENTS	**vii**
	LIST OF TABLES	**xiii**
	LIST OF FIGURES	**xv**
	LIST OF ABBREVIATION	**xviii**
	LIST OF APPENDICES	**xx**
1	**INTRODUCTION**	**1**
	1.1 Overview	1
	1.2 Introduction	2
	1.2.1 Microbial cell factory for protein productions overview	2
	1.2.2 Metabolic engineering	3
	1.2.2.1 Ant Colony Optimization Algorithm	4
	1.2.2.2 Flux Variability Analysis	5
	1.3 Background of the Problem	6
	1.4 Problem Statement	8
	1.5 Research Aim	9

1.6 Objectives 10
1.7 Scope of Study 10
1.8 Significance of Study 11
1.9 Report Organization 12

2 LITERATURE REVIEW 14
2.1 Overview 14
2.2 *In Silico* Studies 15
2.3 Metabolic Engineering 15
2.4 Genome-Scale Metabolic Network (GSMN) 19
2.5 Constraints- Based Modelling Approach 20
2.5.1 Flux Balance Analysis (FBA) 20
2.5.2 Flux Variability Analysis (FVA) 23
2.5.3 Dynamic Flux Balance Analysis (dFBA) 24
2.5.4 Minimization of Metabolic Adjustment (MOMA) 25
2.5.5 Regulatory On/Off Minimization (ROOM) 26
2.5.6 Discussion of Constraints-based Modelling Approaches 27
2.5.7 Comparison between Metabolic Network Analyses Approaches 28
2.6 Reaction Knockout to Improve Metabolite Production 32
2.6.1 OptKnock 33
2.6.2 OptFlux 34
2.6.3 ReacKnock 35
2.6.4 FastPros 36
2.6.5 OptORF 36
2.6.6 Comparison between Modelling Framework Approaches 38
2.7 Swarm Intelligence Based Optimization Algorithms 42
2.7.1 Ant Colony Optimization Algorithm (ACO) 42
2.7.2 Genetic Algorithm (GA) 44

2.7.3 Particle Swarm Optimization (PSO) 45
2.7.4 Bees Algorithm (BA) 46
2.7.5 Artificial Bee Colony (ABC) 47
2.7.6 Discussion of Swarm Intelligence Based Optimization Algorithm 49
2.7.7 Comparison between Swarm Intelligence Based Optimization Algorithms 50
2.8 Hybridization of Algorithms 54
2.8.1 Ant Colony Optimization and Flux Balance Analysis (ACOFBA) 55
2.8.2 Continuous Bees Algorithm and Flux Balance Analysis (CBAFBA) 55
2.8.3 Ant Colony Optimization and Minimization of Metabolic Adjustment (ACOMOMA) 56
2.8.4 Artificial Bee Colony and Minimization of Metabolic Adjustment (ABCMOMA) 56
2.8.5 Comparison between Existing Hybrid Algorithms 58
2.9 Discussion on Trends of Algorithms Used in *In Silico* Metabolic Engineering 61
2.9 Challenge 63

3 RESEARCH METHODOLOGY 64
3.1 Overview 64
3.2 Research Framework 65
3.3 Phase 1: Modelling of a proposed hybrid algorithm 68
3.3.1 Literature review of target domain 68
3.3.2 Collect sources of references 70
3.3.3 Identify problem, aim, objectives, and scope 70
3.3.4 Collection of Dataset Model 71
3.3.4.1 Microorganisms 75
3.3.4.1.1 Yeast *Saccharomyces cerevisiae* 75

3.3.4.1.1 Cyanobacteria *Synechocystis* 76
3.3.4.2 Metabolites 77
3.3.4.2.1 L-Phenylalanine 77
3.3.4.2.2 Biohydrogen 79
3.3.5 Model the proposed hybrid of Ant Colony Optimization and Flux Variability Analysis 80
3.3.5.1 Model Pre-processing 81
3.3.5.2 A Hybrid of Ant Colony Optimization and Flux Variability Analysis (ACOFVA) 82
3.3.5.2.1 Steps In ACOFVA 83
3.3.5.2.2 Ant Representation 85
3.3.5.2.3 Initializing population 86
3.3.5.2.4 Solution construction 87
3.3.5.2.5 Scoring fitness of an individual 88
3.3.5.2.6 Pheromone update 92
3.4 Implementation of the proposed hybrid algorithm 93
3.4.1 Development the hybrid algorithm of ACOFVA 93
3.4.2 Add performance measurement into the hybrid algorithm: production rate, BPCY, and convergence graph 95
3.4.3 Run parameter testing for: Number of Knockout and Number of Population 95
3.4.4 Test the feasibility of the newly developed hybrid algorithm with selected dataset model 96
3.4.5 Collect results of experimental analysis 96
3.5 Phase 3: Evaluation of the results obtained 99
3.5.1 Evaluate the results of this research with results of previous related researches 99
3.5.2 Cross validate the list of reactions and genes suggested to be knocked out 100
3.5.3 Present the convergence graph 101
3.5.4 Conclusion of the results and research 101
3.6 Documentation and verification 102

3.7 Hardware and Software Requirements 102
3.8 Performance Measurement 104
3.9 Summary 107

4 RESULTS 108
4.1 Overview 108
4.2 Pre-processing of the dataset model 109
4.3 Parameter Setting 109
4.4.1 Parameter Setting for Number of Knockout 110
4.4.2 Parameter Setting for Number of Population 113
4.4.3 Parameter Setting for ACOFVA 114
4.4 Experimental Results and Discussions 116
4.5.1 Production rate of L-Phenylalanine and biohydrogen 117
4.5.1.1 Production of L-Phenylalanine in *Saccharomyces cerevisiae* 117
4.5.1.2 Production of Biohydrogen in *Synechocystis* sp. PCC6803 120
4.5.2 Biomass Product Coupled Yield of L-Phenylalanine and biohydrogen 122
4.5.2.1 Biomass Product Coupled Yield of L-Phenylalanine in *Saccharomyces cerevisiae* 123
4.5.2.2 Biomass Product Coupled Yield of Biohydrogen in *Synechocystis* sp PCC6803 126
4.6 Discussion 128
4.7 Summary 129

5 CONCLUSION 130
5.1 Conclusion 130
5.2 Limitations 132
5.3 Future Works 132

8	**REFERENCES**	134-144
9	**APPENDIX A (DATASET)**	145-148
10	**APPENDIX B (Results of 30 Runs)**	149-153

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Comparison of Metabolic Network Analysis Approach	28
2.2	Comparison of the Applications applied in Constraint-based Modelling Approaches	30
2.3	Comparison of the Constraints Incorporated in Constraint-based Modelling Approaches	31
2.4	Comparison of Modelling Framework Approaches	38
2.5	Comparison between Swarm Intelligence Based Optimization Algorithms	51
2.7	Comparison between existing hybrid algorithms	58
3.1	Model information	72
3.2	List of Yeast *S. cerevisiae* model constructed	76
3.3	List of Cyanobacteria *Synechocystis* sp PCC 6803 model constructed	77
3.4	Comparison of ACOFVA and existing hybrid algorithms	82
3.5	Wild-type growth rate, production rate, BPCY value, and maximum theoretical yield for desired metabolites	105
4.1	Comparison between raw and pre-processed model	109
4.2	Comparison of the percentage of optimal solution for different number of knockouts	111
4.3	Parameter setting for ACOFVA	115

4.4	Mean and standard deviation for the growth rate of five knockouts of each metabolite using ACOFVA	116
4.5	Reactions and genes information suggested to be knocked out for L-Phenylalanine in *Saccharomyces cerevisiae*	117
4.6	Comparison between different methods for growth rate and production rate of L-Phenylalanine in *Saccharomyces cerevisiae*	118
4.7	Reactions and genes information suggested to be knocked out for biohydrogen in *Synechocystis sp* PCC6803	120
4.8	Comparison between different methods for growth rate and production rate of biohydrogen in *Synechocystis sp* PCC6803	121
4.9	List of reactions and genes information suggested to be knocked out for L-Phenylalanine in *Saccharomyces cerevisiae*	123
4.10	Comparison between different methods for growth rate and BPCY of L-Phenylalanine in *Saccharomyces cerevisiae*	124
4.11	List of reactions and genes information suggested to be knocked out for biohydrogen in *Synechocystis* sp PCC6803.	126
4.12	Comparison between different algorithms for growth rate and BPCY of biohydrogen in *Synechocystis*	126

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
1.1	Basic Ant Colony Optimization Behavior at Different Time Stamps	5
1.2	Graph theoritical of a complex metabolic network	9
2.1	Overview of Metabolic Engineering	16
2.2	Schematic overview of the metabolic engineering cycle	17
2.3	The outline of the pathway for the synthesis of L-Phenylalanine	18
2.4	The outline of the pathway for the synthesis of biohydrogen	19
2.5	Basis concept in constraint-based modelling	20
2.6	The methodology of Flux Balance Analysis	21
2.7	The optimization principles underlying FVA	24
2.8	Metabolic engineering strategies to enhance the flux through a pathway	33
2.9	Flowchart of a standard ACO	43
2.10	Flowchart of GA	44
2.11	Flowchart of PSO	46
2.12	Flowchart of BA	47
2.13	Flowchart of ABC	48
3.1	Research Framework	67
3.2	Objective reaction for L-Phenylalanine production in Yeast	73

3.3	Objective function for glucose reaction in Yeast	73
3.4	Reaction for the production of biohydrogen in Cyanobacteria	74
3.5	Glucose reaction as the substrate in Cyanobacteria	74
3.6	Overview pathway for L-Phe production	78
3.7	Metabolic pathway for L-Phe production based on KEGG	78
3.8	Biological pathway for biohydrogen production	79
3.9	Flowchart for Overall Model Preprocessing	81
3.10	(a) Bit representation flowchart of the ACOFVA. (b) Flowchart of basic ACO algorithm.	84
3.11	ACO problem representation for individual. A, B, C and D represent nodes.	85
3.12	The representation of reaction and gene of metabolic model.	86
3.13	Overview of solution construction in ACOFVA.	87
3.14	The flow of FVA to obtain the best production rate	89
3.15	The flow of FVA to obtain best BPCY	91
3.16	Flowchart of ACOFVA	94
3.17	Experimental setup for Yeast data	97
3.18	Experimental setup for Cyanobacteria data	98
3.19	Overview of Cobra Toolbox	103
4.1	The graph for average L-Phenylalanine production rate of 1 to 10 knockout in Yeast model	110
4.2	The graph for average biohydrogen production rate of 1 to 10 knockout in Cyanobacteria model	111
4.3	The graph for average L-Phenylalanine production rate of 100 to 250 population size for Yeast model.	113
4.4	The graph for average biohydrogen production rate of 100 to 250 population size for Cyanobacteria model	114
4.5	ACOFVA convergence graph for production of L-Phenylalanine	119

4.6	ACOFVA convergence graph for production of biohydrogen	121
4.7	BPCY convergence graph for the production of L-Phenylalanine	125
4.8	BPCY convergence graph for the production of biohydrogen.	127

LIST OF ABBREVIATION

ABC	-	Artificial Bee Colony
ABCMOMA	-	Artificial Bee Colony and Minimization of Metabolic Adjustment
ACO	-	Ant Colony Optimization
ACOFBA	-	Ant Colony Optimization and Flux Balance Analysis
ACOMOMA	-	Ant Colony Optimization and Minimization of Metabolic Adjustment
BA	-	Bees Algorithm
BAFBA	-	Bees Algorithm and Flux Balance Analysis
BATFBA	-	Bat Algorithm and Flux Balance Analysis
CBA	-	Continuous Bees Algorithm
CBAFBA	-	Continuous Bees Algorithm and Flux Balance Analysis
COBRA	-	Constraint-based Reconstruction and Analysis
dFBA	-	Dynamic Flux Balance Analysis
E.coli	-	*Escherichia coli*
FBA	-	Flux Balance Analysis
FVA	-	Flux Variability Analysis
GA	-	Genetic Algorithm
GACOFBA	-	Genetic Ant Colony Optimization and Flux Balance Analysis
H_2	-	Hydrogen
LP	-	Linear Programing
L-Phe	-	L-Phenylalanine
MATLAB	-	Matrix Laboratory
MILP	-	Mixed Integer Linear Programing

MOMA	-	Minimization of Metabolic Adjustment
PSO	-	Particle Swarm Optimization
QP	-	Quadratic Programing
RAM	-	Random Access Memory
ROOM	-	Regulatory On/Off Minimization
SBML	-	System Biology Mark-up Language
S. cerevisiae	-	*Saccharomyces cerevisiae*
sp.	-	Species

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Dataset	145-148
B	Results of 30 runs	149-153

CHAPTER 1

INTRODUCTION

1.1 Overview

This chapter discusses an overview of the introduction part of this study as an important prior step. The basic concept of metabolic engineering applies in biological systems is introduced in general. A brief explanation about some algorithms that exist in metabolic engineering is also presented. It is important to collect and understand metabolic engineering field because this research focuses on improving the production of metabolites such as L-phenylalanine and biohydrogen from microorganisms through *in silico* study. In addition, the objectives, scopes and justifications are also included for a clearer view of this research.

1.2 Introduction

In this section, the background of microbial cell factory for metabolites productions is briefly discussed. Besides that, a brief explanation about the constraint-based modeling and optimization algorithms that applied in this research are presented.

1.2.1 Microbial cell factory for metabolites productions overview

Nowadays, there is a growing concern about energy generation by fossil fuels to be continued in usage as a source since it is renewable. Biohydrogen or hydrogen (H_2) is a promising fuel that has better energy content and even higher than oil (Demirbas, 2002; Islam *et al.*, 2005). Currently, hydrogen shows the tendency to be an alternative to fossil fuels for transportation (Veziroglu & Macario, 2011). Besides, it is also renewable, efficient and clean, showing that hydrogen production by this bacteria can be generated economically and in an environmental-friendly manner. Over the last decade, production of hydrogen of biological microbes gains more attention. Thus, this study investigates how H_2 production as secondary fuel product in Cyanobacteria *Synechocystis* sp. PCC6803 can be increased and improved for industrial purpose.

The market of an aromatic amino acid, L-phenylalanine (L-Phe) has shown great demand for its commercial value in pharmaceutical and food additives (Liu *et al.*, 2013). According to Bongaerts *et al*., (2001), L-Phenylalanine, is one of the most important commercially produced amino acids. The production of L-phenylalanine is mainly carried out by fermentations. Glucose or sucrose is always used as a carbon source in the process to produce L-Phenylalanine. L-phenylalanine is used as a precursor to vanillin production for food additives (Yin *et al*., 2013). Besides, L-Phenylalanine has been widely used in pharmaceutical industry as a nutritional

supplement. It has been acknowledged that this supplement acts as a precursor for the production of numerous catecholamines which refer to hormones in human body. Therefore, this study targets to optimize the production of L-Phenylalanine in Yeast *Saccharomyces cerevisiae* to meet the high demand of this strain in industries for use in various applications such as flavoring as well as the pharmaceutical field.

The present computational study can identify detailed systems biology and able to simulate metabolisms of bacteria and other microorganisms about their capability in producing metabolites such as H_2 and L-Phe in the mutant strain. This refers to the extensive search for reaction or gene to be added or deleted in order to increase the desired production. The intention of this research is to examine how metabolites production in microorganism can be increased using a systematic *in silico* simulation of metabolic engineering strategy, for instance, constraint-based modeling algorithms. Therefore, to elucidate interesting features of these microorganisms and identify engineering targets to achieve enhanced physiological properties of the strain, metabolic engineering which applies modeling simulation and optimization a with the involvement of reaction deletion strategy is applied.

1.2.2 Metabolic engineering

Metabolic engineering has shown big improvement and is becoming more popular in these recent years. Metabolic engineering has been used to study and manipulate the biological microbial cell metabolism by many researchers in this field. An example of a strategy that is introduced by metabolic engineering is to suggest any genes or reactions from its complex metabolic network to be deleted. This theory has shown many achievements towards microbial fuel cell (MFC) in addressing high yield of product or biochemical. The advancement of metabolic engineering has gained more attention as it is able to improve any desired metabolites strain that can further the process to become valuable products for market industries. With such results, more

developments of quantitative models and algorithms using computational simulation are increasing in recent years. In addition, simulation through computational is powerful and able to save costs and time in manipulating phenotypes to optimize the desired strain model. The goal of metabolic engineering is to develop effective algorithms in order to improve the metabolic capabilities in producing desired metabolites in microorganism for industrial purpose (Wiechert, 2002). In metabolic engineering, there are some recognized algorithms available that are widely in use and which can assist in stimulation to improve metabolite productions in genome-scale model, for example, optimization algorithm, modeling algorithm, and modeling framework as discussed in the next section.

1.2.2.1 Ant Colony Optimization Algorithm

Ant Colony Optimization (ACO) algorithm is a general search technique based on the population size. This algorithm is developed to solve difficult combinatorial problems. ACO is inspired by the behavior of real ants that always tends to find the best shortest path from a food source to the nest. The ant colonies deposited a chemical called pheromone along the trails they used in searching for food from the nest (Lin *et al.*, 2008). The density of pheromone become higher when there is more ants walking through the path, showing that the path is better in the context of distance. This is because ants tend to find the shortest path that leads to a better food source in a short time with promising food quality. The pheromone is used as a guide by other ants. Initially, ants move randomly to the food source while the next ant chooses one path and the higher density of pheromones have a higher probability to be chosen and in a period of time, a common path is finally be developed through positive feedback process. The behavior of this algorithm can be exemplified as illustrated in Figure 1.1.

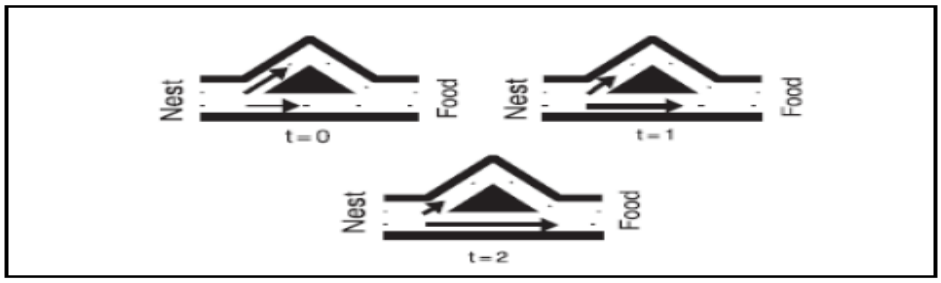

Figure 1.1 Basic Ant Colony Optimization Behavior at Different Time Stamps (Tavares *et al.*, 2011)

1.2.2.2 Flux Variability Analysis

Flux variability analysis (FVA) is an extension algorithm to further analyze the result obtained by Flux Balance Analysis (FBA) on the genome-scale metabolic network. FVA is a constraint-based algorithm applies for modeling simulation of the metabolic engineering cell. FBA is a linear programming algorithm that searches for the optimal value of the objective function in a genome-scale metabolic network. However, FBA has some limitation on it searching technique. This is caused by the presence of multiple pathways in the solution space which FBA is not set to consider the unique of the solution space. Thus, FVA is used to overcome this constraint. Different from FBA, FVA is able to determine maximum and minimum values of all available fluxes which satisfy the mentioned constraint. Technically, FVA is more advanced than FBA in terms of its modeling simulation process. Moreover, FVA calculates the flux value in a range which reflects to some percentage optimal value restrict. In this research, FVA is used to simulate the growth rate and production rate of microorganisms *Synechocystis* sp. PCC6803 and *Saccharomyces cerevisiae* of both wild-type and modified mutant models.

1.3 Background of the Problem

The advance in metabolic engineering strategies in enhancing many cell factories to produce metabolite of interest has been acknowledged in this era. However, there are some limitations of today's technology that have not been addressed yet. The challenges and problems faced in metabolic engineering are predominantly caused by the complexity of the metabolic network. This complexity issue affects the production of desirable metabolite as there are many competing reactions that are present, hence achieving a desired metabolic state through genetic modification remains difficult (Ohno *et al*., 2013). Apart from that, according to Liu *et al*. (2013) and Hallenbeck and Ghosh (2012) production of metabolites such as L-phenylalanine and biohydrogen from wild-type cell is currently low. This shows the lack of effective genome models which can predict and simulate low production yields as being an important aspect of handling in order to produce higher fitness of the objective function.

The aforementioned issues relate to the algorithmic used by previous researchers to study *in silico* metabolic engineering. With regards to the complexity of the metabolic network, this directly results in low production yield due to the presence of competing for non-desirable compounds. This can relate to the low accuracy of existing algorithms in constructing solutions to acquire the best value. There are numbers of developed algorithms used to simulate flux distribution of particular metabolites from cell factories of the genome-scale model that can directly be applicable for many industrial purposes either by single or hybrid algorithms. Previously, a hybrid of ant colony optimization (ACO) with minimization of metabolic adjustment (MOMA) or flux balance analysis (FBA) have been developed. However existing algorithms constructed are still lacking. This is because MOMA (Raman and Chandra, 2009) and FBA (Lee *et al*., 2006) have certain limitations in their flux distribution methodology. FBA shows no unique and sometimes unrealistic flux distribution value (Megchelenbrink *et al*., 2015) and unreliable in predicting the flux value of by-product (Khannapho *et al*., 2008). On the other hand, MOMA is weak at predicting the final steady state of growth rate and conditionally, unable to represent

the true metabolic state of the organism compared to FBA (Shlomi *et al*., 2005). ACO (Colorni *et al*., 1991) is superior to Continuous Bee Algorithm (CBA) and Artificial Bee Colony (ABC) (Karaboga, 2005) as it has several advantages over other evolutionary algorithms including offering positive feedback resulting in rapid solution finding and having distributed computation which avoids premature convergence (Ab Wahab *et al.*, 2015). Since ABC is a new algorithm, thus it requires new fitness test for the new parameters to improve the performance (Ab Wahab *et al.*, 2015). Bees algorithm also has several parameters that needs to be tuned before it can be used directly (Yuce *et al.*, 2013). These limitations also affect the lack of effective genome model constructed in order to generate better value for the objective function. In addition, wild-type model can only predict low flux value. These issues highlight the need for a new generation of biofuel technology and the strain production of L-phenylalanine (Delucchi, 2010; Melillo *et al.*, 2009; Khamduang *et al*. 2009). Thus, an improved modeling algorithm with new modified genome model is important to highlight. FVA outperform MOMA and FBA as it calculates the full range of flux distribution value while maintaining high growth rate and optimizing the objective function (Mahadevan and Schilling, 2003). In addition, FVA explores alternate solution in the optimal flux space as frequently there is not only one optimal flux distribution can be found (Hay and Schwender, 2014). This indicates that FVA methodology is more detailed and shows realistic flux value in predicting possible by-product production rates under maximal biomass production (Müller and Bockmayr, 2013).

In order to clarify the potential of metabolites strain production, several computational algorithms have been introduced. This is because *in silico* or computational simulation is preferred as less time is required, no labor involved, less research expenditure, and eventually, cost reduction (Salleh *et al*., 2015). In this study, a hybrid of ACO and FVA is proposed to overcome the current limitations faced by other algorithms in order to enhance the microbial secondary metabolites production and strain improvement. ACO is used for the optimization process in order to prevent the solution from being trapped in local optima that can cause premature convergence while FVA is applied to calculate precisely the flux distribution value of the compound in a genome model.

1.4 Problem Statement

As previously discussed, the production of metabolites from microorganisms is in demand and to fulfill the urge of biohydrogen and L-Phenylalanine production. Briefly, the innate potential for higher production of desired metabolites is obscure and relates to the lack of effective genome models. This is mainly due to the fact that in producing metabolites by microorganisms, the presence of interaction among thousands of reactions in the metabolic network caused complex and higher dimensional data size (Liu *et al*., 2010; Price *et al*., 2004; Stephanopoulos and Simpson, 1997; Toya and Shimizu, 2013; Wittmann and Lee, 2012). Complex network is caused by the intracellular relationship between reactions, genes, metabolites. In addition, the presence of competing pathways of the non-desirable product may affect the desired metabolite production. Thus, the best way to enhance the desired production beyond the wild type limit is through the computational simulation implementation. Figure 1.2 shows the complex metabolic network which involves interaction between pathway, genes, and metabolites.

Computational simulation involves modeling, optimization, and simulation are used to perform the modification of the cell network and to obtain an important insight about the metabolic system. Computational time is increasing as the problem size increases, thus some computationally pre-process steps are required, which match the biological theory in having a more suitable and compatible data. Other than that, the models need to undergo an optimization process in order to prevent the solution from being trapped in local optima, causing premature convergence. ACO is used as an optimization algorithm as it is better at exploring and constructing a good solution in a short time. This algorithm is efficient in avoiding premature convergence compared to other existing algorithms. FVA as a modeling simulation algorithm for simulating the entire flux distribution of a cell in the range of maximum to minimum of the solutions constructed by the optimization algorithm of a particular objective function. The hybrid of both algorithms with applying reactions deletion strategy is applied for this research to overcome current limitations in metabolic engineering study and to enhance the metabolites production of a strain model.

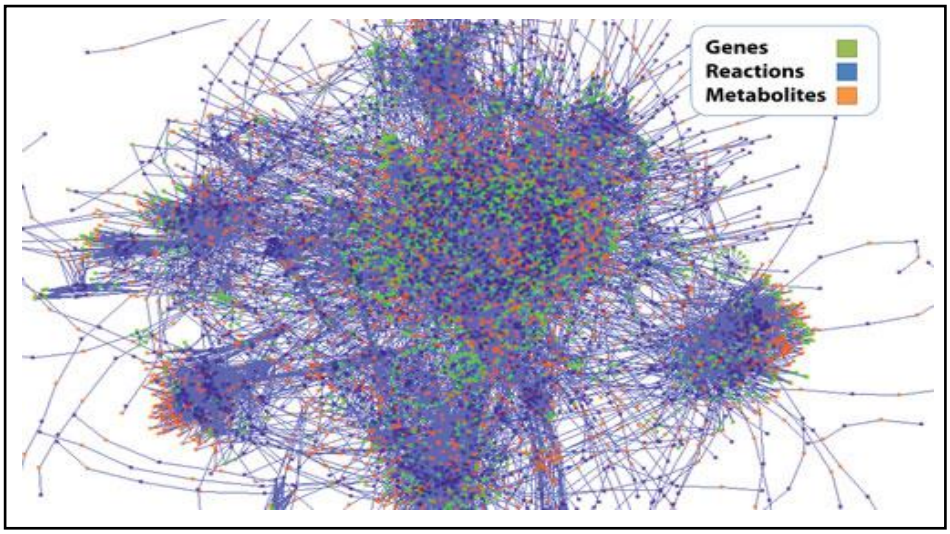

Figure 1.2 Graph theoretical of a complex metabolic network (Genetic Engineering and Biotechnology News, 2016).

1.5 Aim

The aim of this research is to propose a hybrid of ACO and FVA to reduce the complexity of metabolic network by identifying a near-optimal combination of reactions to be knocked out in order to improve the production of L-Phenylalanine and biohydrogen.

1.6 Objectives

Several objectives have been identified in order to achieve the aim:

1. To model a hybrid algorithm of Ant Colony Optimization and Flux Variability Analysis for metabolites production in a genome-scale metabolic model of microorganisms.

2. To develop a hybrid algorithm of Ant Colony Optimization and Flux Variability Analysis to solve the complex metabolic network issue by identifying a list of reactions to be knocked out for improving the production rate and Biomass-Product Coupled Yield (BPCY) of L-phenylalanine and biohydrogen in microbial strains.

3. To evaluate the experimental results (BPCY, production rate, and growth rate) of each metabolites with previous works and validate the list of reactions and genes deletion through biological database.

1.7 Scopes of Study

The scopes of this study are listed as below:

i. L-Phenylalanine in Yeast and Biohydrogen in Cyanobacteria are the products being focused.

ii. Two datasets going to be used are:

a) Model yeast4.05.xml downloaded from published literature by *Nogales et al.*, (2010) derived from http://sourceforge.net/projects/yeast/files/yeast_4.05.zip/download
b) Model *iJN678.xml* from published literature by Sohn *et al*., (2010) derived from bigg.ucsd.edu/models/iJN678

iii. The format of the dataset is in System Biology Markup Language (SBML). SBML is representing models of biological processes based on XML, which is a readable language machine. SBML contains metabolic networks, cell-signaling pathways, regulatory networks, and many kinds of systems. It is a standard language for exchanging information, storing and also fitting parameters for any given experimental data. These features play important roles in the modification of accurate genome model.

iv. The software used in this research is Constraints Based Reconstruction and Analysis (COBRA) Toolbox for constraint-based modeling which is implemented in MATLAB to perform the hybrid algorithm.

v. A hybrid of Ant Colony Optimization and Flux Variability Analysis (ACOFVA) is used in this research which involves and consider reaction deletion strategy and regulatory networks information.

1.8 Significance of study

This research is conducted to simulate and analyze the effect of genome metabolic modification of microorganism in order to improve the production of desired biohydrogen and L-Phenylalanine by implementing a computational simulation. The following points shows the significant of conducting this research:

i. Investigate the potential improvement of metabolites production in the microbial cell factories.

ii. Give a clear insight of biohydrogen and L-Phenylalanine production from microbial cell factory by using computational modeling and analysis that can provide a better understanding of cellular level function.

iii. Developments of *in silico* modeling of microbial cell factory that enable the optimization of interest production yield with better prediction.

iv. Researchers and biologists can use this information to do laboratory experiment using the constructed metabolic model as references towards a more promising production in time effective manner.

1.9 Thesis Organization

i) Chapter 1 presents a brief introduction to metabolites production and metabolic engineering included Ant Colony Optimization and Flux Variability Analysis. The background of the problems which refers to the existing issue in the related field and also problem statements of this study are described in detail. Some important points for this study included the aim, objectives, and scopes are also presented precisely.

ii) Chapter 2 presents some literature reviews retrieved from published journals and other available sources on the existing algorithms that are used in analyzing the genome-scale metabolic model. Besides, some details about metabolic engineering that consist of some different groups related to it such as constraint-based analysis, optimization algorithm, and modeling framework are discussed comprehensively. Reading materials

that relate to this research topic with beneficial and helpful information, such as journals, articles, and conference working papers are listed, too.

iii) Chapter 3 discusses the research methodology as a planning form used to conduct this research. The detailed descriptions of activities involved are presented and divided by particular phases for easy following. The information about the data set chosen to be used is clarified. Besides, basic requirements of hardware and software and performance measurement that is used for this research are clearly described. The proposed algorithm is presented in this chapter.

iv) Chapter 4 explains and discusses the flow of implementation of the proposed hybrid algorithm, which is ACOFVA. Pre-processing step and preparation of the two data sets chosen are performed in this chapter. The designed steps involved in ACOFVA is also discussed. The results of the proposed hybrid algorithm with the explanation about the reactions and genes suggested to be deleted in order to increase the production of metabolites are also included in this presentation.

v) Chapter 5 summarizes the contents of all previously discussed chapters. Conclusion, contributions, and limitations of this research are also being discussed.

REFERENCES

Ab Goldstein, Y., and Bockmayr, A. (2015). Double and multiple knockout simulations for genome-scale metabolic network reconstructions. *Algorithms for Molecular Biology*, *1*(10), 1-10.

Ab Wahab, M. N., Nefti-Meziani, S., and Atyabi, A. (2015). A comprehensive review of swarm optimization algorithms. *PloS one*, *10*(5), e0122827.

Adachi, E., Torigoe, M., Sugiyama, M., Nikawa, J.-I., and Shimizu, K. (1998). Modification of metabolic pathways of Saccharomyces cerevisiae by the expression of lactate dehydrogenase and deletion of pyruvate decarboxylase genes for the lactic acid fermentation at low pH value. *Journal of fermentation and bioengineering*, *86*(3), 284-289.

Alper, H., Jin, Y.-S., Moxley, J., and Stephanopoulos, G. (2005). Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli. *Metabolic engineering*, *7*(3), 155-164.

Baart, G. J., and Martens, D. E. (2012). Genome-scale metabolic models: reconstruction and analysis. *Neisseria meningitidis: Advanced Methods and Protocols*, 107-126.

Becker, S. A., Feist, A. M., Mo, M. L., Hannum, G., Palsson, B. Ø., and Herrgard, M. J. (2007). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature protocols*, *2*(3), 727-738.

Bonarius, H. P., Hatzimanikatis, V., Meesters, K. P., de Gooijer, C. D., Schmid, G., and Tramper, J. (1996). Metabolic flux analysis of hybridoma cells in different culture media using mass balances. *Biotechnology and bioengineering*, *50*(3), 299-318.

Bongaerts, J., Krämer, M., Müller, U., Raeven, L., and Wubbolts, M. (2001). Metabolic engineering for microbial production of aromatic amino acids and derived compounds. *Metabolic engineering, 3*(4), 289-300.

Borodina, I., Kildegaard, K. R., Jensen, N. B., Blicher, T. H., Maury, J., Sherstyk, S., et al. (2015). Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine. *Metabolic engineering, 27*, 57-64.

Boyle, N. R., and Morgan, J. A. (2009). Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC systems biology, 3*(1), 1.

Brochado, A. R., Matos, C., Møller, B. L., Hansen, J., Mortensen, U. H., and Patil, K. R. (2010). Improved vanillin production in baker's yeast through in silico design. *Microbial cell factories, 9*(1), 1.

Burgard, A. P., Pharkya, P., and Maranas, C. D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and bioengineering, 84*(6), 647-657.

Bushell, M. E., Sequeira, S. I., Khannapho, C., Zhao, H., Chater, K. F., Butler, M. J., et al. (2006). The use of genome scale metabolic flux variability analysis for process feed formulation based on an investigation of the effects of the zwf mutation on antibiotic production in Streptomyces coelicolor. *Enzyme and Microbial Technology, 39*(6), 1347-1353.

Chemler, J. A., Fowler, Z. L., McHugh, K. P., and Koffas, M. A. (2010). Improving NADPH availability for natural product biosynthesis in Escherichia coli by metabolic engineering. *Metabolic engineering, 12*(2), 96-104.

Chen, Q., Wang, Z., and Wei, D. (2010). Progress in the applications of flux analysis of metabolic networks. *Chinese Science Bulletin, 55*(22), 2315-2322.

Chongsuksantikul, A., Asami, K., Yoshikawa, S., and Ohtaguchi, K. (2015). Dark Anaerobic Hydrogen Production in the Mutants of Synechocystis sp. Strain PCC6803-GT Defective in Lactate Dehydrogenase Activity and/or Alcohol Dehydrogenase Activity, Incubated in Buffer Solutions With or Without Glucose. *International Journal of Biology, 7*(1), 33.

Choon, Y. W., Mohamad, M. S., Deris, S., and Illias, R. M. (2014). A hybrid of bees algorithm and flux balance analysis (BAFBA) for the optimisation of microbial

strains. *International Journal of Data Mining and Bioinformatics 9, 10*(2), 225-238.

Colorni, A., Dorigo, M., and Maniezzo, V. (1991). *Distributed optimization by ant colonies.* Paper presented at the Proceedings of the first European conference on artificial life, 134-142.

Delucchi, M. A. (2010). Impacts of biofuels on climate change, water use, and land use. *Annals of the New York Academy of Sciences, 1195*(1), 28-45.

Demirbaş, A. (2002). Fuel characteristics of olive husk and walnut, hazelnut, sunflower, and almond shells. *Energy Sources, 24*(3), 215-221.

Dobson, P. D., Smallbone, K., Jameson, D., Simeonidis, E., Lanthaler, K., Pir, P., et al. (2010). Further developments towards a genome-scale metabolic model of yeast. *BMC Systems Biology, 4*(1), 1.

Dorigo, M., Maniezzo, V., Colorni, A., and Maniezzo, V. (1991). Positive feedback as a search strategy.

Edwards, J. S., Covert, M., and Palsson, B. (2002). Metabolic modelling of microbes: the flux-balance approach. *Environmental microbiology, 4*(3), 133-140.

Fang, W., Sun, J., Ding, Y., Wu, X., and Xu, W. (2010). A review of quantum-behaved particle swarm optimization. *IETE Technical Review, 27*(4), 336-348.

Genetic Engineering and Biotechnology News (2016). Available from: <http://www.genengnews.com/Media/images/Article/UnivVA_MetabolicNetwork7514635831.jpg>.

Grafahrend-Belau, E., Schreiber, F., Koschützki, D., and Junker, B. H. (2009). Flux balance analysis of barley seeds: a computational approach to study systemic properties of central metabolism. *Plant Physiology, 149*(1), 585-598.

Gudmundsson, S., and Thiele, I. (2010). Computationally efficient flux variability analysis. *BMC bioinformatics, 11*(1), 489.

Haak, D., Gable, K., Beeler, T., and Dunn, T. (1997). Hydroxylation of Saccharomyces cerevisiae ceramides requires Sur2p and Scs7p. *Journal of Biological Chemistry, 272*(47), 29704-29710.

Hallenbeck, P. C., and Ghosh, D. (2012). Improvements in fermentative biological hydrogen production through metabolic engineering. *Journal of environmental management, 95*, S360-S364.

Hay, J., and Schwender, J. (2011). Metabolic network reconstruction and flux variability analysis of storage synthesis in developing oilseed rape (Brassica napus L.) embryos. *The Plant Journal, 67*(3), 526-541.

Höffner, K., Harwood, S., and Barton, P. (2013). A reliable simulator for dynamic flux balance analysis. *Biotechnology and bioengineering, 110*(3), 792-802.

Karaboga, D. (2005). *An idea based on honey bee swarm for numerical optimization*: Technical report-tr06, Erciyes university, engineering faculty, computer engineering departmento. Document Number)

Karaboga, D., and Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of global optimization, 39*(3), 459-471.

Kauffman, K. J., Prakash, P., and Edwards, J. S. (2003). Advances in flux balance analysis. *Current opinion in biotechnology, 14*(5), 491-496.

Kennedy, J. (2011). Particle swarm optimization. In *Encyclopedia of machine learning* (pp. 760-766): Springer.

Khamduang, M., Packdibamrung, K., Chutmanop, J., Chisti, Y., and Srinophakun, P. (2009). Production of L-phenylalanine from glycerol by a recombinant Escherichia coli. *Journal of industrial microbiology and biotechnology, 36*(10), 1267-1274.

Khannapho, C., Zhao, H., Bonde, B. K., Kierzek, A. M., Avignone-Rossa, C. A., and Bushell, M. E. (2008). Selection of objective function in genome scale flux balance analysis for process feed development in antibiotic production. *Metabolic engineering*, 10(5), 227-233.

Kim, J., and Reed, J. (2010). OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. BMC Systems Biology, 4(1), 53.

Kim, B., Kim, W. J., Kim, D. I., and Lee, S. Y. (2015). Applications of genome-scale metabolic network model in metabolic engineering. *Journal of industrial microbiology and biotechnology, 42*(3), 339-348.

Kleessen, S., and Nikoloski, Z. (2012). Dynamic regulatory on/off minimization for biological systems under internal temporal perturbations. *BMC Systems Biology*, 6(1), 16.

Koukourakis, M. I., Giatromanolaki, A., Sivridis, E., Gatter, K. C., Harris, A. L., Tumor, et al. (2005). Pyruvate dehydrogenase and pyruvate dehydrogenase kinase expression in non small cell lung cancer and tumor-associated stroma. *Neoplasia, 7*(1), 1-6.

Kuepfer, L., Sauer, U., and Blank, L. M. (2005). Metabolic functions of duplicate genes in Saccharomyces cerevisiae. *Genome research, 15*(10), 1421-1430.

Lee, J. M., Gianchandani, E. P., and Papin, J. A. (2006). Flux Balance Analysis in the Era of Metabolomics. *Briefings in Bioinformatics*, 7(2), 140-150.

Lin, B.M.T., Lu, C.Y., Shyu, S.J. and Tsai, C.Y. (2008). Development of new features of Ant Colony Optimization for flowshop scheduling. *International Journal of Production Economics*. 112:742-755.

Lincoln, S., Rogers, I., and Srivastava, R. (2015). Metabolic Design And Engineering Through Ant Colony Optimization. In *Proceedings of the 2015 on Genetic and Evolutionary Computation Conference* (225-232).

Liu, L., Agren, R., Bordel, S., and Nielsen, J. (2010). Use of genome-scale metabolic models for understanding microbial physiology. *FEBS letters, 584*(12), 2556-2564.

Liu, S., Xiao, M., Zhang, L., Xu, J., Ding, Z., Gu, Z., and Shi, G. (2013). Production of l-phenylalanine from glucose by metabolic engineering of wild type Escherichia coli W3110. *Process Biochemistry*, 48(3), 413-419.

Long, M. R., Ong, W. K., and Reed, J. L. (2015). Computational methods in metabolic engineering for strain design. *Current opinion in biotechnology, 34*, 135-141.

Lu, S. J., Salleh, A. H. M., Mohamad, M. S., Deris, S., Omatu, S., and Yoshioka, M. (2014). Identification of gene knockout strategies using a hybrid of an ant colony optimization algorithm and flux balance analysis to optimize microbial strains. *Computational biology and chemistry, 53*, 175-183.

Machado, D., and Herrgård, M. (2015). Co-evolution of strain design methods based on flux balance and elementary mode analysis. *Metabolic Engineering Communications*.

Magnani, G., Lomazzi, M., and Peracchi, A. (2013). Completing the folate biosynthesis pathway in Plasmodium falciparum: p-aminobenzoate is produced by a highly divergent promiscuous aminodeoxychorismate lyase. *Biochemical Journal, 455*(2), 149-155.

Mahalik, S., Sharma, A., and Mukherjee, K. (2014). Genome engineering for improved recombinant protein expression in Escherichia coli. *Microb Cell Fact*, 13(1).

Mahadevan, R., and Schilling, C. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic engineering*, *5*(4), 264-276.

Mahadevan, R., Edwards, J. and Doyle, F. (2002). Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli. *Biophysical Journal*, 83(3), 1331-1340.

Mao, L. and Verwoerd, W.S. (2013). Genome-Scale Stoichiometry Analysis to Elucidate the Innate Capability of the Cyanobacterium Synechocystis for Electricity Generation. *Journal of industrial microbiology and biotechnology*, 40(10), 1161-1180.

McAlister, L., and Holland, M. J. (1982). Targeted deletion of a yeast enolase structural gene. Identification and isolation of yeast enolase isozymes. *Journal of Biological Chemistry*, *257*(12), 7181-7188.

McEwen, J. T., Machado, I. M., Connor, M. R., and Atsumi, S. (2012). Engineering *Synechococcus elongatus* PCC7942 to grow continuously in diurnal conditions. *Applied and environmental microbiology*, AEM-03326.

Megchelenbrink, W., Rossell, S., Huynen, M. A., Notebaart, R. A., and Marchiori, E. (2015). Estimating Metabolic Fluxes Using a Maximum Network Flexibility Paradigm. *PloS one*, 10(10), e0139665.

Melillo, J., Reilly, J., Kicklighter, D., Gurgel, A., Cronin, T., and Paltsev, S. et al. (2009). Indirect Emissions from Biofuels: How Important?. Science, 326(5958), 1397-1399.

Mienda, B. S., Shamsir, M. S., Shehu, I., Deba, A. A., and Galadima, I. A. (2014). In silico metabolic engineering interventions of Escherichia coli for enhanced ethanol production, based on gene knockout simulation. *Gene*,*23*(24), 25.

Mienda, B. S., Shamsir, M. S., and Salleh, F. M. (2015). In silico Evaluation of the Effect of pfl Gene Knockout on the Production of D-lactate by Escherichia coli Genome Scale Model using the OptFlux Software Platform. *Indian Journal of Science and Technology*, *8*(2), 172-177.

Montagud, A., Navarro, E., de C'ordoba, P., Urchuegu'ia, J. and Patil, K. (2010). Reconstruction and Analysis of Genome-Scale Metabolic Model of a Photosynthetic Bacterium. *BMC systems biology*, 4(1), 156.

Moradi, S., Fatahi, L., and Razi, P. (2010). Finite element model updating using Bees Algorithm. *Struct Multidisc Optim*. 42:283-291.

Müller, A. C., and Bockmayr, A. (2013). Fast thermodynamically constrained flux variability analysis. *Bioinformatics*, btt059.

Navarro, E., Montagud, A., De Córdoba, P. F., and Urchueguía, J. (2009). Metabolic flux analysis of the hydrogen production potential in Synechocystis sp. PCC6803. *international journal of hydrogen energy*, *34*(21), 8828-8838.

Nielsen, J. (2003). It is all about metabolic fluxes. *Journal of Bacteriology*, *185*(24), 7031-7035.

Nielsen, J. (2013). Production of biopharmaceutical proteins by yeast: advances through metabolic engineering. *Bioengineered*, *4*(4), 207-211.

Niu, K., Zhang, X., Tan, W.-S., and Zhu, M.-L. (2010). Characteristics of fermentative hydrogen production with Klebsiella pneumoniae ECU-15 isolated from anaerobic sewage sludge. *international journal of hydrogen energy*, *35*(1), 71-80.

Nogales, J., Gudmundsson, S., Knight, E., Palsson, B., and Thiele, I. (2012). Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings Of The National Academy Of Sciences*, 109(7), 2678-2683.

Nogales, J., Gudmundsson, S., and Thiele, I. (2013). Toward systems metabolic engineering in cyanobacteria. Bioengineered. 4: 158-163.

Neto, R. T., and Godinho Filho, M. (2013). Literature review regarding Ant Colony Optimization applied to scheduling problems: Guidelines for implementation and directions for future research. *Engineering Applications of Artificial Intelligence*,*26*(1), 150-161.

Norfadzlan,Y., Azlan, M.Z. and Siti Zaiton, M.H. (2012). Evolutionary techniques in optimizing machining parameters: Review and recent applications (2007-2011). Expert Systems with Applications. 39:9909-9927.

Ohno, S., Shimizu, H., and Furusawa, C. (2014). FastPros: screening of reaction knockout strategies for metabolic engineering. *Bioinformatics*, *30*(7), 981-987.

Orth, J. D., Thiele, I. and Palsson, B. (2010). What Is Flux Balance Analysis. *Nature Biotechnology*. 28, 245-248.

Özbakir, L., Baykasoglu, A., and Tapkan, P. (2010). Bees Algorithm for generalized assignment problem. Elsevier. 3782-3795.

Patil, K., Rocha, I., Förster, J. and Nielsen, J. (2005). Evolutionary Programming as a Platform for *In Silico* Metabolic Engineering. *BMC bioinformatics*, 6(1), 308.

Pham, D. T., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S., and Zaidi, M. (2006). The Bees Algorithm-A novel tool for complex optimization problems. *Intelligent Production Machine and Systems.* 454-459.

Pham, D., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S., and Zaidi, M. (2011). *The bees algorithm–a novel tool for complex optimisation.* Paper presented at the Intelligent Production Machines and Systems-2nd I* PROMS Virtual International Conference (3-14 July 2006).

Price, N., Papin, J., Schilling, C. and Palsson, B. (2003). Genome-Scale Microbial In Silico Models: The Constraints-Based Approach. *Trends in biotechnology*, 21(4), 162-169.

Price, N. D., Reed, J. L., and Palsson, B. Ø. (2004). Genome-scale models of microbial cells: evaluating the consequences of constraints. Nature Reviews Microbiology, 2(11), 886-897.

Raman, K. and Chandra, N. (2009). Flux Balance Analysis of Biological Systems: Applications and challenges. *Briefings in Bioinformatics*. 10(4): 435-449.

Ramírez-Morales, J. E., Tapia-Venegas, E., Toledo-Alarcón, J., and Ruiz-Filippi, G. (2015). Simultaneous production and separation of biohydrogen in mixed culture systems by continuous dark fermentation. *Water Science and Technology*, *71*(9), 1271-1285.

Randazzo, A. (2012). Swarm Optimization Methods in Microwave Imaging. *International Journal of Microwave Science and Technology*, 2012, 1-12.

Rocha, I., Maia, P., Evangelista, P., Vilaça, P., Soares, S., Pinto, J. P., ... and Rocha, M. (2010). OptFlux: an open-source software platform for in silico metabolic engineering. *BMC systems biology*, *4*(1), 45.

Roeva, O., Fidanova, S., and Paprzycki, M. (2015). Population size influence on the genetic and ant algorithms performance in case of cultivation process modeling. *In Recent Advances in Computational Optimization* (pp. 107-120).

Salleh, A. H. M., Mohamad, M. S., Deris, S., Omatu, S., Fdez-Riverola, F., and Corchado, J. M. (2015). Gene knockout identification for metabolite

production improvement using a hybrid of genetic ant colony optimization and flux balance analysis. *Biotechnology and Bioprocess Engineering*. 20: 685-693.

San Chua, P., Salleh, A. H. M., Mohamad, M. S., Deris, S., Omatu, S., and Yoshioka, M. (2015). Identifying a gene knockout strategy using a hybrid of the bat algorithm and flux balance analysis to enhance the production of succinate and lactate in Escherichia coli. *Biotechnology and Bioprocess Engineering*, *20*(2), 349-357.

Schuster, S., Pfeiffer, T., and Fell, D. A. (2008). Is maximization of molar yield in metabolic networks favoured by evolution? *Journal of theoretical biology*, *252*(3), 497-504.

Segre, D., Vitkup, D. and Church, G. (2002). Analysis of Optimality in Natural and Perturbed Metabolic Networks. *Proceedings of the National Academy of Sciences*, 99(23), 15112-15117.

Shastri, A. A., and Morgan, J. A. (2005). Flux balance analysis of photoautotrophic metabolism. *Biotechnology progress*, *21*(6), 1617-1626.

Shlomi, T., Berkman, O. and Ruppin, E. (2005). Regulatory on/off Minimization of Metabolic Flux Changes After Genetic Pertubations. *Proc Natl Acad Sci USA*. 102: 7695-7700.

Sivanandam, S.N. and Deepa, S.N. (2008). Introduction to Particle Swarm Optimization and Ant Colony Optimization. *Introduction to Genetic Algorithms*. 1:403-424.

Stephanopoulos, G., and Simpson, T. W. (1997). Flux amplification in complex metabolic networks. *Chemical Engineering Science*, *52*(15), 2607-2627.

Steuer, R., Knoop, H., and Machne, R. (2012). Modelling cyanobacteria: from metabolism to integrative models of phototrophic growth. *Journal Of Experimental Botany*, *63*(6), 2259-2274.

Tanniche, I., Senger, R. S., Yen, J. Y., Fisher, A. K., Gillaspy, G. E., and Bevan, D. R. (2015). Designing metabolic engineering strategies with genome-scale metabolic flux modeling. *Advances in Genomics & Genetics*, *5*.

Tavares Neto, R.F. and Godinho Filho, M. (2011). An Ant Colony Optimization approach to a permutational flowshop Scheduling Problem with outsourcing allowed. *Computers and Operations Research*. 38:1286-1293.

Tokuhiro, K., Ishida, N., Nagamori, E., Saitoh, S., Onishi, T., Kondo, A., et al. (2009). Double mutation of the PDC1 and ADH1 genes improves lactate production in the yeast Saccharomyces cerevisiae expressing the bovine lactate dehydrogenase gene. *Applied microbiology and biotechnology, 82*(5), 883-890.

Toya, Y., and Shimizu, H. (2013). Flux analysis and metabolomics for systematic metabolic engineering of microorganisms. *Biotechnology advances, 31*(6), 818-826.

Urade, H.S. and Patel, R. (2011). Study and analysis of Particle Swarm Optimization: A review. *IJCA Proceedings on 2nd National Conference on Information and Communiication Technology NCICT*. 4:1-5.

van Buuringen, N., Janssen, S., Man, M., and Noorlander, L. Flux-balance analysis of vinblastine production. *Radboud Honours Academy FNWI*, 17.

Varma, A., and Palsson, B. O. (1994). Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. *Bio/technology*, *12*.

Veziroglu, A., and Macario, R. (2011). Fuel cell vehicles: State of the art with economic and environmental concerns. *International Journal of Hydrogen Energy*, 36(1), 25-43.

Walker, M. E., Val, D. L., Rohde, M., Devenish, R. J., and Wallace, J. C. (1991). Yeast pyruvate carboxylase: identification of two genes encoding isoenzymes. *Biochemical and biophysical research communications, 176*(3), 1210-1217.

Wiechert, W. (2002). Modeling and Simulation: Tools for Metabolic Engineering. *Journal of Biotechnolog.* 94(1): 37-63.

Wittmann, C., and Lee, S. Y. (2012). *Systems metabolic engineering*: Springer Science & Business Media.

Yan, C., and Xu, X. (2008). Bifunctional enzyme FBPase/SBPase is essential for photoautotrophic growth in cyanobacterium Synechocystis sp. PCC 6803. *Progress in Natural Science*, 18(2), 149-153.

Yang, X. S., Deb, S., Loomes, M., and Karamanoglu, M. (2013). A framework for self-tuning optimization algorithm. *Neural Computing and Applications*, 23(7-8), 2051-2057.

Yin, L. H., Choon, Y. W., Chai, L. E., Chong, C. K., Deris, S., Illias, R. M., et al. (2013). Prediction of vanillin production in yeast using a hybrid of continuous

bees algorithm and flux balance analysis (CBAFBA). In *Advances in Biomedical Infrastructure 2013* (pp. 101-116): Springer.

Xu, Z., Zheng, P., Sun, J., and Ma, Y. (2013). ReacKnock: Identifying Reaction Deletion Strategies for Microbial Strain Optimization Based on Genome-Scale Metabolic Network. *Plos ONE*, 8(12), e72150.

Xue, X.D., Cheng, X.D., Xu, B., Wang, H.L. and Jiang, C.P. (2010). The basic principle and application of Ant Colony Optimization algorithm. *Artificial Intelligence and Education (ICAIE) International Conference*. 358-360.

Yaseen, S. G., and Al-Slamy, N. M. (2008). Ant colony optimization. *IJCSNS*,*8*(6), 351.

Yin, L., Choon, Y., Mohamad, M., Chai, L., Chong, C., and Abdullah, A. et al. (2013). Prediction of Vanillin and Glutamate Productions in Yeast Using a Hybrid of Continuous Bees Algorithm and Flux Balance Analysis (CBAFBA). *CBIO*, 9(3), 284-294.

Yuce, B., Packianather, M. S., Mastrocinque, E., Pham, D. T., and Lambiase, A. (2013). Honey bees inspired optimization method: the Bees Algorithm. *Insects*, *4*(4), 646-662.

Zaremberg, V., and McMaster, C. R. (2002). Differential partitioning of lipids metabolized by separate yeast glycerol-3-phosphate acyltransferases reveals that phospholipase D generation of phosphatidic acid mediates sensitivity to choline-containing lysolipids and drugs. *Journal of Biological Chemistry*, *277*(41), 39035-39044.

Zhao, Q. and Kurata, H. Comparison of the Prediction Abilities of FBA, MOMA and ROOM for a pykF Mutant of E. coli. *17th International Conference on Genome Informatics. December 18-20. Pacifico Yokohama, Japan*, 107-1.