

**METABOLIC NETWORK ANALYSIS OF MYCO-BACTERIUM
TUBERCULOSIS (MTB)**

BY

ACHAS MOSES JOY

13PCG00482

**AN MSc RESEARCH PROJECT REPORT SUBMITTED TO THE
DEPARTMENT OF COMPUTER AND INFORMATION SCIENCE,
SCHOOL OF POST GRADUATE STUDIES,
COVENANT UNIVERSITY, OTA,
OGUN STATE.**

**IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE
AWARD OF MASTER OF SCIENCE (M.Sc.) DEGREE IN COMPUTER
SCIENCE.**

2016

CERTIFICATION

We, the undersigned certify that this project work was carried out by **ACHAS MOSES JOY** in partial fulfilment of the requirement for the award of Master of Science (Honours) Degree in Computer Science, Covenant University, Ota, Nigeria

Dr O. J. Oyelade
(Supervisor)

Date

Dr. A. A. Adebiyi
(HOD, Computer and Information Sciences)

Date

DECLARATION

It is hereby declared that this research was undertaken by ACHAS MOSES JOY which is based on the original study in the department of Computer and Information Science, College of Science and Technology, Covenant University, Ota. Under the supervision of Dr O.J. Oyelade. Ideas and views in this research work are products of the original research undertaken by ACHAS Moses Joy and the views of other researchers used in this research has been referenced and acknowledged.

Dr. O.J Oyelade

Supervisor

Signature.....

Date:

DEDICATION

This work is dedicated to the Almighty God, the very purpose of my existence, the one who was, who is and is to come for the privilege to begin and finish this phase of my life.

ACKNOWLEDGEMENTS

Foremost, I want to express my deepest appreciation to my Supervisor, Dr O.J. Oyelade and his assistant Mrs. Itunuoluwa Isewon for their patience, motivation, and support from the beginning to the end of this research work; they've shown the attitude and the substance of a Mentor.

A special thanks to my family members, my parents Pst and Mrs Atunde for their prayers and support. I express my sincere gratitude to the light of Hope Orphanage family for their support.

I appreciate the Vice Chancellor Covenant University, Prof C.K Ayo, Head of Department Computer and Information Science, Dr A.A Adebiyi and every member of CUBRE, Covenant University for their support during my stay at Covenant University.

In addition, I thank the staff of Bells University of Technology , College of Information and Communications Technology, for their support and encouragement towards the success of this Masters Program. I also appreciate my colleagues, Mrs Adeyinka Bolaji, Mr Salami Fatai, Miss Paula Fiddi and Mr Bayo Omotosho who kept giving all the support they could both in kind, effort and in counsel.

TABLE OF CONTENTS

Title Page	i
Certification	ii
Declaration	iii
Dedication	iv
Acknowledgements	v
Table of Content	vi
List of Figures	ix
List of Table	x
List of Acronyms	xi
Abstract	xiv
Chapter One: Introduction	
1.1 Background	1
1.1.1 Biochemical Metabolic Network	1
1.1.2 Flux Balance Analysis	2
1.1.3 Underlying Structure of Flux Balance Analysis	4
1.1.4 Importance of Flux Balance Analysis	5
1.2 Statement of the Problem	6
1.3 Aim and Objective of The Study	7
1.4 Research Methodology	9
1.5 Significance of Study	10
1.6 Organization of Thesis	10

Chapter Two: Literature Review

2.1 Overview of Metabolic Pathways and Metabolic Networks	11
2.2. Reason for Modelling Metabolic Network	12
2.2.1 Approaches to Modelling Metabolic Networks	13
2.3 Metabolic Network Reconstruction	15
2.4 Overview of Flux Balance Analysis	15
2.4.1 Formulation of Flux Balance Analysis	16
2.4.2 Evolution of Flux Balance Analysis	19
2.4.3 Flux Balance Analysis Is Constraints Based	20
2.5 Optimizing Phenotypes	22
2.6 Constraints of Flux Balance Analysis	24
2.7 Stoichiometric Matrix	28
2.8 Flux Balance Analysis of Metabolic System	31

Chapter Three: Research Methodology

3.1 Reaction Network Formulation Stage	35
3.2 Flux Balance Analysis Formulation Stages	37
3.3 Extreme Pathways Algorithm	42
3.3.1 Algorithm Notations	48
3.4 Hypothetical Flux Distribution Stage	49
3.5 The Simplex Method	49

Chapter Four: System Implementation

4.1 Implementation Tool	52
4.2 Flux Balance Analysis Implementation	52
4.2.1 Performing Extreme Pathway Analysis	53
4.2.2 Flux Optimization	53
4.3 Result	54
4.3.1 Extreme Pathways Result	59
4.3.2 Flux Optimization	59
4.3.2.1 Unique Active Metabolites	61
4.3.2.2 Common Active Metabolite between Biomass and Optimization Result	63
4.3.2.3 De-Activated Objective Flux Metabolites	63

Chapter Five: Conclusion and Recommendation

5.1 Conclusion	67
5.2 Recommendation	67
5.3 Future Work	68
References	69
Appendix	73
Biomass Component used for Optimization Test	73
Optimization Result	75
Source Code	78

List of Figures

Figure 1.1: Overview of a flux balance system	8
Figure 2.1: A visual representation of a small reaction network	14
Figure 2.2: A sketch out of key steps entailed in FBA	18
Figure 2.3: FBA development time line	21
Figure 2.4: The basic concept of constraint based modelling	25
Figure 2.5: Genome scale metabolic stoichiometrically balanced Biochemical reactions transformed into Matrix	26
Figure 2.6: A system of linear equations at steady state	27
Figure 2.7: A biological objective phenotype	27
Figure 2.8: System optimization using linear programming	30
Figure 3.1: Procedure to develop an fba analysis model	37
Figure 3.2: How FBA applies to a metabolic network	41
Figure 3.3: A sample metabolic network with 5 metabolites and 10 reactions	42
Figure 3.4: Stoichiometric matrix of the biological in 3.5	43
Figure 3.5: Flow chart of a simplex method approach linear problem	51
Figure 4.1: Interface of developed system	55
Figure 4.2: Basic constraints setting interface	56
Figure 4.3: Result of Extreme pathway output interface	57
Figure 4.4: optimization using extreme pathways and objective metabolites	58
Figure 4.5: Graphical view of result	66

List of Tables

Table 4.1: List of Unique active metabolites	61
Table 4.2: List of common active metabolites between biomass and active flux	63
Table 4.3: List of deactivated object flux metabolites	65

List of Acronyms

TB	Tuberculosis
FBA	Flux Balance Analysis
Ac1PIM1	acyl phosphatidylinositol mannoside
Ac1PIM2	acyl phosphatidylinositol mannoside di-mannose
Ac1PIM3	acyl phosphatidylinositol mannoside tri-mannose
Ac1PIM4	acyl phosphatidylinositol mannoside tetra-mannose
Ac2PIM2	di-acyl phosphatidylinositol mannoside di-mannose
PIM1	phosphatidylinositol mannoside
PIM2	phosphatidylinositol mannoside di-mannose
PIM3	phosphatidylinositol mannoside tri-mannose
PIM4	phosphatidylinositol mannoside tetra-mannose
PIM5	phosphatidylinositol mannoside penta-mannose
PIM6	phosphatidylinositol mannoside hexa-mannose
arabinanagalfragund	arabinan-arabinofuranose-galactofuranosyl(30)-rhamanosyl-N-acetylglucosamyl-undecaprenyl diphosphate
arach	Arachidic acid
clpn160190	cardiolipin (dihexadecanoyl, dimethylstearoyl)
peptido_TB1	peptidoglycan subunit
peptido_TB2	peptidoglycan subunit

fcmbtt	iron(III) chelated carboxymycobactin T
hdca	Hexadecanoate
hdcea	Hexadecenoate
hexc	hexacosanoate
kmycolate	kmycolate(c) keto mycolate (2 cyclopropanated rings)
mbhn	methyl behenic acid
mcbs	mycobactin S
mfrppdima	phenolic glycolipid
mmmycolate	methoxy mycolate (1 cyclopropanated ring)
mycolate	mycolate (2 cyclopropanated rings)
modca	10-methylstearic acid
mkmycolate	keto mycolate (1 cyclopropanated rings)
ocdca	octadecanoate
ocdcea	octadecenoate
pa160	1,2-dihexadecanoyl-sn-glycerol 3-phosphate
pa160190	1,2-sn-glycerol 3-phosphate
pa190190	1,2-sn-glycerol 3-phosphate
pdima	phthiocerol dimycocerosate A (Mtb)
pe160	phosphatidylethanolamine (dihexadecanoyl,)
pg160	Phosphatidylglycerol (dihexadecanoyl,)

pg160190	Phosphatidylglycerol (hexadecanoyl, methylstearoyl)
pg190	Phosphatidylglycerol (dimethylstearoyl,)
ppdima	phenol phthiocerol dimycoerolate (Mtb)
tmha1	tetramycolyl hexaarabinoside (tdm1 + tdm2 + tre)
tmha2	tetramycolyl hexaarabinoside (tdm1 + tdm3 + tre)
tmha3	tetramycolyl hexaarabinoside (tdm1 + tdm4 + tre)
tmha4	tetramycolyl hexaarabinoside (tdm2 + tdm3 + tre)
ttdca	tetradecanoate
phdca	phenol palmitic acid

ABSTRACT

Tuberculosis is a multisystem disorder that is characterized by the formation of hematomas, a type of swelling that is filled with blood that is caused due to breakage in the wall of a blood vessel. This hematomas occurs in different organ of the infected victim has claimed the life of most of its victims. This disease is caused by a bacterium known as *Mycobacterium Tuberculosis (MTB)* which can be represented as a metabolic system. Every biological system is made up a metabolites which include genes, proteins and enzymes that are inter-connected which define the function, features and characteristics of the biological system. These biological systems can be analysed using different computational techniques among which is flux balance analysis. Flux balance analysis is a constraint based approach to metabolic network analysis. It's based on the steady state assumption of $S.v = 0$. A more grounded understanding of this features, characteristics and nature of this bacterium will lead to better approaches to reduce the damage of the disease.

The flux balance analysis of *MTB* involves the conversion of the metabolic network into a matrix format known as a stoichiometric matrix. This matrix is formed by using the metabolites in the metabolic network as rows and the reactions as the columns. The stoichiometric matrix used in this research is an 828 by 1027 matrix. The analysis of the stoichiometric matrix resulted into a linear problem where the number of unknown is greater than the number of equations. This linear problem was solved using “extreme pathways” and “simplex method” algorithms which makes up a Flux Balance Analysis approach to metabolic network analysis. The extreme pathways algorithm help extract the independent paths in the network while the simplex method is used to optimize the metabolic network to extract metabolites peculiar to an objective function.

At applying the constraint of the steady state assumption, the result showed 1022 distinct pathways instead of the initial 1027 eliminating 5 other reactions. The output from the extreme pathways was used in the optimization process using biomass as the target flux to get metabolites peculiar to biomass production. After the optimization, the result shows 32 new metabolites that become activated when a value of 1 is used to represent the biomass components. The optimization result also shows two categories of metabolites: those that are part of the biomass that become inactive after optimization, those that remain active after the optimization test. The output of this research only focus on the analysis of the metabolic network using biomass as the optimization target.