GENOME-SCALE METABOLIC RECONSTRUCTION OF CRYPTOCOCCUS NEOFORMANS AND DEVELOPMENT OF TREATMENT STRATEGIES

by

Mehmet Yiğit Demirtaş

B.S., Chemical Engineering, Ege University, 2014

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering

Boğaziçi University

2021

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis supervisor Prof. Kutlu Ülgen for her guidance, valuable support and encouragement throughout my study. I would like to thank my thesis committee members, Assoc. Prof. Tunahan Çakır and Assist. Prof. Betül Uralcan, for the time devoted to my thesis and their valuable comments.

I gratefully acknowledge Scientific and Technological Research Council of Turkey (TÜBITAK) for financial support (Project No: 119M923).

I would also like to thank Mustafa Sertbaş for his generous help during this process. I am also very thankful to our modeling team, Recep Güney, İrem İlkay Özbek and Handan Çetin for their support, sharing and friendship. I would also like to thank all of the members of Ulgen Research Group for their support and encouragement.

I would like to express my special thanks to Selma Başıbüyük, Burcu Yeşilırmak, Alper Buğra Dinç, Esen Kaya, Zeynep Gülser Kınalı and Aziz Doğan İlgün for their endless friendship, motivation and encouragement.

Finally, my grateful thanks are extended to my family members, Nakşiye Demirtaş, Mahmut Demirtaş, Kadir Demirtaş and Nesrin Demirtaş for their understanding, love and never-ending support.

ABSTRACT

GENOME-SCALE METABOLIC RECONSTRUCTION OF *CRYPTOCOCCUS NEOFORMANS* AND DEVELOPMENT OF TREATMENT STRATEGIES

Cryptococcus neoformans is a common opportunistic human pathogen that causes fatal infections, especially for immunocompromised individuals. The computational systems biology approach enables the elucidation of the host-pathogen interaction of pathogens and the development of new drug strategies and treatment methods. The aim of this study is to reconstruct a genome-scale metabolic model specific to the *C.neoformans* and to identify potential drug targets by analyzing the metabolic processes of this pathogen with computational methods. The genome-scale metabolic model reconstructed in this study comprises 1267 reactions (1151 internal and 116 exchange reactions). This model also has a total of 1140 metabolites and 649 genes in 8 compartments. The metabolic changes under different environmental conditions were investigated by performing flux balance analysis. The flux distribution obtained is consistent with the literature. The performance and basic properties of the model were tested through MEMOTE and COBRA Toolbox. In the consistency part (stoichiometric consistency, mass/charge balance, and connectivity), the model achieved a score of 97%. The reaction, metabolite and single/double gene deletion analyses were applied to find new drug targets against *C.neoformans* infections. It was determined that 183 out of 1267 reactions were essential reactions. However, 143 genes and 108 metabolites were found to be essential for *C.neoformans*. 57 of the essential genes and 12 of the essential metabolites are not present in the human model. Therefore, these metabolites and genes could be potential drug targets. The drug targets determined by reaction, gene and metabolite deletion analyses are consistent with the literature.

ÖZET

CRYPTOCOCCUS NEOFORMANS'IN GENOM ÖLÇEKLİ METABOLİK MODELİNİN OLUŞTURULMASI VE TEDAVİ STRATEJİLERİNİN GELİŞTİRİLMESİ

Cryptococcus neoformans, özellikle bağışıklığı baskılanmış kişilerde ölümcül enfeksiyonlara sebep olan, yaygın bir oportunistik insan patojenidir. Hesaplamalı sistem biyolojisi yaklaşımı, patojenlerin konakçı-patojen etkileşiminin aydınlatması ve yeni ilaç stratejileri ve tedavi yöntemlerinin geliştirilmesine olanak sağlamaktadır. Bu çalışmanın amacı; *C.neoformans* patojenine özgü genome ölçekli metabolik model oluşturmak ve hesaplamalı yöntemler ile bu patojenin metabolik proseslerini analiz ederek, potansiyel ilaç hedeflerini tespit etmektir. Bu çalışmada geliştirilen genom ölçekli metabolik model, 1267 reaksiyon (1151 içsel, 116 değişim reaksiyonu) içermektedir. Bu model ayrıca, 8 organelde toplam 1140 metabolit ve 649 gene sahiptir. Akı denge analizi uygulanarak, farklı ortam koşullarındaki metabolizma değişimleri incelenmiştir. Elde edilen akı dağılımı, literatür ile uyumlu bulunmuştur. Modelin performans ve temel özellikleri, MEMOTE ve COBRA Toolbox aracılığıyla test edilmilştir. Tutarlılık kısmında (stokiyometrik tutarlılık, kütle/yük dengesi, bağlantısallık) %97 başarı sağlanmıştır. *C.neoformans* enfeksiyonlarına karşı yeni ilaç hedefleri bulmak adına, reaksiyon, metabolit ve tekli/çiftli gen silme analizleri uygulanmıştır. 1267 reaksiyondan 183'ünün zorunlu reaksiyon olduğu tespit edilmiştir. Bununla birlikte, 143 gen ve 108 metabolitin *C.neoformans* için zorunlu olduğu bulunmuştur. Zorunlu genlerin 57'si, zorunlu metabolitlerin 12'si insan modelinde bulunmamaktadır. Bu nedenle bu metabolit ve genlerin potansiyel ilaç hedefi olabileceği anlaşılmıştır. Reaksiyon, gen ve metabolit silme analizleri ile saptanan ilaç hedefleri, literatür ile uyumludur.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

m	Number of metabolites
n	Number of reactions
S	Stoichiometric matrix
v	Flux vector
Z	Objective function

LIST OF ACRONYMS/ABBREVIATIONS

ATP	Adenosine Triphosphate
BiGG	Biochemical Genetic and Genomic
CoA	Coenzyme-A
COBRA	Constraint-Based Reconstruction Analysis
CNS	Central Nervous Systems
EC	Enzyme Commission
FBA	Flux Balance Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
MEMOTE	Metabolic Model Tests
NADH	Nicotinamide Adenine Dinucleotide
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
RNA	Ribonucleic Acid
TCA	Tricarboxylic Acid

1. INTRODUCTION

Systems biology is an approach based on using the biological data of the organism in computational methods to elucidate the organism's structure and metabolic functions. In this approach, the cell is considered as a whole without being divided into small units. Since the reconstruction studies of these models started with the collection of genes and reactions obtained from the whole genome analysis of living things, the increase in genome analyzes has led to an increase in studies and applications on these models. Genome-scale metabolic models provide elucidation of the organism's metabolic processes and prediction of the organism's phenotype based on environmental conditions or genetic perturbation. In this process, constraint-based methods are used. The most widely used constraint-based modeling method is flux balance analysis. In this analysis, the flux distribution is obtained by optimization based on an objective function, using the constraints of the organism.

As a result of the entry of pathogenic microorganisms into the human body and their proliferation in different parts, infectious diseases occur. One of the pathogens that threatens human health is *Cryptococcus neoformans*. *C.neoformans* causes meningitis in immunocompromised patients (such as AIDS patients, organ transplant recipients) [1]. The reconstruction of the genome-scale metabolic model specific to this pathogen is extremely important in terms of understanding both the metabolic activities of the pathogen and its interaction with human tissues.

The objective of this study is to reconstruct the genome-scale metabolic model specific to *C.neoformans* and to elucidate the metabolic processes of this pathogen with a computational systems biology approach. The general information about this subject is presented under the title of 'Background Aspects' in the second chapter. In the third chapter, the processing steps applied for genome-scale model reconstruction of *C.neoformans* in the current study are explained.

Annotated genome data of *C. neoformans var. neoformans* JEC21 in the KEGG database is used for the model reconstruction. The reconstructed model in this study comprises 1267 reactions, 1140 metabolites and 649 genes. This chapter also includes quality and performance tests applied to the model.

The fourth chapter focuses on the results obtained by applying flux balance analysis. Here, the results of the sensitivity of *C.neoformans* pathogen to the presence of oxygen and carbon source are presented. In addition, NADH, NADPH and ATP production and consumption amounts are given in tables to prove the flux consistency of our model. In the second part of this chapter, reaction, gene and metabolite deletion analyses were applied to the reconstructed model to find potential drug targets. Essential reactions, genes and metabolites of the reconstructed model are important for identifying potential drug targets. The results of these analyses are presented in the second part of this chapter. Reactions, genes and metabolites, which are not present in the host but are essential for the pathogen, are critical so that drugs do not harm the host. In this chapter, essentials not found in the human model are also given. The last part of this chapter includes drug targets, virulence factors and infection metabolism found in the literature and their comparison with the results of our model.

The main results of the study and recommendations for future studies are presented under the title of "Conclusion and Recommendations".

2. BACKGROUND ASPECTS

2.1. *Cryptococcus neoformans*

Human pathogens are organisms that pose a threat to human health and cause infectious diseases in both immunocompetent and immunocompromised individuals. *Cryptococcus neoformans* is a fungal pathogen that causes cryptococcosis in humans [2]. *C. neoformans'* spores are eliminated in immunocompetent individuals due to the functionality of their immune system, and rarely causes disease. However, in immunocompromised patients, cryptococcosis causes cryptococcal meningoencephalitis which is responsible for 600,000 deaths annually worldwide [1], [3]. Because of this feature, *C. neoformans* is known as an opportunistic pathogen.

In the first isolation of *C. neoformans*, it was identified as a 'Saccharomyces-like' organism. In the same year, it was isolated from fermenting peach juice and named as *Saccharomyces neoformans*. Seven years later, this organism was renamed as *Cryptococcus neoformans* [4]. There are four serotypes for *C. neoformans* [5]. These are serotype A, B, C, and D. Serotype A and serotype D, which are known as *C. neoformans var. grubii* and *C. neoformans var neoformans*, respectively, cause cryptococcal meningoencephalitis in immunocompromised patients [6]. In addition, *C. neoformans var. gattii* (serotypes B and C) affects immunocompetent individuals.

C. neoformans is not a member of normal human flora, and its habitats are chicken and pigeon dropping matter and soil [5], [7], [8]. *C. neoformans* is acquired by inhalation from the environment in the form of spores or dry yeast and colonizes the airspaces of the lung tissue in the early stages of infection [5], [9], [10]. In the lung, the immune system response is provided by macrophages. However, if the infection is not cleared in the lung, it disseminates to the bloodstream and can affect major organs and systems [9]. As a result of the dissemination of the pathogen to the brain, fatal cryptococcal meningoencephalitis occurs [8].

There are 3 types of drugs that are widely used against *C. neoformans* infection [11], [12]. One of them, amphotericin B, is a drug in the polyene group and targets steroid metabolism [13], [14]. Another antifungal compound that affects steroid metabolism is fluconazole, which is in the azole group. Caspofungin, on the other hand, is a echinocandin group compound and aims to inhibit the cell wall [11]. However, it is known that polyenes have toxic effects, and azoles and echinocandins have drug resistance problems [12].

As a consequence, *C. neoformans* infections continue to be a problem worldwide, especially in developing countries, in terms of high number of cases and deaths, health care and financial problems. Current antifunfal drugs are expensive and have problems such as side effects (toxicity). Therefore, it is necessary to elucidate the metabolic processes of the pathogen and find new drug targets in order to develop new drugs and treatment methods.

2.2. Systems Biology

Systems biology is an approach based on collecting hierarchical information about the organism and transforming it into a predictive mathematical model. Thanks to this approach, it is possible to examine the cell as a whole without dividing it into small units. In this sense, the computational systems biology approach allows to elucidate the metabolic activities of the organism as a whole.

In this approach, the biological components involved in cellular processes and the interactions between these components are studied [15]. After converting the data specific to the organism into a mathematical format, results for the phenotype of the organism are provided under different environmental conditions and/or genetic perturbations. In this sense, component data, networks, *in silico* model and phenotype are obtained, respectively, in systems biology.

2.3. Constraint-Based Modeling

Constraint-based modeling (CBM) is a widely used method to understand the metabolism and responses of complex biological networks under different environmental conditions [16], [17]. This method, which is used in genome-scale metabolic models and works under the assumption of steady-state, contributes to the elucidating of the genotype-phenotype relationship by using some network-specific features as constraints.

2.3.1. Flux Balance Analysis

Flux balance analysis (FBA) is an approach that analyses the flow of metabolites in a metabolic network by means of the network-specific constraints [18], [19]. In this analysis, the reactions in the metabolic network are converted into a mathematical representation, and flux distribution are obtained by the optimization in the presence of network-specific constraints.

It is possible to estimate the organism's growth rate or the rate of production of a specific metabolite in the network by using the flux balance analysis. By applying the flux balance analysis to the metabolic model reconstructed by the data obtained from genome analysis studies, the relationship between an organism's genome and physiology can be understood [20].

There are three main steps in the flux balance analysis: (i) definition of the system with all components, (ii) definition of the constraints, and (iii) optimization. In the step of defining the system, all enzymatic, transport, and exchange reactions in the network, metabolites involved in these reactions are defined in terms of their relations with each other.

A sample system is given in Figure 2.1. This system includes 4 internal (V1, V2, V3, V4) and 3 exchange reactions (Ex1, Ex2, Ex3) for the production, consumption, uptake, and excretion of 3 metabolites (K, L, M).

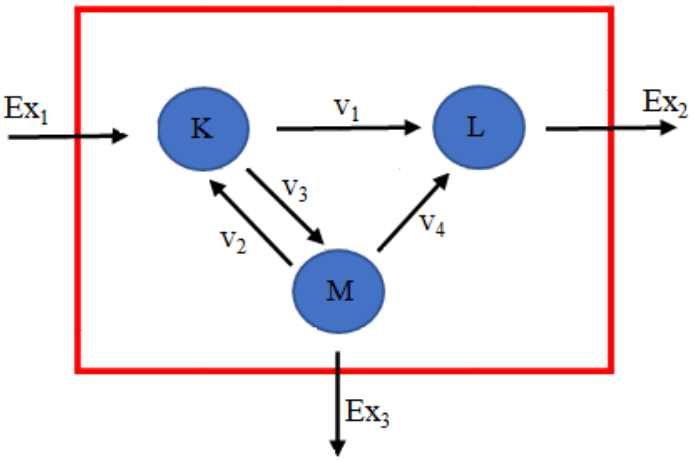

Figure 2.1. Sample system including 4 internal and 3 exchange reactions.

In the step of defining the constraints, the mass balance equation for each reaction is used as a constraint. For the mathematical representation of the reactions in the metabolic network, a stoichiometric matrix (S) with the dimension (m x n) containing the stoichiometric coefficients of the metabolites in the reaction is used. In this matrix, m represents the metabolites and n represents the reactions. The values corresponding to a reaction column give the stoichiometric coefficients of the metabolites in that reaction. A negative stoichiometric coefficient is used for each metabolite consumed in a reaction, while a positive coefficient is used for metabolites produced. Metabolites that do not participate in the reaction are shown as zero. Reaction reversibility is another constraint. This constraint means adding an upper bound and a lower bound for each reaction. While these bounds are in the range between 0 and $+\infty$ for irreversible reactions, $-\infty$ and $+\infty$ for reversible reactions. The mass balance equations for the sample system are given in Equation (2.1).

$$\begin{aligned} \frac{dK}{dt} &= -v_1 + v_2 - v_3 + Ex_1 \\ \frac{dL}{dt} &= v_1 + v_4 - Ex_2 \\ \frac{dM}{dt} &= -v_2 + v_3 - v_4 - Ex_3 \end{aligned} \tag{2.1}$$

The differential equation set is given in Equation (2.2) as a matrix form.

$$\begin{bmatrix} -1 & 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 0 & -1 \end{bmatrix} * \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ Ex_1 \\ Ex_2 \\ Ex_3 \end{bmatrix} = \begin{bmatrix} \frac{dK}{dt} \\ \frac{dL}{dt} \\ \frac{dM}{dt} \end{bmatrix} \tag{2.2}$$

The derivative of the flux of metabolites with respect to time is zero due to the steady-state condition in the system. Therefore, the general expression of the Equation (2.2) is: $S.v = 0$ where S and v represent stoichiometric coefficient matrix and flux vector, respectively. Matrix notation at steady-state condition is given in Equation (2.3).

$$\begin{bmatrix} -1 & 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 0 & -1 \end{bmatrix} * \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ Ex_1 \\ Ex_2 \\ Ex_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.3}$$

In the optimization step, an objective function that satisfies all the constraints must be defined to find a single solution space. In metabolic models, degrees of freedom are not equal to zero since the number of reactions and the number of metabolites are not equal. With optimization, degrees of freedom are reduced, and a single solution space is obtained. Due to the optimization technique based on the linear programming principle, a flux profile that maximizes or minimizes the objective function is obtained.

This objective function may be the maximization of cellular growth or the production of a specific metabolite in metabolic models or the minimization of energy consumption. The objective function equation is: $Z = (c_1v_1+c_2v_2+\ldots+c_nv_n)$ where Z is the objective function and c is a vector representing the weight of each reaction's contribution to the objective function.

2.4. Genome-Scale Metabolic Models

Genome analysis studies on pathogens have been performed since the 1990s. The first genome sequence studies for pathogens were published in 1995 for *Haemophilus influenzae* [21] and *Mycoplasma genetalium* [22]. The analysis of the genomics data gives researchers to obtain pathogen-specific gene, enzyme, and reaction data, which provide the reconstruction of the metabolic network of the pathogen [20]. With this data, genome-scale metabolic models have been curated and converted to mathematical representation by using constrained-based methods in the computational systems biology approach [18], [19].

The pathogen models have become important tools used to elucidate several metabolic activities, such as virulence and host-pathogen interactions [23], [24]. Numerous infection scenarios can be simulated on the genome-scale model. In this way, it is possible to determine virulence factors and to understand how the infection occurs in the host. One of the key benefits of pathogen-specific genome-scale metabolic models is that they allow the identification of the potential drug targets [23], [25], [26]. Determination of genes, reactions, and metabolites essential for the pathogen to maintain its vital activities through genome-scale models allows potential drug targets to be found in relatively short time [24]. In this approach, it is examined how the growth of the pathogen is affected by inhibiting the gene, reaction and metabolites in the model. However, for a target to be a drug target, inhibition of that target should not harm the host. Therefore, the gene, reaction or metabolite identified as a drug target should not be present in the host.

In the process of reconstruction of genome-scale metabolic models, there are important steps such as collecting reactions from databases and literature, determining gene and reaction localization, adding transport and exchange reactions, gene-protein-reaction (GPR) associations, filling the metabolic gaps, and adding the biomass reaction [20]. In the protocol published by Thiele and Palsson [27], the genome-scale model reconstruction process was described in detail .

In the reconstruction process, Kyoto Encyclopedia of Genes and Genomes (KEGG) [28] and Uniprot [29] databases are widely used to collect gene, enzyme, reaction, and pathway information. ModelSEED database [30] provides information about reaction directions and thermodynamic properties. In order to identify the compartments where the reactions took place, Uniprot [29], Predotar [31], Panther [32], PredictProtein [33], and TargetP [34] are most known databases.

Toolboxes such as COBRA Toolbox [35] and RAVEN [36] have been developed to enable the analysis of metabolic networks and their use by researchers easily and in accordance with standards. Due to these toolboxes, the researchers perform model creation and calculation processes using the commands of the relevant functions.

3. METHODS

3.1. Genome-Scale Metabolic Model Reconstruction and Development

A comprehensive protocol for genome-scale model reconstruction has been published by Thiele and Palsson [27] to guide researchers working in this field. The reconstruction of the genome-scale model of *Cryptococcus neoformans* was prepared in accordance with this protocol. The model reconstruction process in this study includes seven main stages: i) collecting the reaction information of the annotated genome of the pathogen, ii) determining the reaction directions, iii) determining the gene localization, iv) addition of transport and exchange reactions, v) determining gene-protein-reaction (GPR) associations, vi) gap-filling, vii) addition of biomass reaction. In the stage of collecting reactions from the annotated genome, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [28] was used, and all reactions were written in BiGG Model [37] notation. Reaction directions and thermodynamic properties were obtained from the ModelSEED database [30]. Uniprot [29], Predotar [31], Panther [32], PredictProtein [33], and TargetP [34] databases were used for the determination of the compartment where the reactions take place.

3.1.1. Collecting Pathogen-Specific Gene, Enzyme and Reaction Data from Databases

In order to reconstruct a genome-scale model for *Cryptococcus neoformans*, a database allowing biochemical reactions to be collect regularly is needed. It was determined that the gene, reaction, and pathway data belonging to *Cryptococcus neoformans var. neoformans* JEC21 [38] were found in the KEGG database. In the KEGG database, each metabolic pathway is represented by a metabolic map. These maps include reaction and gene information obtained from the genome of the organism. Each green box in these maps represents a reaction that is active for the organism. In this way, reactions, metabolites, genes and enzyme numbers can be determined.

In this study, 911 enzymatic reactions were collected from 57 pathways of *C. neoformans*. CHEBI database [39] was used to obtain metabolites' chemical formulas and charges. The pathways and number of reactions in these pathways are given in Table 3.1.

Table 3.1. Pathways and number of reactions.

Pathway	**Number of Reactions**	**Pathway**	**Number of Reactions**
Glycolysis Gluconeogenesis	20	beta-Alanine metabolism	7
Citrate cycle (TCA cycle)	17	Taurine - hypotaurine metabolism	3
Pentose phosphate pathway	14	Glutathione metabolism	9
Pentose - glucuronate interconversions	8	Starch and sucrose metabolism	13
Fructose-mannose metabolism	10	Amino sugar metabolism	11
Galactose metabolism	11	Glycerolipid metabolism	10
Ascorbate and aldarate metabolism	5	Inositol phosphate metabolism	21
Fatty acid biosynthesis	49	Glycerophospholipid metabolism	27
Fatty acid elongation	21	Sphingolipid metabolism	25
Fatty acid degradation	27	Pyruvate metabolism	11

Table 3.1. Pathways and number of reactions. (cont.)

Pathway	**Number of Reactions**	**Pathway**	**Number of Reactions**
Synthesis - degradation of ketone bodies	4	Glyoxylate metabolism	8
Steroid biosynthesis	22	Propanoate metabolism	4
Ubiquinone biosynthesis	10	Butanoate metabolism	4
Arginine biosynthesis	17	C5-Branched dibasic acid metabolism	1
Purine metabolism	68	One carbon pool by folate	11
Pyrimidine metabolism	65	Methane metabolism	2
Alanine, aspartate, glutamate metabolism	9	Thiamine metabolism	9
Glycine, serine, threonine metabolism	37	Riboflavin metabolism	9
Cysteine and methionine metabolism	30	Vitamin B6 metabolism	13
Valine, leucine, isoleucine degradation	36	Nicotinate nicotinamide metabolism	20
Valine, leucine, isoleucine biosynthesis	13	Pantothenate, CoA biosynthesis	11

Table 3.1. Pathways and number of reactions. (cont.)

Pathway	**Number of Reactions**	**Pathway**	**Number of Reactions**
Lysine biosynthesis	10	Biotin metabolism	7
Lysine degradation	5	Folate biosynthesis	14
Arginine and proline metabolism	28	Porphyrin metabolism	15
Histidine metabolism	13	Terpenoid biosynthesis	7
Tyrosine metabolism	11	Nitrogen metabolism	4
Phenylalanine metabolism	6	Sulfur metabolism	5
Phenylalanine, tyrosine, tryptophan biosynthesis	15	Aminoacyl-tRNA biosynthesis	35

3.1.2. Reaction Directionality

ModelSEED database [30] was used to determine the directions of the reactions collected from the KEGG database. The thermodynamic data of the reactions were obtained by searching in the ModelSEED database using KEGG notation. Thus, the Gibbs free energy value for each reaction and directionality were determined. For forward reactions, the lower bound and the upper bound were written as 0 and 1000, respectively. Similarly, for reversible reactions, the lower bound was written as -1000, and the upper bound was written as 1000 in the model.

3.1.3. Arrangement of KEGG Reactions According to BiGG Model Notation

BiGG Model is a database containing more than 70 published genome-scale metabolic network reconstructions [37]. In order to make comparisons with the models of taxonomically similar organisms and to contribute to the technical knowledge, the reactions taken from KEGG were arranged in accordance with the BiGG Model notation.

3.1.4. Gene Localization

Gene localization search was carried out to determine the compartment where the reactions occur. For this purpose, Uniprot [29], Predotar [31], Panther [32], PredictProtein [33] and TargetP [34] databases were used. The organelle containing the gene that regulates the reaction was determined by searching with Gen ID in Uniprot and Panther. In the Predotar, PredictProtein, and TargetP databases, the localization search was made using the gene sequence in FASTA format. The distribution of the reactions collected from the KEGG database into compartments is given in Table 3.2.

Table 3.2. Distribution of reactions in compartments.

Compartment	**Number of Reactions**
Cytosol	660
Mitochondria	196
Peroxisome	28
Nucleus	12
Endoplasmic Reticulum	6
Vacuole	4
Extracellular Space	3
Golgi Apparatus	2

By comparing the results obtained from these five databases, the compartments where the reactions take place were determined. For reactions whose compartments could not be determined based on these databases, the compartment in which the same reaction occurred was taken into account in the iMM904 model developed for *Saccharomyces cerevisiae* [40] and SpoMBEL1693 model developed for *Schizosaccharomyces pombe* [41].

3.1.5. Transport and Exchange Reactions

Transport reactions between compartments were added to ensure the transfer of metabolites required for reactions taking place in different compartments in the cell. In the addition of these reactions, reactions taking place in the compartments and the pathway topology were taken into account. In this way, problems of accumulation of the metabolites or absence of reactants were eliminated. At the same time, metabolites to be taken into the cell or sent out of the cell were determined with the help of the literature. Exchange reactions for these metabolites were added to the model. As a result, 239 transport reactions and 116 exchange reactions were added to the model.

3.1.6. Gene-Protein-Reaction (GPR) Associations

In order to verify the annotated genome data obtained from KEGG database, it is required to determine Gene-Protein-Reaction (GPR) associations in the model. For this purpose, genes in the model and their orthologs in the *Saccharomyces cerevisiae* model (iMM904) were compared to determine how the gene affects the relevant reaction. In this way,

- Proteins forming a protein complex,
- Proteins regulating more than one reactions,
- More than one protein performing the same function were detected [27].

3.1.7. Gap-Filling

In order to detect and fill the gaps between different pathways and/or different reaction groups of the model, the manual gap-filling process was performed. In this process, first of all, dead-end metabolites in the model were detected. The dead-end metabolite is the metabolite that is produced but not consumed in the model or not present in the model although it is a reactant in at least one reaction. In addition to dead-end metabolites, the blocked reactions are essential for the gap-filling process. These reactions do not have flux due to the model topology and/or dead-end metabolites. Therefore, the blocked reactions in the model were detected after finding the dead-end metabolites. In the next step, new reactions were proposed by taking the reactions in the literature, dead-end metabolites, and blocked reactions into consideration. Thus, the number of dead-end metabolites and blocked reactions in the model is reduced. With the completion of the gap-filling, the final model (iYD649) was obtained.

3.1.8. Clusters of Orthologous Groups of Proteins (COG) Classifications

Clusters of Orthologous Groups of proteins (COG) database [42] is an online database used to functionally categorize proteins in the genome. There are functional categories to classify proteins in this database, and these categories are indicated by letters. The COG database was used to make phylogenetic classification of proteins encoded in the *C. neoformans* genome. The class to which the encoded protein belongs was found in searches using Gene ID. As a result of the COG classification, it was found that 21.4% of the genes in the iYD649 model are related to amino acid transport and metabolism. Likewise, 14.58% of the genes in the model are for energy production and conversion, 11.93% for coenzyme transport and metabolism, and 11.74% for carbohydrate transport and metabolism. COG classification results are given in Table 3.3.

Table 3.3. COG results.

Functional Category	**Description**	**Number of Genes**
[A]	RNA processing and modification	0
[B]	Chromatin structure and dynamics	0
[C]	Energy production and conversion	77
[D]	Cell cycle control, cell division, chromosome partitioning	0
[E]	Amino acid transport and metabolism	113
[F]	Nucleotide transport and metabolism	47
[G]	Carbohydrate transport and metabolism	62
[H]	Coenzyme transport and metabolism	63
[I]	Lipid transport and metabolism	51
[J]	Translation, ribosomal structure and biogenesis	31
[K]	Transcription	2
[L]	Replication, recombination and repair	0
[M]	Cell wall/membrane/ envelope biogenesis	10
[N]	Cell motility	0

Table 3.3. COG results. (cont.)

Functional Category	**Description**	**Number of Genes**
[O]	Posttranslational modification, protein turnover, chaperones	3
[P]	Inorganic ion transport and metabolism	15
[Q]	Secondary metabolites biosynthesis, transport and catabolism	14
[R]	General function prediction only	32
[S]	Function unknown	0
[T]	Signal transduction mechanisms	2
[U]	Intracellular trafficking, secretion, and vesicular transport	0
[V]	Defense mechanisms	6
[W]	Extracellular structures	0
[X]	Mobilome	0
[Y]	Nuclear structure	0
[Z]	Cytoskeleton	0

3.1.9. Biomass Reaction

The biomass reaction for the iYD649 model was taken from the iMM904 model [40] developed for *Saccharomyces cerevisiae* which is a similar organism according to phylogenetic classification. The biomass production in the iYD649 model is 0.3940/h when 4 mmol/gDW/h glucose is used as carbon source.

Based on this biomass production, the doubling time was calculated as 1.76 h (106 minutes). Orner and colleagues [43] reported that the doubling time for *C. neoformans* is between 80 minutes and 160 minutes. The doubling time calculation is given in Equation (3.1).

$$Doubling\ Time = \frac{ln2}{0.3940} = 1.76h^{-1}. \tag{3.1}$$

Non-growth-associated maintenance cost means the amount of energy required for the vital activities of the organism except growth. Henson and colleagues [44] reported that this maintenance cost could be taken as 5 mmol/gDW/h for pathogenic organism models. In the iYD649 model, the lower bound of ATP maintenance reaction (ATPM) was set as 5 mmol/gDW/h. If the amount of ATP produced by the carbon source remains below this value, the iYD649 model will not produce any solution.

Metabolites and their coefficients involved in the biomass reaction are given in Table 3.4. The mannan metabolite was deleted from the biomass reaction since there was no information about this metabolite in the annotated genome data from KEGG database.

Table 3.4. Biomass Reaction.

Abbreviation	**Metabolite Name**	**Coefficient**
adp	ADP	59.276
ala_	L-Alanine	0.4588
amp	AMP	0.046
arg_L	L-Arginine	0.1607
asn_L	L-Asparagine	0.1017
asp_L	L-Aspartate	0.2975
atp	ATP	59.276
cmp	CMP	0.0447

Table 3.4. Biomass Reaction.(cont.)

Abbreviation	**Metabolite Name**	**Coefficient**
cys_L	L-Cysteine	0.0066
damp	DAMP	0.0036
dcmp	DCMP	0.0024
dgmp	DGMP	0.0024
dtmp	DTMP	0.0036
gln_L	L-Glutamine	0.1054
glu_L	L-Glutamate	0.3018
gly	Glycine	0.2904
glycogen	Glycogen	0.5185
gmp	GMP	0.046
h	H+	58.70001
h2o	Water	59.276
his_L	L-Histidine	0.0663
ile_L	L-Isoleucine	0.1927
leu_L	L-Leucine	0.2964
lys_L	L-Lysine	0.2862
met_L	L-Methionine	0.0507
phe_L	L-Phenylalanine	0.1339
pi	Phosphate	59.305
pro_L	L-Proline	0.1647
ribflv	Riboflavin	0.00099
ser_L	L-Serine	0.1854
so4	Sulfate	0.02
thr_L	L-Threonine	0.1914
tre	Trehalose	0.0234
trp_L	L-Tryptophan	0.0284
tyr_L	L-Tyrosine	0.102

Table 3.4. Biomass Reaction.(cont.)

Abbreviation	**Metabolite Name**	**Coefficient**
ump	UMP	0.0599
val_L	L-Valine	0.2646
13BDglcn	1,3-beta-D-Glucan	1.1348
ergst	Ergosterol	0.0007
pa_SC	Phosphatidate	6e-06
pc_SC	Phosphatidylcholine	6e-06
pe_SC	Phosphatidylethanolamine	4.5e-05
ps_SC	Phosphatidylserine	1.7e-05
ptd1ino_SC	Phosphatidyl 1D myo inositol	5.3e-05
triglyc_SC	Triglyceride	6.6e-05
zymst	Zymosterol	0.0015

3.1.10. Metabolic Pathways

In the glycolysis/gluconeogenesis pathway, all reactions from glucose (glc_D) to pyruvate (pyr) take place in the cytosol. After the pyruvate is produced, this metabolite passes into mitochondria and turns into acetyl-CoA (accoa). In the KEGG pathway map, this reaction is represented by the sum of four different reactions, R00014, R03270, R02569, and R07618. Instead of these four reactions, PDHm in the BiGG Model was taken as a single reaction during the model reconstruction. In addition, oxaloacetate (oaa), a product of the citrate cycle, transforms into phosphoenolpyruvate (pep) in the glycolysis/gluconeogenesis pathway and participates in pyruvate production.

Galactose is one of the important components of the polysaccharide capsule [45]. According to the galactose metabolism from the KEGG database, D-galactose (gal), D-galactose 1-phosphate (gal1p), and UDPgalactose (udpgal) are converted to each other and participate in the polysaccharide structure required for the capsule.

In the citrate cycle (TCA cycle), oxaloacetate (oaa), produced by an irreversible reaction from pyruvate, converts to citrate (cit) and initiates the cycle that produces isocitrate (icit), 2-oxoglutarate (akg), succinyl-Coa (succoa), succinate (succ), fumarate (fum) and L-malate (mal_L), respectively. The cycle is completed by the conversion of L-malate to oxaloacetate. In KEGG, the conversion of 2-oxoglutarate to succinyl-CoA is represented by the sum of four reactions, R00621, R03316, R02570, and R07618. Instead of these four reactions, AKGDm in the BiGG Model was taken during the model reconstruction.

The pentose phosphate pathway is the pathway where important metabolites for carbon metabolism are produced. Reactions in the pentose phosphate pathway generally take place in the cytosol. D-glucose 6-phosphate (g6p), glyceraldehyde 3-phosphate (g3p), and D-fructose 6-phosphate (f6p), produced in the glycolysis/gluconeogenesis pathway, provide the production of important intermediates in this pathway. These intermediates are 6-phospho-D-glucono-1,5-lactone (6pgl), 6-phospho-D-gluconate (6pgc), D-ribulose 5-phosphate (ru5p_D), D-erythrose 4-phosphate (e4p), D-xylulose 5- phosphate (xu5p_D), D-ribose 5-phosphate (r5p) and sedoheptulose 7-phosphate (s7p).

In the pentose and glucuronate interconversions pathway, D-glucose 1-phosphate (g1p) is converted to UDPglucose (udpg), one of the components of the polysaccharide capsule. All reactions in this pathway take place in the cytosol. Due to the transformation of D-glucose 1-phosphate (g1p) into UDPglucose (udpg), the production of UDPglucose (udpg), one of the components of the polysaccharide capsule, takes place in this pathway.

In the fructose and mannose metabolism pathway, fructose (fru) is produced from sorbitol (sbt_D) and converts to D-fructose 6-phosphate (f6p). Subsequently, D-mannose 1-phosphate (man1p) and D-mannose 6-phosphate (man6p) are produced in this pathway. Also, D-glyceraldehyde (glyald) is derived from D-fructose 1-phosphate (f1p) and turns to glyceraldehyde 3-phosphate (g3p).

In the pathway of synthesis and degradation of ketone bodies, acetoacetyl-CoA (aacoa), which comes from the fatty acid elongation pathway, transforms into acetoacetate (acac). Reactions take place in both cytosol and mitochondria.

The steroid biosynthesis pathway is of great importance as it enables ergosterol production. Besides being in the biomass reaction, ergosterol is also critical for the regulation of membrane fluidity [46], [47]. In this pathway, two consecutive reactions (R00702 and R02872) having the same EC number and the same gene are combined and written as a single reaction as SQLS during the model reconstruction. Then, lanosterol (lanost), zymosterol (zymst), fecosterol (fecost), episterol (epist), and ergosterol (ergst) are produced, respectively. Reactions in this pathway generally take place in the cytosol or endoplasmic reticulum. The organelles where the reactions took place were determined according to the gene localization. Accordingly, transport reactions were also included in the model.

Arginine is one of the essential amino acids found in the biomass reaction of the iYD649 model. Enzymatic reactions in the arginine biosynthesis pathway generally take place in the cytoplasm and mitochondria. As a result of a series of reactions starting with L-glutamate (glu_L), N-acetyl-L-glutamate (acglu), N-acetyl-L-glutamyl 5-phosphate (acg5p), N-acetyl-L-glutamate 5-semialdehyde (acg5sa), N2-acetyl-L-ornithine (acorn), ornithine (orn), L-citrulline (citr_L), L-argininosuccinate (argsuc) and L-arginine (arg_L) are produced respectively. The cycle is completed by reacting L-arginine (arg_L) with water, producing ornithine and urea. Urea turns into carbon dioxide (co2) and ammonia (nh4).

In the starch and sucrose metabolism pathway, the conversion reactions between sucrose (sucr), D-fructose (fru), maltose (malt), D-glucose (glc_D), D-glucose 6- phosphate (g6p), D-fructose 6-phosphate (f6p), trehalose (tre) and trehalose 6- phosphate (tre6p) take place. Glycogen (glycogen) and 1-3 beta D-glucan (13BDglcn), involved in the biomass reaction, are also produced in this pathway. All reactions take place in the cytosol.

Acetyl-CoA (accoa), which comes from the citrate cycle (TCA cycle) pathway, first turns into malonyl CoA (malcoa) and then malonyl-[acyl-carrier protein] (malACP) in the fatty acid biosynthesis pathway. At the same time, acetyl-CoA (accoa) reacts with an acyl carrier protein (ACP) to produce acetyl-ACP (acACP). In the presence of FAS1 and FAS2 enzymes, malonyl- [acyl-carrier protein] (malACP) reacts with acetyl-ACP (acACP). In this way, fatty acid biosynthesis begins. In this pathway, decanoyl-CoA (dcacoa), dodecanoyl-CoA (ddcacoa), hexadecenoyl-CoA (hdcoa), stearoyl-CoA (stcoa), octadecenoyl-CoA (odecoa), octadecynoyl-CoA (ocdycacoa) and tetradecanoyl-CoA (tdcoa) are produced to be consumed in glycerolipid and glycerophospholipid metabolism pathways. Fatty acid biosynthesis reactions take place in the cytosol.

Palmitoyl-CoA (pmtcoa), obtained in the fatty acid biosynthesis pathway, passes to the mitochondria and participates in the fatty acid elongation pathway. The fatty acid elongation pathway takes place in the mitochondria. In this pathway, tetra decanoyl-CoA (tdcoa), dodecanoyl-CoA (ddcacoa), decanoyl-CoA (dcacoa), octanoyl-CoA (occoa), hexanoyl-CoA (hxcoa), butanoyl-CoA (btcoa) and acetyl-CoA (accoa) are produced respectively. The produced acetyl-CoA (accoa) participates in the citrate cycle (TCA cycle) pathway.

The fatty acid degradation pathway is a pathway wherein the fatty acids are broken down. All reactions in this pathway take place in the peroxisome. Reactions starting from Palmitoyl-CoA (pmtcoa) continue until Acetyl-CoA (accoa) production. During the model reconstruction, a lumped reaction (FAO80p in the BiGG Model) between octanoyl-CoA (occoa) and acetyl-CoA (accoa) was also used.

The alanine, aspartate and glutamate metabolism pathway is the pathway wherein important metabolites such as alanine (ala_L), aspartate (asn_L), 4- aminobutanoate (4abut), and L-glutamate (glu_L) are produced. Also, in this pathway, 4- aminobutanoate (4abut) is transformed into the succinic semialdehyde (sucsal), then succinate (succ) is produced. All reactions take place in the cytosol.

Pyrimidine is an important metabolite for the synthesis of the nucleic acids that perform important functions in the cell [48]. In this pathway, reactions that form the pathway known as pyrimidine de novo synthesis take place [49], and important metabolites such as uridine monophosphate (ump), uridine 5'-diphosphate (udp), and uridine triphosphate (utp) are produced. In this synthesis, carbamoyl phosphate (cbp), N-carbamoyl-L-aspartate (cbasp), (S)-dihydroorotate (dhor_S), orotate (orot), orotidine 5'-phosphate (orot5p), uridine monophosphate (ump), uridine 5'-diphosphate (udp) and uridine triphosphate (utp) are produced respectively.

The thiamine metabolism pathway is a pathway including successive reactions. 4-amino-5-hydroxymethyl-2-methylpyrimidine (4ahmmp) is transformed into 4-amino-2-methyl-5-phosphomethylpyrimidine (4ampm) and from this metabolite 2-methyl-4-amino-5-hydroxymethylpyrimidine diphosphate (2mahmp) is produced. In this pathway, 2-methyl-4-amino-5-hydroxymethylpyrimidine diphosphate (2mahmp) is used as a reactant in two reactions (TMPPP_1 and TMPPP). Thiamin monophosphate (thmmp) is produced in both of these reactions. Thiamin monophosphate (thmmp) turns into thiamin (thm). Finally, thiamin (thm) completes the cycle by producing 4-amino-5-hydroxymethyl-2-methylpyrimidine (4ahmmp).

The purine metabolism pathway is one of the pathways with the highest number of reactions in the model. Several currency metabolites are produced in this pathway and transported to different pathways. In terms of the metabolites produced and/or consumed, it is generally linked to the pyrimidine and thiamine pathways. In addition to the reactions in the KEGG map, the sink reaction for 5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxylate (5aizc) and the demand reaction for (S)-2-[5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxamido] succinate (25aics) were added to the model. Reactions generally take place in the cytosol.

There is no usable map of oxidative phosphorylation in the KEGG database. Therefore, the oxidative phosphorylation pathway in the iMM904 model [40] developed for *Saccharomyces cerevisiae* was used in the model reconstruction.

In the ascorbate and aldarate metabolism pathway, Myo-inositol (inost), produced in the inositol phosphate metabolism pathway, is converted to D-glucuronate (glcur). D-glucuronate is important for the production of L-gulonate (guln_L) in the pentose and glucuronate interconversions pathway. Reactions in this pathway usually take place in the cytosol and endoplasmic reticulum, depending on gene localization.

Riboflavin (ribflv) is an important metabolite for cell growth as it is involved in the biomass reaction of the model. In the riboflavin metabolism pathway, 3,4-dihydroxy-2-butanone 4-phosphate (db4p) is produced from D-ribulose 5-phosphate (ru5p_D). After the production of this metabolite, riboflavin (ribflv) is produced in the presence of 4- (1-D-Ribitylamino)-5-aminouracil (4r5au). Since there is no reaction for the production of 4- (1-D-Ribitylamino)-5-aminouracil (4r5au) in the KEGG database, this metabolite was added to the model as a demand reaction (DM_4r5au_c).

In the histidine metabolism pathway, 5-phospho-alpha-D-ribose 1-diphosphate (prpp) from the pentose phosphate pathway initiates the histidine production reactions. As a result of successive reactions, 1-(5-Phosphoribosyl)-ATP (prbatp), 1-(5-Phosphoribosyl)-AMP (prbamp), D-erythro-1-(Imidazol-4-yl)glycerol 3-phosphate (eig3p), 3- (Imidazol-4-yl)-2-oxopropyl phosphate (imacp), L-histidinol phosphate (hisp), L-histidinol (histd) and L-histidine (his_L) are produced respectively. In this pathway, all reactions take place in the cytosol.

In the glycine, serine and threonine metabolism, aspartate (asp_L) is converted to threonine (thr_L). In the production of threonine, intermediates (4-phospho-L-aspartate (4pasp), L-aspartate 4-semialdehyde (aspsa), L-homoserine (hom_L), and O-phospho-L-homoserine (phom)) are also produced. Glycine (gly) is produced from serine (ser_L). Serine, produced in this pathway, participates in sphingolipid synthesis in the sphingolipid pathway. Phosphatidylserine (ps_SC), an important metabolite for glycerophospholipid synthesis, is also produced in this pathway. Reactions generally take place in cytosol and mitochondria.

The ubiquinone and other terpenoid-quinone biosynthesis is a pathway involving reactions where L-tyrosine (tyr_L) turns into 3- (4-hydroxyphenyl) pyruvate (34hpp) in the presence of tyrosine transaminase enzyme. These reactions take place in the cytosol, mitochondria, and peroxisome.

In the lysine biosynthesis pathway, 2-hydroxybutane tricarboxylate (hcit), homo-cis-aconitate (b124tc), homoisocitrate (hicit), oxaloglutarate (oxag), 2 -oxoadipate (2oxoadp), L-2-aminoadipate (L2aadp), L-2-aminoadipate 6-semialdehyde (L2aadp6sa), L- saccharopine (saccrp_L) and L-lysine (lys_L) are produced respectively. All reactions take place in the mitochondria. The lysine degradation pathway is the pathway wherein L-lysine (lys_L) is broken down, in contrast to the lysine biosynthesis pathway. All reactions take place in the mitochondria.

In the phenylalanine, tyrosine and tryptophan biosynthesis pathway, D-erythrose 4-phosphate (e4p), produced in the pentose phosphate pathway, reacts with phosphoenolpyruvate (pep). As a result of a reaction group starting with this reaction, important metabolites for the organism such as 3-dehydroquinate (3dhq), shikimate (skm), chorismate (chor), and prephenate (pphn) are produced.

In the glutathione metabolism pathway, L-glutamate (glu_L) is produced from 5-oxoproline (5oxpro) by the 5-oxoproline amidohydrolase reaction. This reaction starts a cycle in which 5-oxoproline will be produced again. During this cycle, gamma-L-glutamyl-L-cysteine (glucys), reduced glutathione (gthrd), L-cysteinylglycine (cgly), and L-cysteine (cys_L) are produced.

In the nitrogen metabolism pathway, cyanate (cynt) turns into carbamate (cbm). Ammonium (nh4) is produced from carbamate (cbm). Reactions take place in the cytosol. In aminoacyl-tRNA biosynthesis, acyl-tRNA synthesis for amino acids in the iYD649 model is performed. Reactions generally take place in cytosol and mitochondria.

In the amino sugar and nucleotide sugar metabolism, there are conversion reactions between D-fructose 6-phosphate (f6p), D-glucosamine (gam), D-glucosamine 6-phosphate (gam6p), chitin, and chitosan (chitos). Reactions take place in the cytosol.

In the arginine and proline metabolism pathway, L-glutamate (glu_L) and L-glutamate 5-semialdehyde (glu5sa) are produced by the ornithine transaminase reaction. L-glutamate 5-semialdehyde (glu5sa) is first converted to the intermediate product 1-pyrroline-5-carboxylate (1pyr5c), and then proline (pro_L) is produced.

In the terpenoid backbone biosynthesis pathway, metabolites having isoprene units are produced. Depending on the presence of (R)-5-phosphomevalonate (5pmev); (R)-5 diphosphomevalonate (5dpmev), isopentenyl diphosphate (ipdp), dimethylallyl diphosphate (dmpp), geranyl diphosphate (grdp) and farnesyl diphosphate (frdp) are produced. Farnesyl diphosphate, produced in this pathway, participates in the production of ergosterol in the steroid metabolism pathway. Since there is no reaction for the production of (R)-5-phosphomevalonate in the KEGG database, this metabolite was added to the model as a demand reaction (DM_5pmev_c).

In the valine, leucine and isoleucine degradation pathway, OIVD1r (in the cytoplasm) and OIVD1m (in the mitochondria) were added to the model instead of R07601, R07602, R07618, and R04097 in the KEGG map. OIVD2 (in the cytoplasm) and OIVD2m (in the mitochondria) were added to the model instead of R07599, R07600, R07618, and R02662 in the KEGG map. OIVD3 (in the cytoplasm) and OIVD3m (in the mitochondria) were added to the model instead of R07603, R07604, R07618, and R03174 in the KEGG map.

In the phenylalanine metabolism pathway, L-phenylalanine (phe_L), produced from phenylpyruvate (phpyr), is first converted to phenethylamine (peamn) and then to phenylacetaldehyde (pacald). Phenylacetaldehyde is excreted out of the cell by an exchange reaction.

The tyrosine metabolism pathway is the pathway where fumarate (fum) to be used in the citrate cycle and pyruvate (pyr) to be used in pyruvate metabolism are produced. Reactions take place in the cytosol.

The glycerolipid metabolism pathway is the pathway where phosphatidate (pa_SC), triglyceride (triglyc_SC), and 1,2-Diacylglycerol (12dgr_SC) are produced. In the presence of glycerol 3-phosphate (glyc3p), phosphatidate is produced by acyl-CoA metabolites from the fatty acid biosynthesis pathway. The phosphatidate is converted to 1,2-Diacylglycerol. The triglyceride production and conversion to fatty acids also take place in this pathway.

The inositol phosphate metabolism pathway is the pathway where the inositol and its derivatives are produced. D-glucose 6-phosphate (g6p) is converted to 1D-Myo-inositol 3-phosphate (mi3p_D) by an irreversible reaction. Myo-inositol (inost) is derived from the 1D-Myo-inositol 3-phosphate in the presence of 1D-Myo-inositol 3-phosphate phosphohydrolase. Phosphatidyl-1D-myoinositol (ptd1ino_SC) is also produced in this pathway.

In the pathway of glycerophospholipid metabolism, phosphatidylcholine (pc_SC) is produced from phosphatidylethanolamine (pe_SC) as a result of 3 consecutive reactions. At the same time, the conversion of phosphatidylserine (ps_SC) to phosphatidylethanolamine (pe_SC) and the production of CDPdiacylglycerol (cdpdag_SC) from phosphatidate (pa_SC) takes place in this pathway. The reactions take place in the cytoplasm and mitochondria.

2-oxobutanoate (2obut), produced in the glycine, serine and threonine metabolism pathway, passes into the mitochondria and participates in the production of L-isoleucine (ile_L) in the valine, leucine and isoleucine biosynthesis pathway. In this pathway, 3-methyl-2-oxobutanoate (3mob) and 4-methyl-2-oxopentanoate (4mop) are produced in a series of reactions starting with pyruvate (pyr). These metabolites are intermediates for the production of valine (val_L) and leucine (leu_L), respectively.

In the cysteine and methionine metabolism pathway, cysteine (cys_L) and methionine (met_L) are derived from L-cystathionine (cyst_L) and L-glutamate (glu_L), respectively. In this pathway, S-adenosyl-L-methionine (amet) and S-adenosyl- L-homocysteine (ahcys) are also produced and consumed. Methionine (met_L) produced in this pathway is consumed in S-adenosyl- L-methionine (amet) production, which is an important metabolite for steroid biosynthesis. Reactions generally take place in the cytosol.

In the sphingolipid pathway, sphinganine (sphgn) is produced in the presence of L-Serine (ser_L) and palmitoyl-CoA (pmtcoa). Sphinganine (sphgn) is converted into phytosphingosine (psphings). Sphinganine (sphgn) and phytosphingosine (psphings); are intermediate products for the production of metabolites such as dihydroceramide (dhCer), phytoceramide (phCer). Phytoceramide (phCer) is also converted into 1D-myo-Inositol 1-phosphate (mi1p_D) which is important metabolite for the inositol phosphate metabolism. During the reconstruction of this pathway, the studies of Singh and colleagues [50] and Garcia and colleagues [51] were used in addition to the KEGG database.

3.2. Quality Control

Various tests have been applied to check the basic properties of the constructed genome-scale model and to understand whether it has the ability to predict the metabolic processes in *C. neoformans* pathogen. For this purpose, the consistency, metabolites, reactions, GPR associations, and biomass information of the iYD649 model were determined by applying the MEMOTE test. In addition to the MEMOTE test, basic properties including leakage, ATP demand, and energy production from water and/or oxygen were tested.

3.2.1. Metabolic Model Testing (MEMOTE)

MEMOTE is a tool that allows the quality and performance of genome-scale models to be tested [52]. The iYD649 model was tested in MEMOTE for consistency, metabolites, reactions, GPR associations, and biomass. Accordingly, the iYD649 model achieved a total score of 97% in terms of consistency. This score for iMM904 developed for *Saccharomyces cerevisiae* is 96%. Sub-criteria and scores of the consistency are given in Figure 3.1.

Stoichiometric Consistency	99.6%
Mass Balance	92.4%
Charge Balance	95.0%
Metabolite Connectivity	100.0%
Unbounded Flux In Default Medium	91.4%
Sub Total	97%

Figure 3.1. Consistency scores in MEMOTE.

Basic information of the iYD649 model according to MEMOTE is given in Figure 3.2. Metabolic coverage is a parameter that represents the modeling detail. This value is greater than 1 in models with a high level of modeling detail and less than 1 in models with low detail. The metabolic coverage value of the iYD649 model is 1.95.

Model Identifier	COBRA Model
Total Metabolites	1,140
Total Reactions	1,267
Total Genes	649
Total Compartments	8
Metabolic Coverage	1.95

Figure 3.2. Basic information of the iYD649 model in MEMOTE.

In the iYD649 model, 1140 metabolites are distributed in 8 compartments. There is no duplicate metabolite in any compartment in the model. Figure 3.3 shows the metabolite information in MEMOTE.

Unique Metabolites	1,140
Duplicate Metabolites in Identical Compartments	0
Metabolites without Charge	0
Metabolites without Formula	0
Medium Components	8

Figure 3.3. Metabolite information of the iYD649 model in MEMOTE.

iYD649 model comprises 911 purely metabolic (enzymatic) reactions. In addition to this, the number of transport reactions is 239. Figure 3.4 gives information about the reactions of the iYD649 model according to MEMOTE.

Purely Metabolic Reactions	911
Purely Metabolic Reactions with Constraints	4
Transport Reactions	239
Transport Reactions with Constraints	1
Thermodynamic Reversibility of Purely Metabolic Reactions	0.35
Reactions With Partially Identical Annotations	0.00
Duplicate Reactions	0.00
Reactions With Identical Genes	0.49

Figure 3.4. Reaction information of the iYD649 model in MEMOTE.

The number of reactions without GPR is 190, most of which are transport and exchange reactions. The fraction of transport reactions without GPR is 0.72. In addition, there are 47 enzyme complexes in the model. GPR associations of the iYD649 model are given in Figure 3.5.

Reactions without GPR	190
Fraction of Transport Reactions without GPR	0.72
Enzyme Complexes	47

Figure 3.5. GPR associations of the iYD649 model in MEMOTE.

Biomass information of the iYD649 model according to MEMOTE is given in Figure 3.6. In the calculation of biomass consistency, the coefficient of each metabolite in the biomass reaction is multiplied by the molecular weight of that metabolite, and the total value is divided by 1000. MEMOTE calculates the biomass production in 2 different ways, in default medium and complete medium. For the complete medium, all reaction bounds are removed, and FBA is applied.

Biomass Reactions Identified	1
Biomass Consistency	0.83
Biomass Production In Default Medium	0.39
Unrealistic Growth Rate In Default Medium	false
Biomass Production In Complete Medium	61.65
Blocked Biomass Precursors In Default Medium	0
Blocked Biomass Precursors In Complete Medium	0
Ratio of Direct Metabolites in Biomass Reaction	0.02
Number of Missing Essential Biomass Precursors	17

Figure 3.6. Biomass information of the iYD649 model in MEMOTE.

3.2.2. Basic Properties

Various test methods have been developed to check the basic properties of genome-scale models [53]. These methods provide the opportunity to test the functionality of the genome-scale model and allow possible disadvantages to being noticed. For this purpose, various basic property tests were performed. One of the basic property tests is the leakage test. The leakage test is applied to determine whether an exchange metabolite in the model is produced from nothing.

In another leakage test (leakage test 2), it is determined whether there is a leakage or not when demand reactions for each metabolite are added. In this way, the metabolites are evaluated for whether they cause leakage. The energy must be supplied from the carbon source. It should be tested whether the energy is derived from water and/or oxygen. For this purpose, the uptake of all metabolites except water and oxygen is stopped, and energy production is controlled. Similarly, it is tested whether matter production occurs when ATP demand reaction becomes reversible. In order to check the proton balance of the model, demand reactions are added for h[m] and h[c], and FBA is applied by maximizing these reactions. In this way, it is checked whether the model has flux when the h[m] and h[c] demand reactions are maximized.

Another test applied relates to ATP demand under aerobic conditions. In this test, the amount of ATP demand under aerobic conditions is evaluated. For this purpose, the limits of oxygen and water uptakes are removed. Also, the demand reaction for ATP is maximized. It is considered that there is too much ATP demand when the flux value of the ATP demand reaction exceeds the theoretical flux value. The theoretical flux value is 31 mmol/gDW/h. The final test determines whether there are empty columns in the gene-reaction matrix (model.rxnGeneMat) in the model. The results of all tests are given in Figure 3.7 as a MATLAB screenshot.

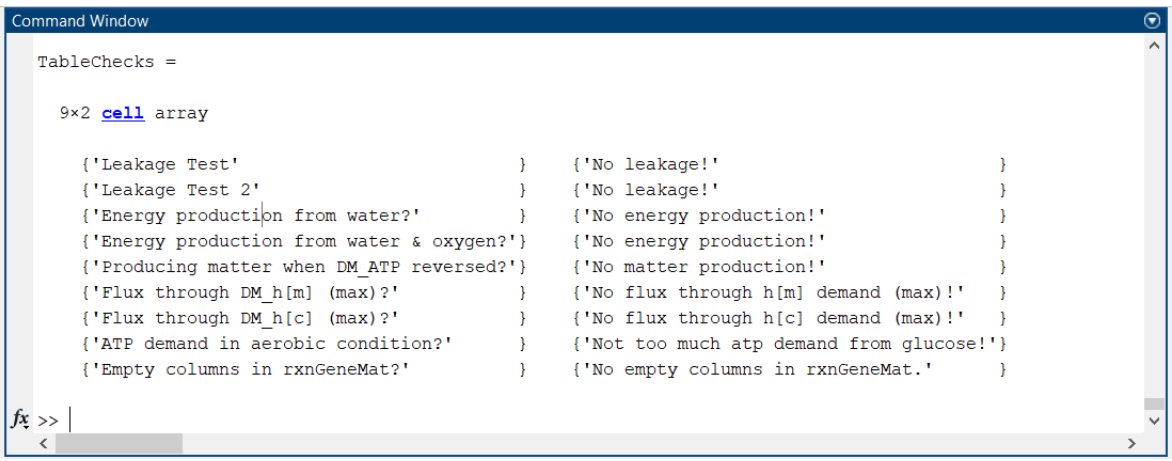

Figure 3.7. Results of basic properties testing.

4. RESULTS AND DISCUSSION

4.1. Flux Balance Analysis of the Reconstructed Metabolic Network

Flux balance analysis is a mathematical optimization method widely used in the analysis of metabolic networks [19]. Network topology, reaction direction, metabolites availability, and transport of metabolites between compartments constitute constraints for the mathematical model. In this study, flux balance analysis was used to optimize the iYD649 model and obtain flux distribution under different conditions. In addition, the sensitivity of biomass production to carbon source and oxygen availability was analyzed by robustness analysis and shadow price analysis.

4.1.1. Robustness Analysis

Robustness analysis is a parameter defined as the sensitivity of objective function in the model to a specific reaction [19]. The growth in the iYD649 model is 0.3940/h when 4 mmol/gDW/h glucose is used as the carbon source. Glucose uptake – growth graph related to the iYD649 model is given in Figure 4.1. In Figure 4.1, the amount of oxygen uptake is unlimited. As the amount of glucose increases, the growth increases linearly since the oxygen uptake is unlimited. When the glucose uptake is 18.5 mmol/gDW/h, the growth reaches about 2/h. However, if the amount of oxygen in the environment is limited, the growth does not increase linearly, even if there is enough glucose. Glucose uptake - growth graph in case of that oxygen uptake is limited to 17 mmol/gDW/h is given Figure 4.2. The slope of the graph decreases after the point where oxygen is insufficient. When the glucose uptake is 18.5 mmol/gDW/h, the growth value remains around 1/h.

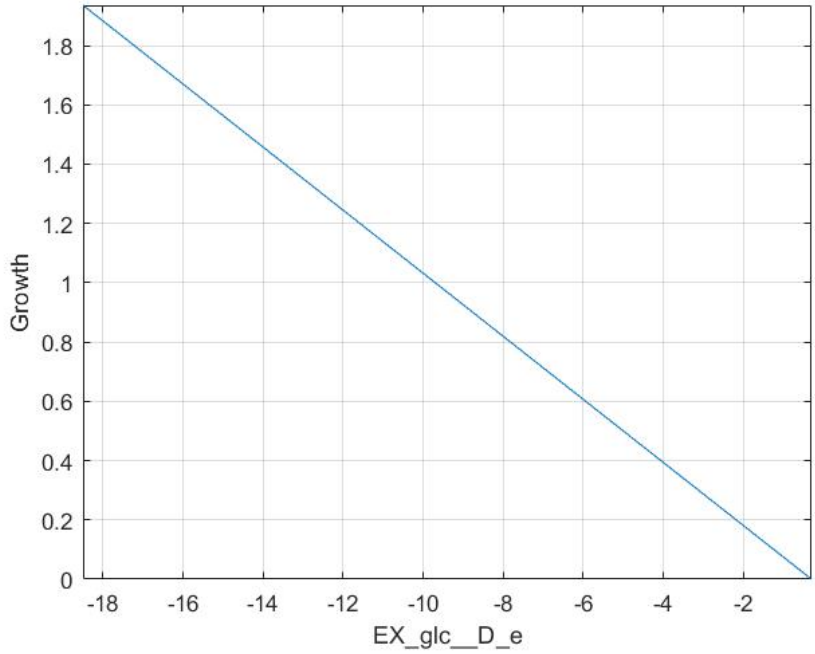

Figure 4.1. Glucose uptake-growth graph in case of unlimited oxygen uptake.

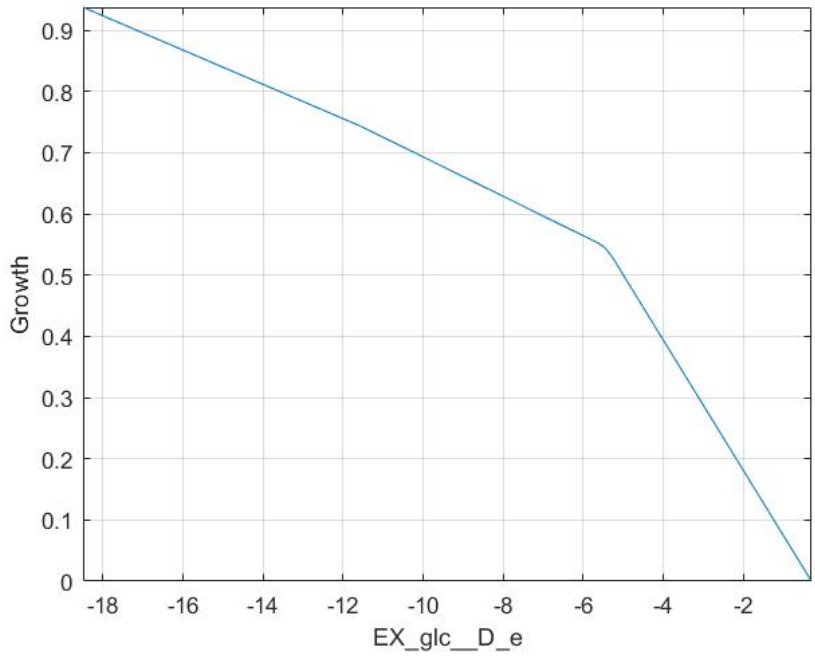

Figure 4.2. Glucose uptake-growth graph with limited oxygen uptake.

4.1.2. Shadow Price

The shadow price of a metabolite in the model is a parameter that shows the sensitivity of the biomass objective function to the change in the amount of this metabolite [19]. All metabolites in the model have a different shadow price value. Additionally, each unique solution in the model creates a different shadow price set. With the shadow price, it can be determined how adding the metabolite affects the biomass objective function. In this section, the change in biomass objective function was determined depending on glucose and oxygen uptake.

In the iYD649 model, the glucose shadow price of each solution calculated based on glucose uptake is given in Figure 4.3. In the graph, the x-axis gives the flux of glucose uptake (EX_glc_D_e). The y-axis on the left side shows the biomass objective function calculated depending on the change in glucose uptake. The y-axis on the right side gives the shadow prices of glucose. During glucose uptake, the amount of oxygen was fixed at 4 mmol/gDW/h.

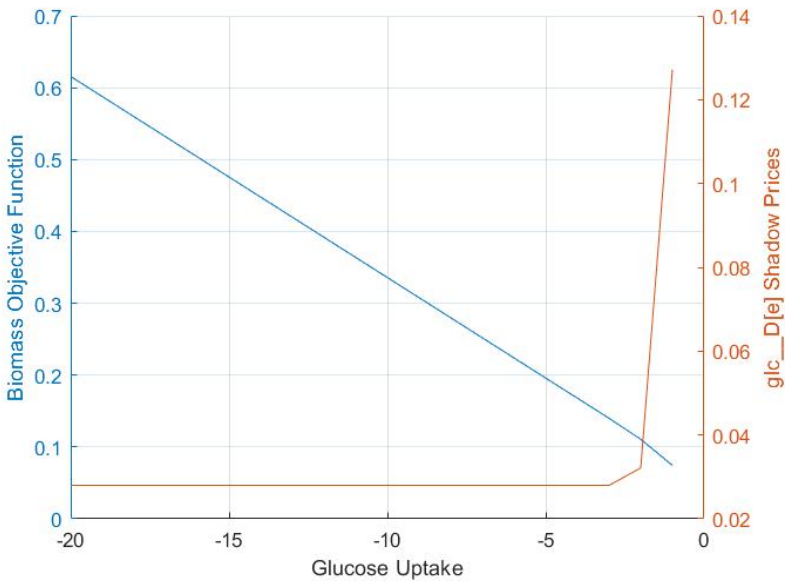

Figure 4.3. Glucose shadow price and biomass production depending on glucose uptake.

According to the graph, as glucose uptake increases, the model produces more biomass. While glucose uptake is in the range of 0-3 mmol/gDW/h, the shadow price of glucose decreases. Then it remains constant as glucose uptake increases. When glucose uptake is 3 mmol/gDW/h, the shadow price of glucose is 0.0280. At this point, the biomass is equal to 0.1399. As 1 mmol/gDW/h glucose is added, the biomass value increases by 0.0280.

In the iYD649 model, the oxygen shadow price of each solution calculated based on oxygen uptake is given in Figure 4.4. In the graph, the x-axis gives the flux of oxygen uptake (EX_o2_e). The y-axis on the left side shows the biomass objective function calculated depending on the change in oxygen uptake. The y-axis on the right side gives the shadow prices of oxygen. During oxygen uptake, the amount of glucose was fixed at 4 mmol/gDW/h.

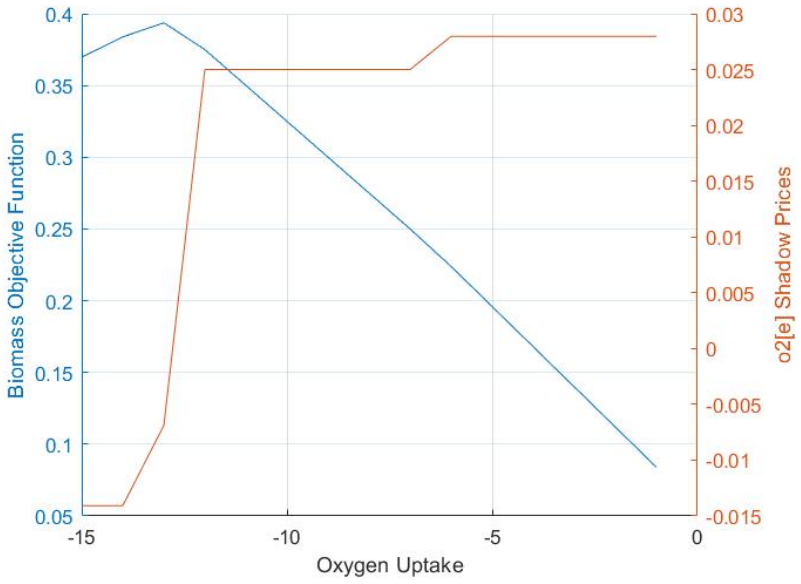

Figure 4.4. Oxygen shadow price and biomass production depending on oxygen uptake.

According to the graph, as the oxygen uptake increases in the range of 1-13 mmol/gDW/h, the biomass increases. After oxygen uptake reaches 12 mmol/gDW/h, the shadow price of oxygen decreases drastically.

4.1.3. NADH, NADPH and ATP Production And Consumption Reactions With Non-Zero Fluxes

NADH, NADPH, and ATP are metabolites involved in several reactions in the iYD649 model. Therefore, it is crucial for the flux consistency of the iYD649 model that the production and consumption amounts of these metabolites are equal. Reactions with non-zero fluxes that produce and consume these metabolites were determined. The production and consumption amounts of these metabolites were calculated and compared. In the iYD649 model, the amount of NADH produced is equal to the amount of NADH consumed. NADH consumption and production reactions are given in Table 4.1 and 4.2, as represented in metabolic models, respectively.

Table 4.1. NADH consumption reactions.

Reaction	**Representation**	**Flux**	**NADH Consumption**
MDH	mal_L[c] + nad[c] $\longleftrightarrow$ h[c] + nadh[c] + oaa[c]	-4.6066	-4.6066
GLUSx	akg[c] + gln_L[c] + h[c] +nadh[c] $\longleftrightarrow$ 2.0 glu_L[c] + nad[c]	2.1009	-2.1009
DHFR2i	dhf[c] + h[c] + nadh[c] $\longleftrightarrow$ nad[c]+ thf[c]	0.0014	-0.0014
NADH2	h[m] + nadh[m] + q6[m] $\longleftrightarrow$ nad[m] + q6h2[m]	20.2384	-20.2384
Total NADH Consumption			-26.9474

Table 4.2. NADH production reactions.

Reaction	**Representation**	**Flux**	**NADH Production**
GAPD	g3p[c] + nad[c] + pi[c] $\longleftrightarrow$ 13dpg[c] + h[c] + nadh[c]	6.3441	6.3441

Table 4.2. NADH production reactions. (cont.)

Reaction	**Representation**	**Flux**	**NADH Production**
C3STDH2	nad[c] + zym_int1[c] ⟷ co2[c] + h[c] + nadh[c] + zym_int2[c]	0.0640	0.0640
IMPD	h2o[c] + imp[c] + nad[c] ⟷ h[c] + nadh[c] + xmp[c]	0.0191	0.0191
IPMD	3c2hmp[c] + nad[c] ⟷ 3c4mop[c] + h[c] + nadh[c]	0.1168	0.1168
SACCD2	h2o[c] + nad[c] + saccrp_L[c] ⟷ akg[c] + h[c] + lys_L[c] + nadh[c]	0.1128	0.1128
HISTD	h2o[c] + histd[c] + 2.0 nad[c] ⟷ 3.0 h[c] + his_L[c] + 2.0 nadh[c]	0.0261	0.0523
PDHm	coa[m] + nad[m] + pyr[m] ⟷ accoa[m] + co2[m] + nadh[m]	4.8074	4.8074
MDHm	mal_L[m] + nad[m] ⟷ h[m] + nadh[m] + oaa[m]	8.9306	8.9306
AKGDm	akg[m] + coa[m] + nad[m] ⟷ co2[m] + nadh[m] + succoa[m]	3.7311	3.7311
HACD1m	aacoa[m] + h[m] + nadh[m] ⟷ nad[m] + 3hbcoa[m]	-2.4904	2.4904
GCCcm	dhlpro[m] + nad[m] ⟷ h[m] + lpro[m] + nadh[m]	0.1661	0.1661
HICITDm	hicit[m] + nad[m] ⟷ h[m] + nadh[m] + oxag[m]	0.1128	0.1128
Total NADH Production			26.9474

In the iYD649 model, the amount of NADPH produced is equal to the amount of NADPH consumed. NADPH consumption and production reactions are given in Table 4.3 and Table 4.4, as represented in metabolic models, respectively.

Table 4.3. NADPH consumption reactions.

Reaction	**Representation**	**Flux**	**NADPH Consumption**
SQLS	2.0 frdp[c] + h[c] + nadph[c] ⟷ nadp[c] + 2.0 ppi[c] + sql[c]	0.1440	0.1440
LNS14DM	2.0 h[c] + 3.0 nadph[c] + 3.0 o2[c] + lanost[c] ⟷ for[c] + 4.0 h2o[c] + 3.0 nadp[c] + 44mctr[c]	0.0640	-0.1919
C14STR	h[c] + nadph[c] + 44mctr[c] ⟷ nadp[c] + 44mzym[c]	0.0640	-0.0640
C4STMO1	3.0 h[c] + 3.0 nadph[c] + 3.0 o2[c] + 44mzym[c] ⟷ 4.0 h2o[c] + 3.0 nadp[c] + 4mzym_int1[c]	0.0640	-0.1919
C3STKR1	h[c] + nadph[c] + 4mzym_int2[c] ⟷ nadp[c] + 4mzym[c]	0.0640	-0.0640
C4STMO2	3.0 h[c] + 3.0 nadph[c] + 3.0 o2[c] + 4mzym[c] ⟷ 4.0 h2o[c] + 3.0 nadp[c] + zym_int1[c]	0.0640	-0.1919
C3STKR2	h[c] + nadph[c] + zym_int2[c] ⟷ nadp[c] + zymst[c]	0.0640	-0.0640
C5STDS	h[c] + nadph[c] + o2[c] + epist[c] ⟷ 2.0 h2o[c] + nadp[c] + ergtrol[c]	0.0003	-0.0003
C22STDS	h[c] + nadph[c] + o2[c] + ergtrol[c] ⟷ 2.0 h2o[c] + nadp[c] + ergtetrol[c]	0.0003	-0.0003
TRDR	h[c] + nadph[c] + trdox[c] ⟷ nadp[c] + trdrd[c]	0.0047	-0.0047
ASAD	aspsa[c] + nadp[c] + pi[c] ⟷ 4pasp[c] + h[c] + nadph[c]	-0.0770	-0.0770

Table 4.3. NADPH consumption reactions. (cont.)

Reaction	**Representation**	**Flux**	**NADPH Consumption**
HSDy	hom_L[c] + nadp[c] ⟷ aspsa[c] + h[c] + nadph[c]	-0.0770	-0.0770
AASAD1	atp[c] + h[c] + nadph[c] + L2aadp[c] ⟷ amp[c] + nadp[c] + ppi[c] + L2aadp6sa[c]	0.1128	-0.1128
SACCD1	glu_L[c] + h[c] + nadph[c] + L2aadp6sa[c] ⟷ h2o[c] + nadp[c] + saccrp_L[c]	0.1128	-0.1128
P5CR	1pyr5c[c] + 2.0 h[c] + nadph[c] ⟷ nadp[c] + pro_L[c]	0.0649	-0.0649
SHK3Dr	3dhsk[c] + h[c] + nadph[c] ⟷ nadp[c] + skm[c]	0.1041	-0.1041
AGPRim	acg5p[m] + h[m] + nadph[m] ⟷ acg5sa[m] + nadp[m] + pi[m]	0.1282	-0.1282
KARA2im	2ahbut[m] + h[m] + nadph[m] ⟷ 23dhmp[m] + nadp[m]	0.0759	-0.0759
KARA1im	alac_S[m] + h[m] + nadph[m] ⟷ 23dhmb[m] + nadp[m]	0.2211	-0.2211
HBCO_m	aacoa[m] + h[m] + nadph[m] ⟷ nadp[m] + 3hbcoa[m]	2.4904	-2.4904
SQLEr	h[r] + nadph[r] + o2[r] + sql[r] ⟷ Ssq23epx[r] + h2o[r] + nadp[r]	0.1440	-0.1440
C24STRer	ergtetrol[r] + h[r] + nadph[r] ⟷ ergst[r] + nadp[r]	0.0003	-0.0003
Total NADPH Consumption			-4.5253

Table 4.4. NADPH production reactions.

Reaction	**Representation**	**Flux**	**NADPH Production**
ICDHyr	icit[c] + nadp[c] ⟷ akg[c] + co2[c] + nadph[c]	1.2170	1.2170
GND	6pgc[c] + nadp[c] ⟷ co2[c] + nadph[c] + ru5p_D[c]	0.1443	0.1443
HMR_1495	nadp[c] + 4mzym_int1[c] ⟷ co2[c] + h[c] + nadph[c] + 4mzym_int2[c]	0.0640	0.0640
PPND2	nadp[c] + pphn[c] ⟷ 34hpp[c] + co2[c] + nadph[c]	0.0402	0.0402
ICDHym	icit[m] + nadp[m] ⟷ akg[m] + co2[m] + nadph[m]	2.9156	2.9156
G6PDH2er	g6p[r] + nadp[r] ⟷ 6pgl[r] + h[r] + nadph[r]	0.1443	0.1443
Total NADPH Production			4.5253

In the iYD649 model, the amount of ATP produced is equal to the amount of ATP consumed. ATP consumption and production reactions are given in Table 4.5 and Table 4.6, as represented in metabolic models, respectively.

Table 4.5. ATP consumption reactions.

Reaction	**Representation**	**Flux**	**ATP Consumption**
HEX1	atp[c] + glc_D[c] ⟷ adp[c] + g6p[c] + h[c]	4.0000	-4.0000
PFK	atp[c] + f6p[c] ⟷ adp[c] + fdp[c] + h[c]	3.1907	-3.1907

Table 4.5. ATP consumption reactions. (cont.)

Reaction	**Reaction**	**Flux**	**ATP Consumption**
PC	atp[c] + hco3[c] + pyr[c] ⟷ adp[c] + h[c] + oaa[c] + pi[c]	0.6407	-0.6407
PRPPS	atp[c] + r5p[c] ⟷ amp[c] + h[c] + prpp[c]	0.0809	-0.0809
FACOAL160	atp[c] + coa[c] + hdca[c] ⟷ amp[c] + pmtcoa[c] + ppi[c]	0.0007	-0.0007
FACOAL100	atp[c] + coa[c] + dca[c] ⟷ amp[c] + dcacoa[c] + ppi[c]	0.0001	-0.0001
FACOAL120	atp[c] + coa[c] + ddca[c] ⟷ amp[c] + ddcacoa[c] + ppi[c]	0.0002	-0.0002
FACOAL161	atp[c] + coa[c] + hdcea[c] ⟷ amp[c] + hdcoa[c] + ppi[c]	0.0004	-0.0004
FACOAL180	atp[c] + coa[c] + ocdca[c] ⟷ amp[c] + ppi[c] + stcoa[c]	0.0001	-0.0001
FACOAL181	atp[c] + coa[c] + ocdcea[c] ⟷ mp[c] + odecoa[c] + ppi[c]	0.0006	-0.0006
FACOAL182	atp[c] + coa[c] + ocdcya[c] ⟷ amp[c] + ocdycacoa[c] + ppi[c]	0.0002	-0.0002
FACOAL140	atp[c] + coa[c] + ttdca[c] ⟷ amp[c] + ppi[c] + tdcoa[c]	0.0003	-0.0003
ARGSS	asp_L[c] + atp[c] + citr_L[c] ⟷ amp[c] + argsuc[c] + h[c] + ppi[c]	0.0633	-0.0633
GK1	atp[c] + gmp[c] ⟷ adp[c] + gdp[c]	0.0009	-0.0009
GMPS	atp[c] + nh4[c] + xmp[c] ⟷ amp[c] + gmp[c] + 2.0 h[c] + ppi[c]	0.0191	-0.0191
ADK1	amp[c] + atp[c] ⟷ 2.0 adp[c]	0.4207	-0.4207

Table 4.5. ATP consumption reactions. (cont.)

Reaction	**Reaction**	**Flux**	**ATP Consumption**
ADNK1	adn[c] + atp[c] ⟷ adp[c] + amp[c] + h[c]	0.0744	-0.0744
ATPM	atp[c] + h2o[c] ⟷ adp[c] + h[c] + pi[c]	5.0000	-5.0000
CBPS	2.0 atp[c] + gln_L[c] + h2o[c] + hco3[c] ⟷ 2.0 adp[c] + cbp[c] + glu_L[c] + 2.0 h[c] + pi[c]	0.1069	-0.2138
NDPK2	atp[c] + udp[c] ⟷ adp[c] + utp[c]	0.6783	-0.6783
UMPK	atp[c] + ump[c] ⟷ adp[c] + udp[c]	0.0200	-0.0200
CTPS1	atp[c] + nh4[c] + utp[c] ⟷ adp[c] + ctp[c] + 2.0 h[c] + pi[c]	0.0176	-0.0176
ASNS1	asp_L[c] + atp[c] + gln_L[c] + h2o[c] ⟷ amp[c] + asn_L[c] + glu_L[c] + h[c] + ppi[c]	0.0401	-0.0401
ASPK	asp_L[c] + atp[c] ⟷ 4pasp[c] + adp[c]	0.0770	-0.0770
HSK	atp[c] + hom_L[c] ⟷ adp[c] + h[c] + phom[c]	0.0770	-0.0770
METAT	atp[c] + h2o[c] + met_L[c] ⟷ amet[c] + pi[c] + ppi[c]	0.0744	-0.0744
AASAD1	atp[c] + h[c] + nadph[c] + L2aadp[c] ⟷ amp[c] + nadp[c] + ppi[c] + L2aadp6sa[c]	0.1128	-0.1128
ATPPRT	atp[c] + prpp[c] ⟷ ppi[c] + prbatp[c]	0.0261	-0.0261
SHKK	atp[c] + skm[c] ⟷ adp[c] + h[c] + skm5p[c]	0.1041	-0.1041

Table 4.5. ATP consumption reactions. (cont.)

Reaction	**Representation**	**Flux**	**ATP Consumption**
FTHFLi	atp[c] + for[c] + thf[c] ⟷ 10fthf[c] + adp[c] + pi[c]	0.0647	-0.0647
PMEVK	atp[c] + 5pmev[c] ⟷ adp[c] + 5dpmev[c]	0.8640	-0.8640
DPMVD	atp[c] + 5dpmev[c] ⟷ adp[c] + co2[c] + ipdp[c] + pi[c]	0.8640	-0.8640
Biomass	59.276 atp[c] + ... ⟷ 59.276 adp[c] + ... + Biomass[c]	0.3940	-23.3576
ATPtm_H	adp[c] + h[c] + atp[m] ⟷ atp[c] + adp[m] + h[m]	29.9426	-29.9426
ACGKm	acglu[m] + atp[m] ⟷ acg5p[m] + adp[m]	0.1282	-0.1282
Total ATP Consumption			-72.4711

Table 4.6. ATP production reactions.

Reaction	**Representation**	**Flux**	**ATP Production**
PGK	3pg[c] + atp[c] ⟷ 13dpg[c] + adp[c]	-6.3441	6.3441
ATPtm_H	adp[c] + h[c] + atp[m] ⟷ atp[c] + adp[m] + h[m]	29.9426	29.9426
NDPK1	atp[c] + gdp[c] ⟷ adp[c] + gtp[c]	-6.0901	6.0901
DGK1	atp[c] + dgmp[c] ⟷ adp[c] + dgdp[c]	-0.0009	0.0009
DADK	atp[c] + damp[c] ⟷ adp[c] + dadp[c]	-0.0014	0.0014

Table 4.6. ATP production reactions. (cont.)

Reaction	**Representation**	**Flux**	**ATP Production**
NDPK3	atp[c] + cdp[c] ⟷ adp[c] + ctp[c]	-0.0094	0.0094
CYTK1	atp[c] + cmp[c] ⟷ adp[c] + cdp[c]	-0.0094	0.0094
URIDK2r	atp[c] + dump[c] ⟷ adp[c] + dudp[c]	-0.0024	0.0024
SUCOASm	atp[m] + coa[m] + succ[m] ⟷ adp[m] + pi[m] + succoa[m]	-3.7311	3.7311
ATPS3m	3.0 h[c] + adp[m] + pi[m] ⟷ atp[m] + h2o[m] + 2.0 h[m]	26.3397	26.3397
Total ATP Production			72.4711

4.2. Metabolic Model Analyses

4.2.1. Gene Deletion Analysis

In order for an organism to grow and reproduce in its environment, it must continue its metabolic activities adapted to environmental conditions. Thus, the organism increases its biomass and reproduces by using the resources of the host organism for its own benefit.

Drug strategies to be developed against the organism focus on metabolic targets that will interrupt or stop the vital metabolic activities of the organism. In this sense, it is of great importance to identify the metabolic reactions that are important for the vital activities of an organism and the genes that catalyse these reactions. For this purpose, single and double gene deletion analyses were performed in the iYD649 model.

The functions of singleGeneDeletion and doubleGeneDeletion in the COBRA Toolbox [35] were used in the gene deletion analyses. In these analyses, glucose is used as the carbon source. The medium also contains aspartate, glutamate, glycine and cysteine as amino acids.

4.2.1.1. Single-Gene Deletion. In the single gene deletion analysis, each of the 649 genes in the iYD649 model was deleted one by one. A new biomass value (grRateKO) was calculated for each gene deleted. grRateKO means the amount of biomass produced when that gene is deleted. In this sense, each gene has its own gRateKO value. In order to understand how the deleted gene is important for the organism, gRateKO was divided by the biomass value of the iYD649 model, and the parameter called grRatio was found for each gene. This parameter is the indicator showing the changes in biomass production due to the deletion of that gene. A threshold value for the grRatio value was determined in the analysis performed, and those below this value were accepted as the essential gene. This threshold value in the analysis was taken as 1e-3.

In addition, all genes that contribute to the increase of the organism's biomass production can be detected due to the single gene deletion analysis. If a gene's grRatio value is equal to 1, it is understood that deletion of this gene does not change the amount of biomass and therefore has no effect on growth. As the grRatio value goes to zero, it is understood that the gene is important for the organism. Based on this principle, HasEffect expression is used to understand the effect of a gene on the organism. If the HasEffect of a gene is TRUE, that gene is affecting biomass production and growth. If the HasEffect is FALSE, that gene has no effect.

As a result of the single-gene deletion analysis, 143 of the 649 genes in the iYD649 model were found to be essential genes for the organism. In addition, 496 genes, including essential genes, were also found to affect biomass production and growth. The genes in the iYD649 model and the corresponding grRateKO, grRatio, HasEffect parameters, and gene essentiality status are given in Appendix C.

The growth rate value obtained by deletion of each gene was plotted in terms of the number of the genes in gene deletion analysis (Figure 4.5). According to Figure 4.5, it is understood that the deletion of more than 150 of 649 genes does not change the growth rate. While deletion of about 45 genes significantly reduces the growth rate, the deletion of 143 genes stops the growth.

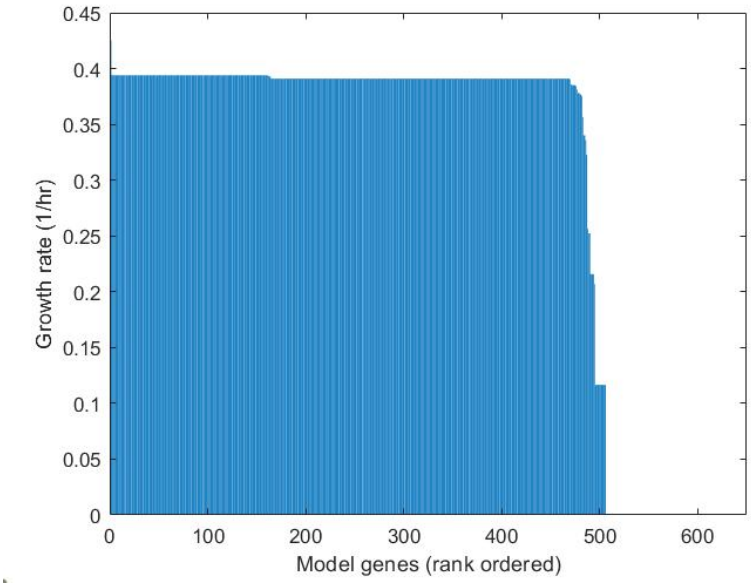

Figure 4.5. The growth rate obtained by deletion of each gene in terms of the number of the genes.

In addition to identifying essential genes of the pathogen, the absence of orthologs of these genes in the human model is critical for finding putative drug targets and understanding the effects of drug applications. Therefore, it was investigated whether they have an ortholog in the human model in single and double gene deletion analyses. In this way, it will be possible to find drug targets that will stop the vital activities of the pathogen without causing any side effects in humans. By searching the orthologs of the essential genes in Panther [32] and Inparanoid [54] databases, it is understood that 57 of 143 essential genes are not found in the human model (Recon3D). The distribution of genes that do not have an ortholog in the human model according to the pathways is given in Table 4.7.

Table 4.7. Essential gene distribution.

Genes Without Ortholog in the Human Model	**Pathway**
CNB03100, CNC04470	Steroid biosynthesis
CNG02210, CNK03240, CNL04470, CNL05510	Oxidative phosphorylation
CNA04370, CND03570, CNA06020	Arginine biosynthesis
CNF01260	Purine metabolism
CNA07120, CNG03730, CNL06550	Pyrimidine metabolism
CNJ02910	Alanine, aspartate and glutamate metabolism
CNA06290, CNC07110, CNI02030, CNA02450, CNJ02040, CNI02930	Glycine, serine and threonine metabolism
CNH03010	Valine, leucine and isoleucine degradation
CNA02570, CNN01460, CNF02480, CNH01530, CNH01520, CNA02270, CNA04310	Valine, leucine and isoleucine biosynthesis
CND01200, CNK00580, CND03850, CNG00170, CND06290	Lysine biosynthesis
CND01510, CND06120, CNB01460, CNA07220, CNH01620, CNB03030 CNC04790	Histidine metabolism

Table 4.7. Essential gene distribution. (cont.)

Genes Without Ortholog in the Human Model	**Pathway**
CNA07970	Phenylalanine metabolism
CNB01990, CNH02650, CNI00560, CNA07880, CNF03410, CNM00820, CNK03330	Phenylalanine, tyrosine and tryptophan biosynthesis
CNG04250	Glycerophospholipid metabolism
CNB02610, CNH03390, CNN02320	Starch and sucrose metabolism
CNE00560, CNF02630, CNC06150	Riboflavin metabolism
CNM00100	Terpenoid backbone biosynthesis
CNJ01640	Transport

4.2.1.2. Double Gene Deletion. The same process was applied by creating a matrix with the genes in the iYD649 model in the double gene deletion analysis. The new biomass value obtained by deleting the genes in the row and column together was recorded in the cell at the intersection of the row and column. If the new biomass value in the cell is less than the threshold value (1e-3), the genes in the row and column corresponding to this cell were considered as essential gene pairs for the iYD649 model. 143 essential genes, found in the single-gene deletion analysis, are growth-stopping genes regardless of which gene they match with in the double gene deletion analysis. Therefore, 143 genes were not included in the double gene deletion analysis. A 506 * 506 matrix containing non-essential genes in terms of the single-gene deletion analysis was created. The deleted gene pairs and new biomass values are given in Figure 4.6.

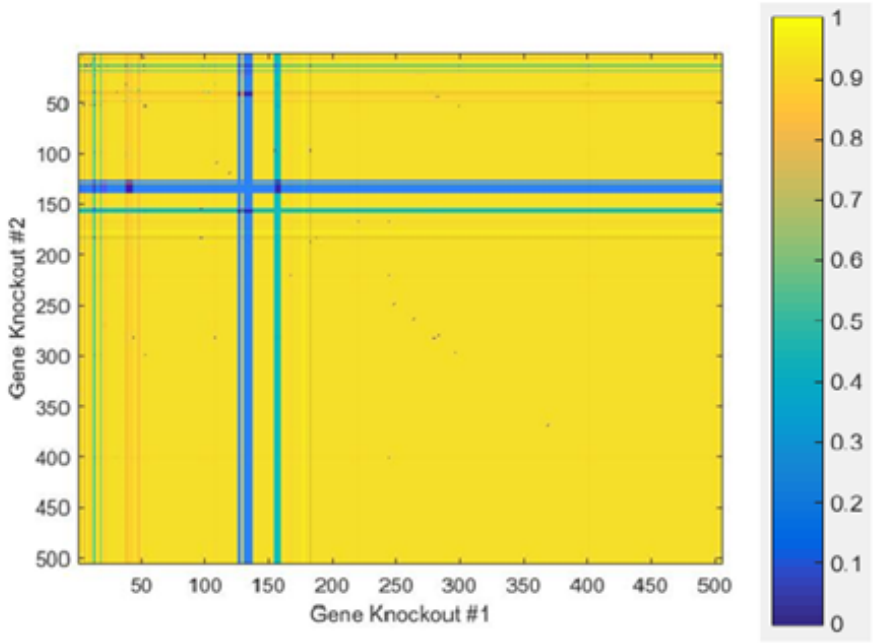

Figure 4.6. Deleted gene pairs and the resulting new biomass values.

As a result, 8 gene pairs that do not have ortholog in the human model were found to be essential for the iYD649 model. The essential gene pairs and reactions regulated by these genes are given in Table 4.8.

Table 4.8. Essential gene pairs and reactions.

Gene Pair No	**Gene ID**	**Reaction**
1	CNH03280	ICL
	CNC04680	THRA2,THRA
2	CNI03590	PPCK
	CNK00280	PGM
3	CNN00260	GLCt1, GALt2
	CNG01480	GLCt1, FRUt2, MANt2
4	CNJ01880	NH4t
	CNA02250	NH4t

Table 4.8. Essential gene pairs and reactions. (cont.)

Gene Pair No	Gene ID	Reaction
5	CNM00800	CYSt2r, TYRt2r, GLNt2r, GLUt2r, ORNt2r, ASPt2r, ARGt2r, GLYt2r, ASNt2r, SERt2r, THRt2r, METt2r, LEUt2r, VALt2r, ILEt2r, LYSt2r, PROt2r
	CNE00270	CYSt2r, TYRt2r, GLNt2r, GLUt2r, ORNt2r, ASPt2r, ARGt2r, GLYt2r, ASNt2r, SERt2r, THRt2r, METt2r, LEUt2r, VALt2r, ILEt2r, LYSt2r, PROt2r
6	CNF04620	UNK3, AATA, PHETA1
	CNB03180	UNK3, AATA, PHETA1
7	CNL06640	DDPA, DDPAm
	CND05120	DDPA
8	CNG00040	G3PCT
	CND01860	G3PCT

4.2.2. Reaction Deletion Analysis

Reaction deletion analysis was applied to determine essential reactions for the organism to maintain its vital activities. For this purpose, the algorithm in the protocol published by Thiele and Palsson [27] was used. In this analysis, similar to the gene deletion analysis, each reaction in the iYD649 model was deleted one by one. A new biomass value was calculated for each reaction deleted. By proportioning the biomass value obtained because of the reaction deletion to the previous biomass value, a parameter called RxnRatio was found.

The parameter RxnRatio for a reaction indicates how the deletion of that reaction affects biomass production. A threshold value for RxnRatio was determined, and reactions having RxnRatio below this threshold value were considered as essential reactions. In the analysis, this threshold value was taken as 1e-3. In this way, reactions that stop or significantly reduce the biomass production in case of the deletion were detected. As a result, 183 of 1267 reactions in the iYD649 model are essential for the organism. Figure 4.7 shows the essential reaction distribution.

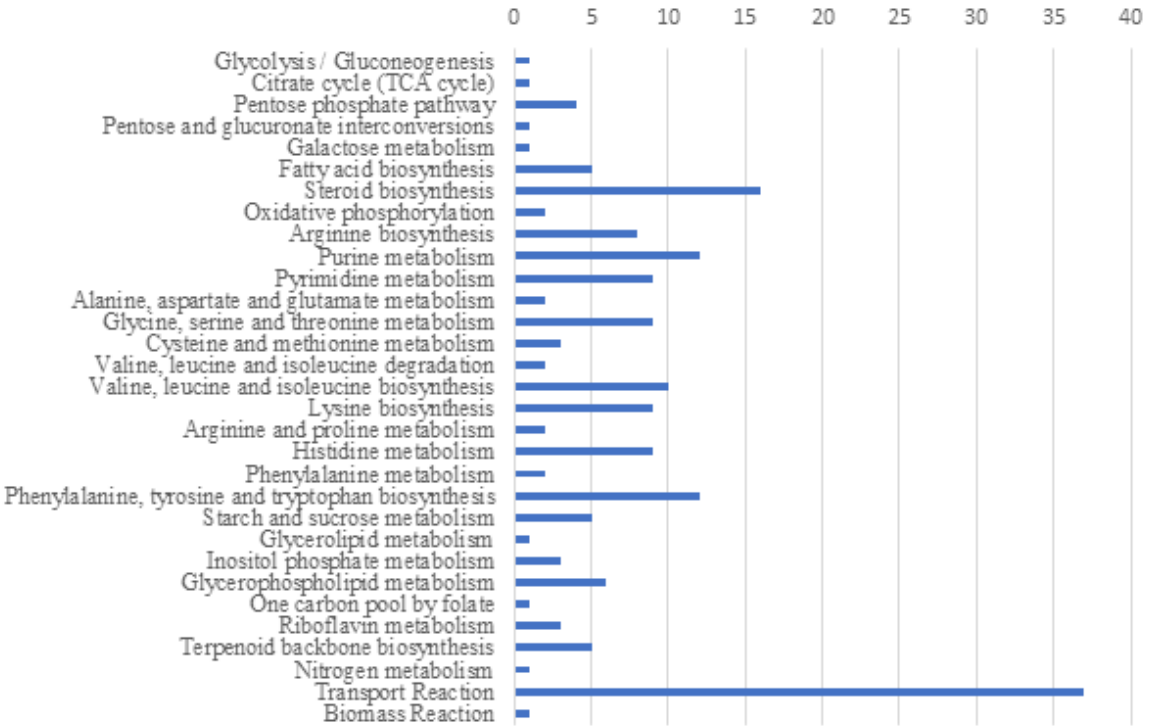

Figure 4.7. Essential reaction distribution.

4.2.3. Metabolite Essentiality Analysis

The metabolite essentiality analysis was performed to identify the metabolites that are essential for the metabolic processes of the iYD649 model. In this analysis, the EMFilter algorithm [25] was used. Firstly, metabolites involved in more than two reactions in the iYD649 model were detected in this analysis. Subsequently, the upper limit and lower limit of the reactions comprising these metabolites were set to zero, and the flux of these reactions was prevented. Thus, the activity of all reactions comprising these metabolites in the iYD649 model was inhibited.

A new biomass value was calculated for each metabolite. Metabolites whose deletions made the new biomass value zero were determined and listed. Then, currency metabolites (metabolites such as water, ATP, NAD, NADH, NADP, NADPH) were removed from these listed metabolites. The remaining metabolites were found to be essential metabolites.

The number of metabolites involved in more than two reactions in the iYD649 model is 379. After preventing the flux of the reactions comprising 379 metabolites and calculating biomass value for each metabolite, it is understood that 177 of these metabolites are essential for biomass production. 69 of the 177 metabolites are currency metabolites. The remaining 108 metabolites were considered as essential metabolites for the iYD649 model. In the essential metabolite analysis, accoa[m], adp[c], adp[m], ahcys[c], akg[c], akg[m], ala_L[c], amet[c], amp[c], gdp[c], atp[c], atp[m], cmp[c], co2[c], co2[m], coa[m], ctp[c], dhf[c], gln_L[c], glu_L[c], glu_L[m], gmp[c], gtp[c], h2o[c], h2o[m], h2o[r], h[c], h[m], h[r], hco3[c], imp[c], nad[c], nad[m], nadh[c], nadh[m], nadp[c], nadp[m], nadp[r], nadph[c], nadph[m], nadph[r], nh4[c], nh4[m], o2[c], pi[c], pi[m], ppi[c], dca[c], ddca[c], pyr[c], pyr[m], q6[m], q6h2[m], so4[c], thf[c], thf[m], trdox[c], trdrd[c], coa[c], o2[m], o2[r], 10fthf[c], h[e], h2o[e], udp[c], udpg[c], ump[c], utp[c] and xmp[c] were taken as the currency metabolites.

108 essential metabolites, found by metabolite essentiality analysis, were compared with the human model (Recon3D). 12 of 108 metabolites were found to be absent in the human model. The distribution of essential metabolites not found in the human model according to the pathways is given in Table 4.9.

Table 4.9. Distribution of essential metabolites.

Metabolite	**Pathway**
13BDglcn[c]	Starch and sucrose metabolism
2dda7p[c]	Phenylalanine, tyrosine and tryptophan biosynthesis

Table 4.9. Distribution of essential metabolites. (cont.)

Metabolite	**Pathway**
3c3hmp[c]	Valine, leucine and isoleucine biosynthesis
3c4mop[c]	Valine, leucine and isoleucine biosynthesis
4r5au[c]	Riboflavin metabolism
aspsa[c]	Glycine, serine and threonine metabolism
chor[c]	Phenylalanine, tyrosine and tryptophan biosynthesis, Folate biosynthesis
epist[c]	Steroid biosynthesis
ergst[c]	Steroid biosynthesis
fecost[c]	Steroid biosynthesis
oxag[m]	Lysine biosynthesis
pphn[c]	Phenylalanine metabolism, Phenylalanine, tyrosine and tryptophan biosynthesis

Metabolites that are not found in humans but are essential for the vital activities of the pathogen are of great importance in terms of the drug target potential. By inhibiting the production of these metabolites in the pathogen, the growth and proliferation of the pathogen are stopped. Since these metabolites are not found in humans, the risk of any side effects to humans is eliminated. In this sense, 12 metabolites, found in the metabolite essentiality analysis, have the potential to be drug targets for the drug strategies to be developed.

4.2.4. Transcriptome Data Integration

Transcriptome data integration is important for studying the metabolic changes of the organism under different conditions. The effect of expression levels of genes on metabolic activities can be understood by integrating transcriptome data. In order to integrate transcriptome data into the iYD649 model, the iMAT algorithm [55] was used. In the model to be created by the integration, it is aimed to maximize the number of reactions having genes with high expression values and to minimize the number of reactions having genes with low expression values.

The results of the transcriptome analysis study conducted by Chen and colleagues [56] were used as transcriptome data. In this study, RNA was extracted from the cerebrospinal fluid (CSF) of two separate patients who had cryptococcal meningitis and had not yet received antifungal treatment. Transcriptome profiling was performed using RNA-seq technology under *in vivo* CSF, *ex vivo* CSF, and laboratory culture (YPD) conditions. For these processes, HC1 and G0 strains were used, and transcriptome data belonging to six different states were obtained. These were named G0_YPD, HC1_YPD, G0_ex_vivo_CSF, HC1_ex_vivo_CSF, G0_in_vivo_CSF, HC1_in_vivo_CSF and added to the Gene Expression Omnibus database (GEO accession: GSE51573).

As a result of integrating the transcriptome data into the iYD649 model, six datasets for six different conditions were obtained. A lower cut-off value is assigned to the biomass objective function for each condition. This value is 0.1. Thus, all six datasets comprises the biomass production reaction (Biomass_Reaction).

G0_YPD dataset (laboratory culture analysis with the G0 strain) obtained by the integration comprises 411 reactions in 38 pathways. In this model, 45 reactions take place in the fatty acid biosynthesis pathway. This pathway is followed by the fatty acid elongation pathway, purine metabolism pathway, and steroid biosynthesis pathway. The numbers of reactions for these pathways are 21, 20, and 16, respectively. The model also has 109 exchange and transport reactions.

HC1_YPD dataset (laboratory culture analysis with HC1 strain) obtained by the integration comprises 397 reactions in 37 pathways. In this model, 45 reactions take place in the fatty acid biosynthesis pathway. This pathway is followed by the purine metabolism pathway, fatty acid elongation pathway, and pyrimidine metabolism pathway. The numbers of reactions for these pathways are 22, 21, and 19, respectively. The model also has 77 exchange and transport reactions.

G0_ex_vivo_CSF dataset (*ex vivo* analysis of cerebrospinal fluid with GO strain) obtained by the integration comprises 431 reactions in 38 pathways. In this model, 45 reactions take place in the fatty acid biosynthesis pathway. This pathway is followed by the purine metabolism pathway, fatty acid elongation pathway, and steroid biosynthesis pathway. The numbers of reactions for these pathways are 22, 20, and 16, respectively. The model also has 121 exchange and transport reactions.

HC1_ex_vivo_CSF dataset (*ex vivo* analysis of cerebrospinal fluid with HC1 strain) obtained by the integration comprises 385 reactions in 38 pathways. In this model, 45 reactions take place in the fatty acid biosynthesis pathway. This pathway is followed by the purine metabolism pathway, fatty acid elongation pathway, and pyrimidine metabolism pathway. The numbers of reactions for these pathways are 20, 19, and 16, respectively. The model also has 87 exchange and transport reactions.

G0_in_vivo_CSF dataset (*in vivo* analysis of cerebrospinal fluid with GO strain) obtained by the integration comprises 439 reactions in 39 pathways. In this model, 45 reactions take place in the fatty acid biosynthesis pathway. This pathway is followed by the purine metabolism pathway, pyrimidine metabolism pathway, and fatty acid elongation pathway. The numbers of reactions for these pathways are 26, 23, and 19, respectively. The model also has 115 exchange and transport reactions.

HC1_in_vivo_CSF dataset (*in vivo* analysis cerebrospinal fluid with HC1 strain) obtained by the integration comprises 330 reactions in 37 pathways. The pathway having highest number of reactions is the purine metabolism pathway in this dataset.

This pathway is followed by the pyrimidine metabolism pathway, steroid biosynthesis pathway and glycine, serine and threonine metabolism pathway. The numbers of reactions for these pathways are 17, 16, and 16, respectively. The model also has 87 exchange and transport reactions.

According to these datasets, lipid metabolism (fatty acid biosynthesis and elongation), purine metabolism, pyrimidine metabolism, and steroid biosynthesis are actively used pathways in case of infection. In addition, it is understood that all of these datasets comprise all essential genes and metabolites not found in the human model when comparing with the essentiality analysis results of the iYD649 model. The number of reactions, metabolites, genes and pathways included in the dataset models are given in Table 4.10.

Table 4.10. Dataset models.

Abbreviation	**Dataset Information**
G0_YPD	411 reactions, 406 metabolites, 278 genes, 38 pathways
HC1_YPD	397 reactions, 392 metabolites 283 genes, 37 pathways
G0_ex_vivo_CSF	431 reactions, 415 metabolites 290 genes, 38 pathways
HC1_ex_vivo_CSF	385 reactions, 389 metabolites 278 genes, 38 pathways
G0_in_vivo_CSF	439 reactions, 412 metabolites 291 genes, 39 pathways
HC1_in_vivo_CSF	330 reactions, 333 metabolites 274 genes , 37 pathways

4.2.5. Flux Coupling Analysis

Flux coupling analysis is important for understanding the relationship between reactions in the metabolic model. F2C2 algorithm was used in the flux coupling analysis. F2C2 is a new tool allowing to determine the dependency of reactions in terms of their fluxes [57]. With this tool, the stoichiometric constraints and thermodynamic constraints are involved in finding out how reactions affect each other.

In this analysis, first of all, the dead-end metabolites and blocked reactions are determined. Then, rows and columns containing these metabolites and reactions are iteratively removed from the stoichiometric matrix. Larhlimi and colleagues [57] reported that removing these rows and columns does not affect the result of the flux coupling analysis. In this way, network reduction is provided. The remaining rows and columns in the matrix are combined with each other. In the last step, the reactions that are coupled with each other are marked on the stoichiometric matrix.

There are three types of reaction pairs in terms of flux coupling. If the flux of one reaction proportionally affects the flux of another reaction, these reactions are defined as fully coupled. If a reaction with zero flux causes the flux of the other reaction to being equal to zero, these reactions are partially coupled reactions. If an irreversible reaction with zero flux implies the flux of another reaction to be zero, these reactions are directionally coupled [57]. The coupled pairs in the iYD649 model are given in Table 4.11.

Table 4.11. Coupled pairs in the iYD649 model

Type of Coupling	**Number of Pairs**
Fully coupled pairs	1389
Partially coupled pairs	274
Directionally coupled pairs	7866

As a result, all essential reactions, found in the reaction essentiality analysis, are coupled with the biomass reaction. There are 16 reactions that are fully coupled with the biomass reaction in the essential reactions. These are IMPD, DADK, TRPS1, ATP-PRT, PRATPP, PRAMPC, PRMICI, IG3PS, IGPDH, HSTPT, HISTP, HISTD, IGPS, PRAIi, ANPRT, EX_biomass. There are 5 reactions that are partially coupled with the biomass reaction in the essential reactions. These are RNDR1, DB4PS, RBFSa, RBFSb, DM_4r5au_c. The remaining part of the essential reactions is directionally coupled with the biomass reaction.

In metabolic models, analysis of the dependency of reactions in terms of their fluxes is provided by the flux coupling analysis. In the iYD649 model, the biomass objective reaction (Biomass_Reaction) is the reaction in which the biomass of the organism is produced. By the production of biomass, the organism grows and reproduces. Determination of reactions affecting the production of biomass is important for the strategies to develop the compounds for antifungal therapy.

In the flux coupling analysis of the iYD649 model, 16 reactions were found to be fully coupled with the biomass reaction. In this sense, the flux of the biomass reaction is directly proportional to the fluxes of these 16 reactions. Perturbating at least one of these reactions will change the amount of biomass production. Similarly, inhibiting the partially or directionally coupled reactions will stop biomass production. All these biomass-coupled reactions have the potential for drug targeting.

4.3. Infection and Drug Targeting

4.3.1. Drugs Therapy for *C. neoformans* Infection

There are three main antifungal drug groups for the treatment of *C. neoformans* infection. These are polyenes, azoles and echinocandins [11]. Polyenes target ergosterol, a lipid component of the membrane, by binding it, the membrane permeability is changed due to channel formation in the membrane [13], [14].

Azoles focus on decreasing the amount of ergosterol by inhibiting the conversion reactions of lanosterol into ergosterol [12]. In this sense, ergosterol, which cannot be produced by mammals, is the target metabolite for polyenes and azoles groups. Echinocandins have the activity to inhibit the synthesis of 1,3-β-glucan which is a component of the cell wall [58].

Ergosterol (ergst) and 1,3-β-glucan (13BDglcn) were two of the 12 metabolites found as essential metabolites in the metabolite essentiality analysis performed using the iYD649 model. These metabolites are also included in the biomass objective reaction of the iYD649 model, and their production contribute to the growth. In addition, episterol (epist) and fecosterol (fecost), which are intermediates in the conversion reactions of lanosterol into ergosterol, were also found as essential metabolites for the iYD649 model.

1,3-β-glucan (13BDglcn) synthesis occurs in the presence of 1,3-β-glucan synthase (gene ID: CNN02320) in the iYD649 model. In the single-gene deletion analysis, this gene (CNN02320) was found to be an essential gene that does not have an ortholog in the human model. The iYD649 model has two essential genes in the steroid biosynthesis pathway. These genes are CNB03100 and CNC04470, and they have no ortholog in the human model. CNB03100 regulates the conversion reaction between fecosterol and zymosterol. CNC04470 regulates the reaction of producing ergosterol from ergtetrol. In this sense, the drug targets of polyenes, azoles and echinocandins are consistent with the findings obtained by iYD649 model.

4.3.2. *C. neoformans* Metabolism During Infection

Although the transcriptome data for *C. neoformans var. grubii* (serotype A) H99 is available in the Gene Expression Omnibus database, there is no transcriptome data for *C. neoformans var. neoformans* JEC21 (serotype D) during the infection. In addition, some genes found in serotype D do not have an ortholog in serotype A.

Due to the lack of the ortholog gene, it is not possible to integrate transcriptome data under infection conditions to obtain a dataset of the same size as the iYD649 model. Instead, the infection environment needs to be simulated for the present model (iYD649). For this purpose, an infection environment was created for the iYD649 model to simulate the situation in macrophage in the host. To apply the conditions in the human lung, the oxygen uptake is limited, and carbon source uptake is increased by the constraints of the iYD649 model. According to the iYD649 model, the growth rate increases from 0.3940/h to 1.3393/h when the infection occurs.

In the iYD649 model under the infection conditions stated above, the fluxes of reactions catalyzed by aconitase and succinate dehydrogenase increase during the infection, consistent with the analysis conducted by Hu and colleagues [59]. The flux of ACONTm (catalyzed by aconitase) increases from 1.66 to 5.69, and the flux of SUCD1m (catalyzed by succinate dehydrogenase) increases from 4.18 to 9.98. Hu and colleagues [59] conducted *in vivo* transcriptome analysis of *C. neoformans* to understand its metabolism during infection. They compared the experimental results with their *in vitro* SAGE libraries. They found that the expression of genes in carbon metabolism and lipid metabolism was elevated, indicating the importance of carbon source utilization in case of the infection. Moreover, the genes that encode aconitase and succinate dehydrogenase in TCA cycle were also found as elevated [59].

Hu and colleagues [59] found that phosphoenolpyruvate carboxykinase catalyzed reaction in gluconeogenesis is important since the glucose is limited in the infected tissue in the early stage of the infection. Phosphoenolpyruvate carboxykinase is an enzyme in the lyase family used in the metabolic pathway of gluconeogenesis. It converts oxaloacetate into phosphoenolpyruvate and carbon dioxide. It is found in two forms, cytosolic and mitochondrial. In the iYD649 model, PPCK (phosphoenolpyruvate carboxykinase) catalyzed reaction has no flux when the carbon source is glucose. However, the flux of PPCK in the iYD649 model varies from 1.27 to 4.11 when acetate, fumarate or succinate is used as the sole carbon source during infection.

Acetyl-CoA synthetase was found to be elevated by Hu and colleagues [59] since the amount of acetate in infected tissues was greater than other metabolites [60]. The acetyl CoA synthetase catalyzes the conversion of acetate to acetyl-CoA. In the iYD649 model, this reaction (ACS) has flux in the presence of acetate.

The flux of the SUCFUMtm (succinate:fumarate antiporter reaction) in the iYD649 model increases from 0.45 to 1.08 under the infection conditions. Likewise, the gene expression of the succinate:fumarate antiporter was also found to be elevated during infection by Hu and colleagues [59].

4.3.3. Virulence factors for *C. neoformans*

Virulence factors for *C. neoformans* infection include phospholipase, urease, growth at body temperature, melanin production, and capsule formation [5], [7], [61], [62]. The findings of the iYD649 model are consistent with the literature in terms of virulence factors. One of the specific enzymes defined as a virulence factor is phospholipase including phospholipase B and lysophospholipase [7]. In the iYD649 model, the genes of CND04180 and CNM00920 encode enzymes in the class of phospholipase. They are responsible for lipid degradation and sn-glycero-3-phosphocholine production. When 4 mmol/gDW/h glucose is used as the carbon source in the model, the amount of sn-glycero-3-phosphocholine produced is 0.0026 mmol/gDW/h. Sabiiti and colleagues [5] reported that phospholipases affect membrane stabilization and lead to infection by suppressing the immune system.

Cox and colleagues [63] defined urease enzyme as a virulence factor. Urease is an enzyme that hydrolyses urea to ammonia. CNH01900 encodes urease enzyme in the iYD649 model. Depending on the increase in this enzyme activity in the iYD649 model, the mitochondrial ATP production and the flux in the oxidative phosphorylation pathway increase. This is consistent with the literature, confirming that invasion can be achieved due to urease enzyme activity.

In the absence of pyruvate kinase, *C. neoformans* has reduced virulence activity [5]. In the iYD649 model, pyruvate kinase is encoded by CNC03080. By this gene activity, phosphoenolpyruvate is converted into pyruvate. The deletion of CNC03080 in the iYD649 model leads to a reduction of biomass production from 0.3940/h to 0.2563/h.

Carbon utilization is vital in terms of growth at body temperature. Price and colleagues [10] reported that glucose utilization plays an important role in *C. neoformans* virulence. They showed that the glycolysis pathway has great importance in persisting in the central nervous system (CNS) and causes virulence. The deficiency in glucose utilization causes the reduced pathogen invasion. Accordingly, even though alternative carbon sources are utilized in case of a defect in glucose usage, the pathogen becomes less virulent.

Biomass production values were calculated for different carbon sources by allowing the uptake of carbon sources in the iYD649 model to understand the effects of carbon sources on biomass production and virulence. Table 4.12 gives data about the biomass production obtained by different carbon sources.

Table 4.12. Biomass production depending on the carbon source.

Carbon Source	**Biomass Production**
Glucose	0.3940
Fructose	0.3631
Acetate	0.0329
Fumarate	0.1640
Galactose	0.3631
Glycerol	0.1757
Sorbitol	0.4222
Succinate	0.1924

These results show that more growth is achieved by glucose compared to other carbon sources. In this sense, the results regarding growth and virulence are consistent with the literature.

Price and colleagues [10] reported that hexose kinase activity in *C. neoformans* is also required for virulence. Hexose kinase I and II are encoded by CNB02660 and CNH01400 in the iYD649 model. HEX1 reaction, converting D-glucose into D-glucose 6-phosphate, takes place in the presence of CNB02660 or CNH01400. Inhibiting hexose kinase activity stops biomass production in the iYD649 model.

It is known that lipids affect virulence and immune responses. The capsule formation of *C. neoformans* is one of its most important virulence factors [64], [65]. Chrisman and colleagues [66] found that phospholipids, especially phosphatidylcholine (pc_SC), have a capsule enlarging effect on *C. neoformans* during infection with macrophages. Among the phospholipids, phosphatidylcholine (pc_SC), phosphatidate (pa_SC), phosphatidylethanolamine (pe_SC), phosphatidylinositol (ptd1ino_SC) provide the capsule enlargement. These metabolites are produced in glycerolipid metabolism and glycerophospholipid metabolism pathways in the iYD649 model. They are also included in the biomass objective function of the iYD649 model.

ERG11 gene was found to contribute to the azole resistance in *Candida albicans* in sterol metabolism [12], [67]. The iYD649 model has an ortholog (CNA00300) for this gene. CNA00300 encodes lanosterol 14 alpha demethylase catalyzing the reaction converting lanosterol into 4,4'-dimethyl cholesta-8,14,24-trienol. CNA00300 was also found as an essential gene in the gene essentiality analysis.

5. CONCLUSION AND RECOMMENDATIONS

5.1. Conclusions

In this study, the genome-scale metabolic model (iYD649) was reconstructed specific to *C.neoformans*. By using the computational systems biology approach and data integration, metabolic processes of this pathogen were attempted to be predicted in different environmental conditions. In addition, potential drug targets for *C.neoformans* infections were investigated by gene, reaction, and metabolite essentiality analyses.

The iYD649 model contains 1267 reactions in 57 pathways. Of these reactions, 239 are transport reactions and 116 are exchange reactions. There are 1140 metabolites and 649 genes in this model.

In the metabolite essentiality analysis, 108 metabolites in the iYD649 model were found to be essential. 12 of these 108 metabolites are not found in the human model. Therefore, the 12 metabolites have the potential to be drug target.

In the single-gene deletion analysis, 143 of the 649 genes in the iYD649 model were found to be essential genes. 57 of 143 essential genes do not have an ortholog in the human model. Therefore, these 57 genes can be considered as potential drug targets. In addition, 8 gene pairs were essential pairs that do not have an ortholog in the human model according to the double-gene deletion analysis performed by using the iYD649 model.

In the reaction essentiality analysis, 183 of 1267 reactions were found as essential reactions in the iYD649 model. All essential reactions were found as coupled with the biomass reaction in the flux coupling analysis. This shows that two different analyses are consistent with each other.

The iYD649 model predicted that during the infection, carbon metabolism was actively used, and the utilization of carbon source may affect the strength of infection. This result is in agreement with the study of Hu and colleagues [59].

The drug targets of main drug groups (polyenes, azoles and echinocandins) used for *C. neoformans* infections are consistent with the findings obtained by the iYD649 model. In the iYD649 model, the essentiality of ergosterol metabolite and ergosterol production gene (gene ID: CNC04470) is consistent with the drug target of the polyenes group drugs [13], [14]. The essentiality of fecosterol and episterol metabolites and fecosterol production gene (gene ID: CNB03100) is consistent with the drug target of the azole group drugs [12], [68] [69]. The essentiality of 1,3-β-glucan metabolite and 1,3-β-glucan production gene (gene ID: CNN02320) is consistent with the drug target of the echinocandins group drugs [58].

Chorismate (chor) is one of the 12 essential metabolites in the iYD649 model. The genes (gene IDs: CNH02650, CNF03410, CNI00560 and CNM00820) that regulate chorismate producing and consuming reactions in the iYD649 model were found to be essential genes in gene essentiality analysis. These genes are not found in the human model. In this sense, chorismate can be a potential drug target, consistent with the study of Ziebart and colleagues [70] reporting that enzymes that utilize the chorismate are important antimicrobial drug targets since they have a central role in survival and virulence.

In the iYD649 model, the reaction producing 2-dehydro-3-deoxy-D-arabino- heptonate 7-phosphate (2dda7p) from phosphoenolpyruvate and erythrose 4-phosphate is an essential reaction. The genes (gene IDs: CND05120 and CNL06640) regulating this reaction were also found as essential genes. Moreover, 2-dehydro-3-deoxy-D-arabino-heptonate 7-phosphate (2dda7p) is an essential metabolite. These results are consistent with Ducati and colleagues [71] reporting that 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase is an antimicrobial drug target since it is important for controlling carbon flow into the shikimate pathway.

L-aspartate 4 - semialdehyde (aspsa) is a metabolite among the 12 essential metabolites. The reaction (reaction abbreviation: ASAD) producing the L-aspartate 4 - semialdehyde is an essential reaction, and the gene (gene ID: CNA02450) that regulates this reaction is an essential gene for the iYD649 model. These results show that L-aspartate 4 - semialdehyde can be a potential drug target, consistent with the study of Dahal and colleagues [72], [73] reporting that aspartate semialdehyde dehydrogenase is a drug target for antifungal drug development.

The iYD649 model has 69 essentials (57 genes and 12 metabolites) not found in the human model. In DrugBank database [74], there are drugs targeting 16 of these essentials. The results of DrugBank research for the potential drug targets of the iYD649 model are given in Table 5.1

Table 5.1. DrugBank results.

Target	**Drugs**	**Organism**
Ergosterol	Candicidin, Nystatin, Butoconazole, Amphotericin B, Natamycin, Clotrimazole	*Candida albicans*
Chorismate	Flavin mononucleotide	*Streptococcus pneumoniae, Helicobacter pylori*
1,3-β-glucan	Ibrexafungerp, Anidulafungin, Caspofungin, Micafungin	*Aspergillus niger*
CNA07880	5-O-phosphono-alpha-D-ribofuranosyl diphosphate	*Erwinia carotovora*
CNF01260	Flavin adenine dinucleotide, Azelaic acid	*Escherichia coli, Staphylococcus aureus*
CNN02320	Ibrexafungerp, Anidulafungin, Caspofungin, Micafungin	*Aspergillus niger*
CNH01520	Alpha-Ketoisovalerate	*Mycobacterium tuberculosis*

Table 5.1. DrugBank results. (cont.)

Target	Drugs	Organism
CNA07120	Dihydroorotic Acid, Orotic acid, Lysine Nz-Carboxylic Acid, N-Carbamoylaspartic acid	*Escherichia coli*
CNG03730	5-O-phosphono-alpha-D-ribofuranosyl diphosphate, Orotic acid	*Salmonella typhimurium*
CNH02650	Flavin mononucleotide	*Streptococcus pneumoniae, Helicobacter pylori*
CNF02630	Dithioerythritol	*Mycobacterium tuberculosis*
CNL06550	6-hydroxyuridine-5'-phosphate, 6-oxouridine 5'-phosphate	*Bacillus subtilis, Escherichia coli*
CNA02450	Nicotinamide adenine dinucleotide phosphate, (4s)-4-{[(2s)-2-Amino-3-Oxopropyl] Sulfanyl}-L-Homoserinate	*Haemophilus influenzae*
CNA02570	Triethylene glycol, Cocarboxylase	*Klebsiella pneumoniae*
CNA04310	3-Isopropylmalic Acid	*Thiobacillus ferrooxidans*
CNA07970	8-Hydroxy-2-oxa-bicyclo[3.3.1] non-6-ene-3,5-dicarboxylic acid	*Escherichia coli*

5.2. Recommendations

With the iYD649 model, a genome-scale model reconstructed for *C.neoformans*, the metabolic processes of this pathogen under different conditions were attempted to be elucidated. This model allows researchers to analyse the changes in the cellular metabolism of this pathogen under different conditions and to use results for their studies.

For further study, combining the iYD649 model with the human tissue models (lung, brain, etc.) will help to examine the pathogen-host metabolism in case of infection with respect to systems biology approach. It will also be possible to investigate the results of manipulations on drug targets in both models.

REFERENCES

1. Park, B. J., K. A. Wannemuehler, B. J. Marston, N. Govender, P. G. Pappas and T. M. Chiller, "Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS", *Aids*, Vol. 23, No. 4, pp. 525–530, 2009.

2. Chayakulkeeree, M. and J. R. Perfect, "Cryptococcosis", *Diagnosis and Treatment of Human Mycoses*, pp. 255–276, 2008.

3. Dambuza, I. M., T. Drake, A. Chapuis, X. Zhou, J. Correia, L. Taylor-Smith, N. LeGrave, T. Rasmussen, M. C. Fisher, T. Bicanic *et al.*, "The Cryptococcus neoformans Titan cell is an inducible and regulated morphotype underlying pathogenesis", *PLoS Pathogens*, Vol. 14, No. 5, e1006978, 2018.

4. Srikanta, D., F. H. Santiago-Tirado and T. L. Doering, "Cryptococcus neoformans: historical curiosity to modern pathogen", *Yeast*, Vol. 31, No. 2, pp. 47–60, 2014.

5. Sabiiti, W. and R. C. May, "Mechanisms of infection by the human fungal pathogen Cryptococcus neoformans", *Future Microbiology*, Vol. 7, No. 11, pp. 1297–1313, 2012.

6. Kim, H., K.-W. Jung, S. Maeng, Y.-L. Chen, J. Shin, J. E. Shim, S. Hwang, G. Janbon, T. Kim, J. Heitman *et al.*, "Network-assisted genetic dissection of pathogenicity and drug resistance in the opportunistic human pathogenic fungus Cryptococcus neoformans", *Scientific Reports*, Vol. 5, No. 1, pp. 1–10, 2015.

7. Buchanan, K. L. and J. W. Murphy, "What makes Cryptococcus neoformans a pathogen?", *Emerging Infectious Diseases*, Vol. 4, No. 1, p. 71, 1998.

8. Rhome, R., A. Singh, T. Kechichian, M. Drago, G. Morace, C. Luberto and M. Del Poeta, "Surface localization of glucosylceramide during Cryptococcus neoformans infection allows targeting as a potential antifungal", *PLoS One*, Vol. 6,

No. 1, e15572, 2011.

9. Waterman, S. R., Y.-D. Park, M. Raja, J. Qiu, D. A. Hammoud, T. V. O'Halloran and P. R. Williamson, "Role of CTR4 in the virulence of Cryptococcus neoformans", *MBio*, Vol. 3, No. 5, e00285–12, 2012.

10. Price, M. S., M. Betancourt-Quiroz, J. L. Price, D. L. Toffaletti, H. Vora, G. Hu, J. W. Kronstad and J. R. Perfect, "Cryptococcus neoformans requires a functional glycolytic pathway for disease but not persistence in the host", *MBio*, Vol. 2, No. 3, e00103–11, 2011.

11. McEvoy, K., T. G. Normile and M. D. Poeta, "Antifungal Drug Development: Targeting the Fungal Sphingolipid Pathway", *Journal of Fungi*, Vol. 6, No. 3, p. 142, 2020.

12. Mota Fernandes, C. and M. Del Poeta, "Fungal sphingolipids: role in the regulation of virulence and potential as targets for future antifungal therapies", *Expert Review of Anti-Infective Therapy*, Vol. 18, No. 11, pp. 1083–1092, 2020.

13. Gray, K. C., D. S. Palacios, I. Dailey, M. M. Endo, B. E. Uno, B. C. Wilcock and M. D. Burke, "Amphotericin primarily kills yeast by simply binding ergosterol", *Proceedings of the National Academy of Sciences*, Vol. 109, No. 7, pp. 2234–2239, 2012.

14. Anderson, T. M., M. C. Clay, A. G. Cioffi, K. A. Diaz, G. S. Hisao, M. D. Tuttle, A. J. Nieuwkoop, G. Comellas, N. Maryum, S. Wang *et al.*, "Amphotericin forms an extramembranous and fungicidal sterol sponge", *Nature Chemical Biology*, Vol. 10, No. 5, p. 400, 2014.

15. Ideker, T., T. Galitski and L. Hood, "A new approach to decoding life: systems biology", *Annual Review of Genomics and Human Genetics*, Vol. 2, No. 1, pp. 343–372, 2001.

16. Haggart, C. R., J. A. Bartell, J. J. Saucerman and J. A. Papin, "Whole-Genome Metabolic Network Reconstruction and Constraint-Based Modeling", *Methods in Enzymology*, Vol. 500, pp. 411–433, 2011.

17. Sun, J., B. Sayyar, J. E. Butler, P. Pharkya, T. R. Fahland, I. Famili, C. H. Schilling, D. R. Lovley and R. Mahadevan, "Genome-scale constraint-based modeling of Geobacter metallireducens", *BMC Systems Biology*, Vol. 3, No. 1, pp. 1–15, 2009.

18. Kauffman, K. J., P. Prakash and J. S. Edwards, "Advances in flux balance analysis", *Current Opinion in Biotechnology*, Vol. 14, No. 5, pp. 491–496, 2003.

19. Orth, J. D., I. Thiele and B. Ø. Palsson, "What is flux balance analysis?", *Nature Biotechnology*, Vol. 28, No. 3, pp. 245–248, 2010.

20. Francke, C., R. J. Siezen and B. Teusink, "The evolutionary relationship between genes", *Trends in Microbiology*, Vol. 11, No. 13, pp. 550–558, 2005.

21. Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick *et al.*, "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd", *Science*, Vol. 269, No. 5223, pp. 496–512, 1995.

22. Fraser, C. M., J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley *et al.*, "The minimal gene complement of Mycoplasma genitalium", *Science*, Vol. 270, No. 5235, pp. 397–404, 1995.

23. Dunphy, L. J. and J. A. Papin, "Biomedical applications of genome-scale metabolic network reconstructions of human pathogens", *Current Opinion in Biotechnology*, Vol. 51, pp. 70–79, 2018.

24. Sertbas, M. and K. O. Ulgen, "Genome-Scale Metabolic Modeling for Unraveling

Molecular Mechanisms of High Threat Pathogens", *Frontiers in Cell and Developmental Biology*, Vol. 8, 2020.

25. Kim, H. U., T. Y. Kim and S. Y. Lee, "Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen Acinetobacter baumannii AYE", *Molecular Biosystems*, Vol. 6, No. 2, pp. 339–348, 2010.

26. Kim, H. U., S. Y. Kim, H. Jeong, T. Y. Kim, J. J. Kim, H. E. Choy, K. Y. Yi, J. H. Rhee and S. Y. Lee, "Integrative genome-scale metabolic analysis of Vibrio vulnificus for drug targeting and discovery", *Molecular Systems Biology*, Vol. 7, No. 1, p. 460, 2011.

27. Thiele, I. and B. Ø. Palsson, "A protocol for generating a high-quality genome-scale metabolic reconstruction", *Nature Protocols*, Vol. 5, No. 1, p. 93, 2010.

28. Kanehisa, M. and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes", *Nucleic Acids Research*, Vol. 28, No. 1, pp. 27–30, 2000.

29. Apweiler, R., A. Bairoch, C. H. Wu, W. C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane *et al.*, "UniProt: the universal protein knowledgebase", *Nucleic Acids Research*, Vol. 32, No. suppl_1, pp. D115–D119, 2004.

30. Seaver, S. M., F. Liu, Q. Zhang, J. Jeffryes, J. P. Faria, J. N. Edirisinghe, M. Mundy, N. Chia, E. Noor, M. E. Beber *et al.*, "The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes", *Nucleic Acids Research*, Vol. 49, No. D1, pp. D575–D588, 2021.

31. Small, I., N. Peeters, F. Legeai and C. Lurin, "Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences", *Proteomics*, Vol. 4, No. 6, pp. 1581–1590, 2004.

32. Mi, H., A. Muruganujan, X. Huang, D. Ebert, C. Mills, X. Guo and P. D. Thomas, "Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v. 14.0)", *Nature Protocols*, Vol. 14, No. 3, pp. 703–721, 2019.

33. Yachdav, G., E. Kloppmann, L. Kajan, M. Hecht, T. Goldberg, T. Hamp, P. Hönigschmid, A. Schafferhans, M. Roos, M. Bernhofer *et al.*, "PredictProtein—an open resource for online prediction of protein structural and functional features", *Nucleic Acids Research*, Vol. 42, No. W1, pp. W337–W343, 2014.

34. Armenteros, J. J. A., M. Salvatore, O. Emanuelsson, O. Winther, G. Von Heijne, A. Elofsson and H. Nielsen, "Detecting sequence signals in targeting peptides using deep learning", *Life Science Alliance*, Vol. 2, No. 5, 2019.

35. Becker, S. A., A. M. Feist, M. L. Mo, G. Hannum, B. Ø. Palsson and M. J. Herrgard, "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox", *Nature Protocols*, Vol. 2, No. 3, p. 727, 2007.

36. Agren, R., L. Liu, S. Shoaie, W. Vongsangnak, I. Nookaew and J. Nielsen, "The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum", *PLoS Comput Biol*, Vol. 9, No. 3, e1002980, 2013.

37. King, Z. A., J. Lu, A. Dräger, P. Miller, S. Federowicz, J. A. Lerman, A. Ebrahim, B. O. Palsson and N. E. Lewis, "BiGG Models: A platform for integrating, standardizing and sharing genome-scale models", *Nucleic Acids Research*, Vol. 44, No. D1, pp. D515–D522, 2016.

38. Loftus, B. J., E. Fung, P. Roncaglia, D. Rowley, P. Amedeo, D. Bruno, J. Vamathevan, M. Miranda, I. J. Anderson, J. A. Fraser *et al.*, "The genome of the basidiomycetous yeast and human pathogen Cryptococcus neoformans", *Science*, Vol. 307, No. 5713, pp. 1321–1324, 2005.

39. Degtyarenko, K., P. De Matos, M. Ennis, J. Hastings, M. Zbinden, A. McNaught, R. Alcántara, M. Darsow, M. Guedj and M. Ashburner, "ChEBI: a database and ontology for chemical entities of biological interest", *Nucleic Acids Research*, Vol. 36, No. suppl_1, pp. D344–D350, 2007.

40. Mo, M. L., B. Ø. Palsson and M. J. Herrgård, "Connecting extracellular metabolomic measurements to intracellular flux states in yeast", *BMC Systems Biology*, Vol. 3, No. 1, pp. 1–17, 2009.

41. Sohn, S. B., T. Y. Kim, J. H. Lee and S. Y. Lee, "Genome-scale metabolic model of the fission yeast Schizosaccharomyces pombe and the reconciliation of in silico/in vivo mutant growth", *BMC Systems Biology*, Vol. 6, No. 1, pp. 1–12, 2012.

42. Tatusov, R. L., M. Y. Galperin, D. A. Natale and E. V. Koonin, "The COG database: a tool for genome-scale analysis of protein functions and evolution", *Nucleic Acids Research*, Vol. 28, No. 1, pp. 33–36, 2000.

43. Orner, E. P., P. Zhang, M. C. Jo, S. Bhattacharya, L. Qin and B. C. Fries, "High-Throughput Yeast Aging Analysis for Cryptococcus (HYAAC) microfluidic device streamlines aging studies in Cryptococcus neoformans", *Communications Biology*, Vol. 2, No. 1, pp. 1–9, 2019.

44. Henson, M. A., G. Orazi, P. Phalak and G. A. O'Toole, "Metabolic modeling of cystic fibrosis airway communities predicts mechanisms of pathogen dominance", *Msystems*, Vol. 4, No. 2, e00026–19, 2019.

45. Li, L. X., A. Ashikov, H. Liu, C. L. Griffith, H. Bakker and T. L. Doering, "Cryptococcus neoformans UGT1 encodes a UDP-Galactose/UDP-GalNAc transporter", *Glycobiology*, Vol. 27, No. 1, pp. 87–98, 2017.

46. Rella, A., A. M. Farnoud and M. Del Poeta, "Plasma membrane lipids and their role in fungal virulence", *Progress in Lipid Research*, Vol. 61, pp. 63–72, 2016.

47. Farnoud, A. M., A. M. Toledo, J. B. Konopka, M. Del Poeta and E. London, "Raft-like membrane domains in pathogenic microorganisms", *Current Topics in Membranes*, Vol. 75, pp. 233–268, 2015.

48. Cox, R., "Macromolecular structure and properties of ribonucleic acids", *Quarterly Reviews, Chemical Society*, Vol. 22, No. 4, pp. 499–526, 1968.

49. de Gontijo, F. A., R. C. Pascon, L. Fernandes, J. Machado Jr, J. A. Alspaugh and M. A. Vallim, "The role of the de novo pyrimidine biosynthetic pathway in Cryptococcus neoformans high temperature growth and virulence", *Fungal Genetics and Biology*, Vol. 70, pp. 12–23, 2014.

50. Singh, A., A. MacKenzie, G. Girnun and M. Del Poeta, "Analysis of sphingolipids, sterols, and phospholipids in human pathogenic Cryptococcus strains", *Journal of Lipid Research*, Vol. 58, No. 10, 2017.

51. Garcia, J., J. Shea, F. Alvarez-Vasquez, A. Qureshi, C. Luberto, E. O. Voit and M. Del Poeta, "Mathematical modeling of pathogenicity of Cryptococcus neoformans", *Molecular Systems Biology*, Vol. 4, No. 1, p. 183, 2008.

52. Lieven, C., M. E. Beber, B. G. Olivier, F. T. Bergmann, M. Ataman, P. Babaei, J. A. Bartell, L. M. Blank, S. Chauhan, K. Correia *et al.*, "MEMOTE for standardized genome-scale metabolic model testing", *Nature Biotechnology*, Vol. 38, No. 3, pp. 272–276, 2020.

53. Heirendt, L., S. Arreckx, T. Pfau, S. N. Mendoza, A. Richelle, A. Heinken, H. S. Haraldsdóttir, J. Wachowiak, S. M. Keating, V. Vlasov *et al.*, "Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0", *Nature Protocols*, Vol. 14, No. 3, pp. 639–702, 2019.

54. Sonnhammer, E. L. and G. Östlund, "InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic", *Nucleic Acids Research*, Vol. 43, No. D1, pp.

D234–D239, 2015.

55. Zur, H., E. Ruppin and T. Shlomi, "iMAT: an integrative metabolic analysis tool", *Bioinformatics*, Vol. 26, No. 24, pp. 3140–3142, 2010.

56. Chen, Y., D. L. Toffaletti, J. L. Tenor, A. P. Litvintseva, C. Fang, T. G. Mitchell, T. R. McDonald, K. Nielsen, D. R. Boulware, T. Bicanic *et al.*, "The Cryptococcus neoformans transcriptome at the site of human meningitis", *MBio*, Vol. 5, No. 1, e01087–13, 2014.

57. Larhlimi, A., L. David, J. Selbig and A. Bockmayr, "F2C2: a fast tool for the computation of flux coupling in genome-scale metabolic networks", *BMC Bioinformatics*, Vol. 13, No. 1, pp. 1–9, 2012.

58. Onishi, J., M. Meinz, J. Thompson, J. Curotto, S. Dreikorn, M. Rosenbach, C. Douglas, G. Abruzzo, A. Flattery, L. Kong *et al.*, "Discovery of novel antifungal (1, 3)-β-D-glucan synthase inhibitors", *Antimicrobial Agents and Chemotherapy*, Vol. 44, No. 2, pp. 368–377, 2000.

59. Hu, G., P.-Y. Cheng, A. Sham, J. R. Perfect and J. W. Kronstad, "Metabolic adaptation in Cryptococcus neoformans during early murine pulmonary infection", *Molecular Microbiology*, Vol. 69, No. 6, pp. 1456–1475, 2008.

60. Himmelreich, U., C. Allen, S. Dowd, R. Malik, B. P. Shehan, C. Mountford and T. C. Sorrell, "Identification of metabolites of importance in the pathogenesis of pulmonary cryptococcoma using nuclear magnetic resonance spectroscopy", *Microbes and Infection*, Vol. 5, No. 4, pp. 285–290, 2003.

61. Hicks, J. K., C. A. D'Souza, G. M. Cox and J. Heitman, "Cyclic AMP-dependent protein kinase catalytic subunits have divergent roles in virulence factor production in two varieties of the fungal pathogen Cryptococcus neoformans", *Eukaryotic Cell*, Vol. 3, No. 1, pp. 14–26, 2004.

62. Huang, W., G. Liao, G. M. Baker, Y. Wang, R. Lau, P. Paderu, D. S. Perlin and C. Xue, "Lipid flippase subunit Cdc50 mediates drug resistance and virulence in Cryptococcus neoformans", *MBio*, Vol. 7, No. 3, e00478–16, 2016.

63. Cox, G. M., J. Mukherjee, G. T. Cole, A. Casadevall and J. R. Perfect, "Urease as a virulence factor in experimental cryptococcosis", *Infection and Immunity*, Vol. 68, No. 2, pp. 443–448, 2000.

64. Doering, T. L., "How sweet it is! Cell wall biogenesis and polysaccharide capsule formation in Cryptococcus neoformans", *Annual Review of Microbiology*, Vol. 63, pp. 223–247, 2009.

65. Zaragoza, O., M. L. Rodrigues, M. De Jesus, S. Frases, E. Dadachova and A. Casadevall, "The capsule of the fungal pathogen Cryptococcus neoformans", *Advances in Applied Microbiology*, Vol. 68, pp. 133–216, 2009.

66. Chrisman, C. J., P. Albuquerque, A. J. Guimaraes, E. Nieves and A. Casadevall, "Phospholipids trigger Cryptococcus neoformans capsular enlargement during interactions with amoebae and macrophages", *PLoS Pathog*, Vol. 7, No. 5, e1002047, 2011.

67. Flowers, S. A., K. S. Barker, E. L. Berkow, G. Toner, S. G. Chadwick, S. E. Gygax, J. Morschhäuser and P. D. Rogers, "Gain-of-function mutations in UPC2 are a frequent cause of ERG11 upregulation in azole-resistant clinical isolates of Candida albicans", *Eukaryotic Cell*, Vol. 11, No. 10, pp. 1289–1299, 2012.

68. Allen, D., D. Wilson, R. Drew and J. Perfect, "Azole antifungals: 35 years of invasive fungal infection management", *Expert Review of Anti-Infective Therapy*, Vol. 13, No. 6, pp. 787–798, 2015.

69. Kathiravan, M. K., A. B. Salake, A. S. Chothe, P. B. Dudhe, R. P. Watode, M. S. Mukta and S. Gadhwe, "The biology and chemistry of antifungal agents: a review",

Bioorganic & Medicinal Chemistry, Vol. 20, No. 19, pp. 5678–5698, 2012.

70. Ziebart, K. T., S. M. Dixon, B. Avila, M. H. El-Badri, K. G. Guggenheim, M. J. Kurth and M. D. Toney, "Targeting multiple chorismate-utilizing enzymes with a single inhibitor: validation of a three-stage design", *Journal of Medicinal Chemistry*, Vol. 53, No. 9, pp. 3718–3729, 2010.

71. Ducati, R., L. A. Basso and D. S. Santos, "Mycobacterial shikimate pathway enzymes as targets for drug design", *Current Drug Targets*, Vol. 8, No. 3, pp. 423–435, 2007.

72. Dahal, G. P., *Development of Selective Inhibitors against Enzymes Involved in the Aspartate Biosynthetic Pathway for Antifungal Drug Development*, Ph.D. Thesis, University of Toledo, 2018.

73. Dahal, G. P., D. Launder, K. M. McKeone, J. P. Hunter, H. R. Conti and R. E. Viola, "Aspartate semialdehyde dehydrogenase inhibition suppresses the growth of the pathogenic fungus Candida albicans", *Drug Development Research*, Vol. 81, No. 6, pp. 736–744, 2020.

74. Wishart, D. S., C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang and J. Woolsey, "DrugBank: a comprehensive resource for in silico drug discovery and exploration", *Nucleic Acids Research*, Vol. 34, No. suppl_1, pp. D668–D672, 2006.

APPENDIX A: REACTION LIST

The content is given in CD.

APPENDIX B: GENE LOCALISATION LIST

The content is given in CD.

APPENDIX C: SINGLE-GENE DELETION LIST

The content is given in CD.