T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

CONSTRAINT-BASED ANALYSIS OF THE GENOME-SCALE METABOLIC NETWORKS FOR KLEBSIELLA PNEUMONIAE TO IDENTIFY NEW PUTATIVE DRUG TARGETS

MÜBERRA FATMA CESUR

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF BIOENGINEERING

GEBZE

2018

T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

CONSTRAINT-BASED ANALYSIS OF THE GENOME-SCALE METABOLIC NETWORKS FOR KLEBSIELLA PNEUMONIAE TO IDENTIFY NEW PUTATIVE DRUG TARGETS

MÜBERRA FATMA CESUR

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF BIOENGINEERING

THESIS SUPERVISOR

ASSIST. PROF. DR. SALİHA DURMUŞ

II. THESIS SUPERVISOR

ASSOC. PROF. DR. TUNAHAN ÇAKIR

GEBZE

2018

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YENİ İLAÇ HEDEFLERİ BELİRLEMEK İÇİN KLEBSIELLA PNEUMONIAE GENOM ÖLÇEKLİ METABOLİK AĞLARININ KISITA-DAYALI ANALİZİ

MÜBERRA FATMA CESUR
YÜKSEK LİSANS TEZİ
BİYOMÜHENDİSLİK ANABİLİM DALI

DANIŞMANI
DR. ÖĞR. ÜYESİ SALİHA DURMUŞ
II. DANIŞMANI
DOÇ. DR. TUNAHAN ÇAKIR

GEBZE
2018

GEBZE TEKNİK ÜNİVERSİTESİ	YÜKSEK LİSANS JÜRİ ONAY FORMU

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 27/06/2018 tarih ve 2018/33 sayılı kararıyla oluşturulan jüri tarafından 05/07/2018 tarihinde tez savunma sınavı yapılan Müberra Fatma CESUR'un tez çalışması Biyomühendislik Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE
(TEZ DANIŞMANI) : Dr. Öğr. Üyesi Saliha DURMUŞ

ÜYE
(2. TEZ DANIŞMANI) : Doç. Dr. Tunahan ÇAKIR

ÜYE : Dr. Öğr. Üyesi Pınar PİR

ÜYE : Dr. Öğr. Üyesi Reaz UDDİN

ÜYE : Dr. Öğr. Üyesi Ali ÇAKMAK

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun

...../...../..... tarih ve ...../..... sayılı kararı.

SUMMARY

Klebsiella pneumoniae is an etiological agent of serious life-threatening nosocomial infections. Conventional treatment approaches are not sufficient to control the *Klebsiella*-mediated infections. Therefore, different approaches must be employed to handle resistant species of this pathogen. Network-based analysis methods provide a comprehensive view to decipher cellular metabolism. Genome-scale metabolic network models (GEMs) are promising platforms that allow analysis of whole metabolic network of a cell. Notably, they are useful to identify novel metabolic drug targets. To date, two metabolic models of *K. pneumoniae* (iYL1228 for MGH 78578 strain and iKp1289 for highly strain KPPR1 strain) have been reconstructed. In the scope of this work, computational systems biology approach based on constraint-based genome-scale metabolic network analysis was used to comparatively analyze the metabolisms of two *Klebsiella* strains and to discover new enzyme-based drug targets. Over 30 essential gene without human homologs were identified through growth simulations of each strain in different host-mimicking conditions. A total of 31 non-homologous genes are found to be druggable. Five of them associated with virulence and show a broad distribution among some popular pathogen species were suggested as drug targets in the study. This putative target list was extended using an updated biomass reaction. Furthermore, three non-homologous genes were also predicted as drug target via a metabolite-centric approach. To our knowledge, this is the first comprehensive effort to elucidate putative drug targets of *K. pneumoniae* strains through the analysis of their GEMs. These findings provide crucial insight for further research.

Keywords: *Klebsiella pneumoniae*, Systems Biology, Drug Targeting, Genome-scale Metabolic Network Model, Constraint-based Analysis, Infection.

ÖZET

Klebsiella pneumoniae yaşamı tehdit eden, ciddi hastane enfeksiyonlarına sebep olmaktadır. Klasik tedavi yaklaşımları *Klebsiella* aracılı enfeksiyonların kontrolünde yetersiz kalmaktadır. Bu yüzden bu patojenle başa çıkabilmek için daha farklı yaklaşımlara ihtiyaç vardır. Ağ-tabanlı analiz yöntemleri, hücresel metabolizmanın çözülebilmesi için kapsamlı çıktılar sunar. Genom-ölçekli metabolik ağ modelleri bir hücreye ait tüm metabolik ağının analizine izin veren önemli platformlardır. Açıkça bu yaklaşım yeni ilaç hedeflerinin belirlemesi için ümit vericidir. Günümüze kadar, *K. pneumoniae* için iki tane genom-ölçekli metabolik ağ modeli (MGH 78578 suşu için iYL1228 ve KPPR1 suşu için iKp1289) oluşturulmuştur. Bu çalışma kapsamında, iki farklı *K. pneumoniae* türünün metabolizmalarının karşılaştırılması ve enzim türevi ilaç hedeflerinin belirlenmesi amacıyla kısıt-tabanlı genom-ölçekli metabolik ağ analizine dayanan hesaplamalı sistem biyolojisi yaklaşımı kullanılmıştır. Böylece, konak ortamını taklit eden iki farklı besiyerinde büyütülen her bir suş için insanda homoloğu bulunmayan 30'un üzerinde hayati gen tespit edilmiştir. Toplamda homolog olmayan 31 gen ilaç molekülüne bağlanabilme özelliğine sahiptir. Bu genlerin virulanslıkla ilişkili olan ve bazı popüler patojen türleri arasında geniş bir yayılım beş tanesi çalışmada ilaç hedefi olarak önerilmiştir. Güncel biyokütle oluşum denklemi kullanılarak bu olası hedef listesi genişletilmiştir. Ayrıca metabolit-odaklı bir yaklaşım kullanılarak belirlenen ve insanda homoloğu olmayan üç gen de olası ilaç hedefi olarak önerilmiştir. Bilgimiz dahilinde, *K. pneumoniae* için oluşturulan genom-ölçekli metabolik ağ modelleri ilk kez bu çalışmada olası ilaç hedeflerinin belirlenmesi amacıyla kullanılmıştır. Bu tez çalışmanın sonuçları, gelecek çalışmalara yönelik önemli bulgular sunmaktadır.

Anahtar Kelimeler: *Klebsiella pneumoniae,* Sistem Biyolojisi, İlaç hedefleme, Genom-ölçekli Metabolik Model, Kısıt-tabanlı Analiz, Enfeksiyon.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, Assist. Prof. Dr. Saliha Durmuş and Assoc. Prof. Dr. Tunahan Çakır for their never-ending patience, excellent guidance and encouragement. I am so glad to have the opportunity to work with them throughout my Master study. I would like to thank the other members of my thesis committee, Assist. Prof. Dr. Pınar Pir, Assist. Prof. Dr. Reaz Uddin and Assist. Prof. Dr. Ali Çakmak for their comments and suggestions on my thesis.

I gratefully acknowledge TUBITAK for financial support (Project Code: 316S005).

I am thankful to the members of Computational Systems Biology Group at Gebze Institute of Technology for providing me a friendly working atmosphere and their very nice collaborations.

I wish to express my profound thanks to my brother Muhammed Raşit Cesur for his continuous support and very valuable comments on my thesis. I am deeply indebted to my mother Zehra Cesur and my father Ali Cesur for their everlasting patience and encouragement. None of this could have been possible without them.

TABLE of CONTENTS

LIST of ABBREVIATIONS and ACRONYMS

Abbreviations and Acronyms		Explanations
β	:	Beta
h	:	Hour
S	:	Stoichiometric Matrix
V	:	Metabolic Reaction Rate Vector
CDG	:	Carbon-D-Glucose
COBRA	:	The Constraint-Based Reconstruction and Analysis
CPS	:	Capsular Polysaccharide
CRE	:	Carbapenem-Resistant *Enterobacteriaceae*
ESBL	:	Extended-Spectrum β-Lactamases
FBA	:	Flux Balance Analysis
FDA	:	Food and Drug Administration
flexoFBA	:	Flexible-Optimality FBA
FVA	:	Flux Variability Analysis
HBF	:	Human Body Fluid
GEM	:	Genome-Scale Metabolic Model
HMDB	:	The Human Metabolome Database
lb	:	Lower Bound
LB	:	Luria-Bertani Broth
LPS	:	Lipopolysaccharide
MBRole	:	Metabolites Biological Role
MDR	:	Multidrug Resistant
NGAM	:	Non-Growth-Associated Maintenance
OMP	:	Outer Membrane Protein
PBIT	:	Pipeline Builder for Identification of Drug Targets
RGM	:	Reaction-Gene Matrix
SM	:	Sputum-Macrophage
TLR	:	Toll-like Receptor
ub	:	Upper Bound
VF	:	Virulence Factor

VFDB	:	Virulence Factor Database
WT	:	Wild-type
XDR	:	Extensive Drug Resistant

LIST of FIGURES

LIST of TABLES

1. INTRODUCTION

Klebsiella pneumoniae is a member of the *Enterobacteriaceae* family and it inhabits a wide range of habitats such as humans, plants, surfaces of various medical devices, water and soil [1, 2].

Although this pathogen usually infects immunocompromised patients, elderly individuals and neonates, hypervirulent strains harboring drug resistance genes can lead to serious diseases (e.g., pneumonia, endophthalmitis, pyogenic liver/neck/lung abscess, and meningitis) in even healthy individuals [3, 4]. Thus, the current notoriety of *K. pneumoniae* is not surprising. It is one of the 'ESKAPE' pathogens (a group of nosocomial pathogens including *Enterococcus faecium*, *Staphylococcus aureus*, *K. pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species) that are capable of escaping the activity of antimicrobial drugs [5, 6]. Even if various therapeutic strategies (e.g., bacteriophage therapy and use of synergistic drug combinations) were proposed to cope with *K. pneumoniae*, these approaches are insufficient to eradicate this pathogen [1, 5–7]. Therefore, more comprehensive strategies are necessary to bring *Klebsiella* infections under control. In this context, it is crucial to examine the behavior of a biological system as a whole in order to detect more efficient drug targets.

Time-consuming, expensive and labor-intensive nature of the high-throughput technologies limit the discovery of new drugs despite the recent advances. Moreover, experimentally generated datasets must be properly interpreted in order to decode the behavior of the systems. Therefore, systems biology has emerged as a promising framework to comprehensively analyze an organism instead of focusing on its isolated parts. This interdisciplinary field generates new hypotheses for wet-lab experiments through *in silico* experiments by considerably narrowing down the solution space. Genome-scale metabolic models (GEMs) are *in silico* representations of the metabolism. To date, GEMs of many biologically important organisms were constructed [8]. The constraint-based flux analysis of the GEMs of pathogens is an extremely promising approach to reveal putative drug targets through rapid systematic perturbations [9]. Thus, the use of computational approaches allows identification of the candidate genes to be targeted for growth inhibition of the pathogens.

Here, two genome scale metabolic network models of different *Klebsiella* strains (i.e., iYL1228 and iKp1289) were analyzed through gene-centric and metabolite-centric approaches to reveal potential metabolic drug targets. To reduce potential side effects and to select more efficient targets, the target candidates were prioritized by applying several criteria based on homology, virulence, druggability, connectivity and broad-spectrum analyses. Thus, over 30 essential genes with no human homologs were determined for each strain in different growth simulations. Particularly five druggable virulence genes (*hldD* (*rfaD*), *kdsA*, *gmhA (lpcA*), *lpxA* and *lpxC*) were predicted to be broad-spectrum targets. Most of these genes were reported as potential drug targets for different pathogens in previous studies [10–14]. This list was extended by using an updated biomass reaction through the gene-centric approach. Thus, 11 additional putative targets were suggested. They are associated with nicotinamide adenine dinucleotide synthesis, pantothenate production, coenzyme A biosynthesis, and the riboflavin synthesis. In addition, three genes (*mrcB, pbpC* and *mrcA*) without human homologs were identified by the metabolite-centric approach. Importantly, the remaining non-homologous, essential genes should be also assessed in detail to extend the putative target list. Therefore, a further work is required to evaluate and validate all possible candidate drug targets.

1.1. Aim and Objectives of the Thesis

Klebsiella pneumoniae is among the etiologic agents of various serious diseases with significant morbidity and mortality worldwide. For instance, hypervirulent strains of this pathogen can lead to a wide variety of diseases such as pneumonia, meningitis, endophthalmitis, bloodstream and urinary tract infections [3, 4].

Despite intensive efforts, *K. pneumoniae* continues to pose a serious public health threat and it was reported among the urgent threat level multidrug resistant pathogens by the World Health Organization, the UK Department of Health and the US Centers for Disease Control and Prevention [15]. Herein, the genome scale metabolic network models of two different *K. pneumoniae* strains were used for a system-level analysis of their metabolisms. The superiority of this study is utilization of multi-level information in the filtering process to reveal the most efficient putative targets.

Briefly, the objective of this study is to reveal the convenient drug targets with a minimal side effect on the host and to gain a better insight into the *K. pneumoniae* metabolism. To date, various proteome- and transcriptome-based studies on the *K. pneumoniae* were reported [83–86]. However, to our knowledge, it is the first effort to identify putative drug targets for *K. pneumoniae* through a comparatively and comprehensive analysis of its GEMs by gene- and metabolite-centric approaches.

2. BACKGROUND ASPECTS

2.1. *Klebsiella pneumoniae*

Klebsiella pneumoniae is a gram-negative, rod-shaped, lactose-fermenting, usually encapsulated and nonmotile pathogen that was first observed in the postmortem lung samples of pneumonia patients in 1882 [16–18].

K. pneumoniae can be traditionally classified according to their capsular and lipopolysaccharide (LPS) serotypes. The strains of this pathogen with the K1 and K2 capsular types within 79 capsular (K antigen) serotypes represent a hypermucoviscous (also known as hypervirulent) phenotype likely owing to the increased biosynthesis of capsular polysaccharide (CPS) [3, 19]. Furthermore, it was suggested that K1 and K2 strains reduce the level of reactive oxygen species released by human neutrophils and the capsule structure including sialic acid may also allow evasion from the immune system by mimicking the host cells. Alternatively, the lack of mannose residue repeats of these strains protects the pathogen against an efficient lectinophagocytosis (an opsonin-independent form of phagocytosis) and against a possible attack by the neutrophiles via blocking the formation of proinflammatory signals that recruit the immune cells [4]. Thus, hypervirulent *Klebsiella* strains lead to serious community-acquired infections. In the similar vein, 9 different O-antigen types were reported for *K. pneumoniae* and O1 serotype is commonly found in clinical isolates [3, 4, 19].

The *Klebsiella* strains can inhabit a wide range of conditions such as water, soil, sewage, medical devices like catheters in addition to humans, plants and other organisms. These pathogens can also be found in the normal flora of intestine, skin, mouth, and nose in humans [1, 2]. They can readily spread to other tissues by colonizing human mucosal surfaces (e.g., gastrointestinal tract, respiratory tract and oropharynx [4, 19]. The colonization rate varies by the type of *Klebsiella* (i.e., community- or hospital-acquired *K. pneumoniae*), tissue and antibiotic treatment. Particularly antibiotic treatment considerably increases the colonization rate [17]. *K. pneumoniae* can also form biofilm on various biotic and abiotic surfaces like wounds and catheters. Thus, they become more resistant to antibiotics and immune response and eventually can lead to chronic infections [19, 20].

2.1.1. Antibiotic Resistance

Drug-resistant infections cause high mortality (over 700,000 deaths/year) and the reduction in the antimicrobial activity of the drugs poses a major therapeutic challenge by affecting global economic cost and public health. A global lose at least US$100 trillion is expected throughout the next few decades if the drug-resistance issue is not addressed [21]. This bad scenario prompts the scientists to investigate the antibiotic resistance mechanisms and novel therapeutic approaches. The recent increase in the publications on antibiotic resistance is presented in Figure 2.1 [22].

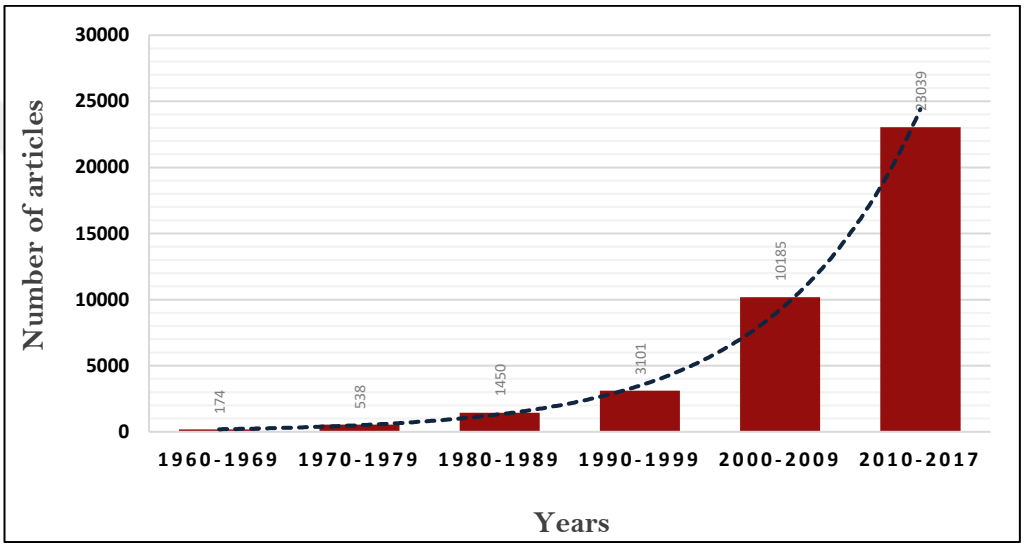

Figure 2.1: Number of scientific publications including “antibiotic resistance” and “antimicrobial resistance” terms in their abstracts from PubMed.

The resistance genes can trigger dissemination of the antibiotic resistance through two major routes including vertical gene transfer (transfer of genetic material with *de novo* mutations from the parent to its daughter cells) or horizontal gene transfer (gene transfer via mobile genetic elements like plasmids through phage transduction, conjugation or transformation) (Figure 2.2) [16, 23]. Resistance mutations support survival of the mutant populations while wild-type populations are eliminated. Thus, the resistant populations predominate [24]. For instance, fluoroquinolones are antibiotics that disrupt the DNA replication by targeting DNA gyrase and topoisomerase IV of the bacterial cell [25]. The resistance against this antibiotic is mainly facilitated through some resistance-conferring mutations on the targets of this antibiotic (i.e., *gyrA*, *gyrB*, *parC*, and *parE*). The affinity of this antibiotic for the

targets depends on the bacterial type and resistant allele formation. If one of these targets is converted into the resistant form, the antibiotic binds to another target. It is important to note that the ratio of fluoroquinolone-resistant *K. pneumoniae* isolates is increasing. For instance, the proportion of these isolates has increased fivefold in Italy between 2005 and 2012 [26].

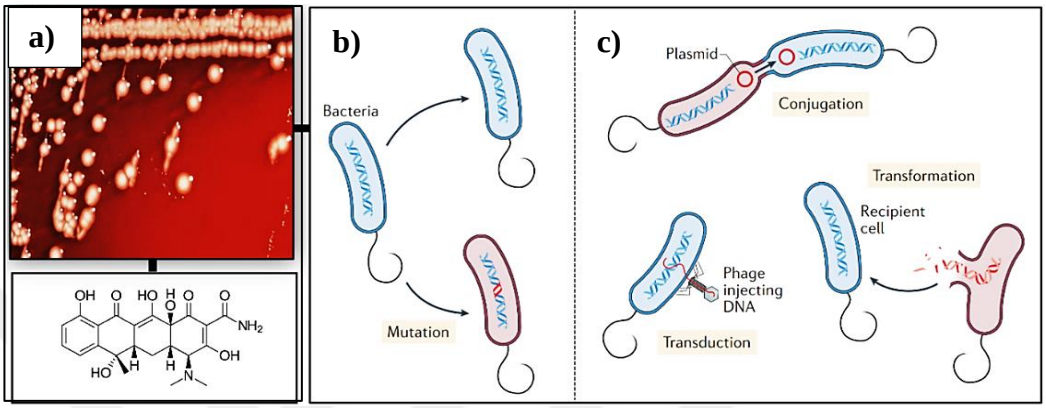

Figure 2.2: Acquisition of antibiotic resistance. a) Antibiotic treatment, b) Verticle gene transfer, c) Horizontal gene transfer.

Especially human gastrointestinal tract provides an ideal condition to trigger transfer of the mobile genetic elements (e.g., resistance-conferring plasmids) due to antibiotic pressure and high bacterial density [27]. Transfer of the resistance genes can be mediated via three main mechanisms including conjugation, natural transformation and transduction (Figure 2.2) [23]. The most common mechanism for horizontal gene transfer is the conjugation. This mechanism provides transmission of conjugative plasmids (self-transmissible plasmids) and integrative conjugative elements (e.g., conjugative transposons) through a conjugative pilus between the donor and recipient bacterial cells [23, 27]. This process can promote conversion of the antibiotic sensitive bacteria to antibiotic resistant microorganisms [28]. As well as the gene transmission between the living cells, the DNAs released from especially lysing bacteria can also be transferred to over 80 different species of naturally transformable bacteria [27, 28]. Moreover, Woo and colleagues (2003) suggested that the horizontal gene exchange may occur through transformation of cell-wall deficient bacteria as a result of administration of antibiotics. This hypothesis is based on the observation that some antibiotics (e.g., beta-lactams and glycopeptides) suppress bacterial cell wall

assembly. Thus, the cell-wall deficient bacteria can acquire the extracellular resistance encoding DNA more easily [29].

We have less information regarding the emergence of antibiotic resistance through transduction when compared to the natural transformation and conjugation. Transduction refers to transfer of the DNA via bacteria infecting viruses (bacteriophages) [27, 28]. After penetration, these viruses can integrate their DNAs into the bacterial genomes to replicate their genetic materials in conjunction with the bacterial genomes. Thus, they may confer resistance to the host due to presence of acquired resistance genes in the virome [23, 27]. Colomer-Lluch and colleagues demonstrated the contribution of the resistance gene bearing phages for the emergence of resistant *E. coli* strains [30]. Furthermore, transfer of the resistance-conferring genes to the other bacteria including *Salmonella* and *Enterococcus* was also shown through *in vitro* experiments [27]. Collectively, continuous exposure to antibiotics leads to resistance development and emergence of multidrug resistant (MDR) and extensively drug resistant (XDR) pathogens by means of vertical or horizontal gene transfer mechanisms. The resistant bacteria have different mechanisms to cope with antimicrobial compounds. Intra and inter-species transmission of these resistance mechanisms is a serious public health threat counteracting the action of available antibiotics.

Antibiotics damage bacterial cells by a) mostly targeting the ribosome in order to block protein synthesis (e.g., aminoglycosides), b) inhibiting metabolic processes (e.g., sulfonamides), c) targeting DNA maintenance (e.g., quinolones) and d) preventing cell-wall synthesis and disrupting structural integrity (e.g., β-lactams) (Figure 2.3) [21, 31]. The resistant bacteria can cope with the attacks of the antibacterial compounds via two major mechanisms: (1) intrinsic resistance and (2) acquired resistance.

Intrinsic resistance is related to the natural, structural and functional characteristics of bacteria regardless of administration of an antibiotic [24]. For instance, outer membrane of the bacteria act as a barrier against large antibiotics such as vancomycin. *Pseudomonas* does not include any sensitive targets for the biocide triclosan (i.e., bearing of only Triclosan-insensitive allele of *fabI*). Modification of the membrane regulated by GraRS (also known as aps) sensor/regulator system reduces the positive charge in *Staphylococcus aureus* in order to confer resistance. The aerobic bacteria are naturally resistant to metronidazole because the reduction of this antibiotic

to the active form takes place under anaerobic conditions. Removal the antimicrobial compounds by the Sap (sensitivity to antimicrobial peptides) efflux system and by the putative ABC transporters of *S. typhimurium* contributes to the intrinsic resistance of the bacteria [28, 31]. In addition, the inherent efflux system is also crucial for an efficient resistance to fluoroquinolone. Much as the acquired resistance developed via mutations in DNA gyrase and topoisomerase IV confers the bacteria resistance against fluoroquinolone, it was shown this resistance mechanism is not sufficient per se. This is because the inhibition of the inherent efflux system results in a fluoroquinolone-sensitive population. Therefore, both intrinsic and acquired resistance mecahnisms may be significant to cope with fluoroquinolone antibiotic [26].

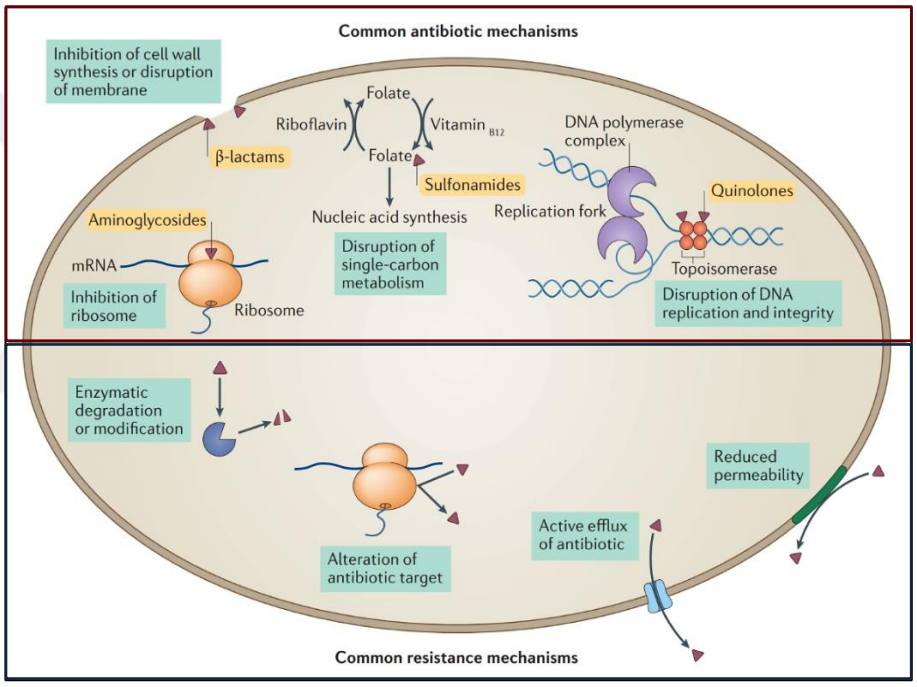

Figure 2.3: Antibiotic resistance mechanisms.

Acquired resistance can occur through various vertical/horizontal gene transfers [28, 31]. As a result of the gene exchange, some strategies (e.g., export of the intracellular antibiotics, modification of the antibiotic targets/antibiotics and degradation of the antibiotics) are employed by these microorganisms to cope with the selective antibiotic pressure (Figure 2.3) [21, 24, 31]. The acquired resistance can emerge only in a subpopulation of the species as the intrinsic resistance emerges in a wide range of bacterial groups [28]. For instance, streptothricin serves as a protein synthesis inhibitor, and the first plasmid-borne resistance to this antibiotic was

detected in *E. coli* isolates after the introduction of streptothricin F [32]. In a similar vein, colistin-resistant *E. coli* isolates harbour a horizontally transferable plasmid bearing the *mcr-1* gene that mediates lipid A modification [33]. Interactions between antibiotics and their targets can be also prevented through mutations of the targets. In this process, some changes in the target structure decrease binding affinity of the antibiotics while the target can maintain its function [31]. An example of this phenomenon was reported for polymyxin-resistant *K. pneumoniae*. Some mutations in *mgrB* gene play a prominent role in the polymyxin resistance by decreasing the interaction of the antibiotic with the bacterial membrane [34–36].

Another strategy adapted by bacteria is the minimization of the antibiotic uptake. High-level carbapenem resistance of *Enterobacter cloacae* is associated with porin alteration. Mutations in the porin genes render this bacteria resistant although they do not have any known carbapenemases [21, 31]. Minimization of the antibiotic concentration can be also supported via the inactivation of the antibiotics by hydrolysis. Resistance of *K. pneumoniae* to the commonly used antibiotics such as β-lactam antibiotics (e.g., penicillins, cephalosporins, carbapenems and monobactams), aminoglycosides and fluoroquinolones have emerged at a rapid pace through transfer of various resistance-conferring plasmids [6].

Commonly shared plasmids across different *Enterobacteriaceae* members including *K. pneumoniae*, *K. oxytoca*, *Escherichia coli*, *Salmonella sp.* and *Enterobacter sp.*, support the dissemination of MDR isolates. Thus, increasing prevelance of the resistant strains is not surprising. Most resistance mechanisms of gram-negative bacteria are also found in *K. pneumoniae* clinical isolates [37]. Some resistance-related *Klebsiell*a genes responsible for the different resistance mechanisms against two popular antibiotic classes (β-lactam and aminoglycoside antibiotics) are illustrated in Figure 2.4 [16].

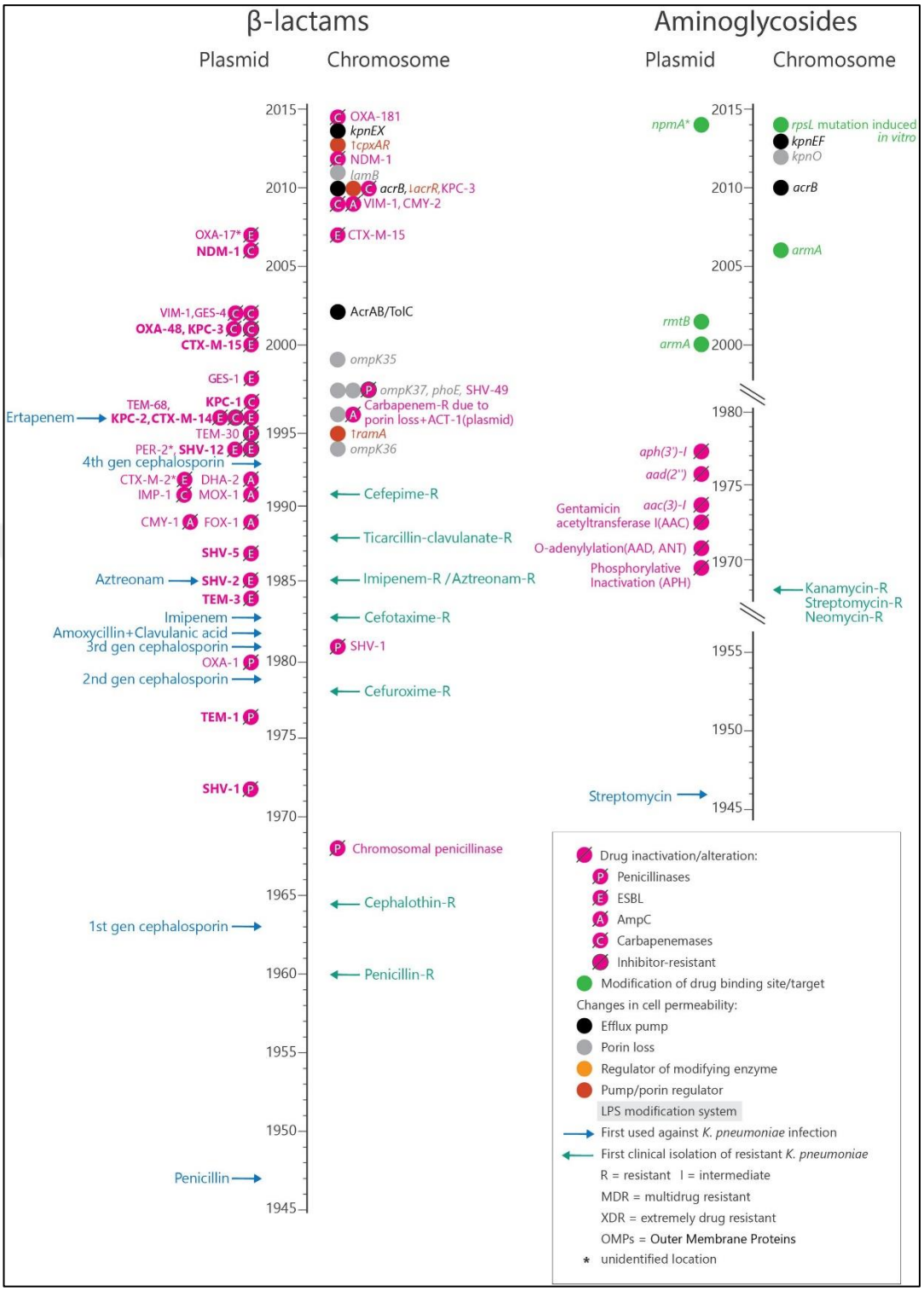

Figure 2.4: Antibiotic resistance genes against β-lactam and aminoglycoside antibiotics.

Particularly integration of some carbapenemase-encoding genes (e.g., *bla*$_{KPC}$) onto the chromosome makes eradication of the resistant *K. pneumoniae* strains

difficult. Therefore, more efficient treatment approaches are necessary to handle these pathogens [16, 38].

2.1.2. Virulence Characteristics of *Klebsiella*

Resistance is directly correlated with the virulence allowing the pathogen to survive in a hostile environment. There are various regulatory factors mediating regulation of both virulence and antimicrobial resistance. A global transcriptional regulator in the *Klebsiella*, RamA, is one of them and it is responsible for the regulation of LPS synthesis and drug resistance [39, 40]. CPS, LPS, fimbriae, outer membrane proteins (OMPs) and siderophores are among the significant virulence factors of *K. pneumoniae* demonstrated in Figure 2.5 [4, 41].

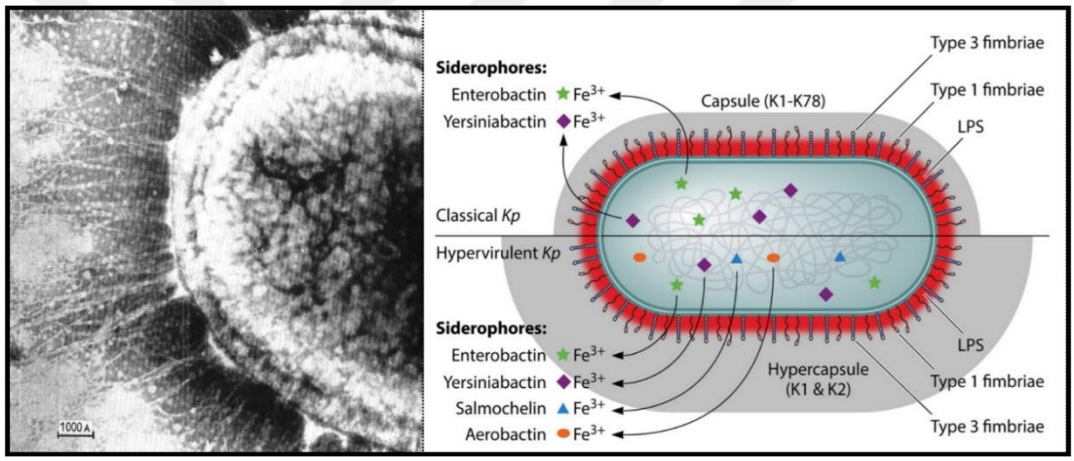

Figure 2.5: Main virulence factors of classical and hypervirulent *K. pneumoniae* strains.

One of the most highly investigated virulence factor of *K. pneumoniae* is the thick polysaccharide capsule which surrounds the microorganism. The capsule of this pathogen includes complex polysaccharides consisting of repeating subunits of sugars and its biosynthesis is regulated by the capsule polysaccharide locus (*cps*) [19, 42]. These gene clusters are essential for synthesis, assembly and transport of the capsule compounds [19]. Moreover, synthesis of the capsule can be enhanced by (1) transcriptional regulators (i.e., regulator of mucoid phenotype A (*rpmA* and *rpmA2*), (2) regulation of the capsule synthesis A and B genes (*rcsA* and *rcsB*) and (3) external clues (e.g., an increase in the glucose concentration) in the hypervirulent *K. pneumoniae* [4]. Expression of the *cps* gene can be stimulated by the antimicrobial

peptides in sublethal concentrations. Bacterial CPS prevents the access of the antimicrobial peptides (e.g., β-defensins) to the pathogen [19, 43]. Thus, capsule structure is a significant indicator determining the virulence of the pathogen and conferring resistance to the antibacterial agents. Moreover, it has a potential function in mediating the bacterial colonization. As well as the above-mentioned functions, CPS is especially essential for evading or suppressing the host immune response. An example of this phenomenon is the phagocytosis-inhibitory role of the capsule. Thus, the pathogen can be protected from the attacks of human neutrophils and macrophages with the help of the capsule [43]. Although it is vital to protect the bacteria against the host immune response, CPS is not sufficient per se. Outer membrane protein A (OmpA) is one of the major components of outer membrane that has a potential function in immune evasion and attenuation of airway epithelial cell-mediated inflammatory responses [19, 43]. This protein was suggested to suppress immune response and to enhance the bacterial resistance against antimicrobial proteins despite some controversial results. It was documented that OmpA-deficient bacteria induced the production of some cytokines (e.g., IL-8 and IL-6) differently from wild-type cells both *in vitro* and *in vivo* (in mouse lungs) [4]. Akin to OmpA, OmpK35 and/or OmpK36 proteins of *K. pneumoniae* are also significant to confer resistance against neutrophil phagocytosis. The ΔompK35/36 mutation leads to a reduced resistance and bacterial fitness [43]. LPS (also known as endotoxin) is one of the major components of the outer membrane. It is another group of virulence factors of *K. pneumoniae* that consists of three main parts (i.e., O-antigen, core oligosaccharide and lipid A) (Figure 2.6) [44]. LPS is a strong inducer of Toll-like receptors (TLRs) such as TLR4, much as it is essential to protect the pathogen from the complement-mediated killing, macrophage/neutrophil phagocytosis and activity of the antimicrobial peptides [4, 19, 43]. Therefore, it was suggested that the pathogens can mask LPS by the capsule components (e.g., K1, K10 and K16 antigens) to prevent the detection by TLRs [4]. Moreover, lipid A modification of the LPS also decreases the activation of inflammatory response (e.g., induction of the nuclear factor κB (NF-κB) signaling) via prevention of the recognition of the lipid A patterns by the immune system [43]. LpxO-dependent lipid A modification was reported to be essential for both survival and colistin resistance of the *Klebsiella* in the lung [39].

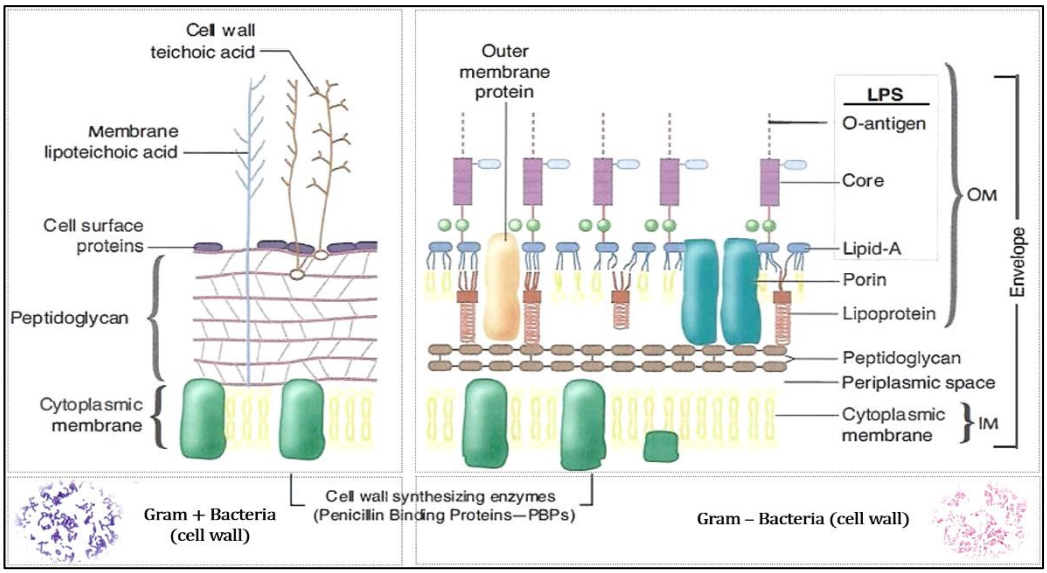

Figure 2.6: Bacterial cell wall structures (OM and IM stand for outer and inner membranes).

Apart from the immune evasion, colonization and biofilm formation is also crucial to persist in the host for long a time. In this context, *K. pneumoniae* includes multiple fimbrial gene clusters responsible for fimbria biosynthesis [42]. Type I fimbriae (a component of the *K. pneumonaiae* cell membrane) contributes to the biofilm synthesis by supporting adhesions to the biotic/abiotic surfaces. Moreover, it was suggested that the fimbria-mediated adherence may contribute to the bacterial colonization and persistence *in vivo*. Thus, this membrane component which is found in the most of the clinical and environmental isolates of *K. pneumoniae* is significant in the virulence of the pathogen. [43]. However, it exhibits tissue-specific expression that is controlled by various external signals. Therefore, it does not contribute the virulence of the bacteria in some tissues. For example, type I fimbriae is not expressed in the gastrointestinal tract and lung in contrast to the urinary tract [4]. Type 3 fimbria encoded by some gene clusters (e.g., *mrkA* gene regarding production of the fimbrial subunit and *mrkD* gene associated with the synthesis of an adhesin polypeptide) is another membrane component which mediates the bacterial interactions with biotic (host cells/tissues) and abiotic surfaces [43]. In summary, the fimbrial structures are essential for adhesion to distinct surfaces and eventually colonization/biofilm formation. Thus, they facilitate the bacterial resistance in a hostile environment.

The battle between the immune system and pathogen is not limited to the suppression of the immune response or immune evasion. Pathogens can also fight to

utilize some host factors. For instance, iron is required for the bacterial growth. However, it is also used by the host immune system for an efficient immune response. Therefore, the host organism sequesters the iron using various iron transporters (e.g., transferrin) to prevent the access of the pathogen [4, 19]. To address this challenge, pathogens secrete small iron-scavenging molecules (e.g., siderophores) with different iron affinities. *K. pneumoniae* can produces the siderophores including enterobactin, salmochelin, yersiniabactin, and aerobactin [4]. Yersiniabactin is one of the virulence-associated siderophores in *Klebsiella* that promotes the lethality in pneumonia mice. It also induces the invasive respiratory tract infections in human as well as supporting the respiratory tract colonization. Another siderophore, enterobactin, has also an crucial role in the growth under iron-limited conditions [43]. Despite these advantages, activity of these small iron-scavenging molecules cannot be sufficient to cope with the iron chelators per se. To escape the blocking activity of the host iron chelators, hypervirulent *K. pneumoniae* strains produce more and more active siderophore combinations than the classical *K. pneumoniae* [4].

Virulence of the *K. pneumoniae* can be triggered by different factors and the hypervirulent strains bearing multiple drug resistance genes can lead to serious diseases (e.g., pyogenic liver, neck or lung abscess, pneumonia, meningitis, endophthalmitis, cellulitis, and so on) in even healthy individuals. On the other hand, the classical *Klebsiella* strains that are etiological agents of some severe infections including pneumonia, urinary tract infections, or bacteremia (primary or secondary bacteremia) in only immunocompromised individuals (e.g., diabetics and patients with malignancies, neonates and the elderly) [3, 4]. Though a high prevalence of mainly *Staphylococcus aureus* was documented in the hospitalized pneumonia patients (pneumonia is the second most common nosocomial infection), some gram-negative bacteria such as *K. pneumoniae* have also drawn a great interest to cope with pneumonia infections in the recent years [45]. Classical *K. pneumoniae* infections can be treated by application of β-lactams and other antibiotics. However, particularly hypervirulent *K. pneumoniae* strains is a significant threat for the community health. Considering the ever-increasing antibiotic resistance and the prevalence of the pathogen (it is first common gram-negative bacteria causing bloodstream infections behind *E. coli*), a deeper understanding of the molecular mechanisms of its virulence and the development of antibiotic resistance is a tremendous need to tackle this pathogen [17, 46]. Recent studies have revealed that the additional virulence factors

are mostly related to the metabolism and the regulation of transcription [4]. Thus, understanding of the pathogen metabolism may hold a great promise to decipher the relationship between virulence and the cellular metabolism and to improve the current treatment strategies via identification of new drug targets and molecular biomarkers.

2.1.3. Therapeutic Approaches for *K. pneumoniae* Infections

The clinical *K. pneumoniae* strains are more resistant to antibiotics than the environmental strains though both strains exhibit a similar virulence and pathogenicity characteristics [17]. The reason of this phenomenon is the accumulation of the antibiotic resistance genes in the hospital-adapted *K. pneumoniae* isolates due to the selective pressure. The MDR pathogen can tolerate different antibiotics mostly through the production of (1) extended-spectrum β-lactamases (ESBLs) and (2) carbapenemases encoded by plasmid-borne bla_{KPC} gene providing resistance to almost all available antibiotics [47, 48]. Ratio of the resistant *K. pneumoniae* isolates based on some countries are presented in Figure 2.7 [49].

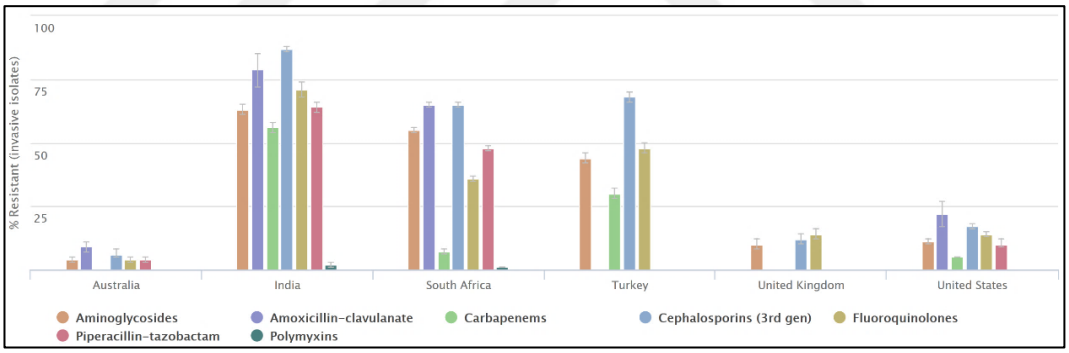

Figure 2.7: Distribution of antibiotic resistance of *Klebsiella pneumoniae* in selected countries.

The carbapenemase-producing *K. pneumoniae* isolates are defined as CRE as other carbapenem-resistant *Enterobacteriaceae*. They were reported among the top three urgent drug-resistant threats in 2013 by Centers for Disease Control and Prevention [47]. Despite nephro- and neurotoxic effect of colistin antibiotic, it has been introduced as the last treatment option to overcome CREs. However, mutations in the chromosomal LPS-modification regulatory genes has caused the emergence of the colistin-resistant strains (e.g., epidemic clone sequence type 258 (ST258)) via the reduction of negative charge on the bacteria surface [15, 48]. Thus, the ability of

carbapenem-resistant *K. pneumoniae* to gain multiclass antibiotic resistance leads to high mortality and important clinical challenges [43]. Collectively, CRE and ESBL producing strains are unarguably much more dangerous than WT strains owing to lack of any effectient therapeutic approaches [4]. Moreover, as mentioned before, easy transfer of the multidrug resistance encoded on plasmids to other strains poses another devastating problem [16, 50]. Therefore, spread of the *Klebsiella* plasmids bearing ESBL and carbapenemase genes is a significant risk factor that likely increases the morbidity and mortality rates at a significant level. Besides, infections by the respiratory pathogens like *K. pneumoniae* are not currently vaccine-treatable [51].

To date, a serious effort has been undertaken to determine novel drug targets and development of combinatorial therapies in order to cope with MDR *K. pneumoniae*. Cefiderocol (S-649266) with a potent activity against a broad range of aerobic gram-negative bacteria (e.g., CREs) is a siderophore cephalosporin in Phase 3 trial [52]. Potential use of cefiderocol against carbapenem-resistant *K. pneumoniae* was suggested for the treatment of lung infections [53]. S649266 (Shionogi Inc.) is a novel siderophore cephalosporin with an antimicrobial activity against beta-lactam and carbapenem resistant pathogens. It has a potential role against multidrug-resistant *K. pneumoniae* infections. As well as individual usage of administration of beta-lactam agents, combination with a beta-lactamase inhibitor is also a significant approach to manage β-lactamase-mediated resistance development. Use of ceftazidime-avibactam (Avycaz) combination in Phase 2 clinical trial is an example of such therapeutic efforts [54]. This drug has been approved by the U.S. Food and Drug Administration (FDA) to manage treatment of complicated intra-abdominal infections and complicated urinary tract infections with limited/no currently available treatment options. Moreover, it is used for the treatment of patients with different types of pneumonia in Europe [55]. Ceftolozane/tazobactam (Zerbaxa) is another drug combination. An FDA approved cephalosporin–beta-lactamase inhibitor combination, Zerbaxa, is useful for intra-abdominal infections and the complicated intra-abdominal infections despite various side effects (e.g., nausea, headache, diarrhea, and fever) [5, 54] However, it does not have any effects on serine group of carbapenamases including KPC and metallo-beta-lactamases [5]. To manage serine carbapenemase-producing *K. pneumoniae* infection, another β-lactamase inhibitor RPX7009 (aka vaborbactam) combined with meropenem (carbavance) was suggested as a drug candidate in a recent study. This drug combination is in phase 3 clinical development [54, 56]. It was

suggested that the carbavance can tackle serious gram-negative 'superbugs' [54]. Alternatively, bacteriophage therapy and identification of novel drug combinations via drug repositioning are among the prominent approaches to preclude the *Klebsiella* infections without emergence of resistance [5, 7]. In a recent study, some synergistic drug combinations out of 25 approved drugs against two drug-resistant *K. pneumoniae* strains were determined. For example, colistin–auranofin–ceftazidime and colistin–auranofin–rifabutin were found to show broad-spectrum activity against some MDR strains of *K. pneumoniae*, *A. baumannii*, *P. aeruginosa*, *C. freundii*, *E. cloacae* and *E. coli* [7].

Currently available treatments cannot completely eradicate the MDR *Klebsiella* and more than half of patients die due to spread of hospital-acquired *K. pneumoniae*. Therefore, it is essential to prevent emergence of resistance via new effective treatment approaches [1, 6].

2.2. Systems Biology Approach

Organisms consist of many complex cellular networks for the homeostatic maintenance. These networks are involved in vital cellular processes (e.g., repair mechanisms, cell division, cellular metabolism, and so on) and/or pathogenesis of the infectious agents. Considering the interconnected and complicated nature of the cellular networks, focusing on a single element of the networks is not sufficient to decipher the whole infection system. Taking this view into account, there was a need for a new field to quantify and analyze all molecular components through an individual network or the integrated networks.

Development of experimental technologies (e.g., tandem affinity purification, yeast two-hybrid, pull-down assay, microarray, RNA sequencing (RNA-seq), dual RNA-Seq, NMR, LC-MS/MS and phage display) and increasingly massive high-throughput data paved the way for the emerge of the concept of 'omics' [57–60]. Many 'omics' research fields based on the analysis of high-throughput datasets (e.g., genome, transcriptome, proteome, metabolome and interactome) have emerged in recent years. Each omic data has a crucial importance to provide an important insight into the infection mechanisms and to develop new threupeutic approaches. However, they have some pros and cons in comparison with each other. For example, genome represents only static information although they are commonly used for classification

and characterization of pathogens, development of novel therapeutic drugs, diagnosis, and detection of virulence factors. On the other hand, some other data sets like transcriptome (non-coding RNA (ncRNA) molecules in addition to protein-coding mRNAs) reflects dynamic nature of the cell [60]. Even if the transcriptome data is much more efficient approach to understand cellular changes in comparison with the genome data, it is also not sufficient per se. This is because high expression level of a gene cannot ensure that amount of its product will be also high or it will be operational [61]. Furthermore, pathogens can utilize a great deal of macromolecules so that they can survive in this host and regulate the gene expression profiles of the host cells. Thus, metabolome data also strongly contribute to our knowledge about the infection mechanisms. As one example, iron is an essential molecule not only for the host cells in terms of its activity for release of the cytokines and regulation of activities of some trascription factors, but it is also vital for pathogens like *Candida albicans* [62, 63]. If a large amount of iron exists in the host cell, the pathogen can uptake it and utilize in the metabolic processes. Therefore, there is a tight regulatory control for iron homeostasis in the host systems [62]. Briefly, use of the different '-omics' datasets is essential to understand the whole picture. As a holistic approach "systems biology" has integrated different disciplines in order to decipher complex biological information via use of '-omics' data at the beginning of the 21st century. Thus, it plays a crucial role in the comprehensive analysis of different datasets and this approach is superior than wet-lab techniques (e.g., RNA silencing or transposon mutagenesis or RNA silencing) in terms of manpower, cost and time.

2.2.1. Constraint-Based Analysis of Genome-Scale Metabolic Network Models

The rapid technologic advancements contributed to the development of omics data collection that provided emergence of metabolic modeling approaches. This systems biology approach is quite useful to reveal the nuances that are not easily realized by the laborious and time consuming experimental techniques and to generate a holistic picture of the cell metabolism. Metabolic models are mathematical representation of the complex systems. They are explicitly useful in some fields like metabolic engineering and medicine because they facilitate various genetic modifications that alter the metabolic phenotypes of the cells [64, 65]. Thus, the

reaction rates can be altered so as to the maximize product of interest [65]. To this end, many metabolic network models have been constructed for a great number of microorganisms such as *Escherichia coli* [66] and *Saccharomyces cerevisia* [67]. Another special focus within the metabolic modeling approach is the modification of the cellular metabolism to predict potential drug targets. It is achieved by investigating the effects of perturbations (e.g., gene knockouts) on the growth of the pathogen [64].

The superiority of the metabolic network modeling approach in terms of time and cost paves the way for the emergence of different modeling techniques that can be broken down into two main categories: (1) kinetic approaches, (2) stoichiometry-based (constraint-based) methods. Even if the kinetic approaches are successful to reflect the real situations via time-dependent dynamic simulations, they need detailed kinetic information [68, 69]. On the other hand, the whole-network metabolic modeling is based on the determination of the constraints, and it is often much easier than identification of the kinetic parameters [70]. Therefore, the constraint-based methods are often preferred in order to analyze a system in the genome-scale manner without challenges of kinetic data acquisition [71, 72].

The stoichiometric models used two major metabolic informations including the metabolic stoichiometry and an objective function (e.g., biomass synthesis). The flux (metabolic reaction rate) distributions were predicted based on the objective function, and there are different stoichiometric modeling techniques to determine metabolic fluxes. Flux balance analysis (FBA) is the most prominent approach [73].

FBA is one of the most frequently used constraint-based modeling methods to predict internal steady-state flux distributions in a genome-scale metabolic network model (GEM) [71, 74]. Thus, it is convenient to reflect only the steady-state growth in continuous cultures or balanced batch cultures [75]. FBA approach applies linear programming to select a metabolic flux vector that satisfies the given objective function (e.g., maximization of the growth reaction) under some constraints (Figure 2.7) [70, 74, 76]. There are three major constraints in the FBA approach: (1) upper and lower bound constraints obtained based on reaction reversibility and flux limitations (information about maximum and minimum reaction rates), (2) equality constraints due to conservation laws for mass (mass and balance) and (3) measurement constraints (Figure 2.8) [70, 73]. The measurement constraints (e.g., uptake rates of the nutrients) are imposed based on the model conditions (e.g., the availability of nutrients) [70]. It should be noted that there are many possible solutions in the underdetermined systems

(more linear equations than unknowns), and these constraints are essential to constrain the solution space.

FBA approach approach is commonly used to improve the production of a useful metabolite [77] and to determine potential drug targets [78, 79]. The formulation of FBA approach is presented in the section 3.3.3. There are different types of FBA approaches such as parsimonious enzyme usage FBA (pFBA) [80], regulatory FBA [81], dynamic FBA (DFBA) [75] and flexible-optimality FBA (flexoFBA) [82] developed in recent years.

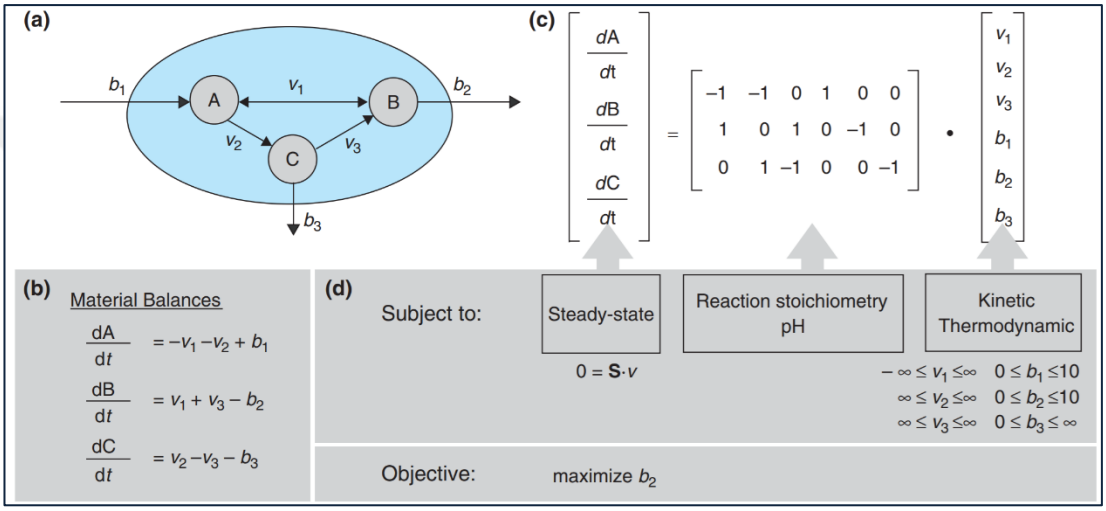

Figure 2.8: A toy metabolic network. a) Metabolic network including three metabolites (A, B and C), b) Mass balance, c) Mathematical representation of constraints, d) Assumptions, constraints, and an objective.

FBA is an explicitly useful technique but it is insufficient to represent the variations in the microbial populations. Both microenvironment and intrinsic factors stimulate a variation in microbial cells. Thus, each single microorganism does not have to exhibit the same growth behaviour. Different flux distributions in a culture due to these differences implying that a microbial cell culture may not grow at the optimal theoretical value. The approach of flexoFBA was developed to reflect the suboptimal growth behaviour of the microorganisms. Even if the microorganisms have a variations in their growth characteristics, they exhibit low flexibility. Therefore, flexoFBA also focuses on enzyme minimization as a secondary objective function. Thus, this approach reduces the solution space more using two consecutive objective functions [82].

In summary, use of FBA (or an FBA-derived approach) enables rapid analysis of the complicated genome-scale metabolic networks. Thus, it is a promising approach to enlighten the pathogen metabolisms and putative drug targets against infectious diseases.

2.2.2. Genome-Scale Metabolic Network Models of *K. pneumoniae*

Hypervirulent *K. pneumoniae* strains are the etiologic agents of some serious diseases such as pneumonia, endophthalmitis, meningitis, bloodstream and urinary tract infections. Hence, they are the leading global causes of morbidity and mortality. Despite the presence of proteome- and transcriptome-based studies on the *K. pneumoniae* [83–86], any comprehensive studies regarding the constraint-based analysis of their metabolic networks were not reported.

To date, two genome scale metabolic network models of different *K. pneumoniae* strains (i.e., MGH 78578 strain and KPPR1 strain) were developed [46, 74]. The GEM of iYL1228 (1,229 genes and 1,658 metabolites involved in 2,262 reactions) was the first reported *K. pneumoniae* model in the literature. This GEM was reconstructed for MGH 78578 strain by mapping from an *E. coli* reconstruction (iAF1260). An experimentally determined biomass equation of the MGH 78578 strain was included in the model. Using FBA approach, the reconstructed model was refined and validated. It was shown that iYL1228 can predict the effects of the carbon sources on the bacterial growth, giving the consistent results with *in vitro* studies. In addition, *in silico* gene deletion analysis was performed to identify essential genes of the microorganism. 118 essential genes were predicted computationally. Eight out of these genes were found as unique when compared to the computational predictions by *E. coli* and *S. typhimurium* models [46].

The second model was derived from iYL1228, designated as iKp1289 (2,474 reactions, 2,000 metabolites and 1,290 genes), through translation of the genes of MGH 78578 into the genes in the KPPR1 genome using KBase platform [87]. KPPR1 strain is a rifampin-resistant human isolate with a higher virulence. This model was also confirmed by comparing the *in silico* and *in vitro* results. The growth assays were designed based on both investigation of the growth profiles on different carbon sources and gene essentiality analysis. Furthermore, both *Klebsiella* models were compared in the study. The researchers revealed that KPPR1 can utilize a larger range of carbon

sources and it can use the nutrients more robustly in comparison with MGH 78578 [46]. Both GEMs were used in the scope of this thesis for a comprehensive system-level analysis of the bacterial metabolisms.

2.2.3. Identification of Drug Targets via Constraint-Based Metabolic Network Modeling

Constraint-based modeling of the pathogens has been a popular approach in order to reveal putative drug targets by dealing with the underlying metabolism [88]. Identification processes of the drug targets can be broken down into two general categories: (1) gene-centric approaches and (2) metabolite-centric approaches. These approaches provide both elucidation of relationships between network components and anti-pathogen target discovery through identification of essential genes/metabolites.

Gene-centric approaches include detection of the novel genes that are putative drug targets in much shorter time in comparison with the labor-intensive and time-consuming experimental techniques. This approach is based on the identification of the genes/reactions essential for survival (Figure 2.9) [89]. Essential genes/reactions can be predicted independent from the condition (regardless of aerobic, anaerobic or facultatively anaerobic) or they can be identified based on a certain culture condition [89]. Many approaches focus on mimicking the real systems by simulating the bacterial growth under conditions similar to host environment for more precise predictions [90]. Mimicking the host environment can be achieved via integration of the host and pathogen metabolism [91] or simulation of the media representing the host environment [92]. Then, FBA approach can be used to identify all essential enzyme targets of the pathogen specified in the GEM that support production of the growth-related metabolites [71, 90]. Thus, many targets can be screened at a low cost in comparison with the high-throughput experimental techniques (e.g., transposon mutagenesis or RNA silencing). Moreover, the results considerably narrow down the targets that will be experimentally validated [71]. Thus, constraint-based approaches provides elucidation of the common and species-specific drug targets by simultaneously screening the genes under the specified conditions [93]. Then, these targets are evaluated based on some criteria such as druggability or minimal side effect to select the most efficient targets with mimimum harm. For instance, Plata and

colleagues (2010) constructed a genome-scale metabolic network of *Plasmodiu falciparum* (1,001 reactions and 616 metabolites) to prevent malaria through identification of potential enzymatic drug targets. Using FBA approach, they suggested 40 candidate drug targets (e.g., nicotinate mononucleotide adenylyltransferase involved in NAD^+ metabolism) that exhibit a low sequence similarity with human proteins [94]. *Clostridium difficile* is another serious pathogen which leads to diarrhea and intestinal inflammation, with a potential to lead to fatal colitis [95]. Genome-scale metabolic network of *C. difficile* strain 630 was reconstructed to identify essential genes of the pathogen and this network consists of 806 genes and 1031 reactions. Thus, Larocque and colleagues (2014) detected 76 essential genes and 39 essential gene pairs that were specific to this strain. Of these putative targets, some genes (e.g., isoprenoid synthase, aspartate-semialdehyde dehydrogenase, UDP-N-acetylenolpyruvoylglucosamine reductase and NH_3-dependent NAD(+) synthetase and diaminopimelate epimerase) are crucial in the amino acid metabolisms [96].

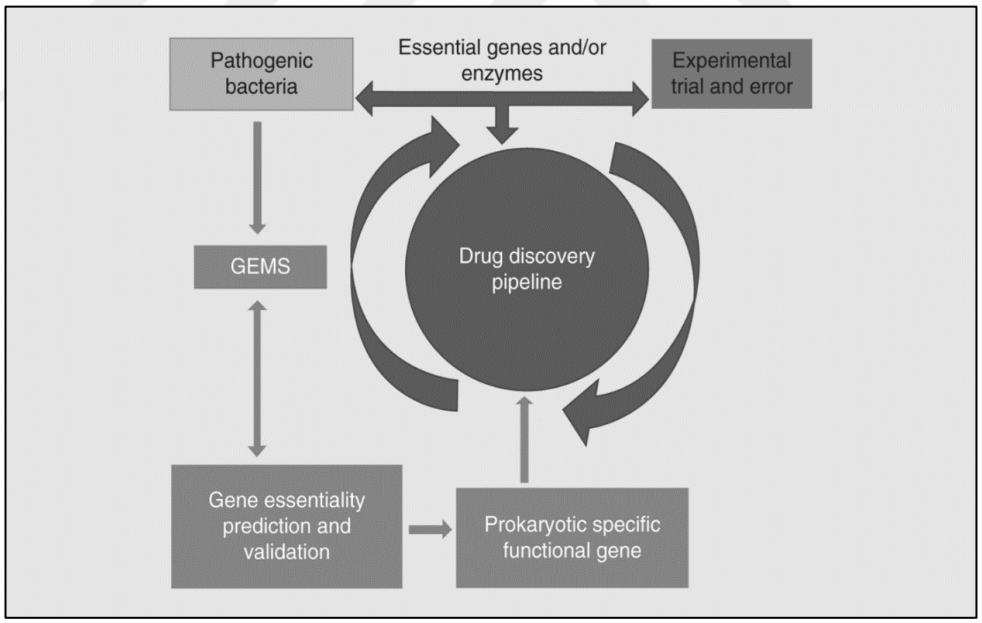

Figure 2.9: General drug development pipeline.

Identification of the drug targets based on a metabolite-based approach is another useful method. The logic underlying this approach is the use of chemical similarity of metabolite structures to the existing drug molecules. Moreover, this approach can delay the resistance development by targeting more than one reaction [97]. Essential

metabolites are identified in the first step of the metabolite-centric approach. After some filtering processes these metabolites can be used for drug screening [97, 98]. Thus, inhibitors may be designed based on the structural resemblance to the selected metabolites. Using this approach, ten candidate metabolites (i.e., lipid II, meso-diaminopimelate, pantothenate, biotin, shikimate, l-aspartyl-4-phosphate, dTDP-α-L-rhamnose, UDP-D-galacto-1, 4-furanose, des-N-acetyl mycothiol/Cys-GlcN-Ins and siroheme) with broad spectrum activity were suggested by Sarker and colleagues (2010). These candidates were selected through literature mining, comparative analysis and pathway analysis mainly based on the following criteria: (1) examination of essentiality of these metabolites for the specified pathogens (*M. tuberculosis*, *K. pneumoniae*, *E. coli*, *P. aeruginosa*, *S. typhirium*, *S. aureus*, *B. Subtilis* and so on) but not involved in human metabolism, (2) determination of biological functions of the metabolites, (3) convenience of the metabolites' physio-chemical characteristics to design the inhibitors, (4) investigation of widely distribution of the metabolites among pathogens, (5) presence of any inhibitors against the enzymes that are associated with the synthesis of the selected metabolites and (6) presence of any knowledges about antimicrobial activity of the metabolite analogs [98]. Similarly, Kim and colleagues (2011) developed EM-Filter approach to select the most efficient essential metabolites in the GEM of *Vibrio vulnificus* CMCP6. This approach has focuses on higher connectivity and lack of the essential metabolites of the pathogen or the related enzymes in human metabolism. Then, they identified the structural analogs of the selected metabolites. An analog of PABA was reported to be a potential inhibitor against *V. vulnificus* [99]. Considering these efforts, evaluation of the essential metabolites can facilitate design of more efficient drugs or screening of the existing structural analogs from chemical compound libraries. Furthermore, this approach can delay resistance development [8].

In the scope of the thesis, putative drug targets were suggested through two approaches discussed in this section. These analyses were performed for both *Klebsiella* strains (MGH 78578 strain and KPPR1 strain) in order to contibute to treatment of the infections by the resistant strains. Another aim of this study is to get insight into the metabolism of both strain. Using the genome-scale metabolic network models, five druggable virulence factors with broad-spectrum distribution among important pathogens were proposed as promising drug target candidates. Moreover, the metabolite-centric approach reveals three metabolic enzymes to be putative drug

targets in order to handle infections caused by both *Klebsiella* strains. It is the first comprehensive study on analysis of the GEMs of MGH 78578 and KPPR1 strains. Further analyses and validation of these putative targets may pave the way for more efficient control of the *Klebsiella* infections in the future.

3. MATERIALS AND METHODS

3.1. *K. pneumoniae* Strains and Their Genome-Scale Metabolic Network Models

Prominent reference strains including a highly virulent, rifampin resistant derivative of *K. pneumoniae* ATCC 43816 (a K2 clinical pneumonia isolate) known as KPPR1 strain (GenBank accession number: CP009208) and relatively avirulent, carbapenem sensitive K52 pneumonia isolate known as MGH 78578 strain (GenBank accession number: CP000647) were used within the scope of this study [43]. Even if these strains exhibit different levels of virulence, the genome of MGH 78578 strain shares highly homology (88%) over the genome of KPPR1 strain. Thus, they share a similar metabolic profile [46].

K. pneumoniae strain MGH 78578 first isolated from the sputum of a pneumonia patient. It contains five plasmids and this MDR pathogen is resistant to ampicillin, ticarcillin, trimethoprim-sulphamethoxazole and gentamicin as it is sensitive to ciprofloxacin, amikacin, and imipenem [46, 100]. Moreover, it harbours some virulence factor-encoding genes regarding biosynthesis of CPS, siderophores, LPS and fimbriae. Expression levels of these genes mostly change based on the growth phase of the bacteria. For instance, most of the genes involved in siderophore and CPS biosynthesis show lower expression levels in the stationary phase [85]. A GEM of this strain with an experimentally defined biomass equation was constructed by Liao et al., in 2011. This model includes 1,229 genes and 1,658 metabolites involved in 2,262 reactions [74].

On the other hand, *K. pneumoniae* strain KPPR1 is a clinical pneumonia isolate that contains a single circular chromosome for a total of 5,374,834 bp in size with 5,191 predicted genes (25 rRNA, 85 tRNA, and 5,081 protein-coding sequences) but it does not have any plasmids. This strain is more virulent in comparison with the MGH 78578 [1, 101]. Lawlor and colleagues observed hypertrophy and hyperplasia at 12 h after pulmonary KPPR1 infection in a mouse model [102]. Hence, this *Klebsiella* strain seems more promising to understand the characteristics of virulence and pathogenesis of this pathogen. A GEM of KPPR1 strain (iKp1289) including 2,474 reactions, 2,000 metabolites and 1,290 genes was recently developed by Henry and

colleagues [46] from closely related model iYL1228 [74]. Both metabolic network models of *K. pneumoniae* were used in this study.

3.2. Softwares and Toolboxes

Herein, MATLAB R2017a platform with some user-friendly toolboxes including Systems Biology Markup Language (SBML) toolbox and Constraint Based Reconstruction and Analysis (COBRA) toolbox was used in the analyses [103]. Fast-SL function developed [104], an in-house developed algorithm (Figure 3.3) and a single gene deletion function in COBRA Toolbox were used to identify essential genes of the strains. Essential metabolites were identified and prioritized by the other in-house MATLAB codes. In addition to FBA approach, flexoFBA was also used as described by Tarlak and colleagues [82] (section 3.3.3).

3.3. Bacterial Growth Under Different Conditions

3.3.1. Bacterial Growth Conditions

In this work, the GEMs of *K. pneumoniae* strains (iYL1228 and iKp1289) were evaluated through a two-step process. First, growth simulations were validated through model-based predictions and measured values in the literature [46, 74]. Second, the validated metabolic networks were used to identify the putative drug targets.

Validation of the models was performed by simulating the growth conditions highlighted in the articles [46, 74] for elucidation of the bacterial growth phenotypes (i.e., M9 or carbon-D-glucose medium (CDG (Table A1.1): its recipe is also available upon KBase [87]) supplemented with different carbon sources [46, 74]) and for the gene essentiality analyses (i.e., M9 minimal medium (M9) or Luria-Bertani (LB) broth medium (Table A1.2): its recipe is also available upon KBase [87]). M9 medium has been already integrated into the iYL1228 by the developers of the model whereas CDG medium was integrated into iKp1289 in this study. Therefore, there is no need to integrate glucose-supplemented mimimal medium condition into iYL1228. Of note, both minimal media include only glucose as the carbon source. Therefore, these media were modified replacing the glucose by each carbon source one by one for each growth simulation.

Host-cell mimicking nutrient environments were constructed based on the literature searches and they were used in the drug target discovery. For instance, *K. pneumoniae* induces production of thick jelly sputum through inflammation in the lungs. This pathogen can be isolated from the sputum and the sputum cultures are prominent for diagnosis [4, 105]. On the other hand, Cano and colleagues proved that *K. pneumoniae* can survive inside alveolar macrophages by blocking phagosome maturation [106]. These conditions including sputum [92] and alveolar macrophage [91] were combined to simulate bacterial growth in this work (Table B1.1 and Table B1.2). In addition to sputum, cultures of other body fluids (e.g., blood, urine, etc.) are frequently tested to reveal the etiological agents of various diseases (e.g., pulmonary diseases, gastrointestinal disorders, etc.), as well [107]. Therefore, human body fluid [108] was also integrated into the models to construct a more comprehensive environment for the simulation of *K. pneumoniae* growth in human body (Table B1.3 and Table B1.4).

3.3.2. Constraining Sytem Boundaries and Stoichiometric Coefficients

Constraining the system boundaries and/or the stoichiometric coefficients is a prerequisite for the growth simulation in a defined condition and refinement of a redundant model.

Herein, each medium was integrated into the model through setting the system boundaries. Lower and upper bounds (lb and ub: maximum and minimum reaction rates) were set based on the medium content (except for M9 medium for iYL1228). Thus, uptake rates of the metabolites were allowed to change between a certain range based on the media used in the study.

Next step of the medium-integration process is to block the use of the other compounds apart from the medium components. This was achieved by allowing uptake of the metabolites only in the growth medium. Therefore, uptake of other metabolites were prevented by setting the corresponding reaction rates to 0. On the other hand, by-product production was not limited. Thus, the bacterial cell could only consume the metabolites in the medium as it could produce any metabolites (by-products) based on its metabolic profile.

In addition to the proper construction of the growth conditions, boundaries of the non-growth-associated maintenance (NGAM) were set to agree with the values from the original papers [46, 74]. NGAM represents the energy (mmolATP/gDCW/h) essential to maintain the cellular processes apart from growth (e.g., motility, cell repair and membrane potential) [109].

Preliminary assessment of the models revealed some redundant reactions (31 duplicated uptake reactions) and metabolites (200 duplicated metabolite names) in the GEM of iKp1289 although this redundancy does not affect simulation results of the model. Nevertheless, the model was refined by ignoring the redundant metabolites and reactions. This step was achieved by setting all non-zero stoichiometric coefficients corresponding to the redundant metabolites to zero; pratically removing the redundant metabolites from the system. For the redundant reactions, the corresponding columns of the stoichiometric matrix were all made zero.

3.3.3. Simulation of Bacterial Growth

The growth simulations were performed using FBA and/or flexoFBA after medium integration process described in the section 3.3.2. As highlighted earlier, FBA is a useful approach to predict the flux distribution in a GEM by assuming the net change in the metabolite concentrations as zero. That is the assumption of steady-state conditions. This powerful technique is based on the mass balance constraint given in the equation (3.1)

$$\sum_{j=1}^{n} \left(S_{ij} V_j\right) = 0 \quad for\ i = 1, 2, \dots, m \tag{3.1}$$

where the stoichiometric matrix with the coefficient of the metabolite *i* (within a total of *m* metabolites) is denoted as S_{ij} and V_j implies the rate of the j^{th} reaction (within a total of *n* reactions).

To constraint the solution space further, the flux boundaries are also defined based on the reaction reversibility (equation (3.2))

$$\mathrm{lb}_j \leq \mathrm{V}_j \leq \mathrm{ub}_j \quad for\ j = 1, 2, \dots, \mathrm{n} \tag{3.2}$$

where lb_j and ub_j are the lower and upper bounds for the j^{th} reaction flux. The fluxes are predicted between an upper bound (+∞) and a lower bound (-∞) for reversible reactions whereas the lower bound is set to zero for irreversible reactions. For modelling purposes, a huge number (i.e., 1000) was used to represent infinity in this study.

FBA approach focuses on prediction of the flux distributions satisfying an optimal objective function under the defined constraints. Maximization of the bacterial growth is commonly used as the biological objective (equation 3.3.a and equation 3.3.b)

$$R_{Biomass}: 0.621\ \text{Alanine} + 0.306\ \text{Arginine} + 0.204\ \text{Asparagine} + \cdots \quad (3.3.a)$$

$$\text{Objective: } max\ (V_{biomass}) \quad (3.3.b)$$

To find out the steady-state flux distributions at the optimal biomass production, this optimization problem (known as a linear programming problem) must be solved using a mathematical optimization software (e.g., open source examples (lp_solve and GLPK) and proprietary versions (CPLEX, LINDO and Gurobi)) [110]. Herein, iYL1228 and iKp1289 were downloaded from 'BiGG Models' [111] and 'KBase' databases. MATLAB platform was used to load the models and to solve corresponding flux distributions using Gurobi based on the FBA approach.

Although FBA is a prominent approach to predict intracellular flux distributions, it ignores the variations within microbial populations. However, microorganisms exhibit a metabolic variability due to the enzymatic differences and intrinsic variations. Therefore, investigation of the microbial metabolisms by considering a strict optimality is not a realistic approach. On the other hand, the experimentally detected replicate fluxes were reported to be similar despite the intrinsic variations. This phenomenon indicates that the organisms may prefer low flexibility. With these assumptions, a relatively new approach (flexoFBA) was developed by Tarlak et al., (2014). This approach takes account of the microbial variations to reflect the slightly suboptimal growth rates [82]. First, it sets the boundaries of the maintenance reaction to a flexible interval (20% range) instead of using a fixed constraint. Then, the optimal growth rate is calculated and is multiplied with a normally distributed (left-tailed) random number with mean 1 and standard deviation 0.05. This step repeats 1000 times.

Thus, the flexibility in the metabolism of microorganisms is reflected. Internal flux distributions are predicted for the objective of minimal enzyme production per cycle. Lastly, the average growth rate is calculated. This method was also used in the study.

3.3.4. Generation of A New Biomass Reaction

Biomass reaction is the mathematical representation of the cell composition and it is prominent to predict essential genes as the putative drug targets accurately. Xavier and colleagues (2017) defined some universally essential organic cofactors (i.e., nicotinamide adenine dinucleotide (NAD), NADH, nicotinamide adenine dinucleotide phosphate (NADP), NADPH, flavin adenine dinucleotide (FAD), coenzyme A (CoA), flavin mononucleotide (FMN), pyridoxal 5′-phosphate (PYDX5P), S-adenosyl-L-methionine (SAM/AMET) and thiamin diphosphate (THMPP)) involved in the biomass compositions of the prokaryotes [112]. These cofactors are missing in the biomass equation of the *Klebsiella* strains. Therefore, a new biomass equation was generated based on these cofactors. This process was achieved by resetting the biomass-associated stoichiometric coefficients of the missing cofactors in the model. In this context, the coefficients of these cofactors were identified based on the biomass compositions of three GEMs of *E. coli* (*i*AF1260 [66], iJO1366 [113] and iJR904 [114]). If a cofactor is absent in the biomass equations of these GEMs, a small coefficient ($1x10^{-5}$) was assigned. Then, 1023 combinations of all cofactor coefficients were integrated to the biomass equation through the *S* matrix. Effect of each combination on the bacterial growth was evaluated to distinguish presence of any cofactors inhibiting the biomass formation. THMPP was removed from the biomass equation due to blocking the bacterial growth. The new biomass equation is illustrated in Figure 3.1.

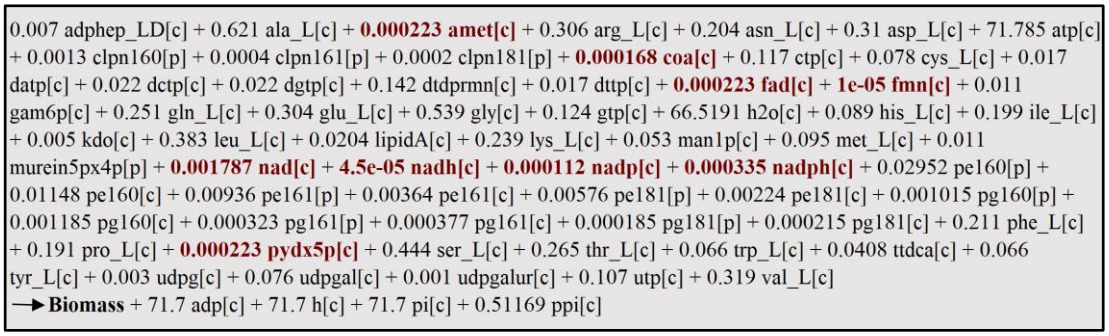

Figure 3.1: Generation of a new biomass reaction. The added cofactors are given in bold and red letters.

3.4. Gene Essentiality Analysis

In silico analysis of gene essentiality was performed under different conditions to validate both models and to determine the putative drug targets. An essential gene implies that the gene is indispensable for survival under a specific growth condition. Therefore, these genes are candidate drug targets. To identify the essential genes for *K. pneumoniae* growth, lb and ub of the reactions related to each gene were set to 0 in the GEM one by one (i.e., removal of each gene by blocking the corresponding reactions). After each change on lb and ub values, the GEM was solved to understand whether the gene has any effects on the bacterial growth or not. If the removal of the reactions associated with any genes resulted in a considerable reduction of the flux regarding the biomass objective function, these genes were accepted to be essential.

There are some common algorithms for the gene deletion in scientific literature. Two of them, used in this study, are based on single gene deletion. One of which is single gene deletion function in COBRA toolbox. This function calls the deletion procedure for each unique gene, mentioned above. The other algorithm is Fast-SL, based on an eliminated reactions list (*eliList*), which excludes ATP maintenance, uptake, exchange, and dead-end reactions from the gene deletion procedure. Thus, it narrows down the search space and performs single or multiple gene deletions faster.

To narrow down the search space and increase the sensitivity, an in-house algorithm was also developed for single gene deletion in the scope of the thesis. This algorithm, first, identifies the combination of gene sets affecting each reaction individually. To determine the gene sets for each reaction, first, gene-reaction associations must be defined in a convenient format. Therefore, a matrix including the gene indices and logical operators (OR/AND) was generated using ‘rules’ matrix in

the SBML model. Then, the logical operators were replaced by 0 or 1 (i.e., 1 instead of "|" and 0 instead of "&") and the blank cells were filled with 0. This new matrix was called reaction-gene matrix (RGM) in this study. It displays the relationships between the reactions and genes. For the sake of clarity, the generation process of the reaction-gene matrix (RGM) is illustrated in two stages (Figure 3.2).

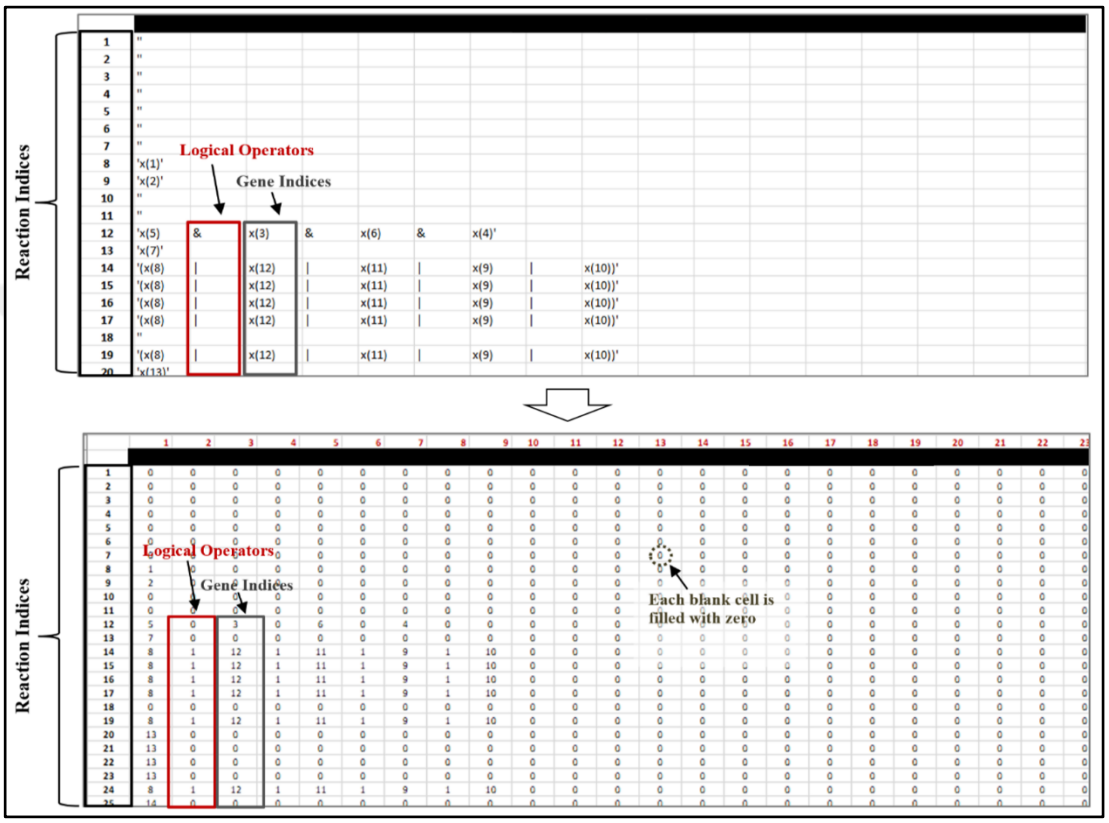

Figure 3.2: Generation of reaction-gene matrix (RGM).

Of the combination of gene sets, one gene set was selected to be deleted for each reaction based on the reaction-gene matrix (RGM). It should be noted that one gene can be associated with multiple reactions. Therefore, all related reactions regarding each gene set were evaluated in the deletion process. If the deletion of the related reactions considerably reduced the bacterial growth (smaller than 1% of the maximum WT growth rate [104]), these gene sets were considered computationally essential. The in-house single gene deletion algorithm is given in Figure 3.3.

```
Input: SBML model of an organism and a special gene-reaction matrix (RGM)
Output: Set of single lethal genes G_sl

(1) Generate an RGM matrix as described in this study

(2) % Do FBA to find maximum growth rate and set the cut-off
    V_bio = max(v_growth);
    cut-off = V_bio x 0.01;

(3) % Single gene deletion
    for each gene i do
        Set an empty matrix R_ind
        Find the reaction indices related to gene i
        for each related reaction j do
        % Find one more than the number of logical OR and the repeat number of each gene
            or_count = number of OR relation + 1 in related row(s) of RGM;
            gene_count = number of elements in related row(s) of RGM;
            if or_count = gene_count then
                Add j to the matrix R_ind
            end if
        end for
            if R_ind is not empty then
                Reset the lower and upper bounds
                Set the upper and lower bounds for the reaction j (V_j) to zero
                % Do FBA to find growth rate for gene knockout
                V_bioKO = max(v_growth);
            end if
    end for

(4) % Determine the essential genes based on the cut-off
    if V_bioKO < cut-off then
        Assign the detected genes to G_sl matrix
    end if
```

Figure 3.3: The algorithm for the gene essentiality analysis.

3.5. Metabolite Essentiality Analysis

Essential metabolites were identified by another in-house developed algorithm. The logic behind this algorithm is to block all outgoing reactions associated with each metabolite and to examine the effect of this perturbation on the bacterial growth. It was achieved by setting the sitoichiometric cofficients of the related reactions to zero. This is the same as setting lb and ub of the outgoing reactions to zero. The same cut-off with gene essentiality analysis (cut-off: 1% of the wild-type growth rate) was used to identify the essential metabolites. Thus, the metabolites whose deletions considerably reduced the bacterial growth were selected.

3.6. Prioritization of Putative Drug Targets

The lists of essential genes and metabolites were filtered to decide more effective drug targets causing no/minimal side-effects in the host. The drug target discovery process is summarized by the following flowchart (Figure 3.4).

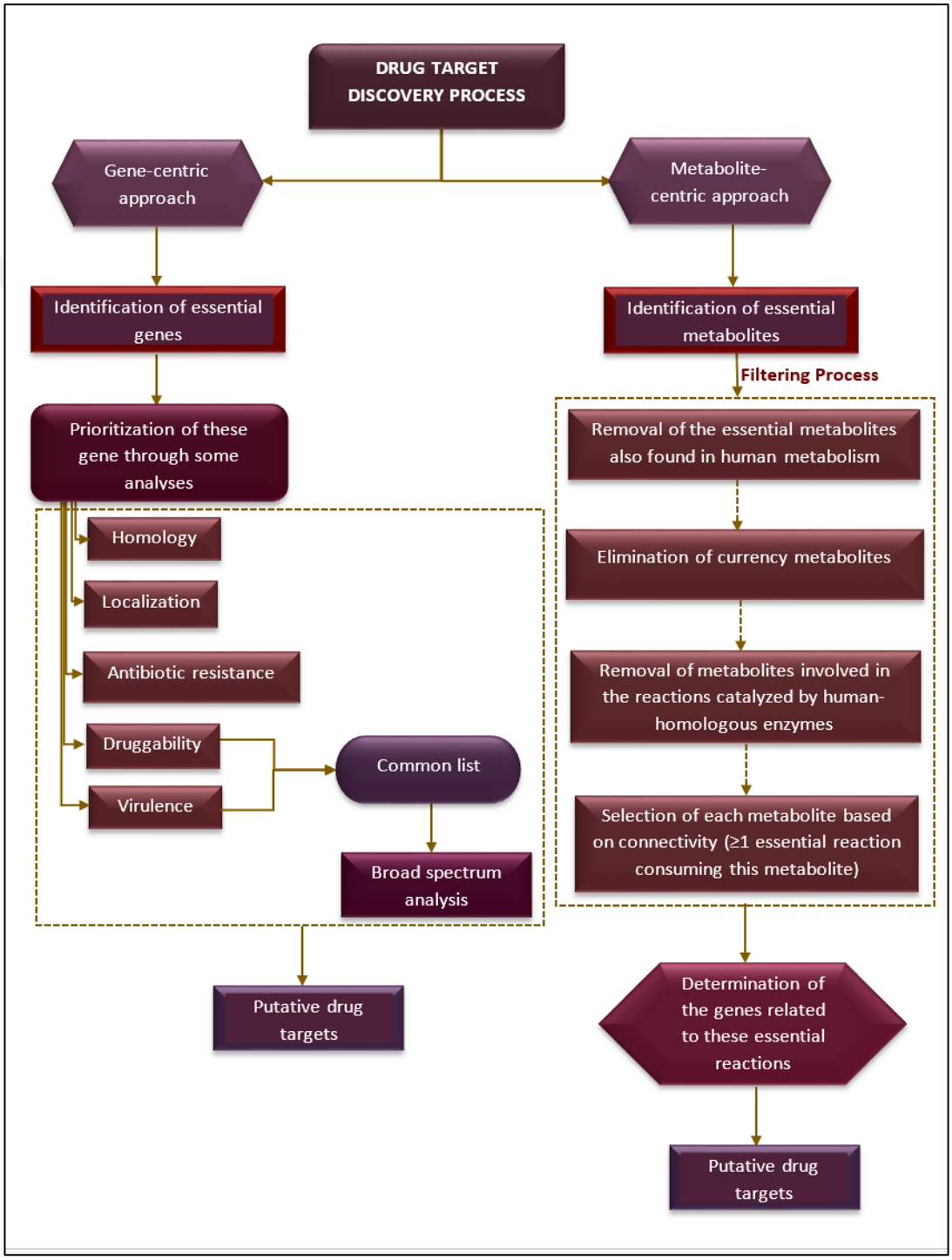

Figure 3.4: Priorization of the drug targets.

Some databases were used to prioritize the essential genes based on (i) homology (BLAST at the NCBI website [115]), (ii) subcellular localization (CELLO2GO [116], PSORTb [117] and iLoc-Gneg [118]), (iii) antibiotic resistance (ARG-ANNOT [119]), (iv) virulence (virulence factor database (VFDB) [120]) and (v) druggability (DrugBank [121]). In addition, prevalence of the selected genes among different pathogens was investigated through PBIT: Pipeline Builder for Identification of drug Targets (PBIT) web browser [122].

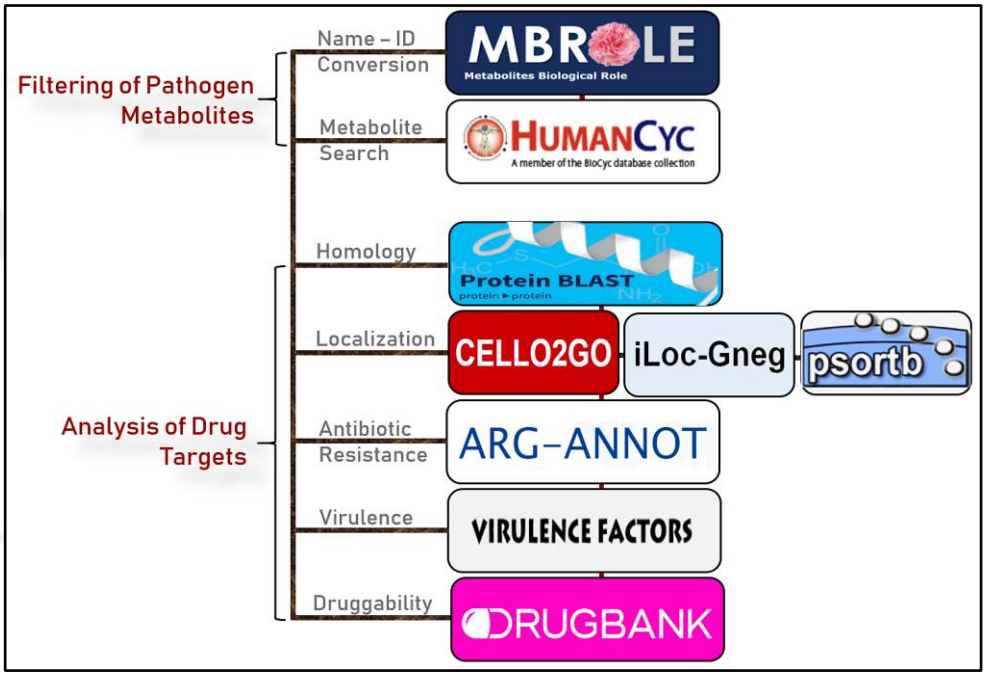

Figure 3.5: Databases used in the scope of this thesis study.

Essential metabolites were also prioritized through (i) elimination of the pathogen metabolites/gene products found in human metabolome/proteome, (ii) removal of the currency metabolites and (iii) selection of the metabolites with the higher connectivity. These steps are discussed in detail in the subsections 3.6.7 and 3.6.8. Two major databases (HumanCyc [123] and Metabolites Biological Role (MBRole) [124]) were used in the prioritization of the essential metabolites. The MBRole server is useful for functional enrichment analysis of metabolites from metabolomics experiments by extracting biological and chemical information from various databases (e.g., KEGG, HMDB and so on) and it contains an ID conversion tool [124]. Thus, it provides IDs of the compounds in the pathogen and human models. HumanCyc includes the manually curated data regarding metabolic pathways and

various enzyme-powered reactions. Likewise, it is useful to detect the metabolites found in human metabolism [123]. HumanCyc also stores a list of compounds in the human metabolism.

The main databases applied in this study are illustrated in Figure 3.5. All analyses used in for priorization of essential genes and metabolites are explained in the following subsections.

3.6.1. Non-Homology Analysis

Non-homology analysis is crucial to reduce the potential hazardous side effects. To identify the genes with no human homologs (non-homologous genes), the essential genes were determined via gene essentiality analysis as highlighted in the section 3.4. These genes were screened in terms of homology. In this step, BLAST search was used against all human protein sequences in the Refseq database [125]. The significantly similar proteins were accepted to be homologous for the expected value (E-value) cut-off $<1x10^{-3}$. The genes having more than 30% sequence identity with their human counterparts were removed [126].

3.6.2. Analysis of the Subcellular Localization

Subcellular localization is an important parameter determining protein functions since the cellular compartments include various compounds necessary for the protein function. Information on the protein localization is also useful to identify potential vaccine candidates and drug targets [127]. To date, many different approaches were developed in order to predict subcellular localization of a protein. Some of them like TargetP is based on the characteristics of the signal sequences to be directed to the certain localizations in a cell [128]. Another approach is to use information of amino acid composition. This method is based on different approaches such as neural network approach or support vector machine (SVM) techniques [129]. A SVM-based method, CELLO, is a prominent web service to predict five main subcellular localizations of the proteins including cytoplasm, periplasm, inner membrane, outer membrane, and extracellular space. It is quite useful to elucidate the localizations of the proteins for both gram-negative and gram-positive bacteria [130]. CELLO2GO server was published in 2014 for more detailed analyses. It provides both localization and GO information by integrating CELLO and BLAST tools [116]. PSORTb is another

commonly used tool that predicts the localizations based on different characteristic properties of the proteins (e.g., inference from the localization of known proteins, amino acid composition, presence of a signal peptide and so on) and Bayesian networks [117]. Unlike CELLO and PSORTb, iLoc-Gneg provides more comprehensive compartment information such as fimbrium, flagellum and nucleoid [118]. In this study, PSORTb (p-value >7.5 [127]), CELLO and iLoc-Gneg were used to predict the subcellular localizations of non-homologous gene products. If at least two of the services lead to the same prediction for a gene product, this location was assumed to be appropriate. Unclear results were validated using UniProt database [131]. Furthermore, some localizations of the gene products were determined using a detailed subcellular localization study of *E. coli* [132].

3.6.3. Antibiotic Resistance Analysis

Recent bioinformatic advances have allowed emergence of various databases for identification of the previously defined or putative antibiotic resistance genes from nucleotide or protein sequences. In this context, Antibiotic Resistance Genes Online (ARGO) has been developed in 2005 as a first attempt. However, it represents limited information about resistance genes. It harbours the sequence information restricted to only lactamase, tetracycline, and vancomycin genes [119, 133]. Antibiotic Resistance Genes Database (ARDB) released in 2009 is proper only if the sequence entry is smaller than 5 kb. Besides, it has not been updated since 2009 [119, 134]. Comprehensive Antibiotic Resistance Database (CARD) includes a manually curated data related to intrinsic, mutation-driven and acquired resistance. This database has been updated in a regular manner. Similarly, Antibiotic Resistance Gene-ANNOTation (ARG-ANNOT) database stores the regularly updated sequences [119, 135]. This versatile database stores the resistance gene sequences against different antibiotic classes (i.e., aminoglycosides, beta-lactamases, fosfomycin, fluoroquinolones, glycopeptides, macrolide-lincosamidestreptogramin, phenicols, rifampicin, sulfonamides, tetracyclines and trimethoprim). Herein, ARG-ANNOT database was used along with a local sequence alignment editor (Bioedit). Identification of the existing and putative new antibiotic resistance genes within the non-homologous gene list was aimed. In this context, the genes with a possible role in

antibiotic resistance were determined using a recently updated data (May, 2018) with E-value threshold of $1x10^{-4}$ via local BLAST in Bioedit without a web interface.

3.6.4. Virulence Analysis

Putative targets related to both survival and virulence of the cell are promising candidates [122]. Virulence factor (VF)-related genes of *Klebsiella* strains within the essential non-homologous gene lists were revealed using Virulence Factor Database (VFDB). This database includes up-to-date, curated data of VFs belonging to 30 popular bacterial pathogens (e.g., *Escherichia*, *Salmonella*, *Pseudomonas*, *Mycoplasma*, *Streptomyces* and so on) [136].

Briefly, VF screening was carried out by browsing amino acid sequences of the essential, non-homologous genes against the VFDB core dataset with E-value cut-off $<1x10^{-4}$, bit score >100 and identity ≥ 65% [137].

3.6.5. Druggability Analysis

The capacity of a target to bind to a drug-like molecule is called druggability. It is a significant parameter in drug target prioritization because all protein structures are not convenient for drug-binding [122]. Since the first release in 2006, DrugBank is widely used up-to-date knowledgebase for drug/target design and it contains comprehensive information regarding drug molecules, prediction of drug metabolism or interaction, docking and so on [121]. To investigate druggability of the selected targets (i.e., essential, non-homologous genes) and to identify the existing drugs for alternate therapeutic aims (drug repositioning/repurposing), BLAST tool was used with E-value of $1x10^{-25}$ against DrugBank database [138].

3.6.6. Broad Spectrum Analysis

In addition to the use of the databases specified in Figure 3.5, PBIT: Pipeline Builder for Identification of drug Targets (PBIT) web browser was used to determine broad spectrum activity of druggable virulence factors. This analysis involves identification of the targets having homologs in other pathogens [122]. Elucidation of widely distributed pathogenic drug targets is important for a comprehensive treatment because a bacterium does not have to lead to the infection process per se. This process

can be managed by multiple distinct bacteria. Broad-spectrum antibacterial target analysis can provide: (1) identification of the targets found in the bacteria causing co-infection, (2) development of broad-spectrum drugs to treat multiple infections and (3) identification of the conserved genes to delay the resistance development. This analysis was performed using BLAST algorithm against protein sequences of 181 pathogenic organisms with E-value cut-off of $1x10^{-25}$, minimum identity of 70% and alignment length cut-off of 90%.

3.6.7. Generation of A Common Metabolite List

Akin to essential genes, essential metabolites must be also selected to avoid any possible side-effects. To eliminate the metabolites involved in both pathogen and human metabolisms, first, 'a common metabolite list' was compiled. The common metabolites were detected using GEMs of Recon 1 (a comprehensive literature-based human metabolic network accounting for 1,905 genes, 2,766 metabolites and 3,741 metabolic reactions) [139] and Recon 2 (updated human metabolic network including 1,789 enzyme-encoding genes, 7,440 reactions and 2,626 unique metabolites) [140] as well as a human metabolite list from HumanCyc [123]. Three main steps used in the study are summarized in Figure 3.6.

Firstly, metabolite names in the models were modified to distinguish the same metabolites with slightly different denominations. The host metabolites from the GEMs of human and HumanCyc database were compared with the names of the compounds in iYL1228 and iKp1289, respectively. The common metabolites were listed. Using MBRole server, both pathogen and human metabolite names were converted to various IDs in databases including KEGG, PubChem, ChEBI and HMDB. The compounds from HumanCyc were also converted to these metabolite IDs by means of MBRole and Metabolite Translation Service of HumanCyc. Thus, the common metabolite lists were extended by detecting the common metabolites of the *Klebsiella* strains with human through comparison of the IDs. Lastly, these lists were combined to obtain a general common metabolite list.

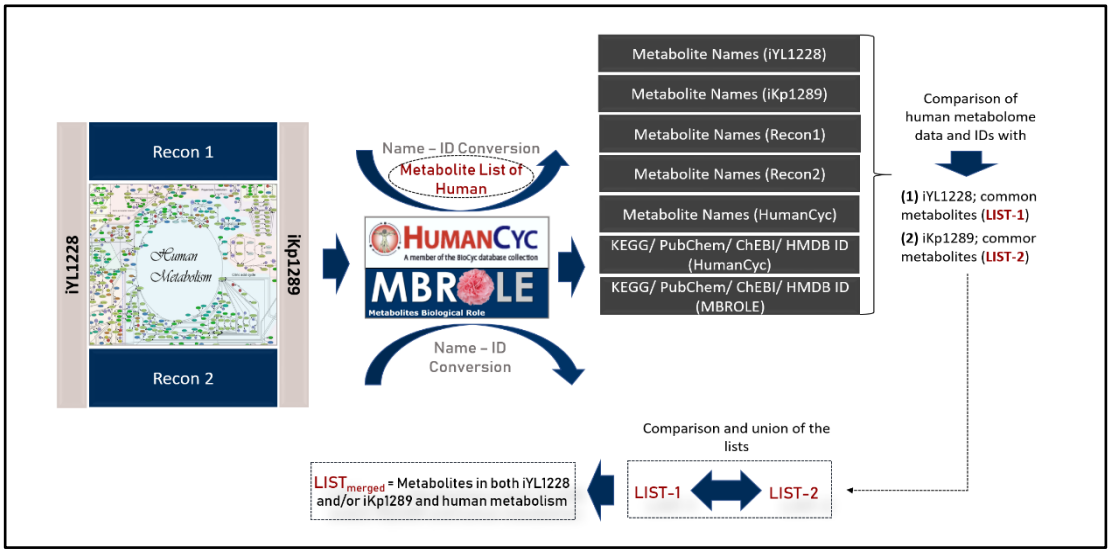

Figure 3.6: Identification of the pathogen metabolites also found in the human metabolism.

3.6.8. Filtering Process to Select Metabolite-Based Targets

Initial set of essential metabolites were filtered based on a modified procedure of Kim and colleagues (2010) (i.e., a modified EMFilter approach) in order to narrow the list of potential targets. This process includes (1) removal of the metabolites involved in the human metabolism, (2) elimination of the currency metabolites, (3) selection of the essential metabolites involved in at least one essential outgoing reaction and (4) elimination of the metabolites associated with human homologous enzymes [141].

To identify metabolites also found human metabolism, all essential pathogen metabolites were screened against the common metabolite list compiled as described in the section 3.6.7. If a pathogen metabolite was detected in the list, it was removed to prevent any distruptions in the host metabolism. Then, currency metabolites were investigated by performing literature search [142]. These metabolites are involved in many reactions in a wide range of different organisms. Therefore, they must be also eliminated to avoid inhibition of general reactions like electron transport in human.

A bit different strategy from the EMFilter approach was executed to evaluate the connectivity of each essential compound [141]. Firstly, essential outgoing reactions were found out by blocking each reaction (one by one) upon the sitoichiometric coefficients. If the inhibition of a reaction considerably reduced the cellular growth (cut-off: 1% of the wild-type growth rate), this reaction was accepted to be essential.

Then, the number of such reactions associated with each essential metabolite was determined.

The last step of the filtering process is homology screening of the genes involved in the outgoing reactions related to each selected metabolite. To this aim, another common list was compiled by comparing the protein sequences of the pathogen with human proteome from the Refseq database [125]. The homologous genes were selected for the E-value (cut-off) $<1x10^{-3}$ and identity $\geq 30\%$ that were added into the 'homologous gene list'. Then, the pathogen metabolites were consumed in the reactions catalyzed by the human-homologous enzymes were eliminated.

The algorithm was developed to determine the number of essential reactions consuming each essential metabolite and to identify the non-homologous genes involved in the outgoing recations associated with these metabolites that is given in Figure 3.7.

```
Input: SBML model of an organism and a special gene-reaction matrix (RGM_ind)
Output: Set single lethal metabolites M_sl
        Set single lethal metabolites R_sl
        Set non-homologous gene matrix D regarding each metabolite

(1) % Do FBA to find maximum growth rate and set the cut-off
    V_bio = max(v_growth);
    cut-off = V_bio x 0.01;

(2) % Find essential metabolites
    for each metabolite i do
        Set all rate coefficients regarding metabolite i in S matrix to 0
        % Do FBA to find growth rate for metabolite deletion
        V_bioKO = max(v_growth);
        if V_bioKO < cut-off then
            Assign the detected genes to M_sl matrix
        end if
    end for

(3) % Find essential reactions
    for each reaction i do
        Set all rate coefficients regarding reaction i in S matrix to 0
        % Do FBA to find growth rate for reaction deletion
        V_bioKO = max(v_growth);
        if V_bioKO < cut-off then
            Assign the detected genes to R_sl matrix
        end if
    end for

(4) % Determine the number of essential outgoing reactions related to each essential metabolite
    for each essential metabolite i do
        for each essential reaction j do
            if S{i,j} < 0 then
                increase relation_count{i} 1
            end if
        end for
    end for

(5) % Find non-homologous gene list related to each essential metabolite
    for each essential metabolite i do
        for each reaction j do
            if S{i,j} ~= 0 then
                % RGM_ind: a matrix includes only indices of gene related to
                % each reaction in RGM
                for each gene k in RGM_ind{j,:}
                    % HGL: human-homologous gene list
                    if each gene{k} in HGL then
                        Add each homologous gene and related essential metabolite i to H matrix
                    else
                        Add each non-homologous gene and essential metabolite i to NH matrix
                    end if
                end for
            end if
        end for
    end for

Set D = difference of all genes in NH matrix and all genes in H matrix
```

Figure 3.7: The algorithm to determine the non-homologous genes and the number of essential outgoing reactions associated with each essential metabolite.

4. RESULTS

4.1. Validation of The Models

Herein, two GEMs of *K. pneumoniae* MGH 78578 (iYL1228) and the more virulent strain *K. pneumoniae* strain KPPR1 (iKp1289) were used to examine the metabolic profiles of these bacteria via investigation of growth profiles of the strains in different media. In addition, accuracy of the models was assessed based on the agreement of model predictions with the *in silico* and *in vitro* growth data under the selected growth media given in the articles [46, 74].

4.1.1. Validation of the Models Based on Growth Phenotypes on Different Carbon Sources

To examine the growth phenotypes of the *Klebsiella* strains, minimal media supplied by different carbon sources were used. The growth in the minimal media supplied with a carbon source (i.e., acetate, citrate, D-xylose, D-glucose, glycerol, L-lactate, L-malate, *myo*-inositol or gluconate) was simulated at the specified uptake rates of carbon sources and oxygen for MGH 78578 strain [74]. In addition, the growth of the time-dependent adapted strain in M9 medium with *myo*-inositol was simulated. The growth rates predicted via FBA and flexoFBA are compatible with those seen in the article [74] (Table 4.1).

Table 4.1: Examination of *in silico* growth phenotypes (growth rates (1/h)) of *K. pneumoniae* MGH 78578 for different carbon sources (Exptl: experimental).

		Growth rate (1/h)			
Carbon source uptake rate (mmol/gDW/h)	**Oxygen uptake rate (mmol/gDW/h)**	**Experimental** (Article)	***In silico*[b]** (Article)	***In silico*[b]** (This study)	***In silico*[c]** (This study)
Acetate (14.291)	14.657	0.293	0.355	0.355	0.346
Citrate (14.017)	21.837	0.570	0.937	0.939	0.914
D-Xylose (6.006)	11.229	0.481	0.479	0.479	0.466
D-Gluconate (17.909)	21.837	0.965	1.264	1.263	1.226

Table 4.1: Continued.

Carbon source uptake rate (mmol/gDW/h)	Oxygen uptake rate (mmol/gDW/h)	Growth rate (1/h)			
		Experimental (Article)	*In silico*[b] (Article)	*In silico*[b] (This study)	*In silico*[c] (This study)
D-Glucose (10.457)	21.744	1.084	1.040	1.040	1.005
Glycerol (10.609)	13.618	0.804	0.599	0.599	0.580
L-Lactate (22.686)	21.837	0.658	0.655	0.658	0.644
L-Lactate (22.686)	21.837	0.658	0.655	0.658	0.644
L-Malate (34.572)	21.837	0.834	1.053	1.054	1.023
myo-Inositol (13.802)	21.837	0.570	1.029	1.029	0.997
myo-Inositol[a] (11.024)	21.837	0.760	0.941	0.940	0.914

[a] Adaptive growth of *K. pneumoniae* MGH 78578 on myo-inositol.

[b] Growth rates predicted using FBA.

[c] Growth rates predicted using flexoFBA.

The same simulations were repeated for the KPPR1 strain. This more virulent pathogen exhibits slightly more efficient metabolism (Table 4.2) as it exhibited higher growth rates in different media despite having the same biomass equation with the MGH 78578 strain.

Table 4.2: Comparison of the growth rates (1/h) of the different *K. pneumoniae* strains in different carbon sources.

Carbon source uptake rate (mmol/gDW/h)	Oxygen uptake rate (mmol/gDW/h)	Growth rate (1/h)	
		In silico[a] (This study)	*In silico*[b] (This study)
Acetate (14.291)	14.657	0.355	0.432
Citrate (14.017)	21.837	0.939	1.072
D-Xylose (6.006)	11.229	0.479	0.570
D-Gluconate (17.909)	21.837	1.263	-

Table 4.2: Continued.

Carbon source uptake rate (mmol/gDW/h)	**Oxygen uptake rate (mmol/gDW/h)**	**Growth rate (1/h)**	
		***In silico*[a]** (This study)	***In silico*[b]** (This study)
D-Glucose (10.457)	21.744	1.040	1.140
Glycerol (10.609)	13.618	0.599	0.692
L-Lactate (22.686)	21.837	0.658	0.884
L-Malate (34.572)	21.837	1.054	1.203
myo-Inositol (13.802)	21.837	1.029	-

[a] Growth rates of *K. pneumoniae* MGH 78578 on different carbon sources.

[b] Growth rates of *K. pneumoniae* KPPR1 on different carbon sources.

To understand whether the differences in the growth rates are due to differences in the minimal growth media (M9 for MGH 78578 and CDG for KPPR1 (Table A.1.1)), medium compositions were compared. The uptake of only two metabolites are different in the models. Ni^{+2} is available in only CDG medium and tungstate consumption is allowed in only iYL1228. However, it was revealed that the reason underlying different bacterial growth rates was not associated with the differences in the medium compositions. Closer inspection of the growth rates calculated by the two models revealed a bit higher uptake rates of some metabolites (e.g., iron) in both media. Therefore, the reason behind the higher growth rate of KPPR1 may be the ability to assimilate particular metabolites.

Effect of another set of carbon sources [46] on the bacterial growth was also investigated. Growth profiles of the strains were validated by a comparison of the *in silico* predicted growth rates with the computationally and experimentally calculated growth rates in the article [46] (Table 4.3). The growth rate of KPPR1 in carbon-L-lysine medium found in this study was inconsistent with the result given in the article. Here, it was found that this strain could not grow in L-lysine. However, its growth rate was reported as 0.4262 h^{-1} in the article. Interestingly, other carbon sources in Table 4.3 were correctly simulated, leading to the same growth rates reported in the article [46].

As for the growth profile of the MGH 78578 strain, the same growth rates were predicted with the article [46] apart from the simulations of two growth media (i.e., carbon-succinic-acid medium and carbon-dulcitol medium). The simulation results are listed in Table 4.3. The table presents a comparison of the growth rates in the article (found through the KBase platform and the an experimental approach) with *in silico* predicted growth rates in this study.

Table 4.3: Comparison of the growth rates (1/h) predicted in iKp1289 and iYL1228 simulations with the article and each other.

Minimal Media	**MGH 78578 strain**			**KPPR1 strain**		
Carbon source	***In silico*[a] Article**	***In silico*[b] This study**	***In vitro*[c] Article**	***In silico*[a] Article**	***In silico*[b] This study**	***In vitro*[c] Article**
None	-	-	-	-	-	-
Carbon-a-D-Glucose	0.5113	0.5113	+	0.5155	0.5155	+
Carbon-L-Arabinose	0.4212	0.4212	+	0.5155	0.5155	+
Carbon-L-Serine	0.1944	0.1944	+	0.2944	0.2944	+
Carbon-L-Proline	0.3387	0.3387	+	0.4478	0.4478	+
Carbon-L-Aspartic-Acid	0.2426	0.2426	+	0.3373	0.3373	+
Carbon-L-Arginine	0.4216	0.4216	+	0.5066	0.5066	+
Carbon-L-Alanine	0.2305	0.2305	+	0.3313	0.3313	+
Carbon-L-Asparagine	0.2428	0.2428	+	0.3376	0.3376	+
Carbon-L-Lysine	0.0000	0.0000	-	0.4262	0.0000	-
Carbon-Succinic-Acid	0.2575	0.0000	+	0.3524	0.3524	+
Carbon-Dulcitol	0.5155	0.5483	+	0.0000	0.0000	+

[a] Growth rates of MGH 78578 and KPPR1 strains predicted using FBA in the
[b] Growth rates of MGH 78578 and KPPR1 strains predicted using FBA in this
[c] Experimentally calculated growth rates of MGH 78578 and KPPR1 strains.

In summary, *in silico* predicted growth rates are compatible with the experimental and computational results reported in the articles [46, 74]. Whole results support the better proliferation capacity of KPPR1 in comparison with MGH 78578.

4.1.2. Validation of the Models Based on Gene Essentiality Analysis

Metabolic profiles of the MGH 78578 and KPPR1 strains were also examined through elucidation of the essential genes in different growth conditions. *In silico* gene knockout results were validated by comparing the predicted results with those given in the articles [46, 74].

For MGH 78578, single gene deletions were performed for *in silico* growth of MGH 78578 under aerobic minimal glucose growth medium (M9). The glucose uptake rate being in the range of 0-10 mmol/gDW/h and the oxygen uptake rate being in the range of 0-20 mmol/gDW/h were used based on the article [74]. Consistent with the article [46], KPPR1 growth was simulated in LB medium (Table A1.2).

The gene essentiality analysis using iYL1228 resulted in the identification of 119 essential genes while 118 essential genes were documented in the article. On the other hand, when Fast-SL algorithm was used, the same 119 genes were identified via *in silico* single gene deletions, further confirming the predictions in this study. The difference may be due to the use of a different solver in this study. GO Enrichment Analysis was also used to characterize the essential genes. It should be noted that, gene annotations of two closely related microorganisms (*E. coli* and *Salmonella*) were used for the GO Enrichment Analysis owing to the lack of *K. pneumoniae* in AmiGO website [143]. Analysis of the biological processes demonstrated that the essential genes found in the study are mainly responsible for fatty acid/lipid A biosynthesis, glutamine metabolic process, leucine/arginine/lysine, biosynthetic process, peptidoglycan production process and cell cycle. Moreover, 8 essential genes were reported as the specific to *K. pneumoniae* (i.e., KPN_02202, KPN_02492, KPN_02493, KPN_03963, KPN_01515, KPN_01093, KPN_04659 and KPN_02834) in comparison with *E. coli* (iAF1260) and *Salmonella* (iRR1083 [144]) in the article [74]. These genes were successfully predicted in this study. A further analyis revealed that the genes are mainly associated with lipopolysaccharide biosynthesis, capsule production and lipid biosynthesis.

As for the model of iKp1289, 58 essential genes (including one unknown gene) were identified through *in silico* gene deletions. The 57 annotated essential genes found in this study are the same with the 57 essential genes in the article. Thus, essential genes were successfully predicted via the single gene deletion algorithm developed in this study. Investigation of the biological processes regarding these essential genes revealed that they are mostly found in the processes supporting the virulence of the bacteria (e.g., extracellular polysaccharide biosynthesis, lipid A biosynthesis, regulation of cell shape and peptidoglycan production).

Taken together, the prediction of essential genes in this study was found to be completely compatible with the published results.

4.2. Identification of Essential Genes for Gene-Centric Drug Target Discovery

It is an expensive and time-consuming task to determine the essential genes experimentally due to the large and complex nature of the high-throughput techniques. On the other hand, this information is crucial to understand the cellular metabolism and to discover novel drug targets. *In silico* deletion of each gene provides comprehensive information on gene essentiality in a time and cost-saving manner. Hence, *in silico* gene essentiality analysis was used to identify the putative drug targets in this work.

Essentiality of each gene was evaluated by simulating growth of the in two host-mimicking media (SM medium (Table B1.1 and Table B1.2) and HBF medium (Table B1.3 and Table B1.4). A gene was considered as essential if its deletion resulted in the growth rate less than a specified cut-off (1% of the maximum WT growth rate). Three different methods (i.e., an in-house function, Fast-SL algorithm, a COBRA Toolbox function) were applied comparatively to perform the single gene deletion.

50 essential genes were identified in growth simulations of *K. pneumoniae* MGH 78578 in the HBF-mimicking condition by own developed algorithm while 51 genes were detected to be essential using COBRA Toolbox and Fast-SL algorithm for the same cut-off (Figure 4.1.a). Of these 51 genes, 50 genes were found by all three algorithms whereas KPN_01284 (*fabI*) was identified by COBRA Toolbox function and Fast-SL algorithm. Therefore, the essentiality of *fabI* was investigated in detail. Using the reaction-gene matrix (RGM), 24 reactions associated with this gene were

detected. All of these reactions are related to enoyl-[acyl-carrier-protein] (ACP) reductase (NADPH) activity necessary for the last step of the fatty acid elongation [145]. This process is reasonable for drug targeting, and its deletion was reported to be lethal for two closely related microorganisms including *S. enterica* sv. Typhimurium [146] and *E. coli* [147] in LB medium. For the further insight into the essentiality of this gene, the reactions associated with *fabI* were specifically blocked. However, the growth rate of *K. pneumoniae* MGH 78578 was not changed in HBF medium when flux boundaries of these 24 reactions were set to zero concurrently. Differently from HBF simulation, *fabI* was found to be essential in SM medium by all methods used in this study. These findings highlight the importance of the simulated condition to determine the gene essentiality.

Similarly, most of the essential genes determined by the three methods were the same for the growth simulation of *K. pneumoniae* MGH 78578 in SM medium. 76 common essential genes were predicted. Of these genes, KPN_00983 (*fabA*) was identified to be essential via our own algorithm and Fast-SL algorithm but not by COBRA Toolbox. On the other hand, KPN_00193 (*fabZ*) was predicted to be essential by only COBRA Toolbox (Figure 4.1.b).

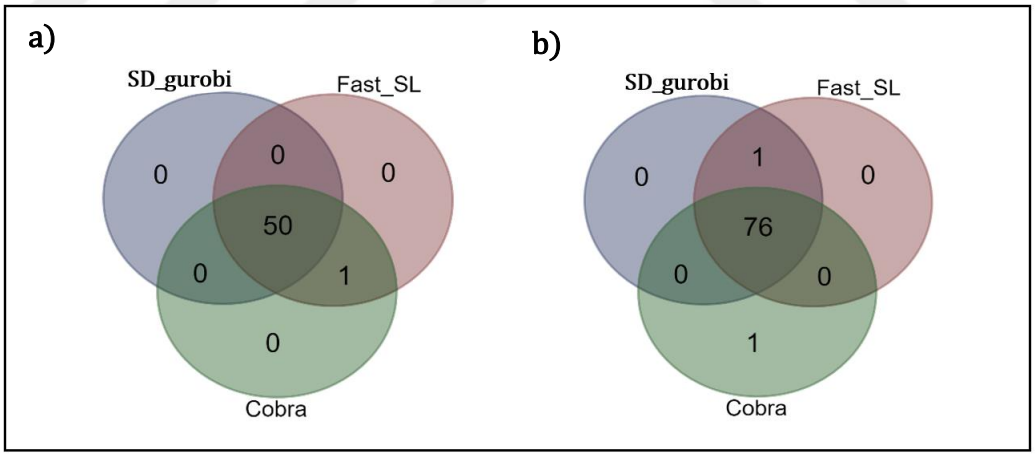

Figure 4.1: Comparison of the number of results from different methods for growth simulations in a) HBF and b) SM (SD_gurobi: in-house algorithm).

Both FabA and its homologue known as FabZ function in fatty acid synthesis by catalyzing the dehydration of β-hydroxyacyl-AcpP in Gram-negative bacteria (Figure 4.2) [148–150]. FabA (β-hydroxyldecanoyl-ACP dehydratase/isomerase) also catalyzes synthesis of unsaturated fatty acids in many bacteria [148]. Despite relatively limited presence of *fabA* genes among bacteria in comparison with FabZ (indicating

FabZ is a more convenient target for broad spectrum antibiotics), FabA may be also a promising drug target [150, 151]. This is because FabA can replace FabZ with its similar substrate binding pocket but not vice versa. Furthermore, lack of any isozymes for FabA is an increasing evidence that supports essentiality of this enzyme [150]. Notably, this phenomenon has also already taken into consideration as designing the GEM of iYL1228. Both FabA and FabZ are responsible for dehydration of beta-hydroxyacyl acyl carrier protein (ACP) to trans 2-enoyl ACP but only FabA can catalyze the conversion of trans-2-decenoyl-ACP to cis-3-dodecenoyl-ACP [148, 150]. Therefore, iYL1228 includes 12 reactions related to *fabA* gene, and 11 out of them were found to be associated with 3-hydroxyacyl-[acyl-carrier-protein] dehydratase activity. These 11 reactions are active in the presence of either FabA or FabZ. Therefore, inhibition of one of them did not affect the bacterial growth. On the other hand, the bacterial growth was not observed due to lack of trans-2-decenoyl-ACP isomerase activity when *fabA* gene was deleted.

To ensure the fabZ is not essential for the bacterial growth in SM medium, effect of deletion of the associated reactions on the bacterial growth was also investigated. Only the previously defined 11 reactions (dehydration of beta-hydroxyacyl acyl carrier protein (ACP) to trans 2-enoyl ACP) were found to be related to FabZ and thus deletion of *fabZ* gene was compensated with FabA activity.

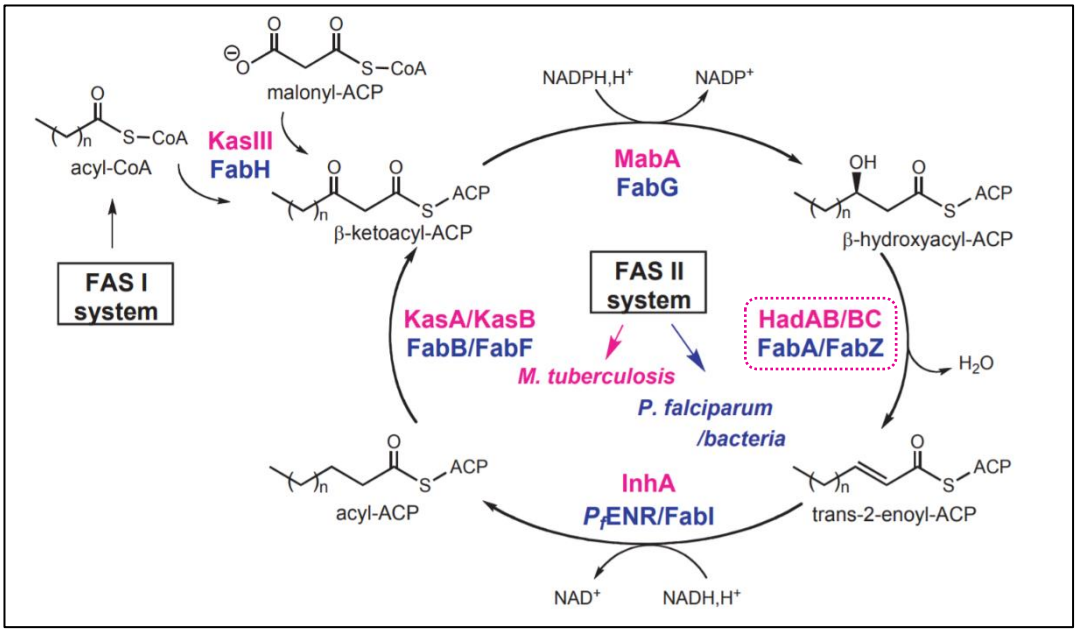

Figure 4.2: Type II fatty acid biosynthesis pathway in various bacteria.

These findings demonstrated that the COBRA Toolbox function or Fast-SL algorithm can report erroneous results in some cases probably by ignoring the compensatory function of the isoenzymes (overlapping functions). Therefore, gene deletion simulations of *K. pneumoniae* KPPR1 growth were performed in host-mimicking medium using only the in-house developed algorithm. 75 essential genes in addition to 1 unannotated gene were predicted for SM medium whilst 49 essential genes and 1 unannotated gene were revealed through the growth simulations in HBF.

4.3. Drug Target Prioritization for the Gene-Centric Candidates

Some of the *in silico* identified drug targets are not convenient candidates to take into consideration for medical use. For instance, a human-homologous gene should be eliminated from the list in order to avoid a possible side effect due to binding of the drug to off-targets in the host cell. Moreover, some of the identified targets may not be convenient for drug-binding [122]. Taken together, further criteria must be applied to evaluate and filter the putative drug targets.

4.3.1. Identification of Essential Non-homologous Pathogen Genes

Homology search was conducted in the first step of drug prioritization to minimize/prevent any possible damages to the human cells. This approach was performed by comparing the amino acid sequences of the pathogen proteins (predicted as essential) with human proteome. 50 genes of MGH 78578 and 49 genes of KPPR1 were predicted to be essential in the HBF simulation (section 4.2). 33 out of these essential genes were found to be non-homologous for MGH 78578 strain while a significant sequence similarity was not found for 34 essential genes of KPPR1 strain. These genes were compared based on their names. Four non-homologous, unannotated genes of KPPR1 were excluded from Figure 4.3. When the remaining genes were compared, 26 common genes with no human homologs were identified. In addition, one common gene (*asd*/*asd2*) was erroneously excluded from the common gene list due to the difference in the nomenclatures. Thus, three non-human homologous genes (*fabD*, *yaeD* and *lpxH*) were determined to be essential only for the KPPR1 strain and six non-human homologous genes (*murE*, *thrB*, *murF*, *uge*, *gmhB* and *ybbF*) were found as essential for MGH 78578 strain (Figure 4.3.a). The same comparison was

performed for the growth simulation in SM medium. 39 out of 43 non-human homologous genes identified through iKp1289 are the annotated genes (with no gene names). This gene list was compared with 42 genes determined via iYL1228 (Figure 4.3.b). The same bacteria-specific genes were detected. Consequently, most of the essential, non-homologous genes are common for both strains. This finding highlights that these pathogenic strains have a similar metabolism.

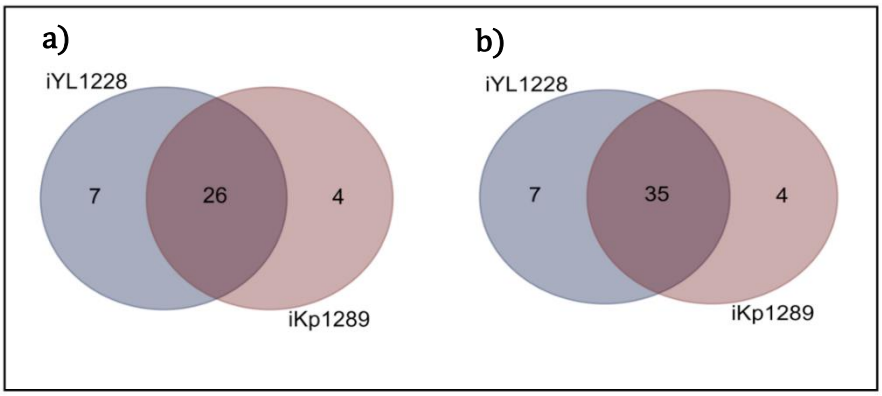

Figure 4.3: Comparison of the number of non-homologous, essential genes in *K. pneumoniae strains* with respect to human proteome, predicted through different simulations: a) HBF simulation and b) SM simulation.

4.3.2. Analysis of Subcellular localizations for Non-homologous Gene Products

Subcellular localization is useful in identification of the protein function and also in determination of the potential drug targets and vaccine candidates [127]. Therefore, the second filtering process was used to search localization of the proteins encoded by the *in silico* predicted essential genes with no homology to human proteins.

Three web services including CELLO, PSORTb and iLoc-Gneg were used to determine the subcellular localizations of non-homologous gene products. If at least two of the services lead to the same prediction for a gene product, this location was accepted. The uncertain results were validated through a literature survey and UniProt database [131]. Apart from the products of *uge* and *murI* genes, a subcellular location was assigned to each gene product. About 92% and 94% of the gene products of the strain MGH 78578 from the growth simulations in SM and HBF were found to be localized in the cytoplasm (Figure 4.4). On the other hand, products of three genes (*fabI*, *mraY* and *murG*) identified through SM simulation and products of two genes

from HBF simulation (excluding *fabI*) were found to be localized in the inner membrane.

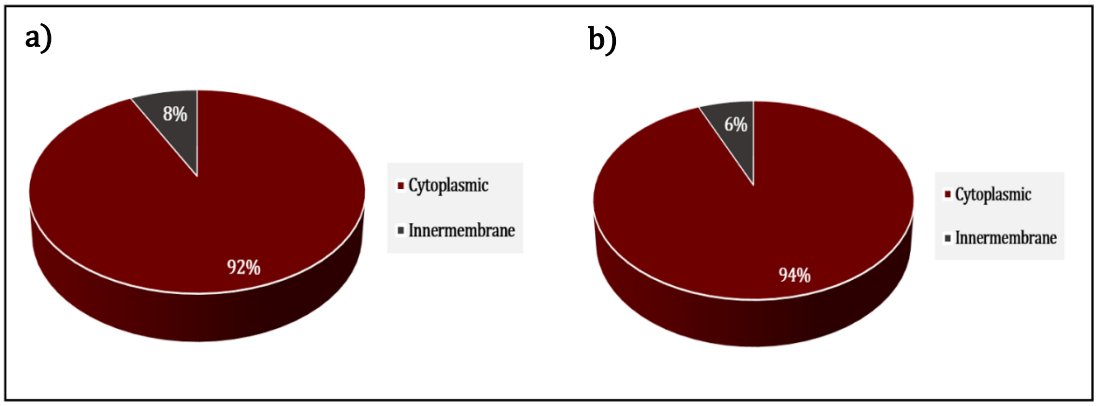

Figure 4.4: Subcellular localizations of the essential, non-human homologous gene products of the strain MGH 78578 : a) SM simulation and b) HBF simulation.

In the similar vein, the cellular components harboring the essential non-homologous gene products of the strain KPPR1 were investigated. High amount of the gene products (almost 88% and 87%) from the growth simulations in SM and HBF were detected to be localized in the cytoplasm (Figure 4.5). On the other hand, five genes (*fabI*, *mraY*, *lpxK*, *murG* as well as a gene encoding 1-acylglycerol-3-phosphate O-acyltransferases domain protein) identified through SM simulation were found be localized in the inner membrane. Four out of these genes (apart from *fabI*) were found in the HBF simulation, as well.

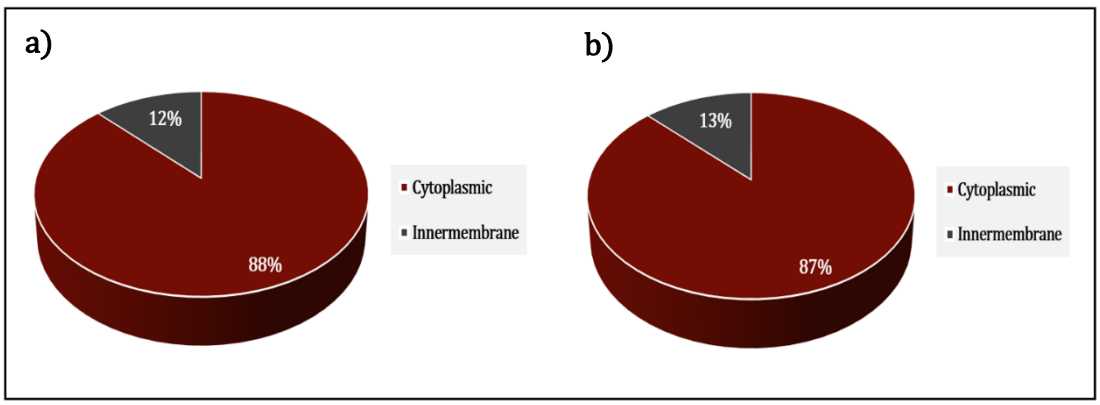

Figure 4.5: Subcellular localizations of the essential, non-human homologous gene products of the strain KPPR1: a) SM simulation and b) HBF simulation.

Mostly cytoplasmic essential proteins were detected in this study. It is important to note that both cytoplasmic and inner membrane proteins can be putative drug targets. On the other hand, potential vaccine candidates are selected among extracellular, the outer membrane and periplasmic proteins [152]. However, any vaccine candidates were not identified in the study.

4.3.3. Investigation of Possible Antibiotic Resistance Genes

Next step is to search whether there are any known or putative antibiotic resistance genes in the non-homologous gene list in order to target antibiotic resistance mechanisms. ARG-ANNOT database [119] includes the sequences of the resistance genes against different antibiotic classes (i.e., macrolide-lincosamide-streptogramin, aminoglycosides, fosfomycin, fluoroquinolones, sulfonamides, glycopeptides, trimethoprim, rifampicin, tetracyclines, phenicols and trimethoprim and beta-lactamases).

ARG-ANNNOT along with Bioedit was used in this work. Only *glmU* gene of was identified as essential in SM and HBF simulations for both strains that exhibited a significantly similarity with resistance-related genes. N-Acetylglucosamine-1-phosphate uridyltransferase (GlmU) is a bifunctional enzyme with uridyltransferase and acetyltransferase activities in both gram-positive and gram-negative bacteria [153, 154]. It plays an important role in the peptidoglycan synthesis and in production of the lipopolysaccharides by catalyzing the UDP-N-acetylglucosamine formation. LPS is indispensable for growth of almost all known bacteria. Considering vital functions of GlmU on the membrane integrity, inhibition of this protein exhibits a bactericidal effect [155]. Therefore, there are significant efforts to design inhibitors againts this protein for different pathogens [154, 155].

This promising non-homologous gene was found as significantly similar to the resistance genes against macrolide-lincosamide-streptogramin and phenicols. However, it has an unacceptable alignment length. Besides, these antibiotic classes are responsible for the inhibition of protein synthesis and a direct relationship of these antibiotics with GlmU protein could not be found in the literature searches. However, an indirect link may be exist. Petráčková and colleagues reported the *glmU* overexpression after 43 h in *E. coli* cells continuously grown in the presence of erythromycin (a member of macrolide antibiotics) [156]. Moreover, such genes can

confer resistance to the antibiotics damaging the cell wall. They can contribute to the reconstruction of the wall even if this hypothesis could not be proved through the ARG-ANNOT database.

4.3.4. Examination of Virulence Gene Profiles

There is an interplay between antibiotic resistance and virulence, and they can be regulated by common global transcriptional regulators [39, 40]. Thus, inhibition of the putative targets associated with both survival and virulence harms the pathogens in many ways. Therefore, the virulence genes of MHG 78578 and KPPR1 were examined in detail.

Virulence-related genes within the non-homologous gene lists were identified through BLAST against VFDB core dataset. To start with, the non-homologous, essential genes from SM simulation of iYL1228 were analyzed. Six genes (*hldD* (*rfaD*), *kdsA*, *gmhA (lpcA*), *lpxD*, *lpxA* and *lpxC*) which were significantly similar to the virulence genes of *Haemophilus influenzae* Rd KW20 were identified. Those six genes were also found to support virulent characteristics of the strain KPPR1 (Table 4.4). This analysis was not repeated for the gene list from HBF simulation because essential, non-homologous gene list from HBF simulations is only a subset of the list of genes from SM simulations.

The identified six genes are associated with lipopolysaccharide synthesis. For example, the *rfaD* gene encoding ADP-L-glycero-D-mannoheptose-6-epimerase catalyzes the generation of ADP-L-glycero-D-manno-heptose through ADP-L-glycero-D-manno-heptose synthetic pathway. This component is crucial for the synthesis of core region of LPS. LPS is the major outer membrane component supporting the outer-leaflet integrity and protection against antimicrobial molecules so *rfaD* product may be a promising drug target [157]. Recently, it was reported that enterohemorrhagic *E. coli* (EHEC) O157:H7 RfaD contributes to the virulence of bacteria via mainly promoting colonization and resistance against some antimicrobial peptides. Thus, RfaD was suggested as a potential drug target [158]. The *lpxA*, *lpxC*, and *lpxD* are other essential genes crucial for the lipid A biosynthesis, and their deletions increase the sensivity to hydrophobic antibiotics (e.g., erythromycin) [159]. Thus, the virulence factors identified in this study were proposed as putative drug targets (Table 4.4).

Table 4.4: Virulence factors identified in both GEMs.

Query (Tag ID)	Gene	VFDB ID	Description	Score	Expect	Identity (%)
VK055_3508 KPN_03963	*rfaD hldD*	VFG000332	(rfaD) ADP-L-glycero-D-mannoheptose-6-epimerase [LOS]	504 bits (1297)	1,00E-144	241/308 (78%)
VK055_0220 KPN_02230	*kdsA*	VFG013465	(kdsA) 2-dehydro-3-deoxyphosphooctonate aldolase [LOS]	480 bits (1235)	1,00E-137	235/283 (83%)
VK055_2330 KPN_00236	*gmhA*	VFG013418	(gmhA/lpcA) phosphoheptose isomerase [LOS]	309 bits (792)	6,00E-86	152/192 (79%)
VK055_2373 KPN_00192	*lpxD*	VFG013374	(lpxD) UDP-3-O-(3-hydroxymyristoyl) glucosamine N-acyltransferase[LOS]	438 bits (1127)	1,00E-124	214/324 (66%)
VK055_2371 KPN_00194	*lpxA*	VFG013390	(lpxA) UDP-N-acetylglucosamine acyltransferase [LOS]	366 bits (940)	1,00E-103	176/262 (67%)
VK055_2471 KPN_00100	*lpxC*	VFG013412	(lpxC) UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosaminedeacetylase [LOS]	483 bits (1243)	1,00E-138	236/304 (77%)

4.3.5. Evaluation of Therapeutic Targets for Druggability

Considering the fact that all protein structures are not convenient for drug-binding, it is crucial to determine the most efficient drug targets [122]. DrugBank database was used to identify druggability properties of the gene products and to reveal the existing drugs targeting these proteins. Evaluation of the existing drugs is a promising strategy that reduces an enormous time and cost. Therefore, it is a promising approach in the drug discovery. Approximately 10 to 17 year process (target discovery and validation: 2-3 years, screening or development of biologically active compounds: 0.5-1 year, candidate optimization: 1-3 years, pre-clinical trials including ADMET evaluation through animal models, clinical trials: 5-6 years and approval: 1-2 years) is necessary in a drug approval pipeline. The cost of all the processes for a new successful drug is almost US$1.78 billion [160]. On the other hand, the efficacies or toxicities of existing drugs were priorly known. Therefore, the existing drugs can be screened in order to detect the drugs having a potential to bind to the products of the putative targets identified by *in silico* simulations of GEMs.

21 common genes out of 31 potential druggable non-homologous targets were identified in KPPR1 and MGH 78578 for E-value cut-off of $1x10^{-25}$. List of the targets and interacting drugs are presented in Table 4.5. These genes are mostly related to pyrimidine synthesis, cell wall structure and fatty acid synthesis. They may be assessed as drug targets for the treatment of the *K. pneumoniae* infection in future. By supporting this assumption, a mutation in *uge* gene encoding UDP galacturonate 4-epimerase negatively affects the virulent characteristics and colonization ability of *K. pneumoniae*. Thus, the mutant cells cannot induce virulence in the septicemia and pneumonia animal models as well as urinary tract infections in rats [161].

Table 4.5: Evaluation of druggability properties of the putative targets and drug repositioning.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
VK055_5030	-	1	Uridine-Diphosphate-N-Acetylgalactosamine	experimental
VK055_2483	-	3	Uridine-5'-Diphosphate-N-Acetylmuramoyl-L-Alanine-D-Glutamate	experimental
			2,6-Diaminopimelic Acid	experimental
			Lysine Nz-Carboxylic Acid	experimental
VK055_2482	-	1	2-CHLORO-N-(3-CYANO-5,6-DIHYDRO-4H-CYCLOPENTA[B]THIOPHEN-2-YL)-5-DIETHYLSULFAMOYL-BENZAMIDE	experimental
VK055_3672 KPN_03799	*asd2 asd*	3	2'-Monophosphoadenosine 5'-Diphosphoribose	experimental
			(4s)-4-{[(2s)-2-Amino-3-Oxopropyl]Sulfanyl}-L-Homoserinate	experimental
			Aspartate Semialdehyde	experimental
VK055_4699 KPN_02812	*dapA*	1	Nz-(1-Carboxyethyl)-Lysine	experimental
VK055_2536 KPN_00039	*dapB*	2	3-Acetylpyridine Adenine Dinucleotide	experimental
			Dipicolinic Acid	experimental
VK055_2386 KPN_00179	*dapD*	5	Pimelic Acid	experimental
			Coenzyme A	investigational, nutraceutical
			2-Aminopimelic Acid	experimental
			Succinyl-Coenzyme A	experimental
			Succinamide-Coa	experimental
VK055_1504 KPN_00983	*fabA*	2	2-Decenoyl N-Acetyl Cysteamine	experimental
			2-Decenoyl N-Acetyl Cysteamine	experimental
VK055_1375	*fabD*	1	3,6,9,12,15-PENTAOXAHEPTADECAN-1-OL	experimental

Table 4.5: Continued.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
VK055_1165 KPN_01284	*fabI*	24	Indole Naphthyridinone	experimental
			3-(6-Aminopyridin-3-Yl)-N-Methyl-N-[(1-Methyl-1h-Indol-2-Yl)Methyl]Acrylamide	experimental
			4-(2-Thienyl)-1-(4-*Methylbenzyl*)-1h-Imidazole	experimental
			3-[(Acetyl-Methyl-Amino)-Methyl]-4-Amino-N-Methyl-N-(1-Methyl-1h-Indol-2-Ylmethyl)-Benzamide	experimental
			1,3,4,9-Tetrahydro-2-(Hydroxybenzoyl)-9-[(4-Hydroxyphenyl)Methyl]-6-Methoxy-2h-Pyrido[3,4-B]Indole	experimental
			Beta-D-Glucose	experimental
			2-(TOLUENE-4-SULFONYL)-2H-BENZO[D][1,2,3]DIAZABORININ-1-OL	experimental
			Triclosan	approved, investigational
			6-METHYL-2(PROPANE-1-SULFONYL)-2H-THIENO[3,2-D][1,2,3]DIAZABORININ-1-OL	experimental
			Triclosan	approved, investigational
			Triclocarban	approved
			Soneclosan	experimental
			Triclosan	approved, investigational
			Triclosan	approved, investigational
			Ethionamide	approved
			Isoniazid	approved, investigational
			C16-Fatty-Acyl-Substrate-Mimic	experimental
			Genz-10850	experimental
			(3S)-N-(3-CHLORO-2-METHYLPHENYL)-1-CYCLOHEXYL-5-OXOPYRROLIDINE-3-CARBOXAMIDE	experimental
			N-(4-METHYLBENZOYL)-4-BENZYLPIPERIDINE	experimental
			(3S)-1-CYCLOHEXYL-5-OXO-N-PHENYLPYRROLIDINE-3-CARBOXAMIDE	experimental
			5-PENTYL-2-PHENOXYPHENOL	experimental
			(3S)-1-CYCLOHEXYL-N-(3,5-DICHLOROPHENYL)-5-OXOPYRROLIDINE-3-CARBOXAMIDE	experimental
			(3S)-N-(3-BROMOPHENYL)-1-CYCLOHEXYL-5-OXOPYRROLIDINE-3-CARBOXAMIDE	experimental
VK055_3340 KPN_04135	*glmU*	5	Coenzyme A	investigational, nutraceutical
			Uridine-Diphosphate-N-Acetylglucosamine	experimental
			2-(N-Morpholino)-Ethanesulfonic Acid	experimental

Table 4.5: Continued.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
			4-chloro-N-(3-methoxypropyl)-N-[(3S)-1-(2-phenylethyl)piperidin-3-yl]benzamide	experimental
			Uridine-Diphosphate-N-Acetylglucosamine	experimental
VK055_2330	*gmhA*	1	D-Glycero-D-Mannopyranose-7-Phosphate	experimental
VK055_0220 KPN_02230	*kdsA*	12	2-phospho-D-glyceric acid	experimental
			Phosphoenolpyruvate	experimental
			Ribose-5-phosphate	experimental
			{[(2,2-Dihydroxy-Ethyl)-(2,3,4,5-Tetrahydroxy-6-Phosphonooxy-Hexyl)-Amino]-Methyl}-Phosphonic Acid	experimental
			1-Deoxy-6-O-Phosphono-1-[(Phosphonomethyl)Amino]-L-Threo-Hexitol	experimental
			2-(Phosphonooxy)Butanoic Acid	experimental
			Arabinose-5-phosphate	experimental
			Erythose-4-Phosphate	experimental
			Phosphoenolpyruvate	experimental
			{[(2,2-Dihydroxy-Ethyl)-(2,3,4,5-Tetrahydroxy-6-Phosphonooxy-Hexyl)-Amino]-Methyl}-Phosphonic Acid	experimental
			3-Fluoro-2-(Phosphonooxy)Propanoic Acid	experimental
			1-Deoxy-Ribofuranose-5'-Phosphate	experimental
KPN_00236	*lpcA*	1	D-Glycero-D-Mannopyranose-7-Phosphate	experimental
VK055_2371 KPN_00194	*lpxA*	2	D-tartaric acid	experimental
			2-HYDROXYMETHYL-6-OCTYLSULFANYL-TETRAHYDRO-PYRAN-3,4,5-TRIOL	experimental
VK055_2471 KPN_00100	*lpxC*	6	Tu-514	experimental
			Palmitoleic Acid	experimental
			3-(heptyloxy)benzoic acid	experimental
			N-{(1S,2R)-2-hydroxy-1-[(hydroxyamino)carbonyl]propyl}-4-{[4-(morpholin-4-ylmethyl)phenyl]ethynyl}benzamide	experimental
			Myristic acid	experimental
			(2R)-N-hydroxy-3-naphthalen-2-yl-2-[(naphthalen-2-ylsulfonyl)amino]propanamide	experimental
VK055_3878 KPN_03599	*murA*	7	Uridine-Diphosphate-N-Acetylglucosamine	experimental
			(S)-2-{Methyl-[2-(Naphthalene-2-Sulfonylamino)-5-(Naphthalene-2-Sulfonyloxy)-Benzoyl]-Amino}-Succinicacid	experimental
			Aminomethylcyclohexane	experimental
			Cyclohexylammonium Ion	experimental
			3'-1-Carboxy-1-Phosphonooxy-Ethoxy-Uridine-Diphosphate-N-Acetylglucosamine	experimental

Table 4.5: Continued.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
			1-Anilino-8-Naphthalene Sulfonate	experimental
			Fosfomycin	approved
VK055_3124 KPN_04350	*murB*	2	Flavin adenine dinucleotide	approved
			(5Z)-3-(4-CHLOROPHENYL)-4-HYDROXY-5-(1-NAPHTHYLMETHYLENE)FURAN-2(5H)-ONE	experimental
VK055_2476 KPN_00095	*murC*	3	Uridine-5'-Diphosphate-N-Acetylmuramoyl-L-Alanine	experimental
			Adenosine-5'-[Beta, Gamma-Methylene]Triphosphate	experimental
			Phosphoaminophosphonic Acid-Adenylate Ester	experimental
VK055_2480 KPN_00092	*murD*	8	Uridine-5'-Diphosphate-N-Acetylmuramoyl-L-Alanine	experimental
			Uridine-5'-Diphosphate-N-Acetylmuramoyl-L-Alanine-D-Glutamate	experimental
			Lysine Nz-Carboxylic Acid	experimental
			N-[(6-butoxynaphthalen-2-yl)sulfonyl]-L-glutamic acid	experimental
			N-[(6-butoxynaphthalen-2-yl)sulfonyl]-D-glutamic acid	experimental
			N-{[6-(PENTYLOXY)NAPHTHALEN-2-YL]SULFONYL}-D-GLUTAMIC ACID	experimental
			N-({6-[(4-CYANOBENZYL)OXY]NAPHTHALEN-2-YL}SULFONYL)-D-GLUTAMIC ACID	experimental
			N-({6-[(4-CYANO-2-FLUOROBENZYL)OXY]NAPHTHALEN-2-YL}SULFONYL)-D-GLUTAMIC ACID	experimental
KPN_00089	*murE*	3	Uridine-5'-Diphosphate-N-Acetylmuramoyl-L-Alanine-D-Glutamate	experimental
			2,6-Diaminopimelic Acid	experimental
			Lysine Nz-Carboxylic Acid	experimental
KPN_00090	*murF*	1	2-CHLORO-N-(3-CYANO-5,6-DIHYDRO-4H-CYCLOPENTA[B]THIOPHEN-2-YL)-5-DIETHYLSULFAMOYL-BENZAMIDE	experimental
VK055_2477 KPN_00094	*murG*	1	Uridine-Diphosphate-N-Acetylgalactosamine	experimental
VK055_2069 KPN_00478	*purE*	1	Citric Acid	approved, nutraceutical, vet_approved
VK055_1392 KPN_01074	*pyrC*	4	Dihydroorotic Acid	experimental
			Orotic Acid	experimental, investigational
			Lysine Nz-Carboxylic Acid	experimental
			N-Carbamoyl-L-Aspartate	experimental

Table 4.5: Continued.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
VK055_3486 KPN_03983	*pyrE*	2	Alpha-Phosphoribosylpyrophosphoric Acid	approved, experimental, investigational
			Orotic Acid	experimental, investigational
VK055_1174 KPN_01277	*pyrF*	3	6-Hydroxyuridine-5'-Phosphate	experimental
			1-(5'-Phospho-Beta-D-Ribofuranosyl)Barbituric Acid	experimental
			Uridine monophosphate	experimental
VK055_3508 KPN_03963	*rfaD hldD*	2	Adenosine-5'-Monophosphate Glucopyranosyl-Monophosphate Ester	experimental
			2'-Monophosphoadenosine 5'-Diphosphoribose	experimental
VK055_1369 KPN_01096	*tmk*	1	P1-(5'-Adenosyl)P5-(5'-Thymidyl)Pentaphosphate	experimental
VK055_5029 KPN_02493	*ugd*	4	Sucrose	approved, experimental, investigational
			Guanosine 5'-(Trihydrogen Diphosphate), P'-D-Mannopyranosyl Ester	experimental
			Udp-Alpha-D-Xylopyranose	experimental
			UDP-alpha-D-glucuronic acid	experimental
KPN_02492	*uge*	2	Uridine-Diphosphate-N-Acetylgalactosamine	experimental
			Uridine diphosphate glucose	experimental

4.3.6. Broad Spectrum Analysis of Druggable Virulence Factors

Six unique virulence factors were proposed as drug targets in the section 4.3.4. The virulence factors found in the druggable, non-homologous gene list (Table 4.5) were determined to evaluate the druggability of the virulence factors.

It should be noted that, *hldD* gene was formerly known as *rfaD* similar to *gmhA* and *lpcA* genes. Therefore, five unique cytosolic virulence factors (*rfaD* (*hldD*), *lpxA*, *gmhA* (*lpcA*), *lpxC* and *ksdA*) were found to be druggable although seven druggable virulence factors are demonstrated in Figure 4.6. Any drug molecules targeting the *lpxD* gene were not identified for the given cut-off in this study. Therefore, a less strict E-value cut-off may be selected in DrugBank to extend the non-homologous druggable gene list.

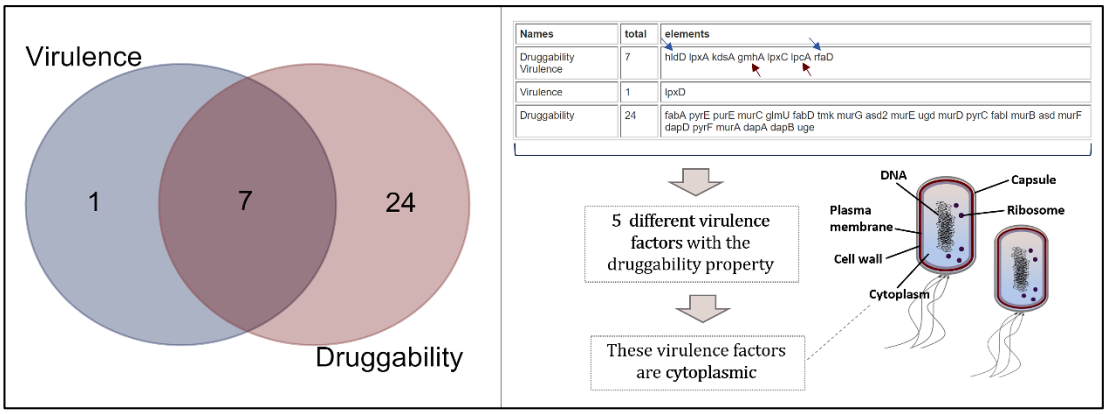

Figure 4.6: Druggable virulence factors (red and blue arrows indicate the same genes).

At least one drug molecule was found through screening all virulence-related genes apart from *lpxD*. Thus, five unique druggable virulence factors were selected for a detailed analysis. In this context, prevalance of these essential genes among other pathogens was investigated using PBIT tool [122]. Broad distribution of a gene is particularly prominent to handle co-infections/multiple infections and to target the conserved genes for delaying the development of antibiotic resistance.

These five druggable, virulence factors were revealed to be involved in the LPS synthesis process. LPS is the outer membrane component surrounding the bacterial cell (Figure 2.6). It contains three main layers including O-antigen, core oligosaccharide and lipid A. Lipid A is the inner layer of LPS that is crucial for both protection and survival of the bacteria [11, 39]. The core oligosaccharide region of LPS consists of inner core (association of 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) and heptose residues) and outer core (association of hexoses and 2-acetoamido-2-deoxy-hexose residues). The virulence-related genes, *gmhA/lpcA* and *hldD/rfaA* are crucial in the synthesis of the heptose precursors while *ksdA* is responsible for the production of the Kdo. Lipid A and core oligosaccharides are linked by Kdo [162]. In addition, the *lpxA* and *lpxC* genes function in lipid A biosynthesis [159].

Broad spectrum analysis has resulted in 26 matches belonging from different pathogens (8 distinct genera except for *Klebsiella*) for the virulence factor *lpxC* (locus tags: VK055_2471 and KPN_00100) under the criteria supporting E-value cut-off $<1x10^{-25}$ and similarity >70% (Table 4.6). Many of these significantly similar genes were found in *Enterobacteriaceae* members when screening for 32 genera in this family using PATRIC database [163] (Table 4.6).

Table 4.6: Broad spectrum analysis of *lpxC* gene.

Query Gene	Gene	Desription	Organism	Score	Expect	identitiy (%)
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Klebsiella pneumoniae* subsp. Pneumoniae (strain ATCC700721/MGH78578)	633 bits (1633)	0	100
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Citrobacter koseri* (strain ATCCBAA-895/CDC4225-83/SGSC4696)	618 bits (1594)	0	96
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Shigella dysenteriae* serotype 1 (strain Sd197)	617 bits (1590)	0	96
lpxC	*ECoA_01349*	UDP-3-0-acyl N-acetylglucosaminedeacetylase	*Escherichia coli* O157:H7 str.1044	617 bits (1590)	0	96
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase	*Shigella flexneri*	617 bits (1590)	0	96
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Shigella flexneri*	617 bits (1590)	0	96
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Salmonella typhimurium* (strain 14028s/SGSC2262)	613 bits (1582)	0	95
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Shigella boydii* serotype 4 (strain Sb227)	613 bits (1581)	0	95
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Salmonella typhi*	611 bits (1576)	0	95
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia enterocolitica* serotype O:8/biotype1B (strain 8081)	595 bits (1534)	0	91
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia pseudotuberculosis* serotype IB (strain PB1/+)	595 bits (1533)	0	91
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia pseudotuberculosis* serotype O:3 (strain YPIII)	595 bits (1533)	0	91
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia pestis* KIMD27	594 bits (1532)	0	91
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia pestis* bv. Antiqua (strain Angola)	594 bits (1532)	0	91

Table 4.6: Continued.

Query Gene	Gene	Desription	Organism	Score	Expect	identitiy (%)
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia pseudotuberculosis* serotype O:1b (strain IP31758)	594 bits (1532)	0	91
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia pestis* (strain PestoidesF)	594 bits (1532)	0	91
lpxC	*lpxC1*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase1	*Yersinia pestis* bv.Antiqua (strain Nepal516)	594 bits (1532)	0	91
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Yersinia pseudotuberculosis*	594 bits (1532)	0	91
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Haemophilus influenzae* (strain PittEE)	503 bits (1294)	1,00E-180	78
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Pasteurella multocida*	502 bits (1293)	2,00E-180	77
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Haemophilus influenzae* (strain PittGG)	501 bits (1291)	3,00E-180	78
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Haemophilus influenzae* (strain 86-028NP)	501 bits (1289)	7,00E-180	77
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Haemophilus somnus* (strain 129Pt)	493 bits (1270)	5,00E-177	75
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Haemophilus somnus* (strain 2336)	492 bits (1267)	1,00E-176	75
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Vibrio parahaemolyticus*	489 bits (1259)	3,00E-175	75
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Vibrio vulnificus* (strain YJ016)	487 bits (1254)	2,00E-174	75
lpxC	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	*Vibrio vulnificus*	487 bits (1254)	2,00E-174	75

Another gene responsible for lipid A biosynthesis, *lpxA* (locus tags: VK055_2371 and KPN_00194), was detected to be significantly similar to 20 genes in distinct pathogens, most of which are members of the *Enterobacteriaceae* family (Table 4.7).

Table 4.7: Broad spectrum analysis of *lpxA* gene.

Query Gene	**Gene**	**Description**	**Organism**	**Score**	**Expect**	**Identitity (%)**
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Klebsiella pneumoniae* subsp. pneumoniae (strain ATCC 700721 / MGH 78578)	533 bits (1373)	0	100
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Shigella dysenteriae* serotype 1 (strain Sd197)	493 bits (1270)	2,00E-178	90
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Shigella flexneri*	493 bits (1270)	2,00E-178	90
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Shigella flexneri*	493 bits (1270)	2,00E-178	90
lpxA	*ECoA_04314*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Escherichia coli* O157:H7 str. 1044	491 bits (1265)	1,00E-177	90
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Shigella boydii* serotype 4 (strain Sb227)	489 bits (1259)	1,00E-176	90
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Salmonella typhimurium* (strain 14028s / SGSC 2262)	482 bits (1241)	6,00E-174	89
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Salmonella typhi*	480 bits (1236)	3,00E-173	88
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia enterocolitica* serotype O:8 / biotype 1B (strain 8081)	463 bits (1192)	2,00E-166	84
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia pestis* KIM D27	451 bits (1159)	2,00E-161	81
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia pseudotuberculosis* serotype IB (strain PB1/+)	451 bits (1159)	2,00E-161	81
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-	*Yersinia pseudotuberculosis*	451 bits (1159)	2,00E-161	81

Table 4.7: Continued.

Query Gene	Gene	Description	Organism	Score	Expect	Identitity (%)
		acetylglucosamine O-acyltransferase	serotype O:3 (strain YPIII)			
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia pestis* bv. Antiqua (strain Angola)	451 bits (1159)	2,00E-161	81
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia pseudotuberculosis* serotype O:1b (strain IP 31758)	451 bits (1159)	2,00E-161	81
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia pestis* (strain Pestoides F)	451 bits (1159)	2,00E-161	81
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia pestis* bv. Antiqua (strain Nepal516)	451 bits (1159)	2,00E-161	81
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Yersinia pseudotuberculosis*	451 bits (1159)	2,00E-161	81
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Pasteurella multocida*	396 bits (1017)	7,00E-140	71
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Haemophilus somnus* (strain 2336)	394 bits (1011)	7,00E-139	71
lpxA	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	*Haemophilus somnus* (strain 129Pt)	392 bits (1008)	2,00E-138	71
lpxA	*CKO_03185*	Putative uncharacterized protein	*Citrobacter koseri* (strain ATCC BAA-895 / CDC 4225-83 / SGSC4696)	305 bits (781)	2,00E-105	91

Broad spectrum distribution of the genes responsible for the synthesis of heptose precursors of LPS including *gmhA/lpcA* (locus tags: VK055_2330 and KPN_00236) and *hldD/rfaA* (locus tags: VK055_3508 and KPN_03963) were also investigated.

The *gmhA*/*lpcA* gene of the *K. pneumoniae* strain MGH 78578/KPPR1 was found to be homologous with the genes in nine different pathogenic genera and a total of 30 hits (by excluding the *K. pneumoniae* strain MGH 78578) were determined via the comparison the amino acid sequences against the proteomes of different pathogens (Table 4.8).

Table 4.8: Broad spectrum analysis of *gmhA*/*lpcA* gene.

Query Gene	Gene	Description	Organism	Score	Expect	Identitiy (%)
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Klebsiella pneumoniae* subsp. pneumoniae (strain ATCC 700721 / MGH 78578)	395 bits (1016)	6,00E-142	100
gmhA/ lpcA	*CKO_0 2963*	Putative uncharacteri zed protein	*Citrobacter koseri* (strain ATCC BAA-895 / CDC 4225-83 / SGSC4696)	375 bits (962)	1,00E-133	94
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Salmonella typhimurium* (strain 14028s / SGSC 2262)	372 bits (955)	1,00E-132	93
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Salmonella typhi*	372 bits (955)	1,00E-132	93
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia pestis* KIM D27	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*YPTS_0 926*	Phosphohep tose isomerase	*Yersinia pseudotuberculosis* serotype IB (strain PB1/+)	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia pseudotuberculosis* serotype O:3 (strain YPIII)	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia pestis* bv. Antiqua (strain Angola)	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia pseudotuberculosis* serotype O:1b (strain IP 31758)	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia pestis* (strain Pestoides F)	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia pestis* bv. Antiqua (strain Nepal516)	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia pseudotuberculosis*	368 bits (945)	4,00E-131	92
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Yersinia enterocolitica* serotype O:8 / biotype 1B (strain 8081)	366 bits (940)	2,00E-130	91
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Shigella boydii* serotype 4 (strain Sb227)	362 bits (929)	1,00E-128	91
gmhA/ lpcA	*ECoA_0 4833*	Phosphohep tose isomerase 1	*Escherichia coli* O157:H7 str. 1044	362 bits (929)	1,00E-128	91
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Shigella flexneri*	362 bits (929)	1,00E-128	91
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Shigella dysenteriae* serotype 1 (strain Sd197)	356 bits (913)	3,00E-126	90
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Aeromonas hydrophila* subsp. hydrophila (strain ATCC 7966 / NCIB 9240)	332 bits (851)	9,00E-117	84

Table 4.8: Continued.

Query Gene	Gene	Description	Organism	Score	Expect	Identitiy (%)
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Shigella dysenteriae* serotype 1 (strain Sd197)	356 bits (913)	3,00E-126	90
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Aeromonas hydrophila* subsp. hydrophila (strain ATCC 7966 / NCIB 9240)	332 bits (851)	9,00E-117	84
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Pasteurella multocida*	324 bits (831)	1,00E-113	80
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Haemophilus influenzae* (strain PittGG)	320 bits (820)	4,00E-112	80
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Haemophilus somnus* (strain 2336)	320 bits (820)	4,00E-112	80
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Haemophilus influenzae* (strain 86-028NP)	319 bits (817)	1,00E-111	79
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Haemophilus influenzae* (strain PittEE)	318 bits (815)	2,00E-111	79
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Haemophilus ducreyi*	318 bits (814)	4,00E-111	79
gmhA/ lpcA	*lpcA*	Phosphohep tose isomerase	*Haemophilus somnus* (strain 129Pt)	316 bits (810)	1,00E-110	81
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Vibrio parahaemolyticus*	303 bits (776)	2,00E-105	80
gmhA/ lpcA	*VV2518*	Phosphohep tose isomerase	*Vibrio vulnificus* (strain YJ016)	303 bits (775)	3,00E-105	81
gmhA/ lpcA	*gmhA*	Phosphohep tose isomerase	*Vibrio vulnificus*	303 bits (775)	3,00E-105	81

Similarly, broad spectrum analysis of *hldD*/*rfaD* gene was performed to understand whether this gene exhibits a broad distribution among the pathogens. It revealed that 28 hits from nine different pathogenic genera (*Salmonella*, *Shigella*, *Citrobacter*, *Haemophilus*, *Vibrio*, *Yersinia*, *Aeromonas*, *Escherichia* and *Pasteurella*) apart from *Klebsiella*. All significant matches are listed in Table 4.9.

Table 4.9: Broad spectrum analysis of *hldD*/*rfaD* gene.

Query Gene	Gene	Description	Organism	Score	Expect	Identity (%)
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Klebsiella pneumoniae* subsp. pneumoniae (strain ATCC 700721 / MGH 78578)	642 bits (1656)	0	100
rfaD/ hldD	*ECoA_05437*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Escherichia coli* O157:H7 str. 1044	627 bits (1618)	0	97
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Shigella dysenteriae* serotype 1 (strain Sd197)	625 bits (1613)	0	97
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Shigella boydii* serotype 4 (strain Sb227)	625 bits (1612)	0	97
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Shigella flexneri*	625 bits (1612)	0	97
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Shigella flexneri*	625 bits (1612)	0	97
rfaD/ hldD	*rfaD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Salmonella typhimurium* (strain 14028s / SGSC 2262)	620 bits (1598)	0	96
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Salmonella typhi*	620 bits (1598)	0	96
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Citrobacter koseri* (strain ATCC BAA-895 / CDC 4225-83 / SGSC4696)	619 bits (1595)	0	95
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia enterocolitica* serotype O:8 / biotype 1B (strain 8081)	547 bits (1410)	0	83
rfaD/ hldD	*rfaD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pestis* KIM D27	546 bits (1406)	0	83
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pseudotuberculosis* serotype IB (strain PB1/+)	546 bits (1406)	0	83
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pseudotuberculosis* serotype O:3 (strain YPIII)	546 bits (1406)	0	83
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pestis* bv. Antiqua (strain Angola)	546 bits (1406)	0	83
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pseudotuberculosis* serotype O:1b (strain IP 31758)	546 bits (1406)	0	83
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pestis* (strain Pestoides F)	546 bits (1406)	0	83
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pestis* bv. Antiqua (strain Nepal516)	546 bits (1406)	0	83
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Yersinia pseudotuberculosis*	546 bits (1406)	0	83

Table 4.9: Continued.

Query Gene	Gene	Description	Organism	Score	Expect	Identity (%)
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Haemophilus influenzae* (strain 86-028NP)	508 bits (1309)	0	79
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Haemophilus influenzae* (strain PittGG)	506 bits (1304)	0	78
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Haemophilus somnus* (strain 2336)	487 bits (1253)	3E-174	75
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Haemophilus somnus* (strain 129Pt)	487 bits (1253)	3E-174	75
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Haemophilus ducreyi*	486 bits (1252)	5E-174	74
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Vibrio parahaemolyticus*	473 bits (1218)	8E-169	75
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Pasteurella multocida*	472 bits (1215)	2E-168	75
rfaD/ hldD	*CGSHiEE_06480*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Haemophilus influenzae* (strain PittEE)	471 bits (1212)	2E-168	77
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Aeromonas hydrophila* subsp. hydrophila (strain ATCC 7966 / NCIB 9240)	472 bits (1214)	4E-168	71
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Vibrio vulnificus*	465 bits (1197)	1E-165	74
rfaD/ hldD	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	*Vibrio vulnificus* (strain YJ016)	463 bits (1192)	8E-165	74

Lastly, homologues of *kdsA* gene (locus tags: VK055_0220 and KPN_02230) were investigated against 181 pathogenic bacteria specified in PBIT web server. This gene was found in ten different pathogenic genera (*Salmonella*, *Shigella*, *Citrobacter*, *Haemophilus*, *Vibrio*, *Yersinia*, *Aeromonas*, *Escherichia*, *Neisseria* and *Pasteurella*) in addition to *Klebsiella*. A total of 31 hits were identified for this gene. They are presented in Table 4.10.

In summary, non-homologous, essential genes predicted in this study have been identified in a number of bacteria. Therefore, they may be considered good potential targets for drug by facilitating design of broad-spectrum drugs.

Table 4.10: Broad spectrum analysis of *ksdA* gene.

Query Gene	Marched Gene	Description	Organism	Score	Expect	Identity (%)
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Klebsiella pneumoniae* subsp. pneumoniae (strain ATCC 700721 / MGH 78578)	580 bits (1495)	0	100
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Shigella dysenteriae* serotype 1 (strain Sd197)	556 bits (1434)	0	95
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Shigella boydii serotype 4* (strain Sb227)	556 bits (1434)	0	95
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Shigella flexneri*	556 bits (1434)	0	95
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Shigella flexneri*	556 bits (1434)	0	95
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Citrobacter koseri* (strain ATCC BAA-895 / CDC 4225-83 / SGSC4696)	556 bits (1433)	0	96
kdsA	*ECoA_02867*	2-Keto-3-deoxy-D-manno-octulosonate-8-phosphate synthase	*Escherichia coli* O157:H7 str. 1044	556 bits (1432)	0	95
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Salmonella typhimurium* (strain 14028s / SGSC 2262)	542 bits (1397)	0	94
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Salmonella typhi*	542 bits (1397)	0	94
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pestis* KIM D27	533 bits (1372)	0	91
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pseudotuberculosis* serotype IB (strain PB1/+)	533 bits (1372)	0	91
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pseudotuberculosis* serotype O:3 (strain YPIII)	533 bits (1372)	0	91
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pestis* bv. Antiqua (strain Angola)	533 bits (1372)	0	91
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pseudotuberculosis* serotype O:1b (strain IP 31758)	533 bits (1372)	0	91
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pestis* (strain Pestoides F)	533 bits (1372)	0	91
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pestis* bv. Antiqua (strain Nepal516)	533 bits (1372)	0	91

Table 4.10: Continued.

Query Gene	Marched Gene	Description	Organism	Score	Expect	Identity (%)
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia pseudotuberculosis*	533 bits (1372)	0	91
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Yersinia enterocolitica* serotype O:8 / biotype 1B (strain 8081)	530 bits (1365)	0	90
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Haemophilus influenzae* (strain 86-028NP)	497 bits (1279)	6,00E-179	83
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Haemophilus influenzae* (strain PittEE)	494 bits (1271)	9,00E-178	83
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Haemophilus influenzae* (strain PittGG)	493 bits (1269)	1,00E-177	83
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Vibrio vulnificus* (strain YJ016)	490 bits (1261)	3,00E-176	84
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Vibrio vulnificus*	490 bits (1261)	3,00E-176	84
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Vibrio parahaemolyticus*	489 bits (1259)	6,00E-176	83
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Haemophilus somnus* (strain 129Pt)	488 bits (1256)	2,00E-175	79
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Haemophilus somnus* (strain 2336)	488 bits (1255)	2,00E-175	79
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Haemophilus ducreyi*	483 bits (1244)	1,00E-173	79
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Pasteurella multocida*	481 bits (1237)	1,00E-172	80
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Aeromonas hydrophila* subsp. hydrophila (strain ATCC 7966 / NCIB 9240)	471 bits (1211)	1,00E-168	79
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Neisseria meningitidis* serogroup C (strain 053442)	402 bits (1033)	1,00E-141	71
kdsA	*kdsA*	2-dehydro-3-deoxyphosphooctonate aldolase	*Neisseria gonorrhoeae* (strain NCCP11945)	402 bits (1033)	1,00E-141	71

4.4. Gene-Centric Approach to Identify Drug Targets Using the Updated Biomass Reaction

Prediction of single-gene essentiality is directly related to the biomass definition. Thereby, the biomass equation of MGH 78578 strain was refined by adding universally

essential cofactors of prokaryotes (NAD, NADH, NADP, NADP, FAD, CoA, FMN, PYDX5P and SAM/AMET) to increase the accuracy of the model and to extend the putative drug target list. Generation and integration processes of the biomass equation are summarized in the section 3.3.4.

Investigation of the essential genes of MGH 788578 strain through SM simulation by using the updated biomass equation has resulted in 102 essential genes but only 61 out of these genes were found to be non-human homologues. These non-homologous genes were compared with 42 non-homologous genes predicted using original biomass equation (section 4.3.1) and 19 new genes (*nadD*, *ribC*, *aceK*, *panD*, *ppnK*, *ribB*, *ribH*, *panB*, *ribF*, *dxs*, *nadE*, *coaD*, *coaA*, *dfp*, *pdxB*, *nadA*, *ribA*, *panC* and *ribD*) were identified. Any virulence-related genes were not determined among these 19 genes. Then, these genes were screened in terms of druggability. 11 genes listed in Table 4.11 were detected as druggable. All of them encode cytosolic proteins. That is any vaccine candidates were not identified.

Table 4.11: Evaluation of druggability properties of the putative targets and drug repositioning for an additional target list.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
KPN_04352	*coaA*	3	Pantothenic acid	approved, nutraceutical, vet_approved
			Coenzyme A	investigational, nutraceutical
			Phosphoaminophosphonic Acid-Adenylate Ester	experimental
KPN_03974	*coaD*	5	Coenzyme A	investigational, nutraceutical
			Dephospho Coenzyme A	experimental
			4'-Phosphopantetheine	experimental
			4'-Phosphopantetheine	experimental
			4'-Phosphopantetheine	experimental
KPN_03979	*dfp*	2	Cytidine-5'-Triphosphate	experimental
			Cytidine-5'-Monophosphate	experimental
KPN_00671	*nadD*	2	Citric Acid	approved, nutraceutical, vet_approved
			Deamido-Nad+	experimental
KPN_01228	*nadE*	6	Gamma-Arsono-Beta, Gamma-Methyleneadenosine-5'-Diphosphate	experimental
			Deamido-Nad^{+}	experimental
			Pyrophosphoric acid	approved, experimental

Table 4.11: Continued.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
			Alpha,Beta-Methyleneadenosine-5'-Triphosphate	experimental
			Deamido-Nad+	experimental
			Gentamicin	approved, vet_approved
KPN_00141	*panB*	2	2-Dehydropantoate	experimental
			Alpha-ketoisovalerate	experimental
KPN_00140	*panC*	5	Tris-Hydroxymethyl-Methyl-Ammonium	experimental
			2,4-Dihydroxy-3,3-Dimethyl-Butyrate	experimental
			Alpha,Beta-Methyleneadenosine-5'-Triphosphate	experimental
			Pantoyl Adenylate	experimental
			Beta-Alanine	experimental
KPN_00139	*panD*	2	Malonic acid	experimental
			S-oxy-L-cysteine	experimental
KPN_02000	*ribC*	1	Riboflavin	approved, investigational, nutraceutical, vet_approved
KPN_00020	*ribF*	3	Flavin mononucleotide	approved, investigational
			Citric Acid	approved, nutraceutical, vet_approved
			7,8-dimethylalloxazine	experimental
KPN_00367	*ribH*	14	5-Nitro-6-Ribityl-Amino-2,4(1h,3h)-Pyrimidinedione	experimental
			6,7-Dioxo-5h-8-Ribitylaminolumazine	experimental
			5-Nitroso-6-Ribityl-Amino-2,4(1h,3h)-Pyrimidinedione	experimental
			3-(7-hydroxy-8-ribityllumazine-6-yl) propionic acid	experimental
			Dithioerythritol	experimental
			4-{2,6,8-Trioxo-9-[(2R,3S,4R)-2,3,4,5-Tetrahydroxypentyl]-1,2,3,6,8,9-Hexahydro-7h-Purin-7-Yl}Butyl Dihydrogen Phosphate	experimental
			D-1,4-dithiothreitol	experimental
			3-{2,6,8-trioxo-9-[(2S,3S,4R)-2,3,4,5-tetrahydroxypentyl]-1,2,3,6,8,9-hexahydro-7H-purin-7-Yl}propyl dihydrogen phosphate	experimental
			(4S,5S)-1,2-dithiane-4,5-diol	experimental

Table 4.11: Continued.

Tag ID	Gene Name	Number of Drug	Interacting Drug	Drug Group
			4-{2,6,8-Trioxo-9-[(2S,3R,4R)-2,3,4,5-Tetrahydroxypentyl]-1,2,3,6,8,9-Hexahydro-7h-Purin-7-Yl}Butyl Dihydrogen Phosphate	experimental
			3-{2,6,8-trioxo-9-[(2R,3S,4R)-2,3,4,5-tetrahydroxypentyl]-1,2,3,6,8,9-hexahydro-7H-purin-7-Yl}propyl dihydrogen phosphate	experimental
			3-{2,6,8-trioxo-9-[(2S,3R,4R)-2,3,4,5-tetrahydroxypentyl]-1,2,3,6,8,9-hexahydro-7H-purin-7-Yl}propyl dihydrogen phosphate	experimental
			3-{2,6,8-trioxo-9-[(2R,3R,4R)-2,3,4,5-tetrahydroxypentyl]-1,2,3,6,8,9-hexahydro-7H-purin-7-Yl}propyl dihydrogen phosphate	experimental
			4-(6-CHLORO-2,4-DIOXO-1,2,3,4-TETRAHYDROPYRIMIDIN-5-YL) BUTYL PHOSPHATE	experimental

4.5. Identification of Essential Metabolites for Metabolite-Centric Drug Target Discovery

Essential metabolites were identified by following the procedure described in the section 3.5. These simulations were performed in the host-mimicking media including SM and HBF. 38 essential metabolites were identified for the growth simulations in SM for the GEM of iYL1228 while 31 essential metabolites were determined for iKp1289. Growth simulation in HBF medium revealed 20 essential metabolites for iYL1228 and 18 essential metabolites for iKp1289, respectively.

To determine more efficient drug-targets, these metabolites were filtered based on presence of them either in the human metabolism or in the reactions catalyzed by human-homologous enzymes. Connectivity was also used as the another parameter in drug target prioritization.

4.6. Drug Target Prioritization for the Metabolite-Centric Candidates

One of the most significant processes in the drug target identification pipeline is to avoid possible side effects in the host and to provide an optimal efficacy in treatment. Two main routes suggested by Kim and colleagues (2010) were followed to eliminate the targets which may contribute to the toxic side effects. Firstly, a common metabolite list was compiled using HumanCyc database and two GEMs (Recon 1 and Recon 2). The essential metabolites were screened against this list and common metabolites were eliminated. A further filtering was performed to eliminate the currency metabolites. Thus, the metabolites which are not found in the human metabolism were considered in the study. Secondly, homologous gene lists were compiled for both models to eliminate the metabolites associated with the reactions catalyzed by human-homologous enzymes. Thus, the metabolites take part in the reactions catalyzed by only non-homologous genes that were selected.

Kim and colleagues (2010) suggested that targeting of many outgoing reactions simultaneously may delay the resistance development. However, they did not consider essentiality of the reactions to be blocked whereas inhibition of many essential reactions is pleusible to damage the metabolism of a pathogen more effectively. Furthermore, all nonessential reactions associated with an essential metabolite must be blocked simultaneously to induce cell death. Therefore, targeting of a single non-essential reaction is not sufficient to completely suppress the bacterial growth. In this context, any mutations in an enzyme catalyzing one of these nonessential reactions may cause loss of this target's effect. On the other hand, inhibition of a single essential reaction prevents survival of the pathogen per se. Moreover, it was shown that essential genes were more conserved in many bacteria by a comparative study [164]. Thus, targeting of the essential reactions may contribute to a lower mutation rate. Therefore, essentiality may be crucial as much as connectivity. Herein, the metabolites involved in the essential outgoing reactions were taken into consideration differently from EMFilter approach [165].

Filtering process resulted in the prediction of a single essential metabolite known as the undecaprenyl-diphospho-N-acetylmuramoyl-(N-acetylglucosamine)-L-ala-D-glu meso-2,6-diaminopimeloyl-D-ala-D-ala (uaagmda). This metabolite is not involved in the human metabolism. It is linked with 3 reactions in the pathogen

metabolism: (1) murein polymerizing transglycosylase reaction, (2) murein polymerizing transglycosylase 2 (three linked units) reaction and (3) UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide)pyrophosphoryl-undecaprenol N-acetylglucosamine transferase reaction. First two reactions consumed this coumpound as the last reaction facilitates production of the metabolite. Of these reactions, the first reaction was predicted as essential in SM and HBF simulations. A total of four genes (*mrcB*, *pbpC*, *mrcA* and *murG*) were detected associated with the outgoing reactions. Three of three of them (*mrcB* (penicillin-binding protein 1b), *pbpC* (penicillin binding protein 1C) and *mrcA* (bifunctional penicillin-binding protein 1a: transglycosylase/ transpeptidase)) are involved in the first essential outgoing reaction.

5. DISCUSSION

Conventional medical treatment approaches are insufficient to handle resistant *Klebsiella pneumoniae* strains, being an etiological agent of serious life-threatening nosocomial infections. Therefore, more comprehensive approaches must be introduced to decipher the complicated cellular pathways of the pathogen and eventually to elucidate new drug targets for the treatment of *Klebsiella*-mediated infections. Constraint-based analysis of the GEMs representing whole metabolism of the pathogens is a quite promising approach from this perspective.

The focus of the thesis is analysis of the GEMs of two strains of *K. pneumoniae* (MGH 78578 and KPPR1 strains) to reveal putative drug targets. To this end, these models were validated as discussed in section 4.1 in the first step. It was revealed that *K. pneumoniae* KPPR1 can reach higher growth rates in comparison with the MGH 78578 strain for the same growth condition. Closer inspection of the growth rates predicted using the models revealed a bit higher uptake rates of some metabolites (e.g., iron) by the KPPR1 strain. Thus, the higher proliferation rate of KPPR1 may be linked to uptake of some compounds in a more efficient manner. This hypothesis seems plausible considering that more virulent strains can adapt to host environment better in order to exploit host sources more efficiently. Therefore, the bacterial pathogenicity is directly correlated with a proficient metabolism. This situation is also linked to the term of 'nutritional virulence' that refers to the adaptation strategy to provide the assimilation of the host sources efficiently [166]. As an example of this phenomenon, *Klebsiella* can utilize the iron by gaining access to this metabolite using small iron-scavenging molecules (e.g., siderophores). Thus, some virulence-related properties of the pathogens may support the bacterial growth by triggering the adaptation into the hostile environment. In other words, the degree of pathogenicity can be associated with efficient bacterial growth.

Comparison of the growth rates were followed to the identify the essential genes, and different approaches (an in-house algorithm developed in this study, a function under COBRA Toolbox and the Fast-SL algorithm) were used in this process. Slightly different results were obtained in simulations through different gene deletion algorithms. As a result of growth simulations in iYL1228, a total of 2 false-positive

and 1 false-negative results were detected by the COBRA function, and 1 false-positive result was found using Fast-SL algorithm.

Homology analysis was performed to eliminate the *Klebsiella* genes manifesting by a significant similarity in the amino acid sequence with human after the essential genes were identified via different simulations. Thus, only protein-coding genes whose homologues are absent on the human genome were selected. In this regard, the sequence similarity search for 30% amino acid identity was used to reveal human-homologues. It should be noted that the higher conservation of the protein structures (e.g., the secondary structure) than the sequence was reported. Therefore, comparison of the secondary structures can be useful particularly to determine the related proteins with divergent sequences (smaller than 30% similarity). Comparison of only protein sequences may fail to identify the related proteins with 20%–30% (or lower) identity [167]. This phenomenon reveals that the sequence similarity (for cut-off of 30%) does not guarantee homology.

Many of the essential genes were found to share no homology with the human in this study. Thus, their products are more conventient to be drug targets. On the other hand, the genes sharing lower than 30% similarity should be evaluated further using their structure information.. A list of the non-homologous essential genes of MGH 78578 strain determined for the 30% similarity and lack of similarity is given in Table 5.1.

Table 5.1: Comparison of the non-homologous essential genes of MGH 78578 based on different sequence similarity criteria.

		SM Simulation		**HBF Simulation**	
Gene	**Description**	**30% similarity**	**No similarity**	**30% similarity**	**No similarity**
fabA	3-hydroxydecanoyl-ACP dehydratase	KPN_00983	-	-	-
accA	acetyl-CoA carboxylase alpha subunit	KPN_00198	-	KPN_00198	-
accD	acetyl-CoA carboxylase beta subunit	KPN_02706	-	KPN_02706	-
accB	acetyl-CoA carboxylase	KPN_03664	KPN_03664	KPN_03664	KPN_03664
purB	adenylosuccinate lyase	KPN_01139	KPN_01139	-	-
rfaD	ADP-L-glycero-D-manno-heptose-6-epimerase	KPN_03963	-	KPN_03963	-
purK	phosphoribosylaminoimidazole carboxylase	KPN_00477	KPN_00477	-	-

Table 5.1: Continued.

		SM Simulation		HBF Simulation	
Gene	**Description**	**30% similarity**	**No similarity**	**30% similarity**	**No similarity**
purE	phosphoribosylaminoimidazole carboxylase catalytic subunit	KPN_00478	-	-	-
asd	aspartate-semialdehyde dehydrogenase	KPN_03799	KPN_03799	KPN_03799	KPN_03799
pyrI	aspartate carbamoyltransferase regulatory subunit	KPN_04656	KPN_04656	-	-
dapF	diaminopimelate epimerase	KPN_04308	KPN_04308	KPN_04308	KPN_04308
dapB	dihydrodipicolinate reductase	KPN_00039	KPN_00039	KPN_00039	KPN_00039
dapA	dihydrodipicolinate synthase	KPN_02812	-	KPN_02812	-
pyrC	dihydroorotase	KPN_01074	-	-	-
tmk	thymidylate kinase	KPN_01096	KPN_01096	KPN_01096	KPN_01096
fabI	enoyl-(acyl carrier protein) reductase	KPN_01284	-	-	-
glmU	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase/glucosamine-1-phosphate acetyltransferase	KPN_04135	KPN_04135	KPN_04135	KPN_04135
uge	uridine diphosphate galacturonate 4-epimerase	KPN_02492	-	KPN_02492	-
murI	glutamate racemase	KPN_04256	KPN_04256	KPN_04256	KPN_04256
gmhB	hypothetical protein	KPN_00214	KPN_00214	KPN_00214	KPN_00214
thrB	homoserine kinase	KPN_00003	KPN_00003	KPN_00003	KPN_00003
kdsA	2-dehydro-3-deoxyphosphooctonate aldolase	KPN_02230	KPN_02230	KPN_02230	KPN_02230
lpxB	lipid-A-disaccharide synthase	KPN_00195	KPN_00195	KPN_00195	KPN_00195
pyrF	orotidine 5'-phosphate decarboxylase	KPN_01277	KPN_01277	-	-
pyrE	orotate phosphoribosyltransferase	KPN_03983	-	-	-
mraY	phospho-N-acetylmuramoyl-pentapeptide-transferase	KPN_00091	KPN_00091	KPN_00091	KPN_00091
pssA	phosphatidylserine synthase	KPN_02908	KPN_02908	KPN_02908	KPN_02908
gmhA	phosphoheptose isomerase	KPN_00236	KPN_00236	KPN_00236	KPN_00236
lpxK	tetraacyldisaccharide 4'-kinase	KPN_00942	KPN_00942	KPN_00942	KPN_00942
dapD	2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase	KPN_00179	KPN_00179	KPN_00179	KPN_00179

Table 5.1: Continued.

		SM Simulation		HBF Simulation	
Gene	**Description**	**30% similarity**	**No similarity**	**30% similarity**	**No similarity**
lpxD	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	KPN_00192	-	KPN_00192	-
murE	UDP-N-acetylglucosamine acyltransferase	KPN_00089	KPN_00089	KPN_00089	KPN_00089
lpxA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	KPN_00194	KPN_00194	KPN_00194	KPN_00194
murA	N-acetylglucosaminyl transferase	KPN_03599	KPN_03599	KPN_03599	KPN_03599
murG	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	KPN_00094	-	KPN_00094	-
murD	UDP-N-acetylmuramate--L-alanine ligase	KPN_00092	KPN_00092	KPN_00092	KPN_00092
murC	UDP-N-acetylenolpyruvoylglucosamine reductase	KPN_00095	KPN_00095	KPN_00095	KPN_00095
murB	UDP-glucose dehydrogenase	KPN_04350	KPN_04350	KPN_04350	KPN_04350
ugd	D-alanine:D-alanine-adding enzyme	KPN_02493	-	KPN_02493	-
murF	UDP-3-O-acyl N-acetylglucosamine deacetylase	KPN_00090	KPN_00090	KPN_00090	KPN_00090
lpxC	UDP-2,3-diacylglucosamine hydrolase	KPN_00100	KPN_00100	KPN_00100	KPN_00100
ybbF		KPN_00480	KPN_00480	KPN_00480	KPN_00480

The non-homologous essential genes of the KPPR1 strain identified for the 30% and 0% similarity criteria. They are listed in Table 5.2.

Table 5.2: Comparison of the non-homologous essential genes of KPPR1 based on different sequence similarity criteria.

		SM Simulation		HBF Simulation	
Gene	**Description**	**30% similarity**	**No similarity**	**30% similarity**	**No similarity**
fabA	beta-hydroxyacyl-(acyl-carrier-protein) dehydratase FabA	VK055_1504	VK055_1504	-	-
accA	acetyl-CoA carboxylase, carboxyl transferase, alpha subunit	VK055_2367	-	VK055_2367	-

Table 5.2: Continued.

		SM Simulation		HBF Simulation	
Gene	**Description**	**30% similarity**	**No similarity**	**30% similarity**	**No similarity**
accB	acetyl-CoA carboxylase, biotin carboxyl carrier protein	VK055_3812	VK055_3812	VK055_3812	VK055_3812
accD	acetyl-CoA carboxylase, carboxyl transferase, beta subunit	VK055_4807	-	VK055_4807	-
purB	adenylosuccinate lyase	VK055_1323	VK055_1323	-	-
rfaD	ADP-glyceromanno-heptose 6-epimerase	VK055_3508	-	VK055_3508	-
-	1-acylglycerol-3-phosphate O-acyltransferases domain protein	VK055_4038	-	VK055_4038	-
purK	phosphoribosylaminoimidazole carboxylase, ATPase subunit	VK055_2070	VK055_2070	-	-
purE	phosphoribosylaminoimidazole carboxylase catalytic subunit	VK055_2069	-	-	-
asd2	aspartate-semialdehyde dehydrogenase	VK055_3672	VK055_3672	VK055_3672	VK055_3672
pyrI	aspartate carbamoyltransferase, regulatory subunit	VK055_2808	VK055_2808	-	-
dapF	diaminopimelate epimerase	VK055_3167	VK055_3167	VK055_3167	VK055_3167
dapB	dihydrodipicolinate reductase	VK055_2536	VK055_2536	VK055_2536	VK055_2536
dapA	dihydrodipicolinate synthase	VK055_4699	-	VK055_4699	-
pyrC	dihydroorotase, homodimeric type	VK055_1392	-	-	-
tmk	thymidylate kinase	VK055_1369	VK055_1369	VK055_1369	VK055_1369
fabI	NADH-dependent enoyl-(acyl carrier protein) reductase	VK055_1165	-	-	-
glmU	UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase	VK055_3340	VK055_3340	VK055_3340	VK055_3340
-	polysaccharide biosynthesis family protein	VK055_5030	-	VK055_5030	-
murI	glutamate racemase	VK055_3220	VK055_3220	VK055_3220	VK055_3220
yaeD	D,D-heptose 1,7-bisphosphate phosphatase	VK055_2352	VK055_2352	VK055_2352	VK055_2352
kdsA	3-deoxy-8-phosphooctulonate synthase	VK055_0220	VK055_0220	VK055_0220	VK055_0220
lpxB	lipid-A-disaccharide synthase	VK055_2370	VK055_2370	VK055_2370	VK055_2370

Table 5.2: Continued.

		SM Simulation		HBF Simulation	
Gene	**Description**	**30% similarity**	**No similarity**	**30% similarity**	**No similarity**
fabD	malonyl CoA-acyl carrier protein transacylase	VK055_1375	VK055_1375	VK055_1375	VK055_1375
pyrF	orotidine 5prime-phosphate decarboxylase	VK055_1174	VK055_1174	-	-
pyrE	orotate phosphoribosyltransferase	VK055_3486	-	-	-
mraY	phospho-N-acetylmuramoyl-pentapeptide-transferase	VK055_2481	VK055_2481	VK055_2481	VK055_2481
pssA	CDP-diacylglycerol--serine O-phosphatidyltransferase	VK055_4593	VK055_4593	VK055_4593	VK055_4593
gmhA	phosphoheptose isomerase	VK055_2330	VK055_2330	VK055_2330	VK055_2330
lpxK	tetraacyldisaccharide 4'-kinase	VK055_1542	VK055_1542	VK055_1542	VK055_1542
dapD	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	VK055_2386	VK055_2386	VK055_2386	VK055_2386
lpxD	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	VK055_2373	-	VK055_2373	-
-	UDP-N-acetylmuramyl tripeptide synthase	VK055_2483	VK055_2483	VK055_2483	VK055_2483
lpxA	acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase	VK055_2371	VK055_2371	VK055_2371	VK055_2371
murA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	VK055_3878	VK055_3878	VK055_3878	VK055_3878
murG	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	VK055_2477	-	VK055_2477	-
murD	UDP-N-acetylmuramoylalanine--D-glutamate ligase	VK055_2480	VK055_2480	VK055_2480	VK055_2480
murC	UDP-N-acetylmuramate--alanine ligase	VK055_2476	VK055_2476	VK055_2476	VK055_2476
murB	UDP-N-acetylenolpyruvoylglucosamine reductase	VK055_3124	VK055_3124	VK055_3124	VK055_3124
ugd	nucleotide sugar dehydrogenase family protein	VK055_5029	-	VK055_5029	-

Table 5.2: Continued.

		SM Simulation		HBF Simulation	
Gene	**Description**	**30% similarity**	**No similarity**	**30% similarity**	**No similarity**
-	UDP-MurNAc-L-Ala-D-Glu-L-Lys:D-Ala-D-Ala ligase	VK055_2482	VK055_2482	VK055_2482	VK055_2482
lpxC	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	VK055_2471	VK055_2471	VK055_2471	VK055_2471
lpxH	UDP-2,3-diacylglucosamine hydrolase	VK055_2067	VK055_2067	VK055_2067	VK055_2067

Another filtering strategy is to reveal subcellular localizations of proteins. The protein locations were investigated through three prediction tools (CELLO, PSORTb and iLoc-Gneg), literature survey and UniProt database. It was found that the products of all predicted essential, non-homologous genes are localized in either the cytoplasm or the inner membrane. It indicates that these gene products may be putative drug targets (but not vaccine candidates). Another approach in the drug target prioritization process is to investigate presence of any defined or putative antibiotic resistance genes within the non-homologous gene list. ARG-ANNOT database was used in the screening process. A single gene (i.e., *glmU*) was identified as a potential resistance-related gene for the aforementioned cut-off value. N-Acetylglucosamine-1-phosphate uridyltransferase (GlmU) is a bifunctional enzyme with uridyltransferase and acetyltransferase activities in both gram-positive and gram-negative bacteria [153, 154]. It is crucial in peptidoglycan and lipopolysaccharide synthesis [155]. Thus, *glmU* gene is essential for the membrane integrity, supporting that it may be a potential drug target. On the other hand, a direct relationship was not detected between the predicted antibiotics (macrolide-lincosamide-streptogramin and phenicol antibiotics) and *glmU* gene although there was a significant similarity between the *glmU* gene and the matched resistance genes. Therefore, an indirect link may exist between these genes or it may be false positive considering the short alignment lenght. Even if ARG-ANNOT database [119] is prominent thanks to its flexibility to find out distantly related genes, this flexibility may result in emergence of false positives. Therefore, considering the alignment lenght as an additional filetering parameter may produce more accurate results.

There is a clear relationship between the antibiotic resistance and virulence. In this context, inhibition of the putative targets responsible for both survival and virulence of the cell may severely damage the pathogen in many ways. Therefore, virulence factors were identified through the bacterial virulence analysis. It should be noted that VFDB database does not include any data for *K. pneumoniae*. However, six non-homologous, essential *Klebsiella* genes (*hldD* (*rfaD*), *kdsA*, *gmhA (lpcA*), *lpxD*, *lpxA* and *lpxC*) were predicted as the virulence-related genes in *Haemophilus influenzae* Rd KW20. The same virulence genes were identified for both *Klebsiella* models. Then, druggability of products of these genes (i.e., virulence factors) was investigated to evaluate the binding capacity to drug molecules. At least one experimental drug was determined for five out of six identified virulence factors. Then, the broad distribution of these genes among different attracting pathogens was screened. All of them are found in at least seven distinct pathogenic genera (*Shigella*, *Escherichia*, *Salmonella*, *Yersinia*, *Pasteurella*, *Citrobacter* and *Haemophilus*). Most of these pathogenic genera are members of *Enterobacteriaceae* family. The relatively high prevalence of these genes among different popular pathogens prompted us for the further investigation of these genes. All of these genes were found to be associated with the LPS synthesis.

As mentioned before, LPS is the outer membrane component that consists of three main parts including O-antigen, core oligosaccharide and lipid A. Lipid A is essential for the protection (e.g., resistance against a great variety of antibiotics) and survival of the bacteria [11, 39]. Its biosynthesis is managed through nine enzyme-catalyzed reactions, and *lpxA*, *lpxC*, *lpxD* are among the essential genes in this process. Thereof, deletion of these genes triggers the sensivity to hydrophobic antibiotics such as erythromycin [159]. In the beginning of the lipid A synthesis, LpxA (UDP-GlcNAc acyltransferase) catalyzes production of UDP-3-(O)-acyl-GlcNAc through acylation of uridine diphosphate (UDP-GlcNAc) in Raetz pathway. Next steps in this pathway are the deacetylation of the GlcNAc and acylation to generate UDP-2,3-(O)-diacyl-GlcN mediated by LpxC and LpxD, respectively. After formation of the UDP-2,3-(O)-diacyl-GlcN, Lipid A is synthesized via additional reactions catalyzed by other enzymes [12, 168]. All of the enzymes involved in the Raetz pathway have a potential to be drug targets for an effective therapy [168]. In a consistent manner, LpxD was predicted as a drug target for *K. pneumoniae* in both this study and a proteome-based *in silico* study by Georrge and colleagues [83]. Inhibitor design against especially

LpxC enzyme is a popular approach owing to its conservation within gram-negative bacteria. However, essentiality of this gene depends on its level in *E. coli* cells. Another enzyme in the Raetz pathway, LpxA, does not exhibit any structural/sequence homology with mammalian enzymes. Therefore, this enzyme was also reported as a promising drug target [11, 12]. Recently, the crystal structure of *Moraxella catarrhalis* LpxA was elucidated by Pratap and colleagues (2017). The researchers suggested potential LpxA inhibitors via the computational methods by asserting that these inhibitors may also bind to the LpxA enzymes from other gram-negative bacteria [10].

One of the identified putative targets, *hldD* (formerly known as *rfaD*) gene, encodes ADP-L-glycero-D-mannoheptose-6-epimerase. This enzyme catalyzes the generation of ADP-L-glycero-D-manno-heptose (heptose) essential for the LPS core domain biosynthesis in gram-negative bacteria [157, 158]. Mammalian cells do not have the heptose contrary to bacteria. Moreover, it has been shown that LPS of *rfaD* mutants (i.e., core-defective/heptoseless mutants) loses its barrier characteristics. Thus, the mutated gene leads to a poorer survival of the pathogen in the host due to the reduced pathogenic characteristics and impairments in the intrinsic resistance, causing an enhanced antibiotic susceptibility [157]. More recently, researchers have demonstrated the contribution of Enterohemorrhagic *E. coli* (EHEC) O157:H7 RfaD to the virulence of bacteria and they suggested that this protein can be a suitable drug target as highlighted in the section 4.3.4 [158]. Furthermore, this enzyme was predicted to be essential in *Yersinia pestis* through a computational approach [93]. Thus, it may be a potential target to handle the infection triggered by different pathogens. Similarly, *kdsA* gene was reported as essential in *Francisella tularensis* and *Y. pestis* in the same study [93]. In addition to *F. tularensis* and *Y. Pestis*, *kdsA* gene was identified to be a candidate drug target for *P. aeruginosa* [13] and *Leptospira interrogans* [14]. 3-deoxy-D-manno-octulosonate 8-phosphate synthetase (Kdo-8-phosphate synthetase) enzyme encoded by *kdsA* gene is responsible for the production of one of the major components of LPS known as 3-deoxy-D-manno-octulosonic acid (Kdo). Kdo plays an important role in linking lipid A with the core oligosaccharides [162]. *In silico* analysis of the *kdsA* gene revealed that it does not have any human homologues and it is essential for the survival of *P. aeruginosa* [13]. Essentiality of this gene in *P. aeruginosa* PAO1 was also validated through *in vitro* approaches [169]. Thus, development of inhibitors against the KdsA may enable to cope with antibiotic-resistant *K. pneumoniae* strains.

Another druggable virulence gene involved in the synthesis of heptose precursors of the LPS inner core (ADP-L-glycero-D-manno heptose) is *gmhA* (*lpcA*) encoding a sedoheptulose 7-phosphate isomerase. It was reported that inhibition of GmhA has increased the susceptibility of *Fusobacterium nucleatum* to various antimicrobial agents (e.g., novobiocin) in even low concentrations by causing emergence of heptoseless mutants [170].

Collectively, all identified druggable virulence factors may be evaluated as drug targets to handle an infection by *K. pneumoniae* strain MGH 78578 or strain KPPR1. However, the genes involved in lipid A and Kdo biosynthesis may be particularly promising because inhibition of these processes leads to a lethal effect. On the other hand, inhibition of the genes responsible for the biosynthesis of heptoses causes an enhanced antibiotic sensitivity and a dramatic reduction of virulence. It is important to note that novel strategies should be introduced to design inhibitors against these targets to overcome the difficulties related to substrate availability and the chemically complicated structures of these enzymes [171].

The target list identified in the study was also extended by addition of an updated biomass equation reaction into iYL1228. A total of 102 essential genes was identified via SM simulation and 11 non-homologous, druggable genes could not be determined via the old biomass reaction. These 11 genes (*nadE*, *nadD*, *panB*, *panC*, *panD*, *coaA*, *dfp*, *coaD*, *ribH*, *ribF* and *ribD*) are mainly related to NAD synthesis, pantothenate production, coenzyme A biosynthesis and riboflavin synthesis. Nicotinic acid mononucleotide (NaMN) adenylyltransferase (*nadD*) and NAD synthetase (*nadE*) are responsible in the last two steps of NAD biosynthesis process. These steps are broadly conserved while the early steps vary across bacteria [172, 173]. On the other hand, they have substantial differences compared with the human counterparts [173]. These ezymes are crucial in the maintenance of NAD^+ pool and so cell survival [174]. In this context, there are many studies supporting the potenial of these enzymes to be antibacterial drug targets and/or investigating potential inhibitors against these enzymes in distinct bacterial pathogens [172, 174–176].

Coenzyme A is another prominent organic cofactor in all organisms which participates in many significant metabolic processes such as degradation and synthesis of the fatty acids, production of nonribosomal proteins and biosynthesis of phospholipids [177, 178]. CoA biosynthesis pathway was reported as a potential target for novel inhibitors [179]. This essential acyl carrier was synthesized through a

reaction system begining with formation of β-alanine from aspartate by aspartate-1-decarboxylase (encoded by *panD* gene). Then, ketopantoate hydroxymethyltransferase enzyme (encoded by *panB* gene) catalyzes production of pantoate that is essential for the bacterial growth [177, 180]. After the pantoate was reduced by ketopantoate reductase (encoded by *panE* gene), pantothenate synthetase (the *panC* gene product) catalyzes the condensation of the intermediates (β-alanine and pantoate) to the pantothenic acid (vitamin B5) [177]. The genes of *panB, panC* and *panD* were suggested to be essential, non-homologous candidate drug targets in this work. The pantothenic acid synthesis pathway seems a promising target to handle *Klebsiella* infection because production of this vitamin is crucial for biosynthesis of CoA. In this context, prevention of the pantothenic acid synthesis may cause lethal effect on the MGH 78578 strain due to inhibition of CoA synthesis and eventually suppression of many vital cellular processes. Nevetheless, pantothenic acid synthesis pathway as a drug target should be investigated further because it was reported that suppression of this process was not lethal in *E. coli* owing to uptake of exogenous pantothenate [181]. Transportation of this vitamin into the bacterial cells was managed by a sodium-dependent permease (encoded by *panF* gene) and this enzyme is involved in the metabolism of MGH 78578 strain. Therefore, this process is nonessential if this pathogen can utilize the pantothenic acid (HMDB ID: HMDB0000210) from the host environment. Although *panB, panC* and *panD* were predicted to be essential in the SM medium, they were not essential for the survival in the HBF. This is because the HBF includes this compound. This phenomenon shows that selection of the growth medium is crucial for accurate essentiality predictions.

Synthesis or uptake of pantothenate is essential in the five-step universal pathway of CoA biosynthesis. To aid this process, pantothenate kinase (the *coaA* gene product) mediates conversion of the pantothenate to 4'-phosphopantetheine in in most organisms [177, 178, 182]. This rate limiting step is essential for the bacterial growth. It is a potential target for designing antimicrobials against pathogenic bacteria due to the structural differences with the human counterpart. The structure of PanK enzyme from *K. pneumoniae* was solved by co-crystallization with N-[2-(1,3-benzodioxol-5-yl)ethyl] pantothenamide in 2014 and it may be useful for the target based drug design [183]. After the PanK-catalysed step of the CoA biosynthesis, condensation of 4′-phosphopanthenate with cysteine and decarboxylation of the product are mediated by a bifunctional enzyme (the product of *coaBC*/*dfp* gene in iYL1228/iKp1289) [177]. It

was proposed as a target for *K. pneumoniae* infections in this work. This enzyme (phosphopantothenoylcysteine synthase/decarboxylase) is found in various bacteria and mostly shares a homology with *E. coli coaBC* gene. Phosphopantothenoylcysteine synthase/decarboxylase of *Mycobacterium tuberculosis* was reported as essential for *in vitro* and *in vivo* conditions. Therefore, the silencing of *coaBC* manifests a bactericidal effect and it was also suggested as a new bactericidal drug target for *M. tuberculosis* infections by Evans and colleagues [184]. Next step of the CoA synthesis is formation of dephospho-CoA catalyzed by 4'-phosphopantetheine adenylyltransferase (the *coaD* gene product) and eventually formation of CoA via phosphorylation of the product by dephospho-CoA kinase (the *coaE* gene product) [177]. 4'-phosphopantetheine adenylyltransferase is a prominent target due to its structurally conservation among bacteria and a distant relationship with the human counterpart. Therefore, intracellular activity of CoaD inhibitors were screened for both gram-positive and gram-negative bacteria to support the design of efficient compounds in different studies [179, 185]. Collectively, the transporter PanF can likely compensate the suppression of the pantothenic acid synthesis in *K. pneumoniae* so it may not be an efficient target. However, other genes related to coenzyme A synthesis are promising targets. Nevertheless, a further analysis of these targets is necesary to validate the suppressive effect on the *Klebsiella*-mediated infections.

Riboflavin (vitamin B2) is an essential precursor of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) that can be synthesized by many microorganisms and plants [186, 187]. Animals and humans cannot synthesize this vitamin so they must obtain it through the diet [187]. Broad spectrum antibiotics acting as a potent inhibitor of the riboflavin synthesis were reported [188]. Biosynthesis of the riboflavin is managed by GTP cyclohydrolase II (the *ribA* gene product), 3,4-dihydroxy-2-butanone-4-phosphate (DHBP) synthase (the *ribB* gene product), riboflavin synthase α subunit (the *ribC* gene product), riboflavin deaminase/reductase (the *ribD* gene product), and ribityl lumazine synthase (the *ribE* gene product) in *E. coli* [187]. Riboflavin synthesis is essential for the gram-negative bacteria including *E. coli* and *Salmonella* sp. because they do not have any uptake systems for flavins. Furthermore, *ribC* gene was also reported to be essential in *Haemophilus influenzae* Rd strain KW20 [189]. Thus, the enzymes responsible for riboflavin biosynthesis may be drug targets depending on the presence of uptake system [187]. Three essential genes (*ribC*, *ribE* and *ribH*) from SM simulation of iYL1228 were related to the

riboflavin synthesis in *Klebsiella* and they were found to be non-homologous and druggable in this study.

In addition to the gene-centric approach, some putative drug targets were also elucidated based on the metabolite essentiality in this study. All simulations of both GEMs predicted a single essential metabolite (uaagmda) involved in one essential outgoing reaction (murein polymerizing transglycosylase reaction) in addition to two nonessential reactions (UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase and the murein polymerizing transglycosylase 2 (three linked units) reaction. Then, three essential genes (*mrcB*, *pbpC* and *mrcA*) associated with this essential reaction were identified.

The murein (peptidoglycan) is a cross-linked polymer in the periplasm that protects the gram-negative bacteria from a high intracellular pressure [190, 191]. Transfer of peptidyl disaccharide subunit from lipid II to the growing glycan chain is a necessary step for synthesis of the cell wall peptidoglycan, and this polymerization process is managed by glycosyltransferases. Another essential enzyme involved in the peptidoglycan biosynthesis is transpeptidase, and it functions in the production of cross-linked peptidoglycan from lipid intermediates [192]. Penicillin-binding proteins (PBPs) are divided into two classes: (1) bifunctional enzymes (class A) responsible for both polymerization and transpeptidation and monofunctional enzymes (class B) with only transpeptidation activity. There are many antibiotics (e.g., penicillin) targeting transpeptidases to prevent peptidoglycan synthesis because inhibition of its synthesis results in the bacterial death via lysis [193]. Therefore, these targets seem to be convenient to cope with the *Klebsiella* infections. Moreover, the drugs identified through DrugBank can be evaluated or a structural analog of ‘uaagmda’ can be designed to block the related enzymes.

Whole gene list which includes the suggested putative drug targets in the scope of this thesis is given in Table 5.3. Herein, the initial findings regarding metabolisms of two different *Klebsiella* strains were reported and eventually the potential drug targets for these strains were suggested. Essential, druggable and non-homologous targets including virulence factors, cofactor synthesis-related genes and the putative targets from a metabolic-centric approach were mainly evaluated in the study. Investigation of the other non-homologous, druggable, essential genes may pave the way for the detection of additional target candidates. Of these genes, some of them such as *murA*, *murG* and *murD* were revealed as the putative targets by Georrge and

colleagues via an *in silico* proteome-based study [83]. Therefore, there is a need for a further evaluation of essential, druggable targets predicted in this study.

Table 5.3: Drug target list suggested in this study.

Method	**Gene**	**Description**	**Function**
Gene-centric approach (original biomass)	*hldD* (*rfaD*)	ADP-L-glycero-D-manno-heptose-6-epimerase	LPS biosynthesis
	lpxA	UDP-N-acetylglucosamine acyltransferase	
	lpxC	UDP-3-O-acyl N-acetylglucosamine deacetylase	
	lpcA (g*mhA*)	Phosphoheptose isomerase	
	kdsA	2-dehydro-3-deoxyphosphooctonate aldolase	
Gene-centric approach (updated biomass)	*nadD*	Nicotinic acid mononucleotide adenyltransferase	Cofactor production
	nadE	NAD(+) synthetase	
	coaA	Pantothenate kinase	
	dfp (*coaBC*)	Phosphopantothenoylcysteine synthase/decarboxylase	
	coaD	Phosphopantetheine adenylyltransferase	
	panB	3-methyl-2-oxobutanoate hydroxymethyltransferase	
	panC	Pantoate-beta-alanine ligase	
	panD	Aspartate 1-decarboxylase precursor	
	ribC	Riboflavin synthase subunit alpha	
	ribF	Hypothetical protein	
	ribH	Riboflavin synthase subunit beta	
Metabolite-centric approach	*mrcA*	Bifunctional penicillin-binding protein 1a: transglycosylase/transpeptidase	Peptidoglycan synthesis
	mrcB	Penicillin-binding protein 1b	
	pbpC	Penicillin-binding protein 1C (PBP 1C)	

6. CONCLUSIONS

A system-level analysis of the bacterial metabolisms is the requisite to decipher the tremendous, complicated systems and to discover effective therapeutic approaches. Herein, putative drug targets for two different *K. pneumoniae* strains were determined through two approaches: gene- and metabolic-centric methods.

Single gene deletion was taken into consideration in the gene-centric approach and it was based on the completely deletion of each gene to detect the impact of this intervention on the bacterial growth. In other words, the whole related reactions were completely blocked in this approach. On the other hand, high drug levels may be required to completely inhibit these enzymes. Therefore, if the pathogen can be killed through a partly reduction of an enzymatic activity (but not completely inhibition of the related reactions), the gene encoding this enzyme may be more efficient drug target. Moreover, targeting such genes may reduce drug dose. A reduction in the drug concentration may decrease a negative impact of the drug on the gut microbiota. Protection of these bacteria is significant to combat antibiotic-resistant pathogens. It should be noted that the targets of *K. pneumoniae* suggested in this study share a homology with many gut bacteria based on the 'non-homology analysis against gut microbiota proteomes' through PBIT [122]. Thus, new approaches are necessay to reduce any serious damages to gut microbiota. In addition to gut microbiota, the host organism may also have structurally similar proteins. Thus, reduction of the drug dose may decrease any damages due to off-target effect. Therefore, the proposed targets should be also evaluated in terms of structural homology.

Identification of the proper putative drug targets depends on simulation of the real growth conditions of the pathogens. To construct a more realistic environment, host-mimicking media were used in the study. In addition, an important insight into condition-specific and more realistic behivour of the pathogen can be provided by integration of the expression data. More recently, an expression data (GEO accession: GSE110628) of some pathogens including *K. pneumoniae* strain MGH 78578 exposed to pulmonary surfactant has been published. The pulmonary surfactant contributes colonization and pathogenesis of the bacteria (e.g., lipopolysaccharide modification, biofilm formation, antibiotic resistance and capsule production) via induction of virulence gene expression [86]. Thus, this data can help to elucidate the interactions

between pathogen and host at the infection site. Moreover, it may support to identify additional drug targets.

Importantly, most of the putative targets in MGH 78578 and KPPR1 strains were found to be common in this work. Therefore, examination of these strains (with the different virulence levels) in terms of their virulence characteristics may provide more promising results to distinguish their metabolic backgrounds. In this context, virulence factor synthesis may be selected as the objective function in the simulations [194]. Considering the virulence-linked targets can delay the spread of the antibiotic resistance, this approach may also reveal more conserved drug targets. Of note, synthesis of the virulence factors is highly connected to the growth phase of the bacteria. In this regard, the growth phase dependent-expression data of *K. pneumoniae* MGH 78578 was published in 2012 (GEO accession: GSE35926) may be guiding in future studies [85].

Briefly, introduction and use of alternative approaches may provide discovery of more efficient and selective alternative targets in future studies.

REFERENCES

[1] Bachman M. A., Breen P., Deornellas V., Mu Q., Zhao L., Wu W., Cavalcoli J. D., (2015), "Genome-Wide Identification of Klebsiella pneumoniae Fitness Genes during Lung Infection", MBio, 6(3), 1–9.

[2] Tang Y.-W., Sussman M., Liu D., Poxton I., Schwartzman J., (2015), "Molecular Medical Microbiology", Academic press.

[3] Doorduijn D. J., Rooijakkers S. H. M., van Schaik W., Bardoel B. W., (2016), "Complement resistance mechanisms of Klebsiella pneumoniae", Immunobiology, 221(10), 1102–1109.

[4] Paczosa M. K., Mecsas J., (2016), "Klebsiella pneumoniae: Going on the Offense with a Strong Defense", Microbiol Mol Biol Rev, 80(3), 629–661.

[5] Taneja N., Kaur H., (2016), "Insights into Newer Antimicrobial Agents against Gram-negative Bacteria", Microbiol Insights, 9, 9–19.

[6] Santajit S., Indrawattana N., (2016), "Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens", Biomed Res Int, 2016.

[7] Sun W., Weingarten R. A., Xu M., Southall N., Dai S., Shinn P., Sanderson P. E., Williamson P. R., Frank K. M., Zheng W., (2016), "Rapid antimicrobial susceptibility test for identification of new therapeutics and drug combinations against multidrug-resistant bacteria", Emerg Microbes Infect, 5(11), e116.

[8] Kim H. U., Sohn S. B., Lee S. Y., (2012), "Metabolic network modeling and simulation for drug targeting and discovery", Biotechnol J, 7(3), 330–342.

[9] Chandra N., (2011), "Computational approaches for drug target identification in pathogenic diseases", Expert Opin Drug Discov, 6(10), 975–979.

[10] Pratap S., Kesari P., Yadav R., Dev A., Narwal M., Kumar P., (2017), "Acyl chain preference and inhibitor identification of Moraxella catarrhalis LpxA : Insight through crystal structure and computational studies", Int J Biol Macromol, 96, 759–765.

[11] Barb A. W., Zhou P., (2008), "Mechanism and Inhibition of LpxC: an essential zinc-dependent deacetylase of bacterial lipid A synthesis", Curr Pharm Biotechnol, 9(1), 9–15.

[12] Joo S. H., (2015), "Lipid A as a Drug Target and Therapeutic Molecule",23(6), 510–516.

[13] Perumal D., Lim C. S., Sakharkar M. K., (2007), "In silico Identification of Putative Drug Targets in Pseudomonas aeruginosa Through Metabolic Pathway Analysis". In: Pattern Recognition in Bioinformatics, Springer Berlin

Heidelberg, pp 323–336.

[14] Amineni U., Pradhan D., Marisetty H., (2010), "In silico identification of common putative drug targets in Leptospira interrogans", J Chem Biol, 3(4), 165–173.

[15] Kidd T. J., Mills G., Sá-Pessoa J., Dumigan A., Frank C. G., Insua J. L., Ingram R., Hobley L., Bengoechea J. A., (2017), "A *Klebsiella pneumoniae* antibiotic resistance mechanism that subdues host defences and promotes virulence", EMBO Mol Med, 9(4), 430–447.

[16] Navon-Venezia S., Kondratyeva K., Carattoli A., (2017), "Klebsiella pneumoniae: A major worldwide source and shuttle for antibiotic resistance", FEMS Microbiol Rev, 41(3), 252–275.

[17] Martin R. M., Bachman M., (2018), "Colonization, Infection, and the Accessory Genome of Klebsiella pneumoniae", Front Cell Infect Microbiol, 8(January), 4.

[18] Yayan J., Ghebremedhin B., Rasche K., (2015), "No carbapenem resistance in pneumonia caused by klebsiella species", Med (United States), 94(6), 1–10.

[19] Li B., Zhao Y., Liu C., Chen Z., Zhou D., (2014), "Molecular pathogenesis of Klebsiella pneumoniae.", Future Microbiol, 9(9), 1071–81.

[20] Chen P., Seth A. K., Abercrombie J. J., Mustoe T. A., Leung K. P., (2014), "Activity of imipenem against Klebsiella pneumoniae Biofilms in vitro and in vivo", Antimicrob Agents Chemother, 58(2), 1208–1213.

[21] Crofts T. S., Gasparrini A. J., Dantas G., (2017), "Next-generation approaches to understand and combat the antibiotic resistome", Nat Publ Gr, 15(7), 422–434.

[22] Web 1, (2018), https://www.ncbi.nlm.nih.gov/pubmed/, (Erişim Tarihi: 03.06.2016).

[23] Sommer M. O. A., Munck C., Toft-Kehler R. V., Andersson D. I., (2017), "Prediction of antibiotic resistance: Time for a new preclinical paradigm?", Nat Rev Microbiol, 15(11), 689–696.

[24] Munita J. M., Arias C. A., (2016), "Mechanisms of Antibiotic Resistance", Microbiol Spectr, 4(2).

[25] Hooper D. C., (2000), "Mechanisms of Action and Resistance of Older and Newer Fluoroquinolones", Clin Infect Dis, 31, S24–S28.

[26] Redgrave L. S., Sutton S. B., Webber M. A., Piddock L. J. V, (2014), "Fluoroquinolone resistance: Mechanisms, impact on bacteria, and role in evolutionary success", Trends Microbiol, 22(8), 438–445.

[27] Huddleston J. R., (2014), "Horizontal gene transfer in the human

gastrointestinal tract: Potential spread of antibiotic resistance genes", Infect Drug Resist, 7, 167–176.

[28] Bhattacharjee M. K., (2016), "Chemistry of Antibiotics and Related Drugs". Springer, Switzerland, pp 27–49.

[29] Woo P. C. Y., To A. P. C., Lau S. K. P., Yuen K. Y., (2003), "Facilitation of horizontal transfer of antimicrobial resistance by transformation of antibiotic-induced cell-wall-deficient bacteria", Med Hypotheses, 61(4), 503–508.

[30] Colomer-Lluch M., Jofre J., Muniesa M., (2011), "Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples", PLoS One, 6(3).

[31] Blair J. M. A., Webber M. A., Baylay A. J., Ogbolu D. O., Piddock L. J. V, (2015), "Molecular mechanisms of antibiotic resistance", Nat Rev Microbiol, 13(1), 42–51.

[32] Van Hoek A. H. A. M., Mevius D., Guerra B., Mullany P., Roberts A. P., Aarts H. J. M., (2011), "Acquired antibiotic resistance genes: An overview", Front Microbiol, 2(203).

[33] Andersson D. I., Hughes D., Kubicek-Sutherland J. Z., (2016), "Mechanisms and consequences of bacterial resistance to antimicrobial peptides", Drug Resist Updat, 26, 43–57.

[34] Poirel L., Jayol A., Bontron S., Villegas M. V., Ozdamar M., Türkoglu S., Nordmann P., (2015), "The mgrB gene as a key target for acquired resistance to colistin in Klebsiella pneumoniae", J Antimicrob Chemother, 70(1), 75–80.

[35] Pragasam A. K., Shankar C., Veeraraghavan B., Biswas I., Nabarro L. E. B., Inbanathan F. Y., George B., Verghese S., (2017), "Molecular mechanisms of colistin resistance in klebsiella pneumoniae causing bacteremia from INDIA-A first report", Front Microbiol, 7(2135).

[36] Olaitan A. O., Morand S., Rolain J. M., (2014), "Mechanisms of polymyxin resistance: Acquired and intrinsic resistance in bacteria", Front Microbiol, 5(643).

[37] Kumar V., Sun P., Vamathevan J., Li Y., Ingraham K., Palmer L., Huang J., Brown J. R., (2011), "Comparative genomics of Klebsiella pneumoniae strains with different antibiotic resistance profiles", Antimicrob Agents Chemother, 55(9), 4267–4276.

[38] Huang W., Wang G., Sebra R., Zhuge J., Yin C., Fallon J. T., (2017), "Emergence and Evolution of Multidrug- Resistant Klebsiella pneumoniae with both blaKPC and blaCTX-M Integrated in the Chromosome", Antimicrob Agents Chemother, 61(7).

[39] Llobet E., Martínez-Moliner V., Moranta D., Dahlström K. M., Regueiro V., Tomás A., Cano V., Pérez-Gutiérrez C., Frank C. G., Fernández-Carrasco H.,

Insua J. L., Salminen T. A., Garmendia J., Bengoechea J. A., (2015), "Deciphering tissue-induced *Klebsiella pneumoniae* lipid A structure", Proc Natl Acad Sci, 112(46), E6369–E6378.

[40] Trevor F. J., Snow G. A., (2005), "Biochemistry and Molecular Biology of Antimicrobial Drug Action", Springer, Boston, MA.

[41] Thornley M., Horne R., (1962), "Electron microscope observations on the structure of fimbriae, with particular reference to Klebsiella strains, by the use of the negative staining technique", J Gen Microbiol, 28, 51–56.

[42] Clegg S., Murphy C. N., (2016), "Epidemiology and Virulence of Klebsiella pneumoniae", Microbiol Spectr, 4(1).

[43] Gomez-Simmonds A., Uhlemann A.-C., (2017), "Clinical Implications of Genomic Adaptation and Evolution of Carbapenem-Resistant Klebsiella pneumoniae", J Infect Dis, 215, S18–S27.

[44] Web 2, (2018), https://microbeonline.com/general-and-differential-characteristics-of-gram-positive-and-gram-negative-bacteria/%0A, (Erişim tarihi: 12.06.2018).

[45] Sader H. S., Castanheira M., Flamm R. K., (2017), "Antimicrobial activity of ceftazidime-avibactam against Gram-negative bacteria isolated from patients hospitalized with pneumonia in U.S. medical centers, 2011 to 2015", Antimicrob Agents Chemother, 61(4), 1–10.

[46] Henry C. S., Rotman E., Lathem W. W., Tyo K. E. J., Hauser A. R., Mandel M. J., (2017), "Generation and validation of the iKp1289 Metabolic model for klebsiella pneumoniae KPPR1", J Infect Dis, 215, S37–S43.

[47] Web 3, (2013), https://www.cdc.gov/drugresistance/threat-report-2013/. (Erişim tarihi: 22.09.2017).

[48] Jana B., Cain A. K., Doerrler W. T., Boinett C. J., Fookes M. C., Parkhill J., Guardabassi L., (2017), "The secondary resistome of multidrug-resistant Klebsiella pneumoniae", Sci Rep, 7, 1–10.

[49] Web 4, (2018), https://resistancemap.cddep.org/, (Erişim tarihi: 12.06.2018).

[50] Eftekhar F., Naseh Z., (2015), "Extended-spectrum beta-lactamase and carbapenemase production among burn and non-burn clinical isolates of Klebsiella pneumoniae", Iran J Microbiol, 7(3), 144–149.

[51] Nguyen T. K. P., Tran T. H., Roberts C. L., Graham S. M., Marais B. J., (2017), "Child pneumonia – focus on the Western Pacific Region", Paediatr Respir Rev, 21, 102–110.

[52] Bassetti M., Giacobbe D. R., Giamarellou H., Viscoli C., Daikos G. L., Dimopoulos G., De Rosa F. G., Giamarellos-Bourboulis E. J., Rossolini G. M.,

Righi E., Karaiskos I., Tumbarello M., Nicolau D. P., Viale P. L., Poulakou G., (2018), "Management of KPC-producing Klebsiella pneumoniae infections", Clin Microbiol Infect, 24(2), 171–174.

[53] Matsumoto S., Singley C. M., Hoover J., Nakamura R., Echols R., Rittenhouse S., Tsuji M., Yamano Y., (2017), "Efficacy of Cefiderocol against Carbapenem-Resistant Gram-Negative Bacilli in Immunocompetent-Rat Respiratory Tract Infection Models Recreating Human Plasma Pharmacokinetics", Antimicrob Agents Chemother, 61(9), e00700-17.

[54] Fernandes P., Martens E., (2017), "Antibiotics in late clinical development", Biochem Pharmacol, 133, 152–163.

[55] Sader H. S., Castanheira M., Flamm R. K., (2017), "crossm Avibactam against Gram-Negative Bacteria Isolated from Patients", Antimicrob Agents Chemother, 61(4), 1–10.

[56] Castanheira M., Rhomberg P. R., Flamm R. K., Jones R. N., (2016), "Effect of the β-lactamase inhibitor vaborbactam combined with meropenem against serine carbapenemase-producing Enterobacteriaceae", Antimicrob Agents Chemother, 60(9), 5454–5458.

[57] Bunnik E. M., Le Roch K. G., (2013), "An Introduction to Functional Genomics and Systems Biology.", Adv wound care, 2(9), 490–498.

[58] Durmuş S., Ülgen K., (2013), "Systems biology of pathogen-host interaction: Networks of protein-protein interaction within pathogens and pathogen-human interactions in the post-genomic era", Biotechnol J, 8(1), 85–96.

[59] Dix A., Vlaic S., Guthke R., Linde J., (2016), "Use of systems biology to decipher host-pathogen interaction networks and predict biomarkers", Clin Microbiol Infect, 22(7), 600–606.

[60] Durmus S., Çakir T., Özgür A., Guthke R., (2015), "A review on computational systems biology of pathogen-host interactions", Front Microbiol, 6:235.

[61] Zhang W., Li F., Nie L., (2010), "Integrating multiple “omics” analysis for microbial biology: Application and methodologies", Microbiology, 156(2), 287–301.

[62] Lengeling A., Pfeffer K., Balling R., (2001), "The battle of two genomes: Genetics of bacterial host/pathogen interactions in mice", Mamm Genome, 12(4), 261–271.

[63] Forst C. V, (2006), "Host – pathogen systems biology", Handb Stat Syst Biol, 11(5–6), 220–7.

[64] Oberhardt M. A., Yizhak K., Ruppin E., (2013), "Metabolically re-modeling the drug pipeline", Curr Opin Pharmacol, 13(5), 778–785.

[65] Nielsen J., Keasling J. D., (2016), "Engineering Cellular Metabolism", Cell, 164(6), 1185–1197.

[66] Feist A. M., Henry C. S., Reed J. L., Krummenacker M., Joyce A. R., Karp P. D., Broadbelt L. J., Hatzimanikatis V., Palsson B., (2007), "A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information", Mol Syst Biol, 3(121), 1–18.

[67] Herrgård M. J., Swainston N., Dobson P., Dunn W. B., Yalçin K., Arvas M., Blüthgen N., Borger S., Costenoble R., (2014), "A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology", Nat Biotechnol, 26(10), 1155–1160.

[68] Töpfer N., Kleessen S., Nikoloski Z., (2015), "Integration of metabolomics data into metabolic networks"6, 1–13.

[69] Stanford N. J., Lubitz T., Smallbone K., Klipp E., Mendes P., (2013), "Systematic Construction of Kinetic Models from Genome- Scale Metabolic Networks"8(11).

[70] Terzer M., Maynard N. D., Covert M. W., Stelling J., (2009), "Genome-scale metabolic networks", Wiley Interdiscip Rev Biol Med, 1(3), 285–297.

[71] Huthmacher C., Hoppe A., Bulik S., Holzhütter H., (2010), "Antimalarial drug targets in Plasmodium falciparum predicted by stage-specific metabolic network analysis", BMC Syst Biol, 4:120.

[72] Bazzani S., Hoppe A., Holzhütter H., (2012), "Network-based assessment of the selectivity of metabolic drug targets in Plasmodium falciparum with respect to human liver metabolism", BMC Syst Biol, 6, 118.

[73] García Sánchez C. E., Torres Sáez R. G., (2014), "Comparison and analysis of objective functions in flux balance analysis", Biotechnol Prog, 30(5), 985–991.

[74] Liao Y., Huang T., Chen F., Charusanti P., Hong J. S. J., Chang H., Tsai S., Palsson B. O., Hsiung C. A., (2011), "An experimentally validated genome-scale metabolic reconstruction of Klebsiella pneumoniae MGH 78578, iYL1228", J Bacteriol, 193(7), 1710–1717.

[75] Henson M. A., Hanly T. J., (2014), "Dynamic flux balance analysis for synthetic microbial communities", IET Syst Biol, 8(5), 214–229.

[76] Wintermute E. H., Lieberman T. D., Silver P. A., (2013), "An objective function exploiting suboptimal solutions in metabolic networks", BMC Syst Biol, 7(1), 1.

[77] Alper H., Jin Y. S., Moxley J. F., Stephanopoulos G., (2005), "Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli", Metab Eng, 7(3), 155–164.

[78] Raman K., Rajagopalan P., Chandra N., (2005), "Flux Balance Analysis of Mycolic Acid Pathway: Targets for Anti-Tubercular Drugs", PLoS Comput Biol, 1(5), 0349–0358.

[79] Plata G., Hsiao T.-L., Olszewski K. L., Llinás M., Vitkup D., (2010), "Reconstruction and flux-balance analysis of the Plasmodium falciparum metabolic network", Mol Syst Biol, 6:408.

[80] Lewis N. E., Hixson K. K., Conrad T. M., Lerman J. A., Charusanti P., Polpitiya A. D., Adkins J. N., Schramm G., Purvine S. O., Lopez-Ferrer D., Weitz K. K., Eils R., König R., Smith R. D., Palsson B., (2010), "Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models", Mol Syst Biol, 6(390).

[81] Covert M. W., Schilling C. H., Palsson B., (2001), "Regulation of gene expression in flux balance models of metabolism", J Theor Biol, 213(1), 73–88.

[82] Tarlak F., Sadıkoğlu H., Çakır T., (2014), "The role of flexibility and optimality in the prediction of intracellular fluxes of microbial central carbon metabolism", Mol Biosyst, 10(9), 2459.

[83] Georrge J. J., Umrania V., (2011), "In silico identification of putative drug targets in Klebsiella pneumonia MGH78578", Indian J Biotechnol, 10(4), 432–439.

[84] Khan A., Sharma D., Faheem M., Bisht D., Khan A. U., (2017), "Proteomic analysis of a carbapenem-resistant Klebsiella pneumoniae strain in response to meropenem stress", J Glob Antimicrob Resist, 8, 172–178.

[85] Seo J. H., Hong J. S., Kim D., Cho B. K., Huang T. W., Tsai S. F., Palsson B. O., Charusanti P., (2012), "Multiple-omic data analysis of Klebsiella pneumoniae MGH 78578 reveals its transcriptional architecture and regulatory features", BMC Genomics, 13(1).

[86] Willsey G. G., Ventrone S., Schutz K. C., Wallace A. M., Ribis J. W., Suratt B. T., Wargo M. J., (2018), "Pulmonary surfactant promotes virulence gene expression and biofilm formation in Klebsiella pneumoniae", Infect Immun, 86(7).

[87] Allen B., Drake M., Harris N., Sullivan T., (2017), "Using KBase to assemble and annotate prokaryotic genomes", Curr Protoc Microbiol, 46, 1E.13.1-1E.13.18.

[88] Cesur, Müberra Fatma Abdik E., Güven-Gülhan, Ünzile Durmuş S., Çakır T., (2018), "Metabolic Interaction in Infection". In: Silvestre R, Torrado E (eds). Springer, Cham.

[89] Mienda B. S., Salihu R., Adamu A., Idris S., (2018), "Genome-scale metabolic models as platforms for identification of novel genes as antimicrobial drug targets", Future Microbiol, 13(4), 455–467.

[90] Milne C., Kim P., Eddy J., Price N., (2011), "Accomplishments in genome-scale in silico modeling for industrial and medical biotechnology", Biotechnol J, 4(12), 1653–1670.

[91] Bordbar A., Lewis N. E., Schellenberger J., Palsson B. Ø., Jamshidi N., (2010), "Insight into human alveolar macrophage and M . tuberculosis interactions via metabolic reconstructions", Mol Syst Biol, 6(422).

[92] Oberhardt M., Goldberg J., Hogardt M., Papin J., (2010), "Metabolic Network Analysis of Pseudomonas aeruginosa during Chronic Cystic Fibrosis Lung Infection", J Bacterio, 192(20), 5534–5548.

[93] Ahn Y., Lee D., Burd H., Blank W., Kapatral V., (2014), "Metabolic Network Analysis-Based Identification of Antimicrobial Drug Targets in Category A Bioterrorism Agents", PLoS One, 9(1), e85195.

[94] Plata G., Hsiao T. L., Olszewski K. L., Llinás M., Vitkup D., (2010), "Reconstruction and flux-balance analysis of the Plasmodium falciparum metabolic network", Mol Syst Biol, 6(408).

[95] Steinway S., Biggs M., Jr L. T., JA P., Albert R., (2015), "Inference of Network Dynamics and Metabolic Interactions in the Gut Microbiome", PLoS Comput Bio, 11(5), e1004338.

[96] Larocque M., Chénard T., Najmanovich R., (2014), "A curated C. difficile strain 630 metabolic network: prediction of essential targets and inhibitors", BMC Syst Bio, 8:117.

[97] Kim H. U., Kim T. Y., Lee S. Y., (2010), "Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen Acinetobacter baumannii AYE", Mol BioSyst, 6(2), 339–348.

[98] Sarker M., Chopra S., Mortelmans K., Kodukula K., Talcott C., (2011), "In Silico Pathway Analysis Predicts Metabolites that are Potential Antimicrobial Targets", J Comput Sci Syst Biol, 4(2), 21–26.

[99] Kim H. U., Kim S. Y., Jeong H., Kim T. Y., Kim J. J., Choy H. E., Yi K. Y., (2011), "Integrative genome-scale metabolic analysis of Vibrio vulnificus for drug targeting and discovery", Mol Syst Biol, 7:460.

[100] Ramirez M. S., Xie G., Marshall S. H., Hujer K. M., Chain P. S. G., Bonomo R. A., Tolmasky M. E., (2012), "Multidrug-resistant (MDR) Klebsiella pneumoniae clinical isolates: A zone of high heterogeneity (HHZ) as a tool for epidemiological studies", Clin Microbiol Infect, 18(7), E254–E258.

[101] Broberg C. A., Wu W., Cavalcoli J. D., Miller V. L., Bachman M. A., (2014), "Complete Genome Sequence of Klebsiella pneumoniae Strain ATCC 43816 KPPR1, a Rifampin-Resistant Mutant Commonly Used in Animal, Genetic, and Molecular Biology Studies", Genome Announc, 2(5), e00924-14-e00924-14.

[102] Lawlor M. S., Hsu J., Rick P. D., Miller V. L., (2005), "Identification of Klebsiella pneumoniae virulence determinants using an intranasal infection model", Mol Microbiol, 58(4), 1054–1073.

[103] Keating S. M., Bornstein B. J., Finney A., Hucka M., (2006), "SBMLToolbox: An SBML toolbox for MATLAB users", Bioinformatics, 22(10), 1275–1277.

[104] Pratapa A., Balachandran S., Raman K., (2015), "Fast-SL: an efficient algorithm to identify synthetic lethal sets in metabolic networks", Bioinformatics, 31(20), 3299–305.

[105] Fukuyama H., Yamashiro S., Kinjo K., Tamaki H., Kishaba T., (2014), "Validation of sputum Gram stain for treatment of community-acquired pneumonia and healthcare-associated pneumonia: A prospective observational study", BMC Infect Dis, 14(1).

[106] Cano V., March C., Insua J. L., Aguiló N., Llobet E., Moranta D., Regueiro V., Brennan G. P., Millán-Lou M. I., Martín C., Garmendia J., Bengoechea J. A., (2015), "Klebsiella pneumoniae survives within macrophages by avoiding delivery to lysosomes", Cell Microbiol, 17(11), 1537–1560.

[107] Goroll A. H., Mulley A. G. J., (2009), "Primary care medicine: office evaluation and management of the adult patient", Philadelphia: Wolters Kluwer/Lippincott Williams and Wilkins Health.

[108] Hadi M., Marashi S.-A., (2014), "Reconstruction of a generic metabolic network model of cancer cells.", Mol Biosyst, 10(11), 3014–21.

[109] Goyal N., Padhiary M., Karimi I. A., Zhou Z., (2015), "Flux measurements and maintenance energy for carbon dioxide utilization by Methanococcus maripaludis", Microb Cell Fact, 14(145).

[110] Haggart C. R., Bartell J. A., Saucerman J. J., Papin J. A., (2008), "Whole-genome metabolic network reconstruction and constraint-based modeling", Methods Enzym, 29(10), 1883–1889.

[111] King Z. A., Lu J., Miller P., Federowicz S., Lerman J. A., Ebrahim A., Palsson B. O., Lewis N. E., (2016), "BiGG Models: A platform for integrating, standardizing and sharing genome-scale models", Nucleic Acids Res, 44, 515–522.

[112] Xavier J. C., Patil K. R., Rocha I., (2017), "Integration of Biomass Formulations of Genome-Scale Metabolic Models with Experimental Data Reveals Universally Essential Cofactors in Prokaryotes", Metab Eng, 39, 200–208.

[113] Orth J. D., Conrad T. M., Na J., Lerman J. A., Nam H., Feist A. M., Palsson B. Ø., (2011), "A comprehensive genome-scale reconstruction of Escherichia coli metabolism — 2011", Mol Syst Biol, 7(535).

[114] Reed J. L., Vo T. D., Schilling C. H., Palsson B. O., (2003), "An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR)", Genome Biol, 4(9), R54.

[115] Coordinators N. R., (2015), "Database resources of the National Center for Biotechnology Information", Nucleic Acids Res, 43(D1), D6–D17.

[116] Yu C.-S., Cheng C.-W., Su W.-C., Chang K.-C., Huang S.-W., Hwang J.-K., Lu C.-H., (2014), "CELLO2GO: A Web Server for Protein subCELlular LOcalization Prediction with Functional Gene Ontology Annotation", PLoS One, 9(6), e99368.

[117] Yu N. Y., Wagner J. R., Laird M. R., Melli G., Rey S., Lo R., Dao P., Cenk Sahinalp S., Ester M., Foster L. J., Brinkman F. S. L., (2010), "PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes", Bioinformatics, 26(13), 1608–1615.

[118] Xiao X., Wu Z. C., Chou K. C., (2011), "A multi-label classifier for predicting the subcellular localization of gram-negative bacterial proteins with both single and multiple sites", PLoS One, 6(6).

[119] Gupta S. K., Padmanabhan B. R., Diene S. M., Lopez-Rojas R., Kempf M., Landraud L., Rolain J. M., (2014), "ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes", Antimicrob Agents Chemother, 58(1), 212–220.

[120] Chen L., Yang J., Yu J., Yao Z., Sun L., Shen Y., Jin Q., (2005), "VFDB: A reference database for bacterial virulence factors", Nucleic Acids Res, 33(DATABASE ISS.), 325–328.

[121] Wishart D. S., Knox C., Guo A. C., Cheng D., Shrivastava S., Tzur D., Gautam B., Hassanali M., Tg C., (2008), "DrugBank: a knowledgebase for drugs, drug actions and drug targets", Nucleic Acids Res, 36(Database issue), D901–D906.

[122] Shende G., Haldankar H., Barai R. S., Bharmal M. H., Shetty V., Idicula-Thomas S., Hancock J., (2017), "PBIT: Pipeline builder for identification of drug targets for infectious diseases", Bioinformatics, 33(6), 929–931.

[123] Trupp M., Altman T., Fulcher C. A., Caspi R., Krummenacker M., Paley S., Karp P. D., (2010), "Beyond the genome (BTG) is a (PGDB) pathway genome database: HumanCyc", Genome Biol, 11, O12.

[124] Chagoyen M., Pazos F., (2011), "MBRole: Enrichment analysis of metabolomic data", Bioinformatics, 27(5), 730–731.

[125] Pruitt K. D., Tatusova T., Maglott D. R., (2007), "NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes , transcripts and proteins"35(Database issue), 61–65.

[126] Presta L., Bosi E., Mansouri L., Dijksho L., Fani R., (2017), "Constraint-based modeling identifies new putative targets to infections", 1–12.

[127] Uddin R., Saeed K., Khan W., Azam S., Wadood A., (2015), "Metabolic pathway analysis approach: identification of novel therapeutic target against methicillin resistant Staphylococcus aureus", Gene, 556(2), 213–226.

[128] Emanuelsson O., Brunak S., von Heijne G., Nielsen H., (2007), "Locating proteins in the cell using TargetP, SignalP and related tools", Nat Protoc, 2(4), 953–971.

[129] Yu C.-S., Lin C.-J., Hwang J.-K., (2004), "Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions.", Protein Sci, 13(5), 1402–6.

[130] Yu C., Chen Y., Lu C., Hwang J., (2006), "Prediction of Protein Subcellular Localization", Proteins Struct Funct Bioinforma, 64(3), 643–651.

[131] Wu C., Apweiler R., Bairoch A., Natale D., Barker W., Boeckmann B., Ferro S., Gasteiger E., Huang H., Lopez R., Magrane M., Martin M., Mazumder R., O'Donovan C., Redaschi N., Suzek B., (2006), "The Universal Protein Resource (UniProt): an expanding universe of protein information Cathy", Nucleic Acids Res, 34(90001), D187–D191.

[132] Lopez-Campistrous A., Semchuk P., Burke L., Palmer-Stone T., Brokx S. J., Broderick G., Bottorff D., Bolch S., Weiner J. H., Ellison M. J., (2005), "Localization, Annotation, and Comparison of the *Escherichia coli* K-12 Proteome under Two States of Growth", Mol Cell Proteomics, 4(8), 1205–1209.

[133] Scaria J., Chandramouli U., Verma S., (2005), "Antibiotic Resistance Genes Online (ARGO): A Database on vancomycin and β-lactam resistance genes", Bioinformation, 1(1), 5–7.

[134] Liu B., Pop M., (2009), "ARDB - Antibiotic resistance genes database", Nucleic Acids Res, 37, 443–447.

[135] Jia B., Raphenya A. R., Alcock B., Waglechner N., Guo P., Tsang K. K., Lago B. A., Dave B. M., Pereira S., Sharma A. N., Doshi S., Courtot M., Lo R., Williams L. E., Frye J. G., Elsayegh T., Sardar D., Westman E. L., Pawlowski A. C., Johnson T. A., Brinkman F. S. L., Wright G. D., McArthur A. G., (2017), "CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database", Nucleic Acids Res, 45(D1), D566–D573.

[136] Chen L., Zheng D., Liu B., Yang J., Jin Q., (2016), "VFDB 2016: Hierarchical and refined dataset for big data analysis - 10 years on", Nucleic Acids Res, 44(D1), D694–D697.

[137] Gawade P., Ghosh G. driven approach for identification of novel therapeutic targets in S. enterica P., (2018), "Genomics driven approach for identification of novel therapeutic targets in Salmonella enterica", Gene, 668(2017), 211–220.

[138] Holman A. G., Davis P. J., Foster J. M., Carlow C. K., Kumar S., (2009), "Computational prediction of essential genes in an unculturable endosymbiotic bacterium, Wolbachia of Brugia malayi", BMC Microbiol, 9(243).

[139] Duarte N. C., Becker S. A., Jamshidi N., Thiele I., Mo M. L., Vo T. D., Srivas R., Palsson B. O., (2007), "Global reconstruction of the human metabolic network based on genomic and bibliomic data", Proc Natl Acad Sci, 104(6), 1777–1782.

[140] Thiele I., Swainston N., Fleming R., Hoppe A., Sahoo S., Aurich M., Haraldsdottir H., Mo M., Rolfsson O., Stobbe M., Thorleifsson S., Agren R., Bölling C., Bordel S., Chavali A., Dobson P., Dunn W., Endler L., Hala D., Hucka M., Hull D., Jameson D., Jamshidi N., Jonsson J., (2013), "A community-driven global reconstruction of human metabolism", Nat Biotechnol, 31(5), 419–425.

[141] Kim H. U., Kim T., Lee Y. S., (2010), "Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen Acinetobacter baumannii AYE", Mol Biosyst, 6(2), 339–348.

[142] Samal A., Martin O. C., (2011), "Randomizing genome-scale metabolic networks", PLoS One, 6(7).

[143] Carbon S., Ireland A., Mungall C. J., Shu S., Marshall B., Lewis S., Lomax J., Mungall C., Hitz B., Balakrishnan R., Dolan M., Wood V., Hong E., Gaudet P., (2009), "AmiGO: Online access to ontology and annotation data", Bioinformatics, 25(2), 288–289.

[144] Raghunathan A., Reed J., Shin S., Palsson B., Daefler S., (2009), "Constraint-based analysis of metabolic capacity of Salmonella typhimurium during host-pathogen interaction", BMC Syst Biol, 3:38.

[145] Yao J., Rock C. O., (2016), "Resistance mechanisms and the future of bacterial enoyl-acyl carrier protein reductase (FabI) antibiotics", Cold Spring Harb Perspect Med, 6(3), 1–11.

[146] Knuth K., Niesalla H., Hueck C. J., Fuchs T. M., (2004), "Large-scale identification of essential Salmonella genes by trapping lethal insertions", Mol Microbiol, 51(6), 1729–1744.

[147] Baba T., Ara T., Hasegawa M., Takai Y., Okumura Y., Baba M., Datsenko K. A., Tomita M., Wanner B. L., Mori H., (2006), "Construction of Escherichia coli K-12 in-frame , single-gene knockout mutants: the Keio collection", 2(1).

[148] Finzel K., Nguyen C., Jackson D. R., Gupta A., Tsai S. C., Burkart M. D., (2015), "Probing the Substrate Specificity and Protein-Protein Interactions of the E. coli Fatty Acid Dehydratase, FabA", Chem Biol, 22(11), 1453–1460.

[149] Delaine T., Bernardes-Génisson V., Quémard A., Constant P., Cosledan F.,

Meunier B., Bernadou J., (2012), "Preliminary Investigations of the Effect of Lipophilic Analogues of the Active Metabolite of Isoniazid Toward Bacterial and Plasmodial Strains", Chem Biol Drug Des, 79(6), 1001–1006.

[150] Moynié L., Hope A. G., Finzel K., Schmidberger J., Leckie S. M., Schneider G., Burkart M. D., Smith A. D., Gray D. W., Naismith J. H., (2016), "A Substrate Mimic Allows High-Throughput Assay of the FabA Protein and Consequently the Identification of a Novel Inhibitor of Pseudomonas aeruginosa FabA", J Mol Biol, 428(1), 108–120.

[151] Zhang Y. M., White S. W., Rock C. O., (2006), "Inhibiting bacterial fatty acid synthesis", J Biol Chem, 281(26), 17541–17544.

[152] Sharma O. P., Kumar M. S., (2016), "Essential proteins and possible therapeutic targets of Wolbachia endosymbiont and development of FiloBase-a comprehensive drug target database for Lymphatic filariasis", Sci Rep, 6(October 2015), 1–11.

[153] Mochalkin I., Lightle S., Zhu Y., Ohren J. F., Spessard C., Chirgadze N. Y., Banotai C., Melnick M., Mcdowell L., (2007), "Characterization of substrate binding and catalysis in the potential antibacterial target N-acetylglucosamine-1-phosphate uridyltransferase (GlmU)", Protein Sci, 16(12), 2657–2666.

[154] Rani C., Mehra R., Sharma R., Chib R., Wazir P., Nargotra A., Khan I. A., (2015), "High-throughput screen identifies small molecule inhibitors targeting acetyltransferase activity of Mycobacterium tuberculosis GlmU", Tuberculosis, 95(6), 664–677.

[155] Buurman E. T., Andrews B., Gao N., Hu J., Keating T. A., Lahiri S., Otterbein L. R., Patten A. D., Stokes S. S., Shapiro A. B., (2011), "In vitro validation of acetyltransferase activity of GlmU as an antibacterial target in Haemophilus influenzae", J Biol Chem, 286(47), 40734–40742.

[156] Petráčková D., Janeček J., Bezoušková S., Kalachová L., Techniková Z., Buriánková K., Halada P., Haladová K., Weiser J., (2013), "Fitness and proteome changes accompanying the development of erythromycin resistance in a population of Escherichia coli grown in continuous culture", Microbiologyopen, 2(5), 841–852.

[157] Deacon A. M., Ni Y. S., Coleman W. G., Ealick S. E., (2000), "The crystal structure of ADP-L-glycero-D-mannoheptose 6-epimerase: Catalysis with a twist", Structure, 8(5), 453–462.

[158] Kuo C.-J., Chen J.-W., Chiu H.-C., Teng C.-H., Hsu T.-I., Lu P.-J., Syu W.-J., Wang S.-T., Chou T.-C., Chen C.-S., (2016), "Mutation of the Enterohemorrhagic Escherichia coli Core LPS Biosynthesis Enzyme RfaD Confers Hypersusceptibility to Host Intestinal Innate Immunity In vivo", Front Cell Infect Microbiol, 6(August), 1–14.

[159] Clements J. M., Coignard F., Johnson I., Palan S., Waller A., Wijkmans J.,

Hunter M. G., Chandler S., (2002), "Antibacterial Activities and Characterization of Novel Inhibitors of LpxC", Antimicrob Agents Chemother, 46(6), 1793.

[160] Li Y. Y., Jones S. J. M., (2012), "Drug repositioning for personalized medicine", Genome Med, 4(3), 1–14.

[161] Regué M., Hita B., Pique N., Izquierdo L., Merino S., Fresno S., Benedi V. J., Tomás J. M., (2004), "A Gene, uge, Is Essential for Klebsiella pneumoniae Virulence", Infect Immun, 72(1), 54–61.

[162] Strohmaier H., Remler P., Renner W., Hogenauer G., (1995), "Expression of genes kdsA and kdsB involved in 3-deoxy-D-manno-octulosonic acid metabolism and biosynthesis of enterobacterial lipopolysaccharide is growth phase regulated primarily at the transcriptional level in Escherichia coli K-12", J Bacteriol, 177(15), 4488–4500.

[163] Wattam A. R., Abraham D., Dalay O., Disz T. L., Driscoll T., Gabbard J. L., Gillespie J. J., Gough R., Hix D., Kenyon R., MacHi D., Mao C., Nordberg E. K., Olson R., Overbeek R., Pusch G. D., Shukla M., Schulman J., Stevens R. L., Sullivan D. E., Vonstein V., Warren A., Will R., Wilson M. J. C., Yoo H. S., Zhang C., Zhang Y., Sobral B. W., (2014), "PATRIC, the bacterial bioinformatics database and analysis resource", Nucleic Acids Res, 42(D1), 581–591.

[164] Luo H., Gao F., Lin Y., (2015), "Evolutionary conservation analysis between the essential and nonessential genes in bacterial genomes", Sci Rep, 5(July), 1–8.

[165] Kim H. U., Kim T. Y., Lee S. Y., (2010), "Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen Acinetobacter baumannii AYE", Mol Biosyst, 6(2), 339–348.

[166] Peyraud R., Cottret L., Marmiesse L., Gouzy J., Genin S., (2016), "A Resource Allocation Trade-Off between Virulence and Proliferation Drives Metabolic Versatility in the Plant Pathogen Ralstonia solanacearum", PLOS Pathog, 12(10), e1005939.

[167] Geourjon C., Combet C., Blanchet C., Deléage G., (2001), "Identification of related proteins with weak sequence identity using secondary structure information", Protein Sci, 10(4), 788–797.

[168] Smith E. W., Zhang X., Behzadi C., Andrews L. D., Cohen F., Chen Y., (2015), "Structures of Pseudomonas aeruginosa LpxA reveal basis for substrate selectivity", Biochemistry, 54(38), 5937–5948.

[169] Perumal D., Sakharkar K. R., Tang T. H., Chow V. T. K., Lim C. S., Samal A., Sugiura N., Sakharkar M. K., (2010), "Cloning and targeted disruption of two lipopolysaccharide biosynthesis genes, kdsA and waaG, of pseudomonas aeruginosa PAO1 by site-directed mutagenesis", J Mol Microbiol Biotechnol,

19(4), 169–179.

[170] Kumar A., Thotakura P. L., Tiwary B. K., Krishna R., (2016), "Target identification in Fusobacterium nucleatum by subtractive genomics approach and enrichment analysis of host-pathogen protein-protein interactions", BMC Microbiol, 16(1), 1–12.

[171] De Leon G. P., Elowe N. H., Koteva K. P., Valvano M. A., Wright G. D., (2006), "An In Vitro Screen of Bacterial Lipopolysaccharide Biosynthetic Enzymes Identifies an Inhibitor of ADP-Heptose Biosynthesis", Chem Biol, 13(4), 437–441.

[172] Sorci L., Pan Y., Eyobo Y., Rodionova I., Huang N., Kurnasov O., Zhong S., MacKerell A. D., Zhang H., Osterman A. L., (2009), "Targeting NAD Biosynthesis in Bacterial Pathogens: Structure-Based Development of Inhibitors of Nicotinate Mononucleotide Adenylyltransferase NadD", Chem Biol, 16(8), 849–861.

[173] Rodionova I. A., Zuccola H. J., Sorci L., Aleshin A. E., Kazanov M. D., Ma C. T., Sergienko E., Rubin E. J., Locher C. P., Osterman A. L., (2015), "Mycobacterial Nicotinate Mononucleotide Adenylyltransferase structure, mechanism, and implications for drug discovery", J Biol Chem, 290(12), 7693–7706.

[174] Wang X., Ahn Y. M., Lentscher A. G., Lister J. S., Brothers R. C., Kneen M. M., Gerratana B., Boshoff H. I., Dowd C. S., (2017), "Design, synthesis, and evaluation of substituted nicotinamide adenine dinucleotide (NAD+) synthetase inhibitors as potential antitubercular agents", Bioorganic Med Chem Lett, 27(18), 4426–4430.

[175] Rodionova I. A., Schuster B. M., Guinn K. M., Sorci L., Scott D. A., Li X., Kheterpal I., Shoen C., Cynamon M., Locher C., Rubin E. J., Osterman A. L., (2014), "Metabolic and bactericidal effects of targeted suppression of NadD and NadE enzymes in mycobacteria", MBio, 5(1), 1–9.

[176] Huang N., Kolhatkar R., Eyobo Y., Sorci L., Rodionova I., Osterman A., Mackerell A., Zhang H., (2010), "Complexes of bacterial nicotinate mononucleotide adenylyltransferase with inhibitors: implication for structure-based drug design and improvement", J Med Chem, 53(14), 5229–39.

[177] Leonardi R., Jackowski S., (2007), "Biosynthesis of Pantothenic Acid and Coenzyme A", EcoSal Plus, 2(2).

[178] Spry C., Kirk K., Saliba K. J., (2008), "Coenzyme A biosynthesis: An antimicrobial drug target", FEMS Microbiol Rev, 32(1), 56–106.

[179] Rath C. M., Benton B. M., De Vicente J., Drumm J. E., Geng M., Li C., Moreau R. J., Shen X., Skepper C. K., Steffek M., Takeoka K., Wang L., Wei J. R., Xu W., Zhang Q., Feng B. Y., (2018), "Optimization of CoaD Inhibitors against Gram-Negative Organisms through Targeted Metabolomics", ACS Infect Dis,

4(3), 391–402.

[180] Begley T. P., Kinsland C., Strauss E., (2001), "The biosynthesis of coenzyme A in bacteria.", Vitam Horm, 61, 157–171.

[181] Gerdes S. Y., Scholle M. D., Souza M. D., Bernal A., Baev M. V, Farrell M., Kurnasov O. V, Daugherty M. D., Mseeh F., Polanuyer B. M., Campbell J. W., Anantha S., Shatalin K. Y., Chowdhury S. A. K., Fonstein M. Y., Osterman A. L., (2002), "From Genetic Footprinting to Antimicrobial Drug Targets : Examples in Cofactor Biosynthetic Pathways", J Bacteriol, 184(16), 4555–4572.

[182] Yang K., Eyobo Y., Brand L. A., Martynowski D., Tomchick D., Strauss E., Zhang H., (2006), "Crystal structure of a type III pantothenate kinase: Insight into the mechanism of an essential coenzyme A biosynthetic enzyme universally distributed in bacteria", J Bacteriol, 188(15), 5532–5540.

[183] Hughes S. J., Antoshchenko T., Kim K. P., Smil D., Park H. W., (2014), "Structural characterization of a new N-substituted pantothenamide bound to pantothenate kinases from Klebsiella pneumoniae and Staphylococcus aureus", Proteins, 82(7), 1542–1548.

[184] Evans J. C., Trujillo C., Wang Z., Eoh H., Ehrt S., Schnappinger D., Boshoff H. I. M., Rhee K. Y., Barry C. E., Mizrahi V., (2016), "Validation of CoaBC as a Bactericidal Target in the Coenzyme A Pathway of Mycobacterium tuberculosis", ACS Infect Dis, 2(12), 958–968.

[185] De Jonge B. L. M., Walkup G. K., Lahiri S. D., Huynh H., Neckermann G., Utley L., Nash T. J., Brock J., San Martin M., Kutschke A., Johnstone M., Laganas V., Hajec L., Gu R. F., Ni H., Chen B., Hutchings K., Holt E., McKinney D., Gao N., Livchak S., Thresher J., (2013), "Discovery of inhibitors of 4'-phosphopantetheine adenylyltransferase (PPAT) to validate PPAT as a target for antibacterial therapy", Antimicrob Agents Chemother, 57(12), 6005–6015.

[186] Gutiérrez-Preciado A., Torres A. G., Merino E., Bonomi H. R., Goldbaum F. A., García-Angulo V. A., (2015), "Extensive identification of bacterial riboflavin transporters and their distribution across bacterial species", PLoS One, 10(5).

[187] Fassbinder F., Kist M., Bereswill S., (2000), "Structural and functional analysis of the riboflavin synthesis genes encoding GTP cyclohydrolase II (ribA), DHBP synthase (ribBA), riboflavin synthase (ribC), and riboflavin deaminase/reductase (ribD) from Helicobacter pylori strain P1", FEMS Microbiol Lett, 191(2), 191–197.

[188] Long Q., Ji L., Wang H., Xie J., (2010), "Riboflavin biosynthetic and regulatory factors as potential novel anti-infective drug targets:", Chem Biol Drug Des, 75(4), 339–347.

[189] Saeed-kothe A., Yang W., Mills S. D., (2004), "Use of the riboflavin synthase gene (ribC) as a model for development of an essential gene disruption and complementation system for Haemophilus influenzae", Appl Env Microbiol 2004, 70(7), 4136–4143.

[190] Vollmer W., Bertsche U., (2008), "Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli", Biochim Biophys Acta - Biomembr, 1778(9), 1714–1734.

[191] Vollmer W., Höltje J. V., (2000), "A simple screen for murein transglycosylase inhibitors", Antimicrob Agents Chemother, 44(5), 1181–1185.

[192] Mesleh M. F., Rajaratnam P., Conrad M., Chandrasekaran V., Liu C. M., Pandya B. A., Hwang Y. S., Rye P. T., Muldoon C., Becker B., Zuegg J., Meutermans W., Moy T. I., (2016), "Targeting Bacterial Cell Wall Peptidoglycan Synthesis by Inhibition of Glycosyltransferase Activity", Chem Biol Drug Des, 87(2), 190–199.

[193] Derouaux A., Sauvage E., Terrak M., (2013), "Peptidoglycan glycosyltransferase substrate mimics as templates for the design of new antibacterial drugs", Front Immunol, 4(78).

[194] Bartell J. A., Blazier A. S., Yen P., Thøgersen J. C., Jelsbak L., Goldberg J. B., Papin J. A., (2017), "Reconstruction of the metabolic network of Pseudomonas aeruginosa to interrogate virulence factor synthesis", Nat Commun, 8(14631).

BIOGRAPHY

Müberra Fatma CESUR was born in Uşak in July 1, 1991. She graduated from Gebze Technical University Department of Molecular Biology and Genetics in 2015. She has a master's degree from the same department in 2017. She has been working as a Research Assistant in Bioengineering Department since 2015. She is also a Msc. student in Bioinformatics and Systems Biology Program under Bioengineering Department at Graduate School of Natural and Applied Sciences in the same university.

APPENDICES

Appendix A: Media Used in Model Validation

Minimal growth medium was used in the model validation by *in silico* prediction of the growth rates . Composition of carbon-D-glucose (CDG) medium was imported from KBase platform and this medium was integrated into the iKp1289 to simulate the growth phenotype of the bacteria in the minimal medium. Other minimal media used in this study were generated by replacing the carbon source (glucose) in the CDG medium for iKp1289.

Table A1.1: *In silico* CDG Medium (iKp1289).

Recipe for CDG Medium:	**Metabolite in iKp1289:**	**Lower bound:**	**Upper bound:**
H_2O	cpd00001[e0]	-100	1000
O_2	cpd00007[e0]	-10	1000
Phosphate	cpd00009[e0]	-100	1000
NH_3	cpd00013[e0]	-100	1000
D-Glucose (any carbon sources[a])	cpd00027[e0]	-5	1000
Mn^{2+}	cpd00030[e0]	-100	1000
Zn^{2+}	cpd00034[e0]	-100	1000
Sulfate	cpd00048[e0]	-100	1000
Cu^{2+}	cpd00058[e0]	-100	1000
Ca^{2+}	cpd00063[e0]	-100	1000
H^{+}	cpd00067[e0]	-100	1000
Cl-	cpd00099[e0]	-100	1000
Co^{2+}	cpd00149[e0]	-100	1000
K^{+}	cpd00205[e0]	-100	1000
Ni^{2+}	cpd00244[e0]	-100	1000
Mg	cpd00254[e0]	-100	1000
Na^{+}	cpd00971[e0]	-100	1000
Fe^{2+}	cpd10515[e0]	-100	1000
Fe^{3}	cpd10516[e0]	-100	1000
Molybdate	cpd11574[e0]	-100	1000

[a] Each minimal medium was obtained by changing the carbon source.

Luria-Bertani (LB) broth medium was used in the validation of iKp1289 via the gene essentiality analysis. Composition of LB medium was identified upon online KBase platform.

Table A1.2: *In silico* LB Medium (iKp1289).

Recipe for LB Medium:	**Metabolite in iKp1289:**	**Lower bound:**	**Upper bound:**
H_2O	cpd00001[e0]	-100	100
O_2	cpd00007[e0]	-100	100
Phosphate	cpd00009[e0]	-100	100
NH_3	cpd00013[e0]	-100	100
AMP	cpd00018[e0]	-100	100
L-Glutamate	cpd00023[e0]	-100	100
D-Glucose	cpd00027[e0]	-100	100
Heme	cpd00028[e0]	-100	100
Mn^{2+}	cpd00030[e0]	-100	100
Glycine	cpd00033[e0]	-100	100
Zn^{2+}	cpd00034[e0]	-100	100
L-Alanine	cpd00035[e0]	-100	100
L-Lysine	cpd00039[e0]	-100	100
L-Aspartate	cpd00041[e0]	-100	100
CMP	cpd00046[e0]	-100	100
Sulfate	cpd00048[e0]	-100	100
L-Arginine	cpd00051[e0]	-100	100
L-Serine	cpd00054[e0]	-100	100
Cu^{2+}	cpd00058[e0]	-100	100
L-Methionine	cpd00060[e0]	-100	100
Ca^{2+}	cpd00063[e0]	-100	100
L-Tryptophan	cpd00065[e0]	-100	100
L-Phenylalanine	cpd00066[e0]	-100	100
H^+	cpd00067[e0]	-100	100
L-Tyrosine	cpd00069[e0]	-100	100
UMP	cpd00091[e0]	-100	100
Uracil	cpd00092[e0]	-100	100
Cl^-	cpd00099[e0]	-100	100
L-Leucine	cpd00107[e0]	-100	100
L-Histidine	cpd00119[e0]	-100	100
GMP	cpd00126[e0]	-100	100
L-Proline	cpd00129[e0]	-100	100
Co^{2+}	cpd00149[e0]	-100	100
L-Valine	cpd00156[e0]	-100	100
L-Threonine	cpd00161[e0]	-100	100
Adenosine	cpd00182[e0]	-100	100
Thymidine	cpd00184[e0]	-100	100
K^+	cpd00205[e0]	-100	100

Table A1.2: Continued.

Recipe for LB Medium:	Metabolite in iKp1289:	Lower bound:	Upper bound:
Pyridoxal	cpd00215[e0]	-100	100
Niacin	cpd00218[e0]	-100	100
Riboflavin	cpd00220[e0]	-100	100
HYXN	cpd00226[e0]	-100	100
H2S	cpd00239[e0]	-100	100
Ni^{2+}	cpd00244[e0]	-100	100
Inosine	cpd00246[e0]	-100	100
Uridine	cpd00249[e0]	-100	100
Mg	cpd00254[e0]	-100	100
Guanosine	cpd00311[e0]	-100	100
L-Isoleucine	cpd00322[e0]	-100	100
L-Cystine	cpd00381[e0]	-100	100
Folate	cpd00393[e0]	-100	100
Deoxyadenosine	cpd00438[e0]	-100	100
Hg^{2+}	cpd00531[e0]	-100	100
Lipoate	cpd00541[e0]	-100	100
PAN	cpd00644[e0]	-100	100
Deoxycytidine	cpd00654[e0]	-100	100
Thiamine phosphate	cpd00793[e0]	-100	100
Na^{+}	cpd00971[e0]	-100	100
Cd^{2+}	cpd01012[e0]	-100	100
Arsenate	cpd01048[e0]	-100	100
Vitamin B12	cpd03424[e0]	-100	100
Fe^{2+}	cpd10515[e0]	-100	100
Fe^{3}	cpd10516[e0]	-100	100
Molybdate	cpd11574[e0]	-100	100
Chromate	cpd11595[e0]	-100	100

Appendix B: Host-Mimicking Media Used in the Thesis

Host-mimicking media were integrated into the both models in the drug discovery process. The outgoing reactions, their indices, the uptake rates and the reaction formulas associated with each compound involved in the medium are listed in the following tables. SM medium was developed by combining the sputum and alveolar macrophage environments as highlighed before. Each medium was developed by searching synonyms of each compound owing to lack of the consensus nomenclature in the models. In addition to metabolite names, the metabolite IDs (PubChem, HMDB, ChEBI and KEGG IDs) were also considered by screening each metabolite in both models against HBF.

Table B1.1: *In silico* SM Medium (iYL1228).

Index:	Matched reaction name:	Lower	Upper	Reaction
2179	Sodium exchange	-10	1000	na1[e] <=>
2147	potassium exchange	-10	1000	k[e] <=>
2181	Ammonium exchange	-10	1000	nh4[e] <=>
2172	magnesium exchange	-10	1000	mg2[e] <=>
2186	Nitrate exchange	-10	1000	no3[e] ->
2198	Phosphate exchange	-10	1000	pi[e] <=>
2220	Sulfate exchange	-10	1000	so4[e] <=>
2216	L-Serine exchange	-10	1000	ser_L[e] ->
2229	L-Threonine exchange	-10	1000	thr_L[e] ->
2014	L-Alanine exchange	-10	1000	ala_L[e] ->
2112	Glycine exchange	-10	1000	gly[e] ->
2204	L-Proline exchange	-10	1000	pro_L[e] ->
2141	L-Isoleucine exchange	-10	1000	ile_L[e] ->
2152	L-Leucine exchange	-10	1000	leu_L[e] ->
2253	L-Valine exchange	-10	1000	val_L[e] ->
2024	L-Aspartate exchange	-10	1000	asp_L[e] ->
2111	L-Glutamate exchange	-10	1000	glu_L[e] ->
2196	L-Phenylalanine exchange	-10	1000	phe_L[e] ->
2242	L-Tyrosine exchange	-10	1000	tyr_L[e] ->
2236	L-Tryptophan exchange	-10	1000	trp_L[e] ->
2155	L-Lysine exchange	-10	1000	lys_L[e] ->
2136	L-Histidine exchange	-10	1000	his_L[e] ->
2020	L-Arginine exchange	-10	1000	arg_L[e] ->
2192	Ornithine exchange	-10	1000	orn[e] ->
2049	L-Cysteine exchange	-10	1000	cys_L[e] ->
2169	L-Methionine exchange	-10	1000	met_L[e] ->
2105	D-Glucose exchange	-10	1000	glc_D[e] <=>
2150	L-Lactate exchange	-10	1000	lac_L[e] ->
2074	Fe2+ exchange	-10	1000	fe2[e] <=>
2029	Calcium exchange	-10	1000	ca2[e] <=>
2037	Chloride exchange	-10	1000	cl[e] <=>
2126	H+ exchange	-10	1000	h[e] <=>
2128	H2O exchange	-10	1000	h2o[e] <=>
2187	O2 exchange	-10	1000	o2[e] <=>
2190	octadecenoate (n-C18:1) exchange	-10	1000	ocdcea[e] ->
2110	L-Glutamine exchange	-10	1000	gln_L[e] ->
2238	tetradecanoate (n-C14:0) exchange	-10	1000	ttdca[e] ->
2209	Pyruvate exchange	-10	1000	pyr[e] ->
2189	octadecanoate (n-C18:0) exchange	-10	1000	ocdca[e] ->
2026	Butyrate (n-C4:0) exchange	-10	1000	but[e] ->

Table B1.2: *In silico* SM Medium (iKp1289).

Index:	Matched reaction name:	Lower bound:	Upper bound:	Reaction formula:
2208	EX_Na_plus_(e)	-10	1000	cpd00971[e0] <=>
2325	EX_K_plus_(e)	-10	1000	cpd00205[e0] <=>
2358	EX_NH3(e)	-10	1000	cpd00013[e0] + cpd00013[e0] <=>
2350	EX_Mg(e)	-10	1000	cpd00254[e0] + cpd00254[e0] <=>
2375	EX_Phosphate(e)	-10	1000	cpd00009[e0] <=>
2397	EX_Sulfate(e)	-10	1000	cpd00048[e0] <=>
2393	EX_L_Serine(e)	-10	1000	cpd00054[e0] + cpd00054[e0] <=>
2408	EX_L_Threonine(e)	-10	1000	cpd00161[e0] <=>
2184	EX_L_Alanine(e)	-10	1000	cpd00035[e0] <=>
2296	EX_Glycine(e)	-10	1000	cpd00033[e0] <=>
2382	EX_L_Proline(e)	-10	1000	cpd00129[e0] <=>
2318	EX_L_Isoleucine(e)	-10	1000	cpd00322[e0] <=>
2328	EX_L_Leucine(e)	-10	1000	cpd00107[e0] <=>
2430	EX_L_Valine(e)	-10	1000	cpd00156[e0] <=>
2194	EX_L_Aspartate(e)	-10	1000	cpd00041[e0] <=>
2288	EX_L_Glutamate(e)	-10	1000	cpd00023[e0] <=>
2374	EX_L_Phenylalanine(e)	-10	1000	cpd00066[e0] <=>
2420	EX_L_Tyrosine(e)	-10	1000	cpd00069[e0] <=>
2413	EX_L_Tryptophan(e)	-10	1000	cpd00065[e0] <=>
2333	EX_L_Lysine(e)	-10	1000	cpd00039[e0] <=>
2311	EX_L_Histidine(e)	-10	1000	cpd00119[e0] <=>
2190	EX_L_Arginine(e)	-10	1000	cpd00051[e0] + cpd00051[e0] <=>

Table B1.2: Continued.

Index:	Matched reaction name:	Lower bound:	Upper bound:	Reaction formula:
2369	EX_Ornithine(e)	-10	1000	cpd00064[e0] + cpd00064[e0] <=>
2221	EX_L_Cysteine(e)	-10	1000	cpd00084[e0] <=>
2349	EX_L_Methionine(e)	-10	1000	cpd00060[e0] <=>
2286	EX_D_Glucose(e)	-10	1000	cpd00027[e0] + cpd00027[e0] <=>
2326	EX_L_Lactate(e)	-10	1000	cpd00159[e0] <=>
2251	EX_Fe2_plus_(e)	-10	1000	cpd10515[e0] <=>
2200	EX_Ca2_plus_(e)	-10	1000	cpd00063[e0] <=>
2210	EX_Cl_(e)	-10	1000	cpd00099[e0] <=>
2180	EX_H_plus_(e)	-10	1000	cpd00067[e0] + cpd00067[e0] <=>
2304	EX_H2O(e)	-10	1000	cpd00001[e0] + cpd00001[e0] <=>
2365	EX_O2(e)	-10	1000	cpd00007[e0] + cpd00007[e0] <=>
2362	EX_Nitrate(e)	-10	1000	cpd00209[e0] + cpd00209[e0] <=>
2287	EX_L_Glutamine(e)	-10	1000	cpd00053[e0] <=>
2386	EX_Pyruvate(e)	-10	1000	cpd00020[e0] <=>
2367	EX_octadecenoate(e)	-10	1000	cpd15269[e0] <=>
2415	EX_Myristic_acid(e)	-10	1000	cpd03847[e0] <=>
2198	EX_Butyrate(e)	-10	1000	cpd00211[e0] + cpd00211[e0] <=>

Table B1.3: *In silico* HBF Medium (iYL1228).

Index:	Matched reaction name:	Lower bound:	Upper bound:	Reaction formula:
1989	4-Aminobutanoate exchange	-10	1000	4abut[e] ->
1991	4-Hydroxyphenylacetate exchange	-10	1000	4hphac[e] ->
1996	Acetate exchange	-10	1000	ac[e] ->
1997	Acetaldehyde exchange	-10	1000	acald[e] ->
2000	N-Acetyl-D-glucosamine exchange	-10	1000	acgam[e] ->
2006	Adenine exchange	-10	1000	ade[e] ->
2007	Adenosine exchange	-10	1000	adn[e] ->
2011	2-Oxoglutarate exchange	-10	1000	akg[e] ->
2012	beta-Alanine exchange	-10	1000	ala_B[e] ->
2013	D-Alanine exchange	-10	1000	ala_D[e] ->
2014	L-Alanine exchange	-10	1000	ala_L[e] ->
2017	AMP exchange	-10	1000	amp[e] ->
2019	L-Arabinose exchange	-10	1000	arab_L[e] ->
2020	L-Arginine exchange	-10	1000	arg_L[e] ->
2021	L-Ascorbate exchange	-10	1000	ascb_L[e] ->
2022	L-Asparagine exchange	-10	1000	asn_L[e] ->
2024	L-Aspartate exchange	-10	1000	asp_L[e] ->
2026	Butyrate (n-C4:0) exchange	-10	1000	but[e] ->
2028	Benzoate exchange	-10	1000	bz[e] ->
2029	Calcium exchange	-10	1000	ca2[e] <=>
2034	Cys-Gly exchange	-10	1000	cgly[e] ->
2035	Choline exchange	-10	1000	chol[e] ->
2036	Citrate exchange	-10	1000	cit[e] ->
2037	Chloride exchange	-10	1000	cl[e] <=>
2038	CMP exchange	-10	1000	cmp[e] ->
2039	CO2 exchange	-10	1000	co2[e] <=>
2042	L-Carnitine exchange	-10	1000	crn[e] ->
2043	Cytosine exchange	-10	1000	csn[e] ->
2046	Hydrogen cyanide exchange	-10	1000	cyan[e] ->
2049	L-Cysteine exchange	-10	1000	cys_L[e] ->
2050	Cytidine exchange	-10	1000	cytd[e] ->
2051	Deoxyadenosine exchange	-10	1000	dad_2[e] ->
2055	Deoxycytidine exchange	-10	1000	dcyt[e] ->
2058	Deoxyguanosine exchange	-10	1000	dgsn[e] ->
2062	Deoxyinosine exchange	-10	1000	din[e] ->
2065	Dopamine exchange	-10	1000	dopa[e] ->
2068	Deoxyuridine exchange	-10	1000	duri[e] ->
2071	Ethanol exchange	-10	1000	etoh[e] ->
2074	Fe2+ exchange	-10	1000	fe2[e] <=>
2075	Fe3+ exchange	-10	1000	fe3[e] <=>
2081	Formate exchange	-10	1000	for[e] ->
2082	D-Fructose exchange	-10	1000	fru[e] ->
2085	L-Fucose exchange	-10	1000	fuc_L[e] ->
2094	D-Galactose exchange	-10	1000	gal[e] ->
2102	D-Glucosamine exchange	-10	1000	gam[e] ->

Table B1.3: Continued

Index:	Matched reaction name:	Lower bound:	Upper bound:	Reaction formula:
2104	GDP exchange	-10	1000	gdp[e] ->
2105	D-Glucose exchange	-10	1000	glc_D[e] <=>
2110	L-Glutamine exchange	-10	1000	gln_L[e] ->
2111	L-Glutamate exchange	-10	1000	glu_L[e] ->
2112	Glycine exchange	-10	1000	gly[e] ->
2114	Glycine betaine exchange	-10	1000	glyb[e] ->
2115	Glycerol exchange	-10	1000	glyc[e] ->
2120	GMP exchange	-10	1000	gmp[e] ->
2121	Guanosine exchange	-10	1000	gsn[e] ->
2122	Oxidized glutathione exchange	-10	1000	gthox[e] ->
2123	Reduced glutathione exchange	-10	1000	gthrd[e] ->
2124	GTP exchange	-10	1000	gtp[e] ->
2125	Guanine exchange	-10	1000	gua[e] ->
2126	H+ exchange	-10	1000	h[e] <=>
2128	H2O exchange	-10	1000	h2o[e] <=>
2129	Hydrogen peroxide exchange	-10	1000	h2o2[e] ->
2133	Hexadecanoate (n-C16:0) exchange	-10	1000	hdca[e] ->
2134	Hexadecenoate (n-C16:1) exchange	-10	1000	hdcea[e] ->
2136	L-Histidine exchange	-10	1000	his_L[e] ->
2137	L-Homoserine exchange	-10	1000	hom_L[e] ->
2139	Hypoxanthine exchange	-10	1000	hxan[e] ->
2141	L-Isoleucine exchange	-10	1000	ile_L[e] ->
2142	IMP exchange	-10	1000	imp[e] ->
2144	myo-Inositol exchange	-10	1000	inost[e] ->
2145	Inosine exchange	-10	1000	ins[e] ->
2149	D-Lactate exchange	-10	1000	lac_D[e] ->
2150	L-Lactate exchange	-10	1000	lac_L[e] ->
2151	Lactose exchange	-10	1000	lcts[e] ->
2152	L-Leucine exchange	-10	1000	leu_L[e] ->
2155	L-Lysine exchange	-10	1000	lys_L[e] ->
2159	Maltose exchange	-10	1000	malt[e] ->
2162	Maltotriose exchange	-10	1000	malttr[e] ->
2164	D-Mannose exchange	-10	1000	man[e] ->
2169	L-Methionine exchange	-10	1000	met_L[e] ->
2179	Sodium exchange	-10	1000	na1[e] <=>
2180	Nicotinate exchange	-10	1000	nac[e] ->
2184	Nitric oxide exchange	-10	1000	no[e] ->
2187	O2 exchange	-10	1000	o2[e] <=>
2188	Superoxide anion exchange	-10	1000	o2s[e] ->
2189	octadecanoate (n-C18:0) exchange	-10	1000	ocdca[e] ->
2190	octadecenoate (n-C18:1) exchange	-10	1000	ocdcea[e] ->
2191	octanoate (n-C8:0) exchange	-10	1000	octa[e] ->
2192	Ornithine exchange	-10	1000	orn[e] ->
2196	L-Phenylalanine exchange	-10	1000	phe_L[e] ->
2197	Protoheme exchange	-10	1000	pheme[e] ->

Table B1.3: Continued.

Index:	Matched reaction name:	Lower bound:	Upper bound:	Reaction formula:
2198	Phosphate exchange	-10	1000	pi[e] <=>
2199	(R)-Pantothenate exchange	-10	1000	pnto_R[e] ->
2204	L-Proline exchange	-10	1000	pro_L[e] ->
2209	Pyruvate exchange	-10	1000	pyr[e] ->
2211	Ribitol exchange	-10	1000	rbt[e] ->
2212	D-Ribose exchange	-10	1000	rib_D[e] ->
2215	D-Serine exchange	-10	1000	ser_D[e] ->
2216	L-Serine exchange	-10	1000	ser_L[e] ->
2220	Sulfate exchange	-10	1000	so4[e] <=>
2222	Succinate exchange	-10	1000	succ[e] ->
2223	Sucrose exchange	-10	1000	sucr[e] ->
2226	Taurine exchange	-10	1000	taur[e] ->
2227	Thiocyanate exchange	-10	1000	tcynt[e] ->
2228	Thiamin exchange	-10	1000	thm[e] ->
2229	L-Threonine exchange	-10	1000	thr_L[e] ->
2231	Thymine exchange	-10	1000	thym[e] ->
2232	Thymidine exchange	-10	1000	thymd[e] ->
2235	Trehalose exchange	-10	1000	tre[e] ->
2236	L-Tryptophan exchange	-10	1000	trp_L[e] ->
2237	Thiosulfate exchange	-10	1000	tsul[e] ->
2238	tetradecanoate (n-C14:0) exchange	-10	1000	ttdca[e] ->
2242	L-Tyrosine exchange	-10	1000	tyr_L[e] ->
2249	UMP exchange	-10	1000	ump[e] ->
2250	Uracil exchange	-10	1000	ura[e] ->
2251	Urea exchange	-10	1000	urea[e] ->
2252	Uridine exchange	-10	1000	uri[e] ->
2253	L-Valine exchange	-10	1000	val_L[e] ->
2257	D-Xylose exchange	-10	1000	xyl_D[e] ->
2147	potassium exchange	-10	1000	k[e] <=>
2200	Propionate (n-C3:0) exchange	-10	1000	ppa[e] ->

Table B1.4: *In silico* HBF Medium (iKp1289).

Index:	Matched reaction name:	Lower Bound:	Upper Bound:	Reaction Formula:
2165	EX_GABA_e0	-10	1000	cpd00281[e0] <=>
2166	EX_Acetaldehyde_e0	-10	1000	cpd00071[e0] <=>
2170	EX_N_Acetyl_D_glucosamine_e0	-10	1000	cpd00122[e0] + cpd00122[e0] <=>
2175	EX_Acetate_e0	-10	1000	cpd00029[e0] <=>

Table B1.4: Continued

Index:	Matched reaction name:	Lower Bound:	Upper Bound:	Reaction Formula:
2176	EX_Adenine_e0	-10	1000	cpd00128[e0] <=>
2177	EX_Adenosine_e0	-10	1000	cpd00182[e0] <=>
2180	EX_H_plus__e0	-10	1000	cpd00067[e0] + cpd00067[e0] <=>
2182	EX_2_Oxoglutarate_e0	-10	1000	cpd00024[e0] + cpd00024[e0] <=>
2184	EX_L_Alanine_e0	-10	1000	cpd00035[e0] <=>
2187	EX_AMP_e0	-10	1000	cpd00018[e0] <=>
2189	EX_L_Arabinose_e0	-10	1000	cpd00224[e0] + cpd00224[e0] <=>
2190	EX_L_Arginine_e0	-10	1000	cpd00051[e0] + cpd00051[e0] <=>
2191	EX_L_Ascorbate_e0	-10	1000	cpd00059[e0] <=>
2192	EX_L_Asparagine_e0	-10	1000	cpd00132[e0] <=>
2194	EX_L_Aspartate_e0	-10	1000	cpd00041[e0] <=>
2195	EX_beta_Alanine_e0	-10	1000	cpd00085[e0] <=>
2199	EX_Benzoate_e0	-10	1000	cpd00153[e0] <=>
2200	EX_Ca2_plus__e0	-10	1000	cpd00063[e0] <=>
2205	EX_Cys_Gly_e0	-10	1000	cpd01017[e0] <=>
2206	EX_Choline_e0	-10	1000	cpd00098[e0] <=>
2207	EX_Citrate_e0	-10	1000	cpd00137[e0] <=>
2208	EX_Na_plus__e0	-10	1000	cpd00971[e0] <=>
2210	EX_Cl__e0	-10	1000	cpd00099[e0] <=>
2211	EX_CMP_e0	-10	1000	cpd00046[e0] <=>
2212	EX_CO2_e0	-10	1000	cpd00011[e0] + cpd00011[e0] <=>
2215	EX_Cytosine_e0	-10	1000	cpd00307[e0] <=>

Table B1.4: Continued.

Index:	Matched reaction name:	Lower Bound:	Upper Bound:	Reaction Formula:
2218	EX_HCN_e0	-10	1000	cpd00150[e0] <=>
2221	EX_L_Cysteine_e0	-10	1000	cpd00084[e0] <=>
2222	EX_Cytidine_e0	-10	1000	cpd00367[e0] <=>
2223	EX_D_Lactate_e0	-10	1000	cpd00221[e0] <=>
2225	EX_Deoxyadenosine_e0	-10	1000	cpd00438[e0] <=>
2226	EX_D_Alanine_e0	-10	1000	cpd00117[e0] <=>
2231	EX_Deoxycytidine_e0	-10	1000	cpd00654[e0] <=>
2235	EX_Deoxyguanosine_e0	-10	1000	cpd00277[e0] <=>
2238	EX_Deoxyinosine_e0	-10	1000	cpd03279[e0] <=>
2241	EX_Dopamine_e0	-10	1000	cpd02357[e0] <=>
2242	EX_D_Serine_e0	-10	1000	cpd00550[e0] <=>
2245	EX_Deoxyuridine_e0	-10	1000	cpd00412[e0] <=>
2248	EX_Ethanol_e0	-10	1000	cpd00363[e0] <=>
2251	EX_Fe2_plus__e0	-10	1000	cpd10515[e0] <=>
2254	EX_fe3_e0	-10	1000	cpd10516[e0] + cpd10516[e0] <=>
2258	EX_Formate_e0	-10	1000	cpd00047[e0] <=>
2261	EX_D_Fructose_e0	-10	1000	cpd00082[e0] + cpd00082[e0] <=>
2262	EX_L_Fucose_e0	-10	1000	cpd00751[e0] <=>
2281	EX_GDP_e0	-10	1000	cpd00031[e0] <=>
2286	EX_D_Glucose_e0	-10	1000	cpd00027[e0] + cpd00027[e0] <=>
2287	EX_L_Glutamine_e0	-10	1000	cpd00053[e0] <=>
2288	EX_L_Glutamate_e0	-10	1000	cpd00023[e0] <=>

Table B1.4: Continued.

Index:	Matched reaction name:	Lower Bound:	Upper Bound:	Reaction Formula:
2295	EX_Glycerol_e0	-10	1000	cpd00100[e0] <=>
2296	EX_Glycine_e0	-10	1000	cpd00033[e0] <=>
2297	EX_GMP_e0	-10	1000	cpd00126[e0] <=>
2298	EX_Guanosine_e0	-10	1000	cpd00311[e0] <=>
2299	EX_Oxidized_glutathione_e0	-10	1000	cpd00111[e0] <=>
2300	EX_GSH_e0	-10	1000	cpd00042[e0] <=>
2301	EX_GTP_e0	-10	1000	cpd00038[e0] <=>
2302	EX_Guanine_e0	-10	1000	cpd00207[e0] <=>
2303	EX_H2O2_e0	-10	1000	cpd00025[e0] <=>
2304	EX_H2O_e0	-10	1000	cpd00001[e0] + cpd00001[e0] <=>
2311	EX_L_Histidine_e0	-10	1000	cpd00119[e0] <=>
2312	EX_L_Homoserine_e0	-10	1000	cpd00227[e0] <=>
2313	EX_4_Hydroxyphenylacetate_e0	-10	1000	cpd00489[e0] <=>
2318	EX_L_Isoleucine_e0	-10	1000	cpd00322[e0] <=>
2319	EX_IMP_e0	-10	1000	cpd00114[e0] <=>
2322	EX_Inosine_e0	-10	1000	cpd00246[e0] <=>
2325	EX_K_plus__e0	-10	1000	cpd00205[e0] <=>
2326	EX_L_Lactate_e0	-10	1000	cpd00159[e0] <=>
2328	EX_L_Leucine_e0	-10	1000	cpd00107[e0] <=>
2333	EX_L_Lysine_e0	-10	1000	cpd00039[e0] <=>
2338	EX_Amylotriose_e0	-10	1000	cpd01262[e0] <=>
2340	EX_Maltose_e0	-10	1000	cpd00179[e0] + cpd00179[e0] <=>

Table B1.4: Continued.

Index:	Matched reaction name:	Lower Bound:	Upper Bound:	Reaction Formula:
2344	EX_D_Mannose_e0	-10	1000	cpd00138[e0] + cpd00138[e0] <=>
2349	EX_L_Methionine_e0	-10	1000	cpd00060[e0] <=>
2357	EX_Niacin_e0	-10	1000	cpd00218[e0] <=>
2358	EX_NH3_e0	-10	1000	cpd00013[e0] + cpd00013[e0] <=>
2363	EX_NO_e0	-10	1000	cpd00418[e0] <=>
2365	EX_O2_e0	-10	1000	cpd00007[e0] + cpd00007[e0] <=>
2369	EX_Ornithine_e0	-10	1000	cpd00064[e0] + cpd00064[e0] <=>
2373	EX_Heme_e0	-10	1000	cpd00028[e0] <=>
2374	EX_L_Phenylalanine_e0	-10	1000	cpd00066[e0] <=>
2375	EX_Phosphate_e0	-10	1000	cpd00009[e0] <=>
2378	EX_Propionate_e0	-10	1000	cpd00141[e0] <=>
2382	EX_L_Proline_e0	-10	1000	cpd00129[e0] <=>
2386	EX_Pyruvate_e0	-10	1000	cpd00020[e0] <=>
2389	EX_D_Ribose_e0	-10	1000	cpd00105[e0] <=>
2391	EX_Ribitol_e0	-10	1000	cpd00366[e0] <=>
2393	EX_L_Serine_e0	-10	1000	cpd00054[e0] + cpd00054[e0] <=>
2397	EX_Sulfate_e0	-10	1000	cpd00048[e0] <=>
2399	EX_Succinate_e0	-10	1000	cpd00036[e0] <=>
2400	EX_Sucrose_e0	-10	1000	cpd00076[e0] + cpd00076[e0] <=>
2403	EX_Taurine_e0	-10	1000	cpd00210[e0] <=>
2404	EX_Thiocyanate_e0	-10	1000	cpd01211[e0] <=>
2405	EX_Thymidine_e0	-10	1000	cpd00184[e0] <=>

Table B1.4: Continued

Index:	Matched reaction name:	Lower Bound:	Upper Bound:	Reaction Formula:
2406	EX_Thiamin_e0	-10	1000	cpd00305[e0] <=>
2408	EX_L_Threonine_e0	-10	1000	cpd00161[e0] <=>
2409	EX_Thymine_e0	-10	1000	cpd00151[e0] <=>
2413	EX_L_Tryptophan_e0	-10	1000	cpd00065[e0] <=>
2420	EX_L_Tyrosine_e0	-10	1000	cpd00069[e0] <=>
2426	EX_UMP_e0	-10	1000	cpd00091[e0] <=>
2427	EX_Uracil_e0	-10	1000	cpd00092[e0] <=>
2428	EX_Urea_e0	-10	1000	cpd00073[e0] + cpd00073[e0] <=>
2429	EX_Uridine_e0	-10	1000	cpd00249[e0] <=>
2430	EX_L_Valine_e0	-10	1000	cpd00156[e0] <=>
2464	EX_Acetoacetate_e0	-10	1000	cpd00142[e0] <=>
2208	EX_Na_plus__e0	-10	1000	cpd00971[e0] <=>
2321	EX_L_Inositol_e0	-10	1000	cpd00121[e0] <=>
2198	EX_Butyrate_e0	-10	1000	cpd00211[e0] + cpd00211[e0] <=>
2214	EX_Carnitine_e0	-10	1000	cpd00266[e0] <=>
2278	EX_Galactose_e0	-10	1000	cpd00108[e0] <=>
2309	EX_hexadecenoate_e0	-10	1000	cpd15237[e0] <=>
2364	EX_O2__e0	-10	1000	cpd00532[e0] <=>
2367	EX_octadecenoate_e0	-10	1000	cpd15269[e0] <=>
2368	EX_octanoate_e0	-10	1000	cpd03846[e0] <=>
2435	EX_Xylose_e0	-10	1000	cpd00154[e0] <=>