MAKERERE UNIVERSITY

In silico* identification of candidate B-cell epitopes and Chimeric proteins derived from OmpA, OmpK35, OmpK36, and Pal proteins for control of *Klebsiella pneumoniae

BY

NABUKENYA EDITH

2100715888

21/U/15888/EVE

BLT III

A RESERCH PROJECT REPORT SUBMITTED TO THE COLLEGE OF VETERINARY MEDICINE, ANIMAL RESOURCES AND BISECURITY IN PARTIAL FULFILLMENT OF THE AWARD OF THE DEGREE OF BACHELOR OF BIOMEDICAL LABORATORY TECHNOLOGY OF MAKERERE UNIVERSITY

MAY 2024

DECLARATION

I Nabukenya Edith declare that this research project entitled *In silico* identification of candidate B-cell epitopes and Chimeric proteins derived from OmpA, OmpK35, OmpK36, and Pal proteins for control of *Klebsiella pneumoniae* is original and to the best of my knowledge and has never been submitted to any institution of higher learning for any academic award or publication.

Signature..... Date 5th/06/2024

APPROVAL

This research report has been prepared under the supervision and guidance of the following supervisors;

Dr. Ann Nanteza (PhD)

Senior Lecturer,

Department of Biotechnical and Diagnostic Sciences,

College of Veterinary Medicine, Animal Resources and Biosecurity (COVAB)

Makerere University,

P.O Box 7062, Kampala.

Signature..... Date 5/06/2024

Dr. Peregrine Sebulime (PhD)

Lecturer,

Department of Wildlife and Aquatic Animal Resources,

College of Veterinary Medicine, Animal Resources and Biosecurity (COVAB)

Makerere University,

P.O Box 7062, Kampala.

Signature..... Date 05/06/2024

DEDICATION

To my two aunties, Edith Nabukenya and Dr Harriet Kisembo, who have worked tirelessly to see to it that I reach where I am now.

ACKNOWLEDGEMENT

I acknowledge Dr Ann Nanteza and Dr Sebulime Peregrine who supervised me during the course of carrying out the research and writing this research project report. I also acknowledge Walter Odur, my friend and course mate who also guided me through out.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

OmpA, Outer membrane protein A.

OmpK35, Outer membrane protein K35.

OmpK36, Outer membrane protein K36.

Pal Peptidoglycan-associated lipoprotein.

HvKP Hyper virulent Klebsiella pneumoniae.

cKP classical K. Pneumoniae.

MDR Multidrug-resistant .

MAC Membrane Attack Complex.

NLM National Library of Medicine.

NIH National Institutes of Health.

CDC Centers for Disease Control and Prevention.

WHO World Health Organization.

LPS Lipopolysaccharide.

CPS, Capsular polysaccharide.

BSA, Bovine serum albumin.

KLH, Keyhole limpet hemocyanin.

CRM197, Cross-Reactive Material 197.

QRDR, Quinolone resistance determining regions.

e.g., for example.

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Klebsiella pneumoniae is an opportunistic Gram-negative bacterium that causes nosocomial infection in healthcare settings. Despite the high morbidity and mortality rate associated with these bacterial infections, no effective vaccine is available to counter the pathogen. The application of bioinformatics tools has influenced the development of vaccines by focusing on isolated epitopes capable of eliciting highly targeted immune responses. In this study, bioinformatics tools were used to identify immune-dominant epitopes from four key *Klebsiella pneumoniae* proteins: Outer membrane protein A (OmpA), Outer membrane protein K35 (OmpK35), Outer membrane protein K36 (OmpK36) and Peptidoglycan associated protein (Pal). Twenty-one chimeric proteins were designed by combining these epitopes with flexible and rigid linkers of varying lengths. The chimeras were then thoroughly evaluated for physicochemical properties, antigenicity, and allergenicity. Among the designed chimeras, >Chimera6 was the most promising candidate, exhibiting superior stability, high antigenicity, excellent solubility, and non-allergenic characteristics. These findings suggest that >Chimera6 has the potential to induce a robust and targeted immune response against *Klebsiella pneumoniae* infections. The study findings represent a significant steptowards the rational design of an effective *Klebsiella pneumoniae* vaccine andopens avenues for further research and validation. By combining computational approaches with experimental validation, the development of a successful vaccine that can combat the menace of *Klebsiella pneumoniae* and contribute to global health security will be possible.

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background

Klebsiella pneumoniae is one of the most common causes of nosocomial infections mainly in immunocompromised patients (Martin et al., 2018), yet it is the second leading cause of hospital-acquired gram-negative bloodstream infections (Hafiz et al., 2023). There are over 150 different strains of Klebsiella pneumoniae, and several of these strains can be found in the same hospital or even within the same patient (Doss, 2019). The prevalence of nosocomial infections due to *Klebsiella pneumoniae* is recorded at up to 10% (Asri et al., 2021). Approximately 8.7% of patients worldwide suffer from nosocomial infections, with especially serious outcomes observed among individuals undergoing cancer treatment, organ transplants, or surgical procedures (Ghashghaee et al., 2018). *Klebsiella pneumoniae* is also responsible for diseases like bloodstream infections, surgical site infections, liver abscesses, and meningitis (Bengoechea et al., 2019). It displays a high degree of virulence and antibiotic resistance (Zhu et al., 2021a). In 2019, *Klebsiella pneumoniae* ranked as the second most prevalent pathogen, following *Escherichia coli*. It was accountable for the highest number of deaths attributed to antimicrobial resistance, leading to over 600,000 deaths globally (Hafiz et al., 2023). As of November 2023, the World Health Organization (WHO) reported increased resistance levels of *Klebsiella pneumoniae* to essential antibiotics. This has resulted in a greater reliance on last-resort drugs such as carbapenems.

Factors contributing to virulence mainly include capsules, siderophores, lipopolysaccharide (LPS),s and fimbriae (Zhu et al., 2021a). The polysaccharide capsule of the organism is the most important virulence factor followed by the lipopolysaccharide that covers the outer surface of Gram-negative bacteria. The sensing of the lipopolysaccharide releases an inflammatory cascade in the host organism and has been a major culprit of the sequela in sepsis and septic shock. Fimbriae allows the organism to attach to the host cells whereas siderophores acquire iron from the host to allow propagation of the infecting organism (Riwu et al., 2022). The release of antibiotic-inactivating enzymes, changes in membrane permeability, modification of antibiotic target sites, alteration of metabolic pathways and activation of efflux pump systems are possible resistance mechanisms of *Klebsiella pneumonia* against different classes of antibiotics (Li et al., 2023).

The worldwide prevalence of nosocomial infections by *Klebsiella pneumoniae* is 10% (Asri et al., 2021). The overall prevalence of pathogenic *Klebsiella pneumoniae* in Uganda is 46.3% (Ssekatawa et al., 2021). The study that examined 227 isolates of multi-drug resistant *Klebsiella pneumoniae* collected from four different tertiary hospitals across Uganda to assess their resistance to carbapenem antibiotics discovered a relatively high overall prevalence of phenotypic carbapenem resistance, amounting to 23.3%. This finding aligns with similar studies in Uganda and Tanzania, which reported rates of 22.4% and 24% respectively. In contrast, studies in North and West Africa observed significantly higher resistance rates exceeding 50%, with *Klebsiella pneumoniae* being the most frequently isolated bacteria. A large study encompassing multiple provinces in West Africa revealed alarmingly high phenotypic resistance rates ranging from 47% to 50% for imipenem, meropenem, and doripenem and 84% to 89% for ertapenem (Ssekatawa et al., 2021).

Currently, there is no licensed vaccine for the prevention of *Klebsiella pneumoniae* infections (Zhang et al., 2021). One of the challenges to control *Klebsiella pneumoniae* infections relies on the great genomic diversity of circulating strains (Right et al., 2022). However, various vaccine approaches have been investigated, including Uromune®, a live attenuated vaccine with tonB deletion, which encodes an iron uptake protein, given sublingually; Klebvax®, CPS vaccine, which was not further developed due to the extensive range of clinically relevant CPS types; conjugated vaccines that connect polysaccharides to diverse carrier peptides such as BSA, KLH, and CRM197; and bioconjugate vaccines designed to target the capsule of hypervirulent *Klebsiella pneumoniae* (hvKp) (Moscoso et al., 2022). The application of bioinformatics tools has influenced the development of vaccines by focusing on isolated epitopes capable of eliciting highly targeted immune responses. Through epitope prediction and their incorporation into novel vaccine formulations, this approach offers a more advantageous option compared to traditional whole-pathogen vaccines (Soria-Guerra et al., 2015). In this study, implemented reverse vaccinology approach that encompasses a comprehensive evaluation of vital aspects of the pathogens to explore immunogenic epitopes against OmpA, OmpK35, and OmpK36 and Pal proteins of *Klebsiella pneumoniae.*

1.2 Statement of the problem

Currently, there is no licensed vaccine for the prevention of *Klebsiella pneumoniae* infections (Zhang et al., 2021). While *Klebsiella pneumoniae* is traditionally associated with hospital-acquired infections and tends to target immunocompromised individuals as an opportunistic pathogen, there is a rising occurrence of hypervirulent and multidrug-resistant strains. This is

attributed to the large accessory genome, comprising both plasmids and chromosomally encoded genes (Yang et al., 2023). As a result, the use of antimicrobials in the treatment of infections caused by *Klebsiella pneumonia* has become progressively more difficult (Y. da Silva et al., 2019). Hyper virulent *Klebsiella pneumoniae* (HvKP), a newly identified pathotype alongside classical *Klebsiella pneumoniae* (cKP), has been acknowledged for its heightened pathogenicity and increased mortality rates attributed to its hyper virulent nature (Lin et al., 2020). Unlike cKP caused nosocomial infections, hvKp results in community-acquired infections (Zhu et al., 2021).The worldwide spread of highly virulent strains of *Klebsiella pneumoniae* and the development of antibiotic-resistant variants of this pathogen have limited treatment choices, leading to a renewed focus on the development of vaccines against it (Zargaran et al., 2021). Besides the high rate of morbidity and mortality, it also prolongs hospital stays and treatment costs. The global rise of nosocomial multidrug-resistant *Klebsiella pneumoniae* is anticipated to escalate in the future. Healthcare personnel should be aware of the most effective treatments for both managing and preventing the pathogen (Asri et al., 2021).Hence this research is generating required information that is to contribute to the ongoing search for the vaccine.

1.3 Objectives

1.3.1 General objective

In silico identification of candidate B-cell epitopes and Chimeric proteins derived from OmpA, OmpK35, OmpK36 and Pal proteins for control of *Klebsiella pneumoniae*.

1.3.2 Specific objective

1) To identify candidate B- cell epitopes within OmpA, OmpK35, OmpK36 and Pal proteins that have the potential of eliciting immune response using *in silico* techniques.
2) To determine the antigenicity of chimeric proteins derived from the selected epitopes.
3) To determine physicochemical parameters of candidate chimeric proteins.

1.4 Research questions

1) Which candidate B cells epitopes are within OmpA, OmpK35, and OmpK36 and Pal proteins and have the potential to elicit an immune response?
2) What is the antigenicity of selected chimeric proteins derived from combinations of epitopes?
3) What are the physicochemical parameters of candidate chimeric proteins?

1.5 Justification

Although most infections caused by *Klebsiella pneumoniae* occur in hospital settings or affect individuals with compromised immune systems, there is a growing occurrence of highly aggressive *Klebsiella pneumoniae* infections acquired in the community, known as hypervirulent *Klebsiella pneumoniae* (hvKp) infections(Feldman et al., 2019). By focusing on the outer membrane of *Klebsiella pneumoniae* and identifying the B cell epitopes that are found on the proteins OmpA, OmpK35, OmpK36, and Pal, and designing chimeric proteins from the identified epitopes, this research provided information that can be considered in the ongoing development of a vaccine against *Klebsiella pneumoniae*. Epitope mapping in different target proteins is considered, for it allows us to focus on the selection of the most potent epitopes that could serve as potential targets for the production of epitope-based diagnostics and vaccines (Ahmad et al., 2016). Chimera production becomes known as a strategy to aggregate these potential epitopes into a single improved antigen to maximize efficacy and minimize side effects in comparison to the available vaccines (Montero et al., 2023)**.** This is because incorporating multiple epitopes from different virulence factors allows the vaccine to target a variety of strains and variants of the pathogen and can also induce a stronger and more targeted immune response compared to individual antigens. Antigenicity prediction is used to determine the peptides that have high antigenicity and can be developed as vaccine candidates (Yurina & Adianingsih, 2022).

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Etiology of *Klebsiella pneumoniae*

Klebsiella pneumoniae, a member of the Enterobacteriaceae family (Z. Wang et al., 2022), is a Gram-negative bacterium known for its encapsulated and non-motile nature (Assoni et al., 2021). The bacterium is a normal flora of the mouth and intestines, meaning it typically colonies on human mucosal surfaces of the oropharynx and gastrointestinal tract. (Zhang et al., 2021)However, it is an important opportunistic pathogen infecting critically ill and immune-compromised patients(Dar et al., 2019). Transmission of the bacteria occurs through direct contact, usually via the contamination of ventilator equipment and catheters. Additionally, person-to-person transmission can occur through touch, but no known airborne spread (Naveed et al., 2023). The initial phase in the progression of nosocomial infection involves the opportunistic colonization of the gastrointestinal tract by *Klebsiella pneumoniae* strains. This sets the stage for the potentially spreading bacteria to other tissues, leading to severe and life-threatening infections (Moscoso et al., 2022). The heightened virulence can largely be attributed to elements such as the capsule, siderophores, lipopolysaccharide (LPS), and fimbria.(Parrott et al., 2021).The polysaccharide capsule enables it to evade the host's immune defenses, lipopolysaccharides, found on the bacterial surface, trigger host inflammatory responses, often contributing to sepsis and septic shock. Fimbriae allow the bacterium to attach to host cells, and siderophores are essential for acquiring iron from the host to sustain infection (Ashurst et al., 2022).

2.2 Epidemiology

Klebsiella pneumoniae primarily resides in the human body, with 5% to 38% carrying it in their stool and 1% to 6% in their nasopharynx in the general community. It is commonly found in the gastrointestinal tract and on the hands of the hospital staff, leading to potential outbreaks in healthcare settings. Higher colonization rates are observed in people of Chinese ethnicity and those with chronic alcoholism. In hospitalized patients, *K. pneumoniae* carriage is much higher, with some studies reporting rates as high as 77% been stool samples, often related to the use of antibiotics(Esposito et al., 2018).

Pneumonia caused by *K. pneumoniae* can be categorized as acquired in the community or the hospital. While community-acquired pneumonia is relatively common, *Klebsiella pneumoniae* infections are rare in Western cultures, accounting for about 3% to 5% of cases.

However, in developing countries like Africa, it can make up around 15% of pneumonia cases. Globally, *Klebsiella pneumoniae* contributes to approximately 11.8% of all hospital-acquired pneumoniae cases. Among ventilator-associated pneumonia cases, 8% to 12% are attributed to *Klebsiella pneumoniae*, compared to 7% in non –ventilated patients. The mortality rate varies, with a range of 50% to 100% in individuals with alcoholism and septicemia (Ashurst et al., 2022).

2.3 Pathophysiology

Protection against bacterial invasion primarily relies on two factors: polymorph nuclear granulocytes, which engulf bacteria through phagocytosis, and serum complement proteins, which possess bactericidal properties. In the context of *Klebsiella pneumoniae* infection, the alternate complement activation pathway plays a more prominent role. Neutrophil myeloperoxidase and lipopolysaccharide-binding protein contribute to defense against this specific infection.

Bacteria possess a polysaccharide capsule composed of intricate acidic polysaccharides that influence their pathogenicity. This capsule shields bacteria from both phagocytosis and serum bactericidal proteins. Additionally, the bacteria adhere to host cells through various fimbrial and non-fimbrial adhesions, a crucial step in the infectious process (Ashurst et al., 2022).

2.4 Treatment / Management

Typically, for community-acquired *K. pneumoniae* pneumonia, a 14-day treatment regimen can involve either a third or fourth-generation cephalosporin as a stand-alone therapy, or a respiratory quinolone as monotherapy. Another option is combining one of these regimens with an aminoglycoside. For individuals with a penicillin allergy, aztreonam or respiratory quinolone should be considered (Ashurst et al., 2022). In the case of nosocomial infections, a carbapenem can be used as a single therapy until antibiotic sensitivities are reported (C. Liu & Guo, 2018).

2.5 Prognosis

The prognosis of *Klebsiella pneumoniae* is poor (Seladi-Schulman & Caporuscio, 2020), especially in patients who have alcohol use disorder, diabetes mellitus, nosocomial infection, or septicemia. Mortality from this type of pneumonia exceeds 50% (Ashurst et al., 2022). Swift identification of *Klebsiella pneumoniae* in samples by doctors, followed by the prompt prescription of suitable antibiotics, leads to improved outcomes. However, delays in diagnosis and testing are frequent and can result in a less favorable prognosis. The prognosis

for individuals affected by *Klebsiella pneumoniae* pneumonia is typically grim. Even when physicians select the correct antibiotic treatment, mortality rates range from 30% to 50%. In people with pneumonia from these bacteria, the infection may hinder lung function in the long term, possibly for months (Seladi-Schulman & Caporuscio, 2020).

2.6 Virulence factors of *Klebsiella pneumoniae*

Hyper-virulence is primarily influenced by factors such as the presence of a capsule, the production of siderophores, the composition of lipopolysaccharide (LPS), and the expression of fimbriae (Parrott et al., 2021). Lipopolysaccharides and CPS are the two main factors responsible for the pathogenicity of microorganisms. Lipopolysaccharides contain antigens such as lipid A, core, and O-polysaccharide that are required for microorganisms to repel complement-mediated payoff. CPS is the pathogen's most distant subcaste, and it is primarily involved in resistance to phagocytosis by polymorphonuclear cells by acting as a physical hedge. As a result, both factors are necessary for microorganisms to spread through the blood and cause sepsis (Riwu et al., 2022).

2.7 Antibiotic Resistance

Over time, *Klebsiella pneumoniae* has developed several mechanisms by which they can withstand antimicrobials, rendering the action of these substances ineffective (Y. da Silva et al., 2019). The studies of anti-microbial resistance of *Klebsiella pneumoniae* strains revealed their widespread resistance to aminoglycosides, fluoroquinolones, cephalosporins and carbapenems (Kot et al., 2023). The global emergence and spread of antibiotic-resistant genes in *Klebsiella pneumoniae* isolates, including extended-spectrum beta-lactamase (ESBL) and carbapenemase genes have significant adverse effects on public health(Sharma et al., 2023).

2.7.1 Aminoglycoside resistance gene

Aminoglycoside resistance genes, including those responsible for aminoglycoside-modifying enzymes (AMEs) and 16S ribosomal RNA methyl transferases (16S RMTases), are particularly common in strains of *Klebsiella pneumoniae* that exhibit resistance to carbapenem antibiotics (Liang et al., 2015). The primary resistance mechanism to aminoglycosides is the synthesis of aminoglycoside-modifying enzymes. These enzymes alter 16S rRNA, diminishing its affinity for aminoglycosides, resulting in robust resistance to various aminoglycosides like arbekacin, amikacin, and kanamycin. Presently, seven genes encoding 16S rRNA methylases have been characterized, namely armA, rmtA, rmtB, rmtC, rmtD, rmtE, and npmA (Huang et al., 2022).

2.7.2 Quinolone resistance gene

Fluoroquinolone resistance is brought about through multiple mechanisms, with the primary one involving chromosomal mutations within the quinolone resistance determining regions (QRDR). These mutations affect key genes, specifically DNA gyrases (gyrA and gyrB genes) and topoisomerase IV (parC and parE genes(Geetha et al., 2020) The other mechanism of resistance is plasmid-mediated quinolone resistance (PMQR) , which involves three key mediators known as the qnr proteins (qnrA, qnrB, and qnrS). These proteins provide protection to the enzymes responsible for encoding DNA gyrase and topoisomerase IV (Geetha et al., 2020). Another factor contributing to fluoroquinolone resistance is the presence of the acc(*6*)-Ib-cr gene, which is a modified form of aminoglycoside acetyltransferase. This gene alters ciprofloxacin and norfloxacin, resulting in reduced susceptibility to these antibiotics (Jomehzadeh et al., 2022).

2.7.3 Beta-lactam resistance gene

Ambler's classification categorizes over 1800 described variants of β-lactamases into four types based on their protein sequences: type A (serine penicillinases), Type B (Metallo-β-lactamases), Type C (cephalosporinases), and Type D (oxacillinases). These types exhibit resistance to penicillins, a wide range of β-lactam antibiotics, cephalosporins, and cloxacillin, respectively. (Hussain et al., 2021)Class A and Class D encompass ESBLs and primarily comprise CTX-M, TEM, SHV, and OXA enzymes, while Class B β-lactamases consist of Metallo-β-lactamases like NDM-1. Class C includes AmpC β-lactamases (Hussain et al., 2021). The most current classification is from Bush & Jacoby, based on the activity of the enzyme. This classification divides enzymes into three groups: Group 1 comprises cephalosporinases, which belong to molecular class C by Ambler's classification, and are notably active against cephalosporins, resistant to inhibition by β-lactamase inhibitors like clavulanic acid and tazobactam. Group 2 includes serine β-lactamases from Ambler's groups A and D, representing the largest category of β-lactamases, further categorized into subgroups 2a and 2b according to their specific action spectra. Group 3 encompasses enzymes that require a zinc ion at their active sites to exert antimicrobial activity (Y. da Silva et al., 2019).

2.7.4 Polymyxin resistance gene

Polymyxin disrupts membrane integrity by displacement of extracellular ion (Ca2+/Mg2+) with binding to negatively-charged lipopolysaccharide (LPS), and causing cell lysis (G.

Wang et al., 2020). In *Klebsiella pneumoniae*, resistance to polymixins is mainly due to modification of the lipid A phosphate moieties of the lipopolysaccharide (LPS) with a sugar or ethanolamine, which reduces the electrostatic interaction between the cationic polymixins and anionic LPS (Esposito et al., 2018). The changes are due to mutations in core genes responsible for the maturation of lipid A (lpx M and its regulator ram A) and lipid A neutralization, amino arabinose (pagP, pmrE), an additional combination of phosphoethanolamine (pmrC), or palmitate (pagP). Multiple LPS modified gene regulator, such as phoPQ, pmrA, and pmrD are also involved in Polymyxin resistance. Mutations in one of the two other regulatory genes are also sufficient to cause Polymyxin resistance (G. Wang et al., 2020).

2.7.5 Tigecycline resistance gene

Tigecycline, the first member of the glycylcycline class antibiotic, is considered one of the most effective antibiotics against multidrug-resistant (MDR) infections. However, the emergence and wide distribution of two novel plasmid-mediated tigecycline resistance genes, Tet(X3) and Tet(X4), pose a great threat to the clinical use of tigecycline (Zhai et al., 2022). The known mechanism of tigecycline tolerance is encoded in chromosomes and involves various modifications. These modifications include changes in the 30S and 16S ribosomal units, alterations in cell permeability, and the overall expression of efflux pumps such as AarAB-TolC and OqxAB. Additionally, changes in the expression levels of regulatory proteins like RamA, RamR, RarA, and AcrA can also contribute to tigecycline resistance. Mutations in the RamR gene can enhance the expression of RamA. Furthermore, Lon and rpsJ genes have been associated with tigecycline resistance in K. pneumoniae. Reduced transcript levels of porin ompK35 can lead to increased resistance in K. pneumonia strains (G. Wang et al., 2020). In summary, the RamA and AcrAB pathway, as well as the combined action of RarA with the AcrAB and OqxAB pathways, have primarily been associated with tigecycline resistance in K. pneumoniae (Sheng et al., 2014).

2.8 Overview of the current state of research on vaccine development against *Klebsiella pneumoniae*

No approved vaccine exists for *Klebsiella pneumoniae* yet (Ranjbarian et al., 2023). Lipopolysaccharide (LPS, O-antigen) and capsular polysaccharide (K-antigen) serve as potential targets for a *Klebsiella pneumoniae* vaccine; however, they exhibit evident limitations (Zhang et al., 2021). There are 8 O-antigens and 77 K-antigens to date. O-antigens, however, prove ineffective as targets for a *Klebsiella pneumoniae* vaccine, due to

associated toxic side effects in active immunization. While K-antigens display immunogenicity without toxicity, a vaccine must incorporate a minimum of 24 major K types to encompass 70% of *Klebsiella pneumoniae* strains (Zhang et al., 2021).

In silico, analyses have demonstrated that OMPs are promising candidates for a Klebsiella vaccine. Among the epitopes, those derived from OmpA and outer membrane porins are among the most promising candidates; however, the formulation has still not been tested in vivo (Assoni et al., 2021). Ideally, a vaccine should contain conserved epitopes that in turn can generate specific B-cell and T-cell (CD4 and CD8) responses. This will enable the induction of pathogen-specific immune responses with minimal side effects by incorporating highly conserved epitopes in a vaccine formulation (Dar et al., 2019).

In this study, two of the primary antigens of interest are peptidoglycan-associated lipoprotein (Pal) and outer membrane protein A (OmpA), which are identified as the prominent outer membrane proteins released by gram-negative bacteria in cases of sepsis (Dar et al., 2019). *Klebsiella pneumoniae* produces two classical trimeric porins, OmpK35 and OmpK36 (Sugawara et al., 2016).

2.9 Target proteins

Klebsiella pneumoniae produces three major proteins, OmpK35, OmpK36, and OmpA, in its outer membrane (Sugawara et al., 2016). Outer membrane proteins (OMPs) are crucial for the transport of molecules, maintenance of membrane integrity, and pathogenesis (Choi & Lee, 2019). Outer membrane proteins were selected for epitope prediction because of their high expression in bacterial cell walls, conservation, and active involvement in growth and virulence (Rahmat Ullah et al., 2021). Therefore more likely to be good vaccine candidates as they are readily accessible to the host immune system. Two of the prioritized antigens in this study, peptidoglycan-associated lipoprotein (Pal) and outer membrane protein A (OmpA), are recognized as dominant OMPs released by gram-negative bacteria during sepsis. The proteins OmpA, OmpK35, OmpK36, and Pal are the OMPs to focus on in this study because they are non-human homologs, lie in the outer membrane and extracellular location, essential and most importantly highly virulent (Dar et al., 2019).

2.9.1 Outer membrane protein A

Outer membrane protein A (OmpA) is a class of proteins highly conserved among the Enterobacteriaceae family and throughout evolution; it is one of the best-characterized outer

membrane (OM) proteins (March et al., 2011). *Klebsiella pneumoniae* OmpA is important for immune evasion *in vitro* and *in vivo*. OmpA plays a role in maintaining the position of the peptidoglycan cell wall in the periplasm. Acts as a porin with low permeability that allows slow penetration of small solutes (Hussein et al., 2023).

2.9.2 Peptidoglycan-associated lipoprotein

Klebsiella pneumoniae Pal proteins are important in the maintenance of cell integrity, contribute to virulence, and could be used as attenuated vaccines. *Klebsiella pneumoniae* Pal confers protection not only against serum killing but also against neutrophil phagocytosis and killing (Hsieh et al., 2013). Pal proteins are part of the Tol-Pal system, which plays a role in outer membrane invagination during cell division.

2.9.3 OmpK35 and OmpK36

Porins are important for bacterial survival because of their role in the exchange of substances, including nutrients and toxic metabolites.OmpK35 and OmpK36 are the major outer membrane porins of *Klebsiella pneumoniae*. OmpK35 and OmpK36 provide a channel that allows a wide range of antibiotics to penetrate the *Klebsiella pneumoniae* cell wall. OmpK35 allows for more efficient penetration of cephalosporin than OmpK36 does. OmpK35 and OmpK36 both play dual roles in *Klebsiella pneumoniae* infection (Tsai et al., 2011).

2.10 Epitope mapping concept and relevance

In the field of immunology, epitope mapping involves the experimental determination of the specific location where an antibody binds to its target antigen, typically a protein. This process of identifying and characterizing the antibody-binding site plays a crucial role in advancing the research and development of novel therapeutics, vaccines, and diagnostic tools (Davidson & Doranz, 2014). Epitopes are considered important to both clinical and biomedical researchers because of their potential for vaccine design against variable and rapidly mutating pathogens (Soria-Guerra et al., 2015).

Chimeric protein concept and its relevance

A chimeric protein is a synthetic fusion protein composed of multiple epitopes or antigenic regions from different pathogenic proteins. This unique combination of epitopes allows the chimeric protein to elicit a targeted and robust immune response against various components of the pathogen (M. T. de O. Silva et al., 2021). The concept of using chimeric proteins as vaccine candidates has gained significant attention due to its potential to overcome the limitations of traditional vaccines and enhance protective immunity against infectious agents

(Shanmugaratnam et al., 2012). Incorporating multiple epitopes from different virulence factors allows the vaccine to target a variety of strains and variants of the pathogen and can also induce a stronger and more targeted immune response compared to individual antigens.

2.11 B-cell epitopes

B-lymphocytes have surface receptors that detect B cell epitopes as antigenic elements and differentiate into memory cells and plasma cells (Rahmat Ullah et al., 2021). B-cell epitopes can be defined as groups of amino acids that are exposed on the surface, capable of being recognized by either secreted antibodies or B-cell receptors and have the capacity to trigger either cellular or humoral immune responses(Potocnakova et al., 2016). B-cell epitopes can be classified based on their spatial arrangement into two categories; continuous epitopes, which are also known as linear or sequential epitopes, and discontinuous epitopes, alternatively referred to as nonlinear or conformational epitopes. In the case of discontinuous epitopes, the proximity of amino acid residues is a result of their three-dimensional conformation. Identification of B-cell epitopes is a fundamental step for the development of epitope-based vaccines (Potocnakova et al., 2016).

2.12 *In silico* and bioinformatics concept and relevance

In biology and other experimental sciences, an *in silico* experiment is one performed on a computer or via computer simulation whereas bioinformatics is the use of computational methods and tools to analyze biological data, such as DNA sequences, protein sequences, and other biological information. The *in-silico* approach is the first step towards designing a vaccine for which effective selection of the target protein plays an integral role (Rahmat Ullah et al., 2021), and the most critical step is the selection of suitable antigenic epitopes for the target proteins to design a multi-epitope vaccine construct known as a chimeric protein (Chauhan et al., 2019). Bioinformatics tools have revolutionized the ability to identify potential epitopes without the need to cultivate the specific pathogen under investigation. This approach offers significant advantages compared to traditional vaccine development methods, including quicker results and reduced expenses (Soria-Guerra et al., 2015).

2.13 Immune response to *Klebsiella pneumoniae*

Initial innate immune defenses against *Klebsiella pneumoniae* infection include complement, macrophages, neutrophils, and monocytes; these defenses are primary strategies employed by the host to clear infections (Gonzalez-Ferrer et al., 2021).

2.13.1 The complement pathway

The complement system is a fundamental component of the innate immune system in vertebrates playing a vital role in defending the host against various pathogens, including Klebsiella pneumoniae. This intricate system comprises over 20 proteins distributed in the blood stream and tissue microenvironments, all working in harmony to swiftly combat invading pathogens.

The complement system supports the host's defenses against these invading pathogens through two primary systems:

i. Membrane Attack Complex (MAC): This involves the assembly of the membrane attack complex, composed of C5b-C9 proteins. This complex has the ability to directly rupture the bacterial membrane, resulting in the death of the pathogen.
ii. Opsonization: This process enhances phagocytosis by immune cells. Complement proteins, particularly C3b, coat the surface of pathogens, making them more recognizable to phagocytic cells, such as macrophages and neutrophils.

The complement system can be initiated through three well-defined pathways:

- **Classical Pathway:** Activation of this pathway is triggered by binding of antibodies (IgG or IgM) to the pathogens surface. This antibody-pathogen complex initiates the activation of complement proteins, ultimately leading to the formation of C3 convertases.
- **Lectin Pathway:** in this pathway, pattern recognition molecules, specifically mannose-binding lectin, attach to specific carbohydrate structures on the pathogen's surface. This binding initiates the activation of complement proteins, resulting in the formation of C3 convertases.
- **Alternative Pathway:** this is characterized by its continual low-level activity. It can amplify when the complement proteins spontaneously interact with microbial surfaces or in the presence of specific activators. Similarly, to the pathways this leads to the generation of C3 convertases.

2.13.2 Macrophages

Macrophages are the initial phagocytes that encounter Klebsiella pneumoniae at mucosal sites, such as the lung (Bain & Schridde, 2018).They sense and activate mediators that drive

appropriate immune responses, and typically eliminate Klebsiella pneumoniae following engulfment (Gonzalez-Ferrer et al., 2021).

2.13.3 Neutrophils and Monocytes

They sense and activate mediators that drive appropriate immune responses, and typically eliminate Klebsiella pneumoniae following engulfment (Gonzalez-Ferrer et al., 2021). Rapid recruitment of neutrophils and monocytes is critical for effective control of Klebsiella pn

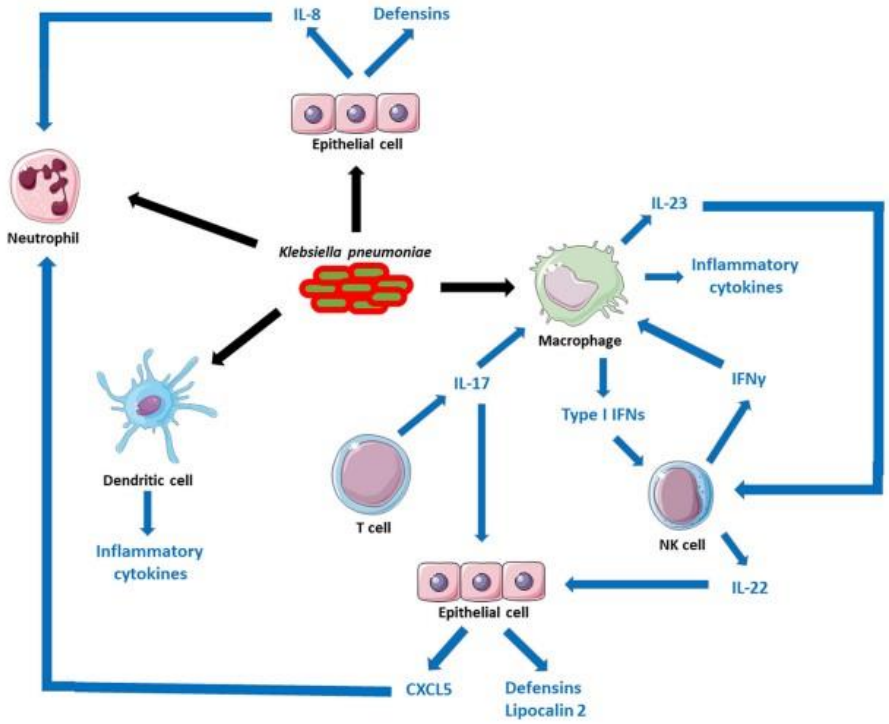

Figure 1: Mechanisms of innate immunity to K. pneumoniae infections.

The figure depicts the cells implicated in containing K. pneumoniae infection. There is conclusive evidence demonstrating the interaction of K. pneumoniae with neutrophils, macrophages (and monocytes [not shown]), dendritic cells and epithelial cells. These interactions are marked with black arrows. The interaction with different subset of T cells, NK cells and other lymphocytes has not been investigated yet, although these cells participate in bacterial clearance. The network of connections between cells, and the role played by different cytokines activating host responses are depicted with blue arrows

2.14 *In silico* databases and resources to be used in this study

A number of tools were used during this study and some of these include the following;

2.14.1 PubMed

PubMed is a free online database and search engine primarily focused on providing access to biomedical and life sciences literature .The United States National Library of Medicine (NLM) operates it, which is part of the National Institutes of Health (NIH). PubMed contains a vast collection of citations, abstracts, and full-text articles from a wide range of sources, including scientific journals, research papers, and other scholarly publications. PubMed allows users to search for specific articles, explore related articles, and access a wealth of information to support their research, clinical practice, or academic studies**.** In this research, PubMed helps in searching for relevant research articles and studies that provide information about epitopes, antigens, and immunological responses related to Klebsiella species and the current state of vaccine production against *Klebsiella pneumoniae*.

2.14.2 UniprotKB (Swiss-prot)

. The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation (How Redundant Are Sequences in UniProtKB?, 2002). UniProtKB/Swiss-Prot (reviewed) is a high quality manually annotated and non-redundant protein sequence database, which brings together experimental results, computed features and scientific conclusions (Boutet et al., 2007). By utilizing UniProtKB, will conveniently access top-tier protein sequences, enriched with essential functional annotations, details about post-translational modifications, and data pertaining to protein-protein interactions.

2.14.3 MEGA 11

This is a widely employed software tool that offers a robust suite of capabilities encompassing sequence alignment, phylogenetic analysis, and evolutionary modeling. Researchers find it to be a potent platform for the examination of genetic and protein variations (Tamura et al., 2021a)**.** In this particular study, MEGA11 will be applied to align the collected protein sequences using the MUSCLE Algorithm.

2.14.4 Simple Consensus maker

In molecular biology and bioinformatics, the consensus sequence is the calculated sequence of most frequent residues; either nucleotide or amino acid, found at each position in a sequence alignment. A simple consensus maker tool is a software that allows users to generate a consensus sequence from multiple sequences. In the context of sequence alignment or bioinformatics, a simple consensus maker tool takes multiple DNA or protein

sequences and generate a consensus sequence that represents the most common or most frequently occurring base or amino acid at each position in the alignment.

A simple consensus maker tool available at the HIV sequence database was employed in this study to generate a consensus sequence from the aligned protein sequences obtained using MEGA11.

2.14.5 National Centre for Biotechnology Information (NCBI)

It is a branch of the United States National Library of Medicine (NLM), that serves as a prominent resource for biological and biomedical information. NCBI provides access to a vast array of databases, tools, and resources related to genetics, genomics, molecular biology, and other areas of life sciences. Some of its well-known services include GenBank (a repository of genetic sequences), PubMed (a database of scientific literature), and BLAST (Basic Local Alignment Search Tool) for sequence analysis, among many others. In this study, the NCBI BLAST (Basic Local Alignment Search Tool) tool was employed to obtain the sequences from different databases, which are homologous to the consensus of retrieved sequences.

2.14.6 MEME Suite (Multiple Em for Motif Elicitation)

The MEME Suite is a comprehensive set of utilities designed for the identification and analysis of short, conserved sequence patterns called motifs within protein sequences (Bailey et al., 2015). These motifs are closely linked to specific protein functions, making them valuable for deducing the biological roles of proteins. This suite encompasses various algorithms, ranging from probabilistic to discrete models, which enable the discovery of motifs. Additionally, it provides tools for scrutinizing the prevalence of established motifs within sets of sequences, gauging motif similarities, and predicting the biological functions and regulatory targets associated with these motifs (Bailey et al., 2015). Motifs themselves are brief, repetitive sequences of amino acids that appear in proteins, typically playing crucial roles in their structure and function. These motifs also serve as useful markers for identifying proteins or protein families that share common characteristics (Zambelli et al., 2013).

2.14.7 The IEDB Analysis resources (Immune Epitope Data Base and Analysis Resources)

This is a comprehensive, publicly available resource for information on the immune epitopes. The IEDB contains data on both linear and conformational epitopes, along with details about the molecules that interact with them, such as antibodies and T cell receptors (Dhanda et al.,

2019). In this study, the IEDB was used to assess various properties of the study protein motifs i.e. their antigenicity, surface accessibility, linearity, beta turns, flexibility and hydrophilicity .The motifs that passed these parameters were considered as potential epitopes.

2.14.8 Epitope linkers

Epitope linkers are short sequences of amino acids used to link different epitopes to form a chimeric protein (Mihara et al., 2009). Epitope linkers can be flexible or rigid linkers. They ensure that the resulting chimeric protein is stable and flexible to minimize epitope interference, and enhance immunogenicity (Patel et al., 2020). Flexible linkers are used when an interaction or movement with a certain degree of the linked domains is demanded (Chen et al., 2013). Rigid linkers separate the functional domains more efficiently than flexible linkers (Gräwe et al., 2020).

2.14.9 Protparam

Protparam is a bioinformatics tool, widely used to analyze the physicochemical properties of the chimeric proteins. The physiochemical parameters include molecular weight, theoretical isoelectric point (pI), and amino acid composition. Protparam also predicts other essential characteristics like extinction coefficient and grand average of hydropathy (GRAVY), which provide an understanding of the protein's stability and solubility. These analyses are crucial in understanding the chimeric protein's overall structure and properties, supporting its suitability as a vaccine candidate (Sever et al., 2016).

2.14.10 VaxiJen v2.0

The VaxiJen tool, a potent immunoinformatics instrument, was utilized to assess the immunogenicity of the engineered chimeric protein. It aims to pinpoint specific regions within the chimeric protein that are likely to trigger a strong immune response. This examination holds significance in the selection of optimal vaccine targets, thereby augmenting the chimeric protein's capacity to evoke protective immunity (Salod & Mahomed, 2022).

2.14.11 ChimeraX

ChimeraX stands out as a robust and adaptable software for visualizing molecular structures, widely utilized by researchers and scientists. It facilitates the 3D visualization and analysis of intricate biological components such as proteins, nucleic acids, and macromolecular complexes (Pettersen et al., 2021). Offering a comprehensive and interactive

perspective of these structures, ChimeraX is equipped with a rich set of features.This software comes with thorough user documentation and is freely available for non-commercial purposes, with compatibility across Windows, Linux, and macOS platforms. Its functionalities encompass various standard molecular modeling applications, including superposition, calculations, measurements, diverse display styles, and the generation of high-resolution images and videos. Continuing the legacy of Chimera, ChimeraX further enhances capabilities in interactive density fitting, map processing, structure analysis, and bridging between sequence and structure(Goddard et al., 2018).

2.14.12 PyMOL

PyMOL, a molecular graphics tool, is extensively employed for the three-dimensional visualization of macromolecules. Its capabilities have been significantly broadened through the integration of numerous plugins, encompassing functionalities such as macromolecular analysis, homology modeling, protein-ligand docking, pharmacophore modeling, virtual screening (VS), and molecular dynamics (MD) simulations. These attributes position PyMOL as a versatile platform for contemporary computational drug design (Yuan et al., 2017). In this research, PyMOL was used to visualize the 3D structure of the chimeric protein.

CHAPTER THREE

3.0 METHODS AND MATERIALS

3.1 Study design

This study utilized a comprehensive suite of bioinformatics tools and databases, including NCBI, MEGA11, HIV consensus sequence maker, MEME Suite, IEDB, ChimeraX, and PyMOL, to conduct an in-depth analysis of protein sequences, identify motifs, and design chimeric proteins. Initially, protein sequences were retrieved from Uniprot KB under NCBI and aligned using MEGA11 to identify conserved regions and generate a consensus sequence using the consensus sequence maker available in the HIV sequence database. The MEME Suite was then employed to discover recurring motifs within these sequences, which were further analyzed for parameters like linearity, surface accessibility, and immunogenicity using IEDB. Based on these insights, chimeric proteins were designed by incorporating critical motifs, followed by their structural and functional analysis through various *in-silico* tools. The 3D structures of these chimeric proteins were modeled using Chimera X and visualized using PyMOL to assess their spatial conformation and potential interactions. Extensive internet utilization facilitated seamless data retrieval and analysis throughout the research, ensuring a robust approach to protein engineering (Rahmat Ullah et al., 2021).

3.2 Retrieval of protein sequences from UniProtKB

The sequences of OmpA, OmpK35, OmpK36, and Pal proteins were obtained from UniprotKB by feeding the individual protein name into the search bar of the website (https://www.uniprot.org/). Downloaded them in Fasta format and saved them in a text editor (notepad)**.**

3.3 Sequence alignment

Sequence alignment is a way of arranging protein (or DNA) sequences to identify regions of similarity that may be a consequence of evolutionary relationships between the sequences Each of the four files was individually imported into MEGA11 and then selected the MUSCLE Algorithm to align the sequences (Tamura et al., 2021b). Then exported in FASTA format all the aligned sequences of all four sequences and saved them on the desktop for easy accessibility.

3.4 Generation of consensus sequences

The consensus sequence serve as a condensed depiction of the predominant amino acid at each position within a set of homologous sequences, offering a simplified overview of the group's diversity (Sternke et al., 2019). To generate this consensus, I employed a consensus maker tool available at the HIV Sequence database through https://www.ebi.ac.uk/Tools/msa/emboss_cons/. Each aligned sequence was inserted individually into the designated text box of the Simple Consensus Maker, chose the "like input" option for the output preferences. I then initiated the consensus generation process by clicking the "Run" button. This procedure was carried out for all four-sequence alignments.

3.5 Selection of homologs

Homologous protein sequences are those with a common evolutionary origin, they share a common ancestor (Pearson, 2013). To identify homologous sequences for each of the four proteins, utilized the BLASTp tool available on the NCBI website https://blast.ncbi.nlm.nih.gov/. The initial step involved inputting the consensus query sequence into the designated text box and then initiating the BLAST search by clicking the corresponding button. Subsequently, I carefully examined the search results and chose nine sequences that exhibited a remarkable degree of similarity with the query, specifically, those with a percentage identity greater than 97.00% and an E-value of 0.0. These selected homologous sequences were then downloaded in Fasta format and saved within a text editor (notepad). This process was replicated for all four proteins.

3.6 Motif discovery

The MEME Suite tool version 5.5.3 https://meme-suite.org/meme was used to identify motifs in the protein sequences (Bailey et al., 2015). The MEME Suite was opened in Google Browse, and then the motif discovery option, which displayed a submission form with an option for motif number. The file containing homologs and consensus sequences combined was chosen as the input file and uploaded.

3.7 Data analysis

Before being deemed as full protein vaccine candidates, the identified motifs within the proteins had to undergo analysis based on six critical parameters. These parameters encompassed the antigenicity of the motifs, their linearity, and their potential as β-turns, their surface exposure, their structural flexibility, and their hydrophobicity (Dhanda et al., 2019).

Specifically, the Linear B-cell epitope prediction tool from IEDB was employed, which was accessed at https://www.iedb.org/, (Martini et al., 2020).

The IEDB encompasses a suite of valuable tools, including Bepipred Linear Epitope Prediction 2.0, Chou & Fasman Beta-Turn Prediction, Emini Surface Accessibility Prediction, Karplus & Schulz Flexibility Prediction, Kolaskar & Tongaonkar Antigenicity, and Parker Hydrophilicity Prediction. Each of the consensus sequences derived from the proteins was copied and pasted into the Antibody Epitope Prediction entry box and subsequently submitted to the six distinct IEDB tools for the prediction of the above-mentioned six parameters (Martini et al., 2020).

3.7.1 Linear B-cell epitope prediction

Employed the BepiPred version 2.0 linear epitope prediction tool available in the IEDB to pinpoint Linear B cell epitopes within the conserved segments of protein sequences such as OmpA, OmpK35, OmpK36, and Pal. Following the analysis of each amino acid, this tool generated an HTML file with a graphical representation, highlighting the linear regions on the entire protein sequence, (Martini et al., 2020). Motifs with an average score exceeding the defined threshold were categorized as linear epitopes, signifying a higher likelihood of their inclusion in an epitope.

3.7.2 Beta-Turn Prediction

The Chou & Fasman Beta-Turn Prediction tool was applied to predict the presence of beta turns, aiming to pinpoint the locations of epitopes. This tool assesses the likelihood of specific amino acids preferring being situated within coil, sheet, or helix structures. Motifs with an average score exceeding the defined threshold were categorized as having the ability to beta-turn.

3.7.3 Surface accessibility prediction

The Emini Surface Accessibility Prediction tool was employed to predict protein regions that are expected to be exposed on the surface. This feature holds significant importance in terms of facilitating the binding of epitopes to B cells and the initiation of a robust immune response (Martini et al., 2020). The selection of motifs was influenced by the threshold value obtained when running consensus sequences in the IEDB, which ensured that the most suitable motifs were chosen.

3.7.4 Structural Flexibility Prediction

The Karplus & Schulz Flexibility Prediction tool was applied to predict the structural flexibility of the protein. This flexibility is of paramount importance for various biological functions, including catalysis, binding, and allostery (Martini et al., 2020). Notably, greater motif flexibility enhances its ability to bind to other molecules and, consequently, increases its interaction potential with antibodies. Motifs with an average score exceeding the defined threshold were categorized as being structurally flexible.

3.7.5 Antigenicity prediction

The Kolaskar & Tongaonkar Antigenicity prediction tool was employed to assess the antigenicity of peptide motifs (Martini et al., 2020). This method employs a semi-empirical approach that draws on the physical and chemical characteristics of amino acid residues, as well as their occurrence in previously recognized epitopes. It boasts an estimated accuracy rate of 75% in identifying antigenic peptides.

3.7.6 Hydrophilicity prediction

The Parker Hydrophilicity Prediction tool was employed to assess the hydrophilic characteristics of peptide motifs (Martini et al., 2020). This tool relies on a predictive algorithm founded on the Parker hydrophilicity scale, which ascribes hydrophilicity values to each of the 22 amino acids based on their inherent chemical properties and their interactions with water. When provided with an input amino acid sequence, the tool computes a hydrophilicity score for each position along the sequence. The resultant hydrophilicity profile furnishes insights into whether the protein sequence as a whole is more hydrophilic or hydrophobic in nature.

3.8 Chimeric protein design

The predicted B cell epitopes from OmpA, OmpK35, OmpK36, and Pal proteins were incorporated to create multiple options of chimeric proteins. In the design process, flexible linkers, rigid linkers, and combined flexible and rigid linkers with varying lengths were used to connect the individual epitopes to form multiple options of the chimeric proteins (Gräwe et al., 2020).

3.9 Physicochemical properties, antigenicity, and allergenicity assessments

The physicochemical properties of different chimeric proteins were evaluated to gain insight into their molecular characteristics. Various parameters, including molecular weight, theoretical (pI), amino acid composition, atomic composition, extinction coefficient,

estimated half-life, instability index, aliphatic index, and Grand Average of Hydropathy (GRAVY), were analyzed using the Protparam server https://web.expasy.org/protparam. These analyses provided essential information on the stability and solubility of different chimeric proteins (Sever et al., 2016). The different chimera options were submitted to the VaxiJen server for antigenicity assessment. It outputs the antigenicity score for the submitted protein sequence which indicates its ability to elicit an immune response (Salod & Mahomed, 2022). These sequences were subsequently subjected to the AlgiPred tool for assessment of allergenicity. AlgPred is a widely used computational tool that predicts allergenicity based on the sequence of a protein. It uses an algorithm to assess the presence of allergenic motifs or sequence patterns associated with known allergens (Dana et al., 2020). The best chimera was selected based on the above parameters.

3.10 Chimeric protein 3D structure modeling

The chimeric protein 3D structure modeling was done using the ChimeraX server, a reliable and widely used tool for predicting protein structures based on homology modeling.

3.11 Visualization of 3D structure model

To visualize the 3D structure model of the chimera, PyMOL molecular graphics system was used (Yuan et al., 2017)**.** The 3D structure model of the chimera obtained from the ChimeraX server was loaded into PyMOL, allowing for a detailed examination of the protein's spatial arrangement and conformation. The visualization provided a valuable understanding of the folding pattern and secondary structure elements (Rahman et al., 2022). PyMOL enabled the generation of high-quality images and interactive representations, facilitating a comprehensive analysis of the chimeric protein's structure (Yuan et al., 2017).

3.12 Ethical statement

The study did not require ethical approval as it was using open access data that are publically accessed.

3.13 Research Limitations

The use of certain software packages and databases was limited in this research study due to financial constraints. As a result, the study relied on free and open-source alternatives, some of which lacked certain features compared to paid options. However, the study utilized the most essential tools and data available through these free open resources to conduct the research successfully and reduce overall expenses

CHAPTER FOUR

4.0 RESULTS

4.1Motif discovery

Multiple Em for Motif Elicitation (MEME Suite) analysis tool was utilized to discover peptide motifs from the consensus sequences of the four proteins. Nine motifs were discovered for OmpA and OmpK36, eight for OmpK35 while five motifs were discovered for Pal protein as shown in Figures 2, 3,4, and 5 below.

4.1.1 The OmpA peptide motifs

The motifs that were identified from OmpA consensus sequences included residues from; 1-29,34-83,84-124,126-175,176-204,205-210,212-261,264-304 and 305-354 .

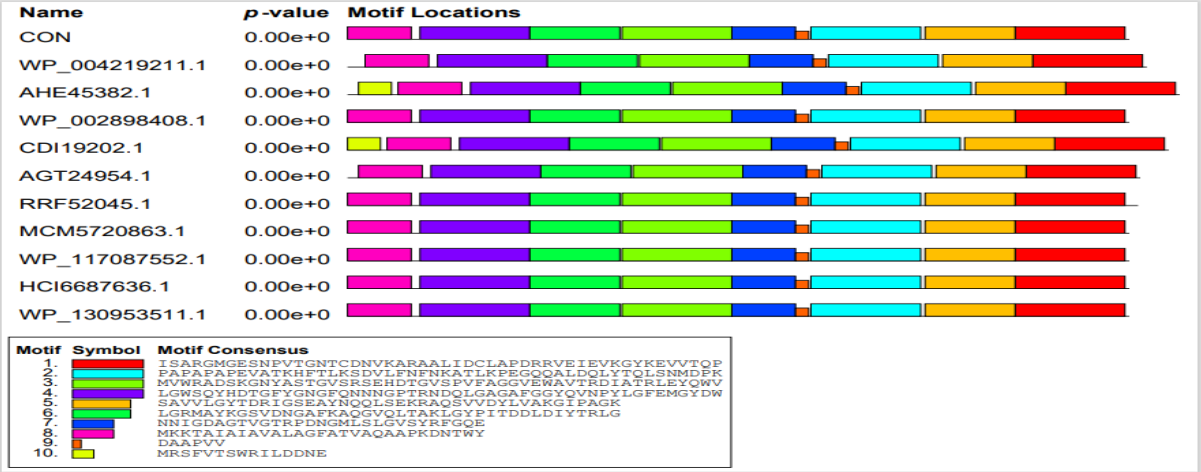

Figure 2: OmpA peptide motifs

4.1.2 The OmpK35 peptide motifs

The motifs that were identified from OmpK35 consensus sequences included residues from; 1-29,34-83,84-133,138-187,189-196,197-246,248-297 and 300-349.

Figure 3: The OmpK35 peptide motifs

4.1.3 The OmpK36 peptide motifs

The motifs that were identified from OmpK36 consensus sequences included residues from; 1-50, 55-104,110-159,160-180,181-230,232-281,282-331,332-360 and 362-367.

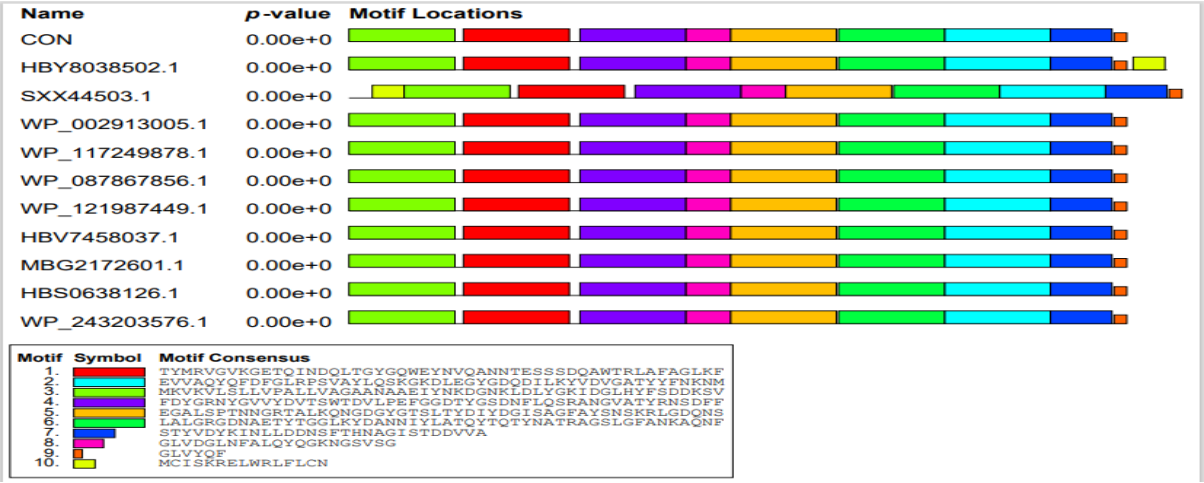

Figure 4: OmpK36 peptide motifs

4.1.4 The Pal peptide motifs

The motifs that were identified from Pal consensus sequences included residues from; 2-12, 15-64,65-70,71-120 and 125-174.

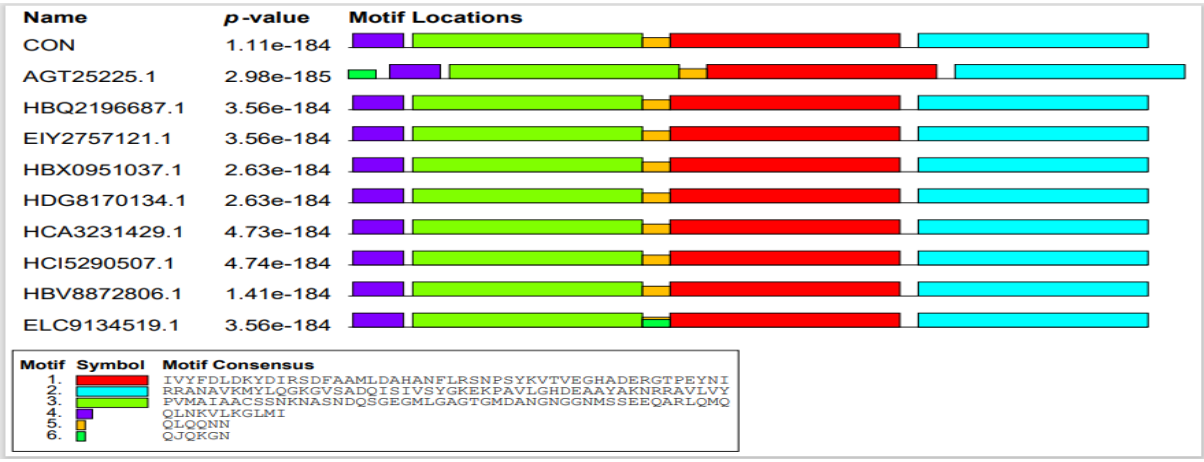

Figure 5:OmpK36 peptide motifs

4.2 Motif analysis

The motifs were assessed using IEDB Analysis tools, and their average scores were computed in Microsoft Excel to gauge their potential to trigger an immune response. Peptide motifs with mean scores surpassing predetermined thresholds were considered, indicating a higher likelihood of being epitopes. The IEDB tools utilized included BepiPred Linear

Epitope Prediction 2.0, Chou & Fasman Beta-Turn Prediction, Emini Surface Accessibility Prediction, Karplus & Schulz Flexibility Prediction, Kolaskar & Tongaonkar Antigenicity, and Parker Hydrophilicity Prediction. The selection criteria for successful epitopes were as follows: (1) Motifs had to exhibit Antigenicity, Surface Accessibility, and Linearity, with a minimum passage of three out of six parameters. (2) For motifs passing at least four out of six parameters, they needed to demonstrate Antigenicity and either Surface Accessibility or Linearity(Mathew et al., 2022). Tables 1, 2, 3, and 4 display motif positions for OmpA, OmpK35, OmpK36, and Pal, along with their respective scores for various IEDB Analysis parameters. Scores highlighted in light blue exceed the threshold, while those in red indicate successful motifs surpassing the threshold.

4.2.1 The OmpA Motif Evaluation

The OmpA Motif at position 212-261 was selected as the potential epitope to be included in the chimeric protein because it passed the preset conditions (Table 1).

Table 1: Evaluation of OmpA peptide motifs using IEDB analysis resources

Motif position	Mean BepiPred (0.5000)	Mean Chou and Fasman (1.021)	Mean Emini (1.000)	Mean Karplus and Schulz (1.000)	Mean Kolaskar and Tangaonkar (1.017)	Mean Parker (1.982)
1-29	0.372	0.857	0.842	0.954	1.045	0.975
34-83	0.547	1.163	1.158	1.0004	0.975	2.098
84-124	0.491	1.0177	0.792	0.995	1.018	1.539
126-175	0.550	1.011	1.038	1.002	1.011	2.191
176-204	0.636	1.106	0.891	1.007	0.989	2.337
205-210	0.622	0.908	O.634	0.971	1.091	2.288
212-261	0.503	0.990	1.234	1.011	1.027	1.789
264-304	0.481	0.983	1.105	1.004	1.049	2.013
305-354	0.510	0.985	0.904	1.009	1.026	2.5353

4.2.2 The OmpK35 motif evaluation

The OmpK35 Motif at positions 248-297 was selected as the potential epitope to be included in the chimeric protein because it passed the preset conditions (Table 2).

Table 2: Evaluation of OmpK35 peptide motifs using IEDB analysis resources

Motif position	Mean BepiPred (0.500)	Mean Chou and Fasman (1.039)	Mean Emini (1.000)	Mean Karplus and Schulz (0.998)	Mean Kolaskar and Tangaonkar (0.996)	Mean Parker (2.320)
1-29	0.315	0.858	0.609	0.949	1.064	0.047
34-83	0.495	1.062	0.988	1.006	0.975	2.490
84-133	0.483	1.009	0.852	0.995	0.997	2.310
138-187	0.538	1.090	1.316	1.009	0.979	2.515
189-196	0.444	1.057	0.543	0.998	0.983	2.995
197-246	0.516	1.071	0.981	1.005	0.992	3.114
248-297	0.561	1.018	1.206	1.002	1.004	2.284
300-349	0.518	1.035	0.993	0.985	1.007	2.104

4.2.3 The OmpK36 motif evaluation

The OmpK36 Motif at positions 282-331 was selected as the potential epitope to be included in the chimeric protein because it passed the preset conditions (Table 3).

Table 3: Evaluation of OmpK36 peptide motifs using IEDB analysis resources

Motif position	Mean BepiPred (0.500)	Mean Chou and Fasman (1.059)	Mean Emini (1.000)	Mean Karplus and Schulz (1.003)	Mean Kolaskar and Tangaonkar (1.012)	Mean Parker (2.126)
1-50	0.404	0.997	0.853	0.979	1.051	1.121
55-104	0.503	1.028	1.121	1.009	0.989	2.459
110-159	0.497	1.105	0.937	1.006	1.009	2.196
160-180	0.569	1.087	0.758	1.008	1.023	1.727
181-230	0.582	1.151	1.178	1.031	0.982	3.077
232-281	0.549	1.045	1.199	1.007	0.998	2.544
282-331	0.500	1.020	1.019	0.984	1.039	1.587
332-360	0.545	1.054	0.773	0.992	1.020	2.045
362-367	0.352	0.834	0.255	0.932	1.161	-2.457

4.2.4 The Pal motif evaluation

The Pal Motif at position 125-174 was selected as the potential epitope to be included in the chimeric protein because it passed the preset conditions (Table 4).

Table 4: Evaluation of Pal peptide motifs using IEDB analysis resources

Motif position	Mean BepiPred (0.500)	Mean Chou and Fasman (1.005)	Mean Emini (1.000)	Mean Karplus and Schulz (0.999)	Mean Kolaskar and Tangaonkar (1.006)	Mean Parker (2.203)
2-12	0.397	0.855	0.402	0.969	1.064	-1.401
15-64	0.597	1.090	1.032	1.019	0.960	3.421
65-70	0.549	0.984	1.108	1.033	1.014	2.269
71-120	0.478	1.007	1.108	0.990	1.012	1.958
125-174	0.501	0.965	1.032	0.993	1.032	2.189

4.3 Chimeric protein design

Immuno-dominant epitopes were combined using both flexible and rigid linkers, with variations in their lengths, resulting in the creation of twenty-one distinct chimera options labeled from >chimera1 to >chimera21. The linkers employed included [(GGGGS)n] where n ranged from 1 to 3, [(G)8], [(G)6], and GSAGSAAGSGEF, which are flexible, along with [A(EAAAK)nA], [(EAAAK)n] where n ranged from 2 to 3, and [(XP)n], where X represents either Ala, Lys, or Gly, with n ranging from 4 to 8.

>chimera1

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGGGGSSQTYNMT
PEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGGGGSEVVAQYQFDFGLRP
SVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGGGGSRRANAVKMYLQGKGVSADQIS
IVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera2

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGGGGSGGGGSSQ
TYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGGGGSGGGGSEVVA
QYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGGGGSGGGGSRRANAV
KMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera3

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGGGGSGGGGSGG
GGSSQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGGGGSGGGG
SGGGGSEVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGGGGSG
GGGSGGGGSRRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera4

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGGGGGGGGSQTY
NMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGGGGGGGGEVVAQYQF
DFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGGGGGGGGRRANAVKMYLQG
KGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera5

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGGGGGGSQTYNM
TPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGGGGGGEVVAQYQFDFGL
RPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGGGGGGRRANAVKMYLQGKGVSAD
QISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera6

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGSAGSAAGSAGE
FSQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGSAGSAAGSAG
EFEVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGSAGSAAGSA
GEFRRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera7

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKAEAAAKASQTYN
MTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRAEAAAKAEVVAQYQFDF
GLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMAEAAAKARRANAVKMYLQGKGV
SADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera8

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKAEAAAKEAAAKA
SQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRAEAAAKEAAAKA
EVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMAEAAAKEAAAKA
RRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera9

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKAEAAAKEAAAKE
AAAKASQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRAEAAAKE
AAAKEAAAKAEVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMAE
AAAKEAAAKEAAAKARRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAV
LVY

>chimera10

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKEAAAKSQTYNMT
PEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSREAAAKEVVAQYQFDFGLRP
SVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMEAAAKRRANAVKMYLQGKGVSADQIS
IVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera11

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKEAAAKEAAAKSQ
TYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSREAAAKEAAAKEVVA
QYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMEAAAKEAAAKRRANAV
KMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera12

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKEAAAKEAAAKEA
AAKSQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSREAAAKEAAA
KEAAAKEVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMEAAAKE
AAAKEAAAKRRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera13

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKAPAPAPAPSQTY
NMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRAPAPAPAPEVVAQYQF
DFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMAPAPAPAPRRANAVKMYLQG
KGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera14

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKAPAPAPAPAPAP
SQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRAPAPAPAPAPAP
EVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMAPAPAPAPAPAP
RRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera15

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKAPAPAPAPAPAP
APAPSQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRAPAPAPAP
APAPAPAPEVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMAPAP
APAPAPAPAPAPRRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera16

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKLPLPLPLPSQTY
NMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRLPLPLPLPEVVAQYQF
DFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMLPLPLPLPRRANAVKMYLQG
KGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera17

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKLPLPLPLPLPLP
SQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRLPLPLPLPLPLP
EVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMLPLPLPLPLPLP
RRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera18

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKLPLPLPLPLPLP
LPLPSQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRLPLPLPLP
LPLPLPLPEVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMLPLP
LPLPLPLPLPLPRRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera19

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGPGPGPGPSQTY
NMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGPGPGPGPEVVAQYQF
DFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGPGPGPGPRRANAVKMYLQG
KGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera20

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGPGPGPGPGPGP
SQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGPGPGPGPGPGP
EVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGPGPGPGPGPGP
RRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

>chimera21

PAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKGPGPGPGPGPGP
GPGPSQTYNMTPEEDNHFAGKTQNFEAVVQYQFDFGLRPSIGYVQTKGKDLQSRGPGPGPGP
GPGPGPGPEVVAQYQFDFGLRPSVAYLQSKGKDLEGYGDQDILKYVDVGATYYFNKNMGPGP
GPGPGPGPGPGPRRANAVKMYLQGKGVSADQISIVSYGKEKPAVLGHDEAAYAKNRRAVLVY

Linkers

Epitopes

4.4 Physicochemical properties, antigenicity and allergenicity assessments

Assessments were conducted on the antigenicity, allergenicity, and physicochemical properties of various designed chimera options. (Table 5).

Table 5: Assessment of physicochemical properties, antigenicity and allergenicity of the twenty-one chimeric proteins

Chimera	Instability index	GRAVY	Estimated half-life	Antigenicity (0.4)	Allergenicity
1	44.50(unstable)	-0.580	>20hours	1.0440	Non allergen
2	47.08(unstable)	-0.573	>20hours	1.3692	Non allergen
3	49.34(unstable)	-0.568	>20hours	1.6543	Non allergen
4	46.81(unstable)	-0.567	>20hours	1.3328	Non allergen
5	44.43(unstable)	-0.572	>20hours	1.1704	Non allergen
6	33.03(stable)	-0.460	>20hours	0.8786	Non allergen
7	37.25(stable)	-0.510	>20hours	0.7765	Non allergen
8	35.52(stable)	-0.503	>20hours	0.8060	Non allergen
9	33.99(stable)	-0.497	>20hours	0.8318	Non allergen
10	38.95(stable)	-0.574	>20hours	0.7709	Non allergen
11	37.07(stable)	-0.563	>20hours	0.8015	Non allergen
12	35.41(stable)	-0.553	>20hours	0.8283	Non allergen
13	56.24(unstable)	-0.514	>20hours	0.7187	Non allergen
14	63.68(unstable)	-0.483	>20hours	0.7252	Non allergen
15	70.41(unstable)	-0.454	>20hours	0.7310	Non allergen
16	47.57(unstable)	-0.407	>20hours	0.7811	Non allergen
17	50.56(unstable)	-0.330	>20hours	0.8745	Non allergen
18	53.26(unstable)	-0.261	>20hours	0.9587	Non allergen
19	36.88(stable)	-0.632	>20hours	0.8420	Non allergen
20	35.51(stable)	-0.650	>20hours	0.9058	Non allergen
21	34.51(stable)	-0.667	>20hours	0.9633	Non allergen

The instability index reveals the degree of instability in a protein, with a value above 40 indicating susceptibility to denaturation or degradation, while a value below 40 suggests stability (Gamage et al., 2019).

GRAVY (Grand Average of Hydropathy) is a measure indicating the overall hydrophobic or hydrophilic nature of a protein or peptide sequence. A positive GRAVY score suggests hydrophobicity, while a negative score indicates hydrophilicity. This score can assess the protein's solubility, with hydrophilic sequences typically more soluble in water and hydrophobic sequences prone to aggregation (Chang & Yang, 2013).

The estimated half-life, measured in hours, predicts the time required for half of the protein to degrade or be cleared from the system. It is crucial for understanding protein stability and the

generation of an immune response, with higher values suggesting prolonged presence and immune stimulation.

Antigenicity, assessed using the VaxiJen server, considers proteins with scores above 0.4 as antigenic, potentially recognized by the immune system. Allergenicity, determined via AlgPred, evaluates the ability of a protein to induce allergic reactions in sensitized individuals, mapping IgE epitopes to predict allergenic potential (Jaan et al., 2022).

>Chimera3 exhibited the highest antigenicity score (1.6543) and solubility, followed by >Chimera2 (1.3692). However, both were highly unstable and susceptible to denaturation. Chimera6 emerged as the preferred chimeric protein due to its stability (33.03), antigenicity (0.8786), solubility (-0.460), extended half-life (>20 hours), and non-allergenic nature. Therefore, >Chimera6 was selected as the optimal chimeric protein

4.5 The 3D structure model of >chimera 6

The 3D structure model of >chimera6 was visualized using PyMOL, with epitopes and linkers colored as shown in Figure 6. Option (a) shows the front view of >chimera6, option (b) the back view, and option (c) an inverted and zoomed-in view.

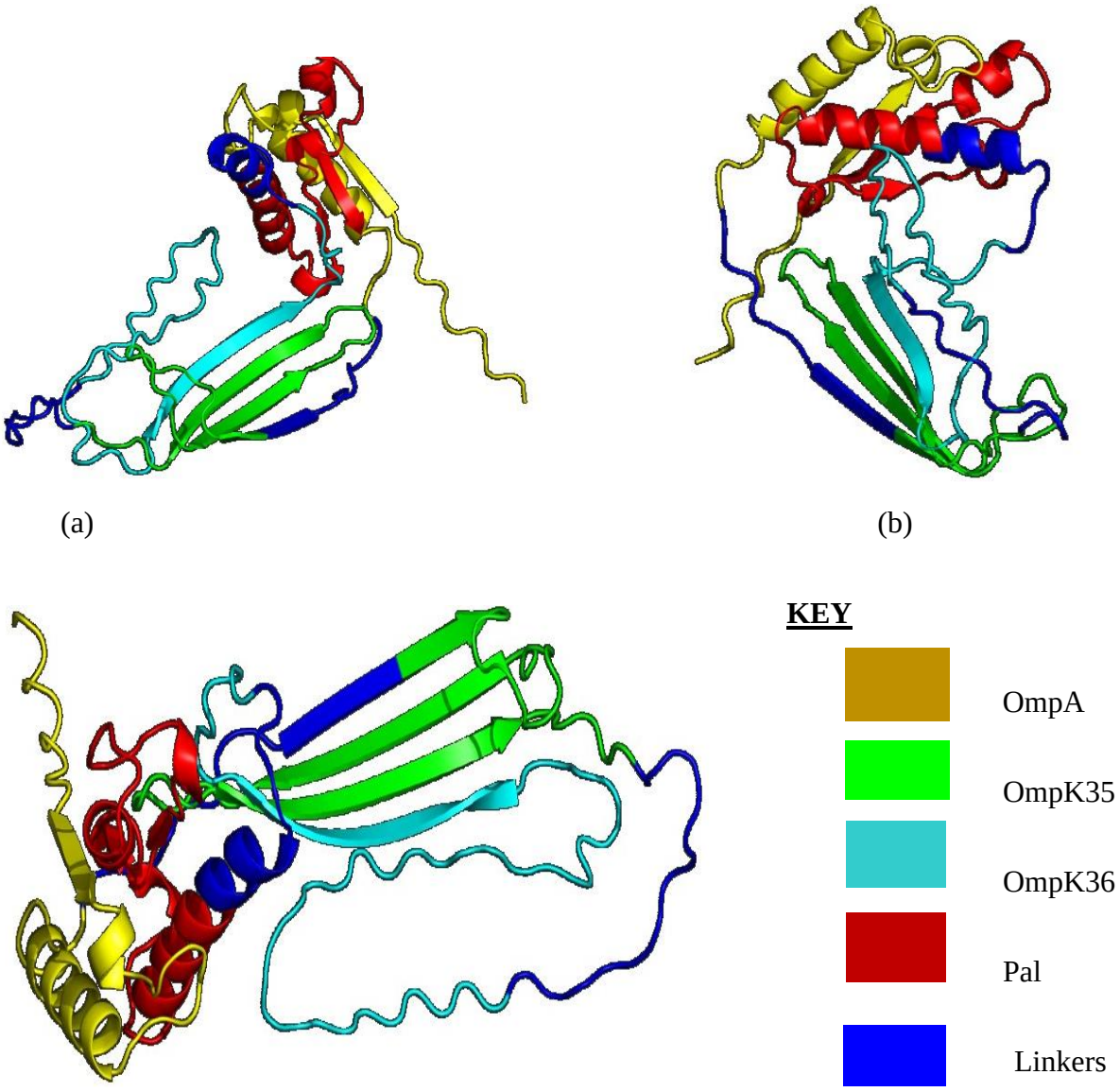

Figure 6: >chimera6 visualized through different orientations using PyMOL

CHAPTER FIVE

5.0 DISCUSSION

While *Klebsiella pneumoniae* has traditionally been associated with infections acquired in hospital settings, particularly affecting immunocompromised individuals as an opportunistic pathogen, there has been a notable increase in the prevalence of highly virulent and multidrug-resistant strains. Consequently, the efficacy of antimicrobial treatments against *Klebsiella pneumoniae* infections has progressively declined (Y. da Silva et al., 2019). The global dissemination of these highly virulent strains, coupled with the emergence of antibiotic-resistant variants, has severely limited treatment options, prompting a renewed emphasis on the development of vaccines (Zargaran et al., 2021). Regrettably, as of now, there is no licensed vaccine available for preventing *Klebsiella pneumoniae* infections (Zhang et al., 2021).

In this research, I employed *in silico* techniques to identify potential B-cell epitopes within OmpA, OmpK35, OmpK36, and Pal proteins that could stimulate an immune response. Subsequently, I designed chimeric proteins incorporating these selected epitopes and evaluated their antigenicity and physicochemical properties.

The initial step in designing chimeric proteins involved leveraging advanced bioinformatics tools for motif discovery and epitope prediction. Specifically, the Multiple Em for Motif Elicitation (MEME Suite) was utilized to identify conserved peptide motifs within the consensus sequences of four *Klebsiella pneumoniae* proteins: OmpA, OmpK35, OmpK36, and Pal. This analysis unveiled multiple motifs for each protein, varying in lengths and positions (Zhao et al., 2015).

Following motif identification, the potential of these motifs to induce an immune response was assessed using IEDB Analysis resources. Selection criteria for successful epitopes encompassed considerations such as antigenicity, surface accessibility, and linearity (El-Manzalawy & Honavar, 2010). Notably, epitopes from OmpA (Position 212-261), OmpK35 (Position 248-297), OmpK36 (Position 284-331), and Pal (Position 125-174) emerged as the most promising candidates for incorporation into chimeric proteins.

To construct chimeric proteins, the selected epitopes were fused using various types of linkers, including both flexible [(GGGGS)n, (G)8, (G)6, and (GSAGSAAGSGEF)] and rigid [A(EAAAK)nA, (EAAAK)n, and (XP)n, where X is either A, L, or G] linkers, with different

lengths. This resulted in the generation of twenty-one distinct chimeras, designated as >chimera1 to >chimera21.

Subsequently, the designed chimeric proteins underwent a comprehensive evaluation of their physicochemical properties, including instability index, GRAVY, estimated half-life, antigenicity, and allergenicity (Kolla et al., 2021). Notably, chimeras 1 to 5 and 13 to 18 exhibited instability indices above 40, indicating susceptibility to denaturation. Conversely, >chimeras 6 to 12 and 19 to 21 displayed stable characteristics, with instability indices ranging from 33.03 to 38.99. >Chimera6 exhibited the most stability, with an instability index of 33.03.

Analysis of GRAVY scores indicated that all chimeric proteins were soluble, with chimera21 displaying particularly high negative GRAVY scores (-0.667), indicating enhanced hydrophilicity and solubility (Rahman et al., 2022).

Further antigenicity analysis using the VaxiJen server identified chimera3 as having the highest antigenic score (1.6543), although its instability nature rendered it unsuitable. Allergenicity assessment via AlgPred indicated that all chimeric proteins were non-allergenic, mitigating concerns regarding potential allergic reactions.

Based on the comprehensive evaluation of physicochemical properties, antigenicity, and allergenicity, >chimera6 emerged as the most promising candidate for *Klebsiella pneumoniae* vaccine development. Its stability, solubility, antigenicity, and non-allergenic nature position it as an ideal immunogen to elicit a robust immune response against *Klebsiella pneumoniae* infections. The successful design and evaluation of >chimera6 signify significant progress in the realm of vaccine development against *Klebsiella pneumoniae* infections. Nonetheless, further research is warranted to optimize its immunogenicity, efficacy, and safety profiles. Future studies should concentrate on *in-vitro* and *in-vivo* validation of >chimera6, encompassing immunogenicity assays, animal model testing, and potential toxicity assessments. Additionally, structural studies and molecular dynamics simulations could offer valuable insights into the stability and conformational dynamics of >chimera6.

CHAPTER SIX

6.0 CONCLUSIONS AND RECOMMENDATION

6.1 Conclusions

In this study, a novel multi-epitope vaccine candidate, containing high-ranked epitopes from all four proteins was constructed and evaluated using immune-informatics approaches. The application of computational tools in vaccine design can significantly enhance the process of vaccine discovery and accomplish this goal in less time and with fewer costs. The designed chimeric protein has suitable physiochemical, structural, and immunological properties that can successfully trigger humoral and cellular immune responses against *Klebsiella pneumoniae*.

6.2 Recommendation

To ascertain the effectiveness and safety of >chimera6, it is imperative to undertake *in vivo* investigations utilizing suitable animal models. Immunization trials in animals will facilitate the evaluation of the immune reaction provoked by the chimeric protein, as well as its ability to safeguard against *Klebsiella pneumoniae* infection. These experiments must adhere strictly to ethical protocols and be conducted under the guidance of proficient researchers (Liu et al., 2017).

Given the primary focus on B-cell epitopes in this study, there is potential for additional research to explore T-cell epitopes within the proteins OmpA, OmpK35, OmpK36, and Pal proteins. Such investigations would contribute to enhancing the immunogenicity of the designed chimera (Wang et al., 2022).

Additional structural investigations employing methodologies like X-ray crystallography or electron microscopy are warranted to unveil the exact three-dimensional configuration of >chimera6. Deciphering its conformation on an atomic scale will facilitate the refinement of its interactions with immune receptors, potentially resulting in enhanced immunogenicity (Altunkaya et al., 2017) .

REFERENCES

Ahmad, T. A., Eweida, A. E., & Sheweita, S. A. (2016). B-cell epitope mapping for the design of vaccines and effective diagnostics. In *Trials in Vaccinology* (Vol. 5, pp. 71–83). Elsevier B.V. https://doi.org/10.1016/j.trivac.2016.04.003

An Overview of Klebsiella pneumoniae A Leading Cause of Hospital Infections. (n.d.).

Ashurst, J. V, & Dawson, A. (n.d.). *Klebsiella Pneumonia Continuing Education Activity*.

Asri, N. A. M., Ahmad, S., Mohamud, R., Hanafi, N. M., Zaidi, N. F. M., Irekeola, A. A., Shueb, R. H., Yee, L. C., Noor, N. M., Mustafa, F. H., Yean, C. Y., & Yusof, N. Y. (2021). Global prevalence of nosocomial multidrug-resistant klebsiella pneumoniae: A systematic review and meta-analysis. In *Antibiotics* (Vol. 10, Issue 12). MDPI. https://doi.org/10.3390/antibiotics10121508

Assoni, L., Girardello, R., Converso, T. R., & Darrieux, M. (2021). Current Stage in the Development of Klebsiella pneumoniae Vaccines. In *Infectious Diseases and Therapy* (Vol. 10, Issue 4, pp. 2157–2175). Adis. https://doi.org/10.1007/s40121-021-00533-4

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Research*, *43*(W1), W39–W49. https://doi.org/10.1093/nar/gkv416

Bain, C. C., & Schridde, A. (2018). Origin, differentiation, and function of intestinal macrophages. *Frontiers in Immunology*, *9*(NOV). https://doi.org/10.3389/fimmu.2018.02733

Bengoechea, J. A., & Sa Pessoa, J. (2019). Klebsiella pneumoniae infection biology: Living to counteract host defences. In *FEMS Microbiology Reviews* (Vol. 43, Issue 2, pp. 123–144). Oxford University Press. https://doi.org/10.1093/femsre/fuy043

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., & Bairoch, A. (2007). UniProtKB/Swiss-Prot. In *Plant Bioinformatics* (pp. 89–112). Humana Press. https://doi.org/10.1007/978-1-59745-535-0_4

Chang, K. Y., & Yang, J. R. (2013). Analysis and Prediction of Highly Effective Antiviral Peptides Based on Random Forests. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0070166

Chauhan, V., Rungta, T., Goyal, K., & Singh, M. P. (2019). Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39299-8

Chen, X., Zaro, J. L., & Shen, W. C. (2013). Fusion protein linkers: Property, design and functionality. In *Advanced Drug Delivery Reviews* (Vol. 65, Issue 10, pp. 1357–1369). https://doi.org/10.1016/j.addr.2012.09.039

Choi, U., & Lee, C. R. (2019). Distinct Roles of Outer Membrane Porins in Antibiotic Resistance and Membrane Integrity in Escherichia coli. *Frontiers in Microbiology*, *10*(APR). https://doi.org/10.3389/fmicb.2019.00953

Consensus sequence. (n.d.). http://db.systemsbiolo

Dana, H., Chalbatani, G. M., Gharagouzloo, E., Miri, S. R., Memari, F., Rasoolzadeh, R., Zinatizadeh, M. R., Zarandi, P. K., & Marmari, V. (2020). In silico analysis, molecular docking, molecular dynamic, cloning, expression and purification of chimeric protein in colorectal cancer

treatment. *Drug Design, Development and Therapy*, *14*, 309–329. https://doi.org/10.2147/DDDT.S231958

Dar, H. A., Zaheer, T., Shehroz, M., Ullah, N., Naz, K., Muhammad, S. A., Zhang, T., & Ali, A. (2019). Immunoinformatics-aided design and evaluation of a potential multi-epitope vaccine against klebsiella pneumoniae. *Vaccines*, *7*(3). https://doi.org/10.3390/vaccines7030088

da Silva, Y., Ferrari, R., Marin, V. A., & Junior, C. A. C. (2019). A Global Overview of β-lactam Resistance Genes in Klebsiella pneumoniae. *The Open Infectious Diseases Journal*, *11*(1), 22–34. https://doi.org/10.2174/1874279301911010022

Davidson, E., & Doranz, B. J. (2014). A high-throughput shotgun mutagenesis approach to mapping B-cell antibody epitopes. *Immunology*, *143*(1), 13–20. https://doi.org/10.1111/imm.12323

Dhanda, S. K., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M. C., Jurtz, V., Andreatta, M., Greenbaum, J. A., Marcatili, P., Sette, A., Nielsen, M., & Peters, B. (2019). IEDB-AR: immune epitope database - analysis resource in 2019. *Nucleic Acids Research*, *47*(W1), W502–W506. https://doi.org/10.1093/nar/gkz452

Doss, W. (n.d.). *Study Explores How Pneumonia Strains Compete*. https://news.feinberg.northwestern.edu/2019/04/23/study-explores-how-pneumonia-strains-compete/#:~:text=There

El-Manzalawy, Y., & Honavar, V. (2010). Recent advances in B-cell epitope prediction methods. In *Immunome Research* (Vol. 6, Issue SUPPL. 2). https://doi.org/10.1186/1745-7580-6-S2-S2

Esposito, E. P., Cervoni, M., Bernardo, M., Crivaro, V., Cuccurullo, S., Imperi, F., & Zarrilli, R. (2018). Molecular epidemiology and virulence profiles of colistin-resistant Klebsiella pneumoniae blood isolates from the hospital agency "Ospedale dei Colli," Naples, Italy. *Frontiers in Microbiology*, *9*(JUL). https://doi.org/10.3389/fmicb.2018.01463

Feldman, M. F., Mayer Bridwell, A. E., Scott, N. E., Vinogradov, E., McKee, S. R., Chavez, S. M., Twentyman, J., Stallings, C. L., Rosen, D. A., & Harding, C. M. (2019). A promising bioconjugate vaccine against hypervirulent Klebsiella pneumoniae. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(37), 18655–18663. https://doi.org/10.1073/pnas.1907833116

Gamage, D. G., Gunaratne, A., Periyannan, G. R., & Russell, T. G. (2019). Applicability of Instability Index for In vitro Protein Stability Prediction. *Protein & Peptide Letters*, *26*(5), 339–347. https://doi.org/10.2174/0929866526666190228144219

Geetha, P. V., Aishwarya, K. V. L., Mariappan, S., & Sekar, U. (2020). Fluoroquinolone Resistance in Clinical Isolates of Klebsiella Pneumoniae. *Journal of Laboratory Physicians*, *12*(02), 121–125. https://doi.org/10.1055/s-0040-1716478

Ghashghaee, A., Behzadifar, M., Azari, S., Farhadi, Z., Bragazzi, N. L., Behzadifar, M., Shahri, S. S. S., Ghaemmohamadi, M. S., Ebadi, F., Mohammadibakhsh, R., Seyedin, H., & Moghadam, M. R. (2018). Prevalence of nosocomial infections in Iran: A systematic review and meta-analysis. *Medical Journal of the Islamic Republic of Iran*, *32*(1). https://doi.org/10.14196/mjiri.32.48

Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*, *27*(1), 14–25. https://doi.org/10.1002/pro.3235

Gonzalez-Ferrer, S., Peñaloza, H. F., Budnick, J. A., Bain, W. G., Nordstrom, H. R., Lee, J. S., & Van Tyne, D. (2021). Finding order in the chaos: Outstanding questions in Klebsiella pneumoniae Pathogenesis. *Infection and Immunity*, *89*(4). https://doi.org/10.1128/IAI.00693-20

Gräwe, A., Ranglack, J., Weyrich, A., & Stein, V. (2020). IFLinkC: An iterative functional linker cloning strategy for the combinatorial assembly and recombination of linker peptides with functional domains. *Nucleic Acids Research*, *48*(4), E24. https://doi.org/10.1093/nar/gkz1210

Hafiz, T. A., Alanazi, S., Alghamdi, S. S., Mubaraki, M. A., Aljabr, W., Madkhali, N., Alharbi, S. R., Binkhamis, K., & Alotaibi, F. (2023). Klebsiella pneumoniae bacteraemia epidemiology: resistance profiles and clinical outcome of King Fahad Medical City isolates, Riyadh, Saudi Arabia. *BMC Infectious Diseases*, *23*(1). https://doi.org/10.1186/s12879-023-08563-8

How redundant are sequences in UniProtKB? (2002). https://www.uniprot.org/help/uniprotkb

Hsieh, P. F., Liu, J. Y., Pan, Y. J., Wu, M. C., Lin, T. L., Huang, Y. Te, & Wang, J. T. (2013). Klebsiella pneumoniae peptidoglycan-associated lipoprotein and murein lipoprotein contribute to serum resistance, antiphagocytosis, and proinflammatory cytokine stimulation. *Journal of Infectious Diseases*, *208*(10), 1580–1589. https://doi.org/10.1093/infdis/jit384

Huang, Y., Rana, A. P., Wenzler, E., Ozer, E. A., Krapp, F., Bulitta, J. B., Hauser, A. R., & Bulman, Z. P. (2022). Aminoglycoside-resistance gene signatures are predictive of aminoglycoside MICs for carbapenem-resistant Klebsiella pneumoniae. *Journal of Antimicrobial Chemotherapy*, *77*(2), 356–363. https://doi.org/10.1093/jac/dkab381

Hussain, H. I., Aqib, A. I., Seleem, M. N., Shabbir, M. A., Hao, H., Iqbal, Z., Kulyar, M. F. e. A., Zaheer, T., & Li, K. (2021). Genetic basis of molecular mechanisms in β-lactam resistant gram-negative bacteria. *Microbial Pathogenesis*, *158*. https://doi.org/10.1016/j.micpath.2021.105040

Hussein, M., Jasim, R., Gocol, H., Baker, M., Thombare, V. J., Ziogas, J., Purohit, A., Rao, G. G., Li, J., & Velkov, T. (2023). Comparative Proteomics of Outer Membrane Vesicles from Polymyxin-Susceptible and Extremely Drug-Resistant Klebsiella pneumoniae. *MSphere*, *8*(1). https://doi.org/10.1128/msphere.00537-22

Jaan, S., Shah, M., Ullah, N., Amjad, A., Javed, M. S., Nishan, U., Mustafa, G., Nawaz, H., Ahmed, S., & Ojha, S. C. (2022). Multi-epitope chimeric vaccine designing and novel drug targets prioritization against multi-drug resistant Staphylococcus pseudintermedius. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.971263

Jomehzadeh, N., Ahmadi, K., & Bahmanshiri, M. A. (2022). Investigation of plasmid-mediated quinolone resistance genes among clinical isolates of Klebsiella pneumoniae in southwest Iran. *Journal of Clinical Laboratory Analysis*, *36*(7). https://doi.org/10.1002/jcla.24342

Kolla, H. B., Tirumalasetty, C., Sreerama, K., & Ayyagari, V. S. (2021). An immunoinformatics approach for the design of a multi-epitope vaccine targeting super antigen TSST-1 of

Staphylococcus aureus. *Journal of Genetic Engineering and Biotechnology*, *19*(1). https://doi.org/10.1186/s43141-021-00160-z

Kot, B., Piechota, M., Szweda, P., Mitrus, J., Wicha, J., Grużewska, A., & Witeska, M. (2023). Virulence analysis and antibiotic resistance of Klebsiella pneumoniae isolates from hospitalised patients in Poland. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-31086-w

Liang, C., Xing, B., Yang, X., Fu, Y., Feng, Y., & Zhang, Y. (2015). Molecular epidemiology of aminoglycosides resistance on Klebsiella pneumonia in a hospital in China. In *Int J Clin Exp Med* (Vol. 8, Issue 1). www.ijcem.com/

Lin, Z. W., Zheng, J. X., Bai, B., Xu, G. J., Lin, F. J., Chen, Z., Sun, X., Qu, D., Yu, Z. J., & Deng, Q. W. (2020). Characteristics of Hypervirulent Klebsiella pneumoniae: Does Low Expression of rmpA Contribute to the Absence of Hypervirulence? *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.00436

Liu, B., Park, S., Thompson, C. D., Li, X., & Lee, J. C. (2017). Antibodies to Staphylococcus aureus capsular polysaccharides 5 and 8 perform similarly in vitro but are functionally distinct in vivo. *Virulence*, *8*(6), 859–874. https://doi.org/10.1080/21505594.2016.1270494

Liu, C., & Guo, J. (2018). Characteristics of ventilator-associated pneumonia due to hypervirulent Klebsiella pneumoniae genotype in genetic background for the elderly in two tertiary hospitals in China. *Antimicrobial Resistance and Infection Control*, *7*(1). https://doi.org/10.1186/s13756-018-0371-8

Li, Y., Kumar, S., Zhang, L., Wu, H., & Wu, H. (2023). Characteristics of antibiotic resistance mechanisms and genes of Klebsiella pneumoniae. In *Open Medicine (Poland)* (Vol. 18, Issue 1). De Gruyter Open Ltd. https://doi.org/10.1515/med-2023-0707

March, C., Moranta, D., Regueiro, V., Llobet, E., Tomás, A., Garmendia, J., & Bengoechea, J. (2011). Klebsiella pneumoniae outer membrane protein A is required to prevent the activation of airway epithelial cells. *Journal of Biological Chemistry*, *286*(12), 9956–9967. https://doi.org/10.1074/jbc.M110.181008

Martini, S., Nielsen, M., Peters, B., & Sette, A. (2020). The Immune Epitope Database and Analysis Resource Program 2003–2018: reflections and outlook. In *Immunogenetics* (Vol. 72, Issues 1–2, pp. 57–76). Springer. https://doi.org/10.1007/s00251-019-01137-6

Martin, R. M., & Bachman, M. A. (2018). Colonization, infection, and the accessory genome of Klebsiella pneumoniae. In *Frontiers in Cellular and Infection Microbiology* (Vol. 8, Issue JAN). Frontiers Media S.A. https://doi.org/10.3389/fcimb.2018.00004

Mathew, S., Fakhroo, A. D., Smatti, M., Al Thani, A. A., & Yassine, H. M. (2022). Immunoinformatics prediction of potential immunodominant epitopes from human coronaviruses and association with autoimmunity. *Immunogenetics*, *74*(2), 213–229. https://doi.org/10.1007/s00251-021-01250-5

Mihara, M., Itoh, T., & Izawa, T. (2009). SALAD database: A motif-based database of protein annotations for plant comparative genomics. *Nucleic Acids Research*, *38*(SUPPL.1). https://doi.org/10.1093/nar/gkp831

Montero, D. A., Garcia-Betancourt, R., Vidal, R. M., Velasco, J., Palacios, P. A., Schneider, D., Vega, C., Gómez, L., Montecinos, H., Soto-Shara, R., Oñate, Á., & Carreño, L. J. (2023). A chimeric protein-based vaccine elicits a strong IgG antibody response and confers partial protection against Shiga toxin-producing Escherichia coli in mice. *Frontiers in Immunology*, *14*. https://doi.org/10.3389/fimmu.2023.1186368

Moscoso, M., Vallejo, J. A., Cabral, M. P., García, P., Fuentes-Valverde, V., Gato, E., Arca-Suárez, J., Aja-Macaya, P., & Bou, G. (2022). A New Live Auxotrophic Vaccine Induces Cross-Protection against Klebsiella pneumoniae Infections in Mice. *Vaccines*, *10*(6). https://doi.org/10.3390/vaccines10060953

NCBI. (n.d.).

OUP accepted manuscript. (2016). *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw1000

Parrott, A. M., Shi, J., Aaron, J., Green, D. A., Whittier, S., & Wu, F. (2021). Detection of multiple hypervirulent Klebsiella pneumoniae strains in a New York City hospital through screening of virulence genes. *Clinical Microbiology and Infection*, *27*(4), 583–589. https://doi.org/10.1016/j.cmi.2020.05.012

Patel, D., Menon, D., & Patel, D. (2020). *Linkers: a synergistic way for chimeric proteins*. https://doi.org/10.22541/au.160199719.94542802/v1

Pearson, W. R. (2013). An introduction to sequence similarity ("homology") searching. *Current Protocols in Bioinformatics*, *SUPPL.42*. https://doi.org/10.1002/0471250953.bi0301s42

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Potocnakova, L., Bhide, M., & Pulzova, L. B. (2016). An Introduction to B-Cell Epitope Mapping and in Silico Epitope Prediction. In *Journal of Immunology Research* (Vol. 2016). Hindawi Limited. https://doi.org/10.1155/2016/6760830

Rahman, M. M., Puspo, J. A., Adib, A. A., Hossain, M. E., Alam, M. M., Sultana, S., Islam, A., Klena, J. D., Montgomery, J. M., Satter, S. M., Shirin, T., & Rahman, M. Z. (2022). An Immunoinformatics Prediction of Novel Multi-Epitope Vaccines Candidate Against Surface Antigens of Nipah Virus. *International Journal of Peptide Research and Therapeutics*, *28*(4). https://doi.org/10.1007/s10989-022-10431-z

Rahmat Ullah, S., Majid, M., Rashid, M. I., Mehmood, K., & Andleeb, S. (2021). Immunoinformatics Driven Prediction of Multiepitopic Vaccine Against Klebsiella pneumoniae and Mycobacterium tuberculosis Coinfection and Its Validation via In Silico Expression. *International Journal of Peptide Research and Therapeutics*, *27*(2), 987–999. https://doi.org/10.1007/s10989-020-10144-1

Ranjbarian, P., Goudarzi, F., Akya, A., Heidarinia, H., Farasat, A., & Rostamian, M. (2023). Finding epitopes of Klebsiella pneumoniae outer membrane protein-K17 (OMPK17) and introducing a 25-mer peptide of it as a vaccine candidate. *Biologia*. https://doi.org/10.1007/s11756-023-01371-0

Right@, C., Paula, A., Arêas, M., & Brigagão, L. A. (n.d.). Ana Paula Mattos Arêas. Kiayia Propst. Klebsiella pneumoniae Infections: A Future Threat to Global Health? *Am J Biomed Sci & Res*, *5*, 2022–2038. https://doi.org/10.34297/AJBSR.2022.16.002273

Riwu, K. H. P., Effendi, M. H., Rantam, F. A., Khairullah, A. R., & Widodo, A. (2022a). A review: Virulence factors of Klebsiella pneumonia as emerging infection on the food chain. In *Veterinary World* (Vol. 15, Issue 9, pp. 2172–2179). Veterinary World. https://doi.org/10.14202/vetworld.2022.2172-2179

Salod, Z., & Mahomed, O. (2022). Mapping Potential Vaccine Candidates Predicted by VaxiJen for Different Viral Pathogens between 2017–2021—A Scoping Review. In *Vaccines* (Vol. 10, Issue 11). MDPI. https://doi.org/10.3390/vaccines10111785

Seladi-Schulman, J., & Caporuscio, J. (2020). *What to know about Klebsiella pneumoniae Medically reviewed*. https://www.medicalnewstoday.com/articles/klebsiella-pneumoniae#prognosis

Sever, W., Garg, V. K., Avashthi, H., Tiwari, A., Jain, A., Ramkete, P. W., Kayastha, A. M., & Singh, K. (2016). Open access Volume 12(2) MFPPI-Multi FASTA ProtParam Interface. *Print) BIOINFORMATION*, *12*(2), 74. http://insilicogenomics.in/mfpcalc/mfppi.html.

Sharma, A., Thakur, A., Thakur, N., Kumar, V., Chauhan, A., & Bhardwaj, N. (2023). Changing Trend in the Antibiotic Resistance Pattern of Klebsiella Pneumonia Isolated From Endotracheal Aspirate Samples of ICU Patients of a Tertiary Care Hospital in North India. *Cureus*. https://doi.org/10.7759/cureus.36317

Sheng, Z. K., Hu, F., Wang, W., Guo, Q., Chen, Z., Xu, X., Zhu, D., & Wang, M. (2014). Mechanisms of tigecycline resistance among Klebsiella pneumoniae clinical isolates. *Antimicrobial Agents and Chemotherapy*, *58*(11), 6982–6985. https://doi.org/10.1128/AAC.03808-14

Silva, M. T. de O., de Pinho, R. B., Bezerra, F. S. B., Scholl, N. R., Moron, L. D., Alves, M. S. D., Woloski, R. dos S., Kremer, F. S., & Borsuk, S. (2021). In silico analyses and design of a chimeric protein containing epitopes of SpaC, PknG, NanH, and SodC proteins for the control of caseous lymphadenitis. *Applied Microbiology and Biotechnology*, *105*(21–22), 8277–8286. https://doi.org/10.1007/s00253-021-11619-x

Soria-Guerra, R. E., Nieto-Gomez, R., Govea-Alonso, D. O., & Rosales-Mendoza, S. (2015). An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. In *Journal of Biomedical Informatics* (Vol. 53, pp. 405–414). Academic Press Inc. https://doi.org/10.1016/j.jbi.2014.11.003

Ssekatawa, K., Byarugaba, D. K., Nakavuma, J. L., Kato, C. D., Ejobi, F., Tweyongyere, R., & Eddie, W. M. (2021). Prevalence of pathogenic Klebsiella pneumoniae based on PCR capsular typing harbouring carbapenemases encoding genes in Uganda tertiary hospitals. *Antimicrobial Resistance and Infection Control*, *10*(1). https://doi.org/10.1186/s13756-021-00923-w

Sternke, M., Tripp, K. W., & Barrick, D. (2019). Consensus sequence design as a general strategy to create hyperstable, biologically active proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *166*(23), 11275–11284. https://doi.org/10.1073/pnas.1816707116

Sugawara, E., Kojima, S., & Nikaido, H. (2016). Klebsiella pneumoniae major porins OmpK35 and OmpK36 Allow more efficient diffusion of β-lactams than their escherichia coli homologs OmpF and OmpC. *Journal of Bacteriology*, *198*(23), 3200–3208. https://doi.org/10.1128/JB.00590-16

Tamura, K., Stecher, G., & Kumar, S. (2021a). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Tsai, Y. K., Fung, C. P., Lin, J. C., Chen, J. H., Chang, F. Y., Chen, T. L., & Siu, L. K. (2011). Klebsiella pneumoniae Outer membrane porins OmpK35 and OmpK36 play roles in both antimicrobial resistance and virulence. *Antimicrobial Agents and Chemotherapy*, *55*(4), 1485–1493. https://doi.org/10.1128/AAC.01275-10

Wang, G., Zhao, G., Chao, X., Xie, L., & Wang, H. (2020). The characteristic of virulence, biofilm and antibiotic resistance of klebsiella pneumoniae. In *International Journal of Environmental Research and Public Health* (Vol. 17, Issue 17, pp. 1–17). MDPI AG. https://doi.org/10.3390/ijerph17176278

Wang, Z., Guo, G., Li, Q., Li, P., Li, M., Zhou, L., Tan, Z., & Zhang, W. (2022). Combing Immunoinformatics with Pangenome Analysis To Design a Multiepitope Subunit Vaccine against Klebsiella pneumoniae K1, K2, K47, and K64. *Microbiology Spectrum*, *10*(4). https://doi.org/10.1128/spectrum.01148-22

Yang, J., Zhang, K., Ding, C., Wang, S., Wu, W., & Liu, X. (2023). Exploring multidrug-resistant Klebsiella pneumoniae antimicrobial resistance mechanisms through whole genome sequencing analysis. *BMC Microbiology*, *23*(1). https://doi.org/10.1186/s12866-023-02974-y

Yuan, S., Chan, H. C. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. In *Wiley Interdisciplinary Reviews: Computational Molecular Science* (Vol. 7, Issue 2). Blackwell Publishing Inc. https://doi.org/10.1002/wcms.1298

Yurina, V., & Adianingsih, O. R. (2022). Predicting epitopes for vaccine development using bioinformatics tools. In *Therapeutic Advances in Vaccines and Immunotherapy* (Vol. 10). SAGE Publications Ltd. https://doi.org/10.1177/25151355221100218

Zambelli, F., Pesole, G., & Pavesi, G. (2013). Motif discovery and transcription factor binding sites before and after the next-generation sequencing era. *Briefings in Bioinformatics*, *14*(2), 225–237. https://doi.org/10.1093/bib/bbs016

Zargaran, F. N., Akya, A., Rezaeian, S., Ghadiri, K., Lorestani, R. C., Madanchi, H., Safaei, S., & Rostamian, M. (2021). B Cell Epitopes of Four Fimbriae Antigens of Klebsiella pneumoniae: A Comprehensive In Silico Study for Vaccine Development. *International Journal of Peptide Research and Therapeutics*, *27*(2), 875–886. https://doi.org/10.1007/s10989-020-10134-3

Zhai, W., Tian, Y., Lu, M., Zhang, M., Song, H., Fu, Y., Ma, T., Sun, C., Bai, L., Wang, Y., Liu, D., & Zhang, Y. (2022). *Presence of Mobile Tigecycline Resistance Gene tet(X4) in Clinical Klebsiella pneumoniae*. https://rast.nmpdr.org/

Zhang, B. Z., Hu, D., Dou, Y., Xiong, L., Wang, X., Hu, J., Xing, S. Z., Li, W., Cai, J. P., Jin, M., Zhang, M., Lin, Q., Li, M., Yuen, K. Y., & Huang, J. D. (2021). Identification and Evaluation of Recombinant

Outer Membrane Proteins as Vaccine Candidates Against Klebsiella pneumoniae. *Frontiers in Immunology*, *12*. https://doi.org/10.3389/fimmu.2021.730116

Zhao, Z., Sun, H. Q., Wei, S. S., Li, B., Feng, Q., Zhu, J., Zeng, H., Zou, Q. M., & Wu, C. (2015). Multiple B-cell epitope vaccine induces a Staphylococcus enterotoxin B-specific IgG1 protective response against MRSA infection. *Scientific Reports*, *5*. https://doi.org/10.1038/srep12371

Zhu, J., Wang, T., Chen, L., & Du, H. (2021a). Virulence Factors in Hypervirulent Klebsiella pneumoniae. In *Frontiers in Microbiology* (Vol. 12). Frontiers Media S.A. https://doi.org/10.3389/fmicb.2021.642484

APPENDIX

Amino acids and their Abbreviations

Amino Acid	3 letter code	1 letter code	Amino Acid	3 letter code	1 letter code
Glycine	Gly	G	Threonine	Thr	T
Alanine	Ala	A	Cysteine	Cys	C
Valine	Val	V	Tyrosine	Tyr	Y
Leucine	Leu	L	Asparagine	Asn	N
Isoleucine	Ile	I	Glutamine	Gln	Q
Methionine	Met	M	Aspartic Acid	Asp	D
Proline	Pro	P	Glutamic Acid	Glu	E
Phenyl alanine	Phe	F	Lysine	Lys	K
Tryptophan	Trp	W	Arginine	Arg	R
Serine	Ser	S	Histidine	His	H