CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Rational Design of Multiepitope based Subunit Vaccine against *Plesiomonas shigelloides*: Subtractive Genomics and Immunoinformatics Based Approach

by

Danish Rasool

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Health and Life Sciences
Department of Bioinformatics and Biosciences

2022

Copyright © 2022 by Danish Rasool

All rights reserved. No part of this thesis may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, by any information storage and retrieval system without the prior written permission of the author.

This study is whole heartedly dedicated to my beloved parents, who have been my source of inspiration and gave me strength when I thought of giving up, which continually provide their moral, spiritual, emotional, and financial support. I dedicated this research to my brother, sisters, friends, and classmates who shared their words of advice and encouragement to finish this study.

CERTIFICATE OF APPROVAL

Rational Design of Multiepitope based Subunit Vaccine against *Plesiomonas shigelloides*: Subtractive Genomics and Immunoinformatics Based Approach

by

Danish Rasool

(MBS203032)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Raees Khan	NUMS Rawalpindi
(b)	Internal Examiner	Dr. Syeda Marriam Bakhtiar	CUST Islamabad
(c)	Supervisor	Dr. Sohail Ahmed Jan	CUST Islamabad

Dr. Sohail Ahmad Jan
Thesis Supervisor
November, 2022

Dr. Syeda Marriam Bakhtiar
Head
Dept. of Bioinfo. and Biosciences
November, 2022

Prof. Dr. Sahar Fazal
Dean
Faculty of Health and Life Sciences
November, 2022

Author's Declaration

I, **Mr. Danish Rasool** hereby state that my MS thesis titled "**"Rational Design of multi-epitope based subunit vaccine against *Plesiomonas shigelloides*: Subtractive genomics and immuno-informatics based approach"**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Danish Rasool)

Registration No:MBS203032

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Rational Design of Multiepitope based Subunit Vaccine against *Plesiomonas shigelloides*: Subtractive Genomics and Immunoinformatics Based Approach**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Danish Rasool)

Registration No:MBS203032

Acknowledgement

All praises are due to Almighty ALLAH, the most merciful and compassionate and His Holy Prophet Muhammad (Peace be upon him), the most perfect and exalted person, who is forever a torch of guidance and knowledge for humanity as a whole. Special appreciation goes to my supervisor **Dr. Sohail Ahmad Jan**, Assistant Professor, Department of Bioinformatics & Biosciences for her inspiring supervision, keen interest, sincere attitude, kind behavior and invaluable assistance. He has made a great contribution to which these few lines can hardly do the justice. His kind remembrance will always remain in my heart. I am grateful to most supportive and kind man to me **Dr. Usman Ali Ashfaq**, Professor, Department of Bioinformatics & Biotechnology from Government College University of Faisalabad for his guidance. I also offer my sincere appreciation for the learning opportunities provided by Head of Department **Dr. Syeda Marriam Bakhtiar**.
I extend heartiest thanks to my friends **Ahtisham Umer, Kiran Shahzadi, Abdur Rahman and Ali Haider** for their sagacious suggestions, guidance and for giving me higher ideas during my entire research project. Special thanks to all other well-wishers who never failed in helping me whenever their assistance was needed.
I offer my sincerest thanks to my all family members, First of all my father, **Mr. Ali Akbar** who inspired me to achieve higher goals in life and my loving, devoted and spiritual mentor my mother **Manzora Bibi** for her sublime and unflagging love for me her hands always raised in prayers for me and **Sana Afzal** gave unconditional support throughout my life and studies. I would like to thank my brother **Muhammad Sharif, Nasir Iqbal, Nabeel Abbas and Muhammad Hassan Mujtaba** especially to my sisters for their endless support, love, prayers and encouragement.

(Danish Rasool)

Abstract

Plesiomonas shigelloides is an anaerobic, facultative, gram-negative rod shaped bacteria that was recently recognized as the only oxidase-positive member of the Enterobacteriaceae family. *P. shigelloides*, like many other Enterobacteriaceae species, can be found in a variety of hosts, including dogs, pigs, cats, cows, and primates. Although *P. shigelloides* is normally considered as freshwater bacteria, it has been correlated to diseases. *P. shigelloides* is spread mostly by environmental media, particularly freshwaters. Ingestion of contaminated water, infected aquatic foods, and contaminated vegetables has been linked to faecal-oral transmission. Furthermore, *P. shigelloides* disease(diarrhea) has been linked to interaction with reptiles, amphibians, and recreational or occupational activities that enhance contact with freshwaters. In the framework of this study, the entire proteome of *P. shigelloides* strains was extracted, revealing a total of 3354 proteins. Six target proteins were isolated from 3354 proteins using a subtractive proteomics approach. A reversed vaccinology-based immunoinformatics technique was utilized to assess the toxicity, antigenicity, and allergenicity of four target proteins used to produce the MEV (multi epitope vaccination) construct. Several techniques employed molecular docking analysis to better understand vaccine binding energy, binding affinities, and interactions. The fact that MEV forms binding contacts with TLR4, MHCI, and MHCII demonstrated that all of these candidate vaccines fit nicely into the MEV's associated protein. Furthermore, MEV construct had higher binding energies of $-51.5 + / - 10.3, -60.5 + / - 9.2, -70.4 + / - 23.7$, and $-79.7 + / - 11.2$ kcal/mol against TLR4, TLR2, MHCII, and MHCI. Simulations were done on docked complexes, validating previous findings. The respective codon of the vaccine design was optimized and then in silico cloned into an *E. coli* expression host to achieve maximum vaccination protein expression. Despite the fact that the in silico analysis utilized in this study produced reliable results, more research is required to confirm the efficacy and performance of the proposed vaccine candidate.

Contents

Author's Declaration iv

Plagiarism Undertaking v

Acknowledgement vi

Abstract vii

List of Figures x

List of Tables xi

Abbreviations xii

1 Introduction 1
- 1.1 Background . . . 1
- 1.2 Problem Statement . . . 9
  - 1.2.1 Aim . . . 9
  - 1.2.2 Objectives . . . 9

2 Literature Review 11
- 2.1 General Overview . . . 11
- 2.2 Cellular structure . . . 13
- 2.3 Ecology . . . 15
- 2.4 Pathogenesis . . . 17
- 2.5 Genome Organization . . . 20
- 2.6 Metabolism . . . 22
- 2.7 Clinical Manifestation . . . 23
- 2.8 Antibiotic resistance . . . 24
- 2.9 Treatment and preventions . . . 26
- 2.10 Subtractive genomics . . . 28
- 2.11 Immuno-informatics . . . 29
- 2.12 Structural analyses . . . 32
- 2.13 Molecular Docking and Simulation . . . 33
- 2.14 Types of Molecular Docking . . . 34

3 Material and Methods **35**
3.1 Retrieval of Proteome Analysis . . . 36
3.2 CTL Epitope Selection and Evaluation . . . 37
3.3 Analysis and Selection of HTL Epitopes . . . 37
3.4 LBL Epitopes Identification and Evaluation . . . 38
3.5 Population Coverage Analysis . . . 38
3.6 Vaccine's Mapping . . . 39
3.7 Structural Analysis . . . 39
3.8 Refinement, Confirmation and Prediction of Tertiary Structure . . . 40
3.9 B-cell Epitopes Screening . . . 41
3.10 MEBSV Stability through Disulfide Engineering . . . 41
3.11 Docking of MHC Molecules and TLR4 Receptor with Constructed Vaccine Disulphide . . . 42
3.12 Molecular Dynamic Simulation . . . 43
3.13 Immune Simulation . . . 43
3.14 In-silico cloning and Optimization of Codons . . . 44

4 Results and Discussion **45**
4.1 Protein's Selection . . . 45
4.2 Epitope CTL Evaluation and Selection . . . 46
4.3 Population Coverage Analysis . . . 48
4.4 Construction of Vaccine . . . 49
4.5 Physiochemical and Immunogenic Profiling . . . 51
4.6 Structural Evaluation . . . 51
4.7 Prediction of tertiary Structure, Validation, and Refinement . . . 51
4.8 Selection of B-cell Epitopes . . . 52
4.9 Disulfide Engineering . . . 53
4.10 Molecular Docking . . . 54
4.11 Molecular Dynamic Simulation . . . 57
4.12 Immune-Simulation . . . 59
4.13 *In-Silico* Cloning . . . 60

5 Conclusion and Future Aspects **65**

Bibliography **67**

List of Figures

2.1 Microscopic representation of *P. Shigelloides*. 14
2.2 The steps followed during subtractive genomics in designing drug against pathogen. 29
2.3 The workflow representation of the Immunoinformatic study for multiepitope vaccine development. 31

3.1 Complete flow chart of whole procedure of research project 35
3.2 Complete graphical abstract 44

4.1 Population coverage 49
4.2 Primary sequence of the MEBSV 50
4.3 Schematic representation of the final MEBSV construct vaccine. . . 50
4.4 Schematic representation of the final MEBSV construct vaccine. . . 52
4.5 Disulphide engineering of MEBSV structure to improve stability. . . 53
4.6 Docked complex of TLR4 receptor with the vaccine receptor indicating the interacting part along with the interacting residue. . . . 54
4.7 Intermolecular binding mode and residue level chemical interactions of MEBSV-MHCI complex. The MEBSV is shown green color while MHC I receptor is represented in cyan color. 55
4.8 Intermolecular binding mode and residue level chemical interactions of MEBSV-MHCI complex. The MEBSV is shown in green color while the MHC II receptor is represented in grey color. 56
4.9 (a, b, c) Simulation of the vaccine-TLR4 complex, showing (a) eigenvalue, (b) deformability, (c) B-factor 57
4.10 (d, e) Simulation of the vaccine-TLR4 complex, showing (d) covariance matrix, and (e) elastic network analysis. 58
4.11 (a & b) C Imm Sim presentation of immune simulation: (a) Immunoglobulin production in different states and B-cell isotypes according to the Simpson index to the antigen. (b) Cytokine level. . . 59
4.12 Cloned sequence of MEBSV (colored as red) into pET-30a (+) expression vector. 60
4.13 Adaptive Immune system 63

List of Tables

4.1 Top proteins with maximum antigenicity nature and extracellular locations. 45
4.2 Epitopes were chosen by CTL for vaccine development against *Plesiomonas Shigelloides.* 47
4.3 HTL epitopes used in the development of a vaccination against *Plesiomonas Shigellodies* 48
4.4 Final selected B-Cell epitopes for vaccine construction against *Plesiomonas Shigelloides* 48
4.5 Selected B cell for screening of vaccine 53
4.6 Docking table indicating the docking scores along different energy values of TlR-4 with vaccine construct 55
4.7 Docking table indicating the docking scores along different energy values of MHC-I with vaccine construct 56
4.8 Docking table indicating the docking scores along different energy values of MHC-II with vaccine construct 57

Abbreviations

Ecc	Enterobacter cloacae complex
HLA	Human Leukocyte Antigen
ICE	Integrative conjugative elements
IEDB	Immune Epitope Database
iMODS	Internal coordinates normal mode analysis
JCat	Java Codon Adaptation tool
MDR	Multiple-drug resistance
MEBSV	Multi-epitope-based subunit vaccine
MHC	Major Histocompatibility complexes
MLST	Multi-locus sequence typing
WGS	Whole genome sequencing

Chapter 1

Introduction

1.1 Background

Plesiomonas shigelloides is a facultative, anaerobic, gram-negative rod shaped bacteria that has been recently identified as the *Enterobacteriaceae*'s first oxidase-positive member of *Enterobacteriaceae* family. *P. shigelloides*, like many other *Enterobacteriaceae* species, can be found in a variety of hosts, such as dogs, pigs, cows, cats and primates [1]. Although *P. shigelloides* is normally considered as freshwater bacteria, it has been correlated with diseases. *P. Shigelloides* is spread mostly by environmental media, particularly freshwaters. Faecal-oral transmission has been associated to drinking polluted water, eating infected aquatic foods, and eating contaminated vegetables. Furthermore, *Plesiomonas shigelloides* disease and infections have been linked to interaction with reptiles, amphibians, and occupational or recreational activities which increase relation with fresh water [2–4]. *P. shigelloides* phylogenetic status has been long a point of contention.

It was previously considered as part of C27 group, along with the main antigen of *P. shigelloides Shigellasonnei* phase I was later discovered to share a set of characteristics with other one bacterial taxa, leading to it being misidentified into the genus the *Vibrionaceae* or *Aeromonas* family. Phylogenetic study using 16S and 5S data, as well as information from MLST (multilocus sequence typing),

has shown this genus belongs to the family of *Enterobacteriaceae*, rather than the Proteus clade [5, 6]. *Plesiomonas shigelloides*, a novel *Enterobacteriaceae* family member, is considered an emerging intestinal pathogen. In humans, it has been linked to both intestinal and extra intestinal illnesses. There have also been instances of *Plesiomonas* involvement in animal episodes of diarrhea. In most aquatic habitats, these microbes can be found [7]. Although these studies were carried out in locations where mixed infections are widespread, none of them looked into coinfections to our knowledge. *P. shigelloides* pathogenicity was evaluated in a province in northwestern Ecuador in the setting of across all age groups and co-infections [8].

Plesiomonas has been blamed for a lot of infections for many years. Gastroenteritis and other forms of diarrhea are among them. Infant meningoencephalitis, sepsis, septic shock and eningitis are examples of extra intestinal diseases that affect the central nervous system, and are of particular concern. Peritonitis, cellulitis, wound and foot infections, endophthalmitis, and keratitis are examples of other infections.

Plesiomonas has been linked to pneumonia, peritonitis, cholangitis, pyosalpinx, pseudo-appendicitis, migratory polyarthritis, cholecystitis and pancreatitis. On the other hand, determining the pathogenic strains and associated virulence markers has been difficult and frustrating. While most standard methods produced equivocal results in terms of the pathogenicity of bacteria, in place found this the microorganism's and its potential for pathogenicity is minimal [9]. *Plesiomonas shigelloides* also belongs to the *Vibrionaceae* family, which has received little attention from the medical world until lately due to its occasional link to human disease. The development of this bacterium as a pathogen of potential enteric of humans and as the infections substances a variety of extra intestinal illnesses is partly responsible for the current spike of scientific interest in it [9–12]. In the *Enterobacteriaceae* family, *P. shigelloides* is a genus with just one specie. It has long been recognized to cause diseases including travelers' diarrhea, gastroenteritis, and serious extra intestinal infections. *P. shigelloides* has also been solely blamed for some foodborne and waterborne outbreaks, based on good microbiological and epidemiological evidence [13]. *P. shigelloides* was first discovered by Henderson

and Ferguson in 1947 and *P. shigelloides* is a genus with just one species. Ingesting contaminated drinking water or foods, as well as coming into contact with infected animals, are all ways for humans to become affected from *P. shigelloides.* *P. shigelloides* is an only source of outer intestinal disorders like as sepsis, septic arthritis, cellulites, osteomyelitis and meningitis. Immunodeficiency, cirrhosis and sickle cell disease are the underlying disorders associated with extra intestinal infections [9].

In China, *P. shigellosis* genotypes were detected in the faeces of a captive YFP kept in semi-natural settings. Whole genome sequencings were performed on strain EE2. Ten separate *P. shigellosis's* samples from human sources, animals and environmental were then matched d to the draught genome. With the Using of Pathosystems Commodity Integration Center (PATRIC) channel, which can provide analytical tools that support biomedical study of infectious bacteria diseases and incorporated data, the goal of this study is to further characterize known pathogenicity as well as find new virulence - associated factors [14].

Although *P. shigelloides* could be determined by a range of available commercially biochemical experiments (Vitex 2 system and Phoenix 100 ID/AST NID card), its low incidence in feces samples and small colony size may cause its presence to be neglected. Only one other *P. shigelloides* draught genome sequence, strain $302-73$, is currently available. However, this sequence is made up of 13 scaffolds that contain 389 contigs.

The entire genome sequence of *P. shigelloides* from the type strain NCTC10360 has been published for the first time [3]. *P. shigelloides* infections have been studied prospectively in Hong Kong and China, and examples of *P. shigelloides* co-infection with viral diarrheal pathogens and bacteria have been recorded. *P. shigelloides* gastroenteritis is widespread in various areas of the world, North America and Europe have lower rates, while Africa and Southeast Asia have higher rates.

Even this, *P. shigelloides* infection is commonly under estimated, in part because it has several clinical symptoms with other infections. Because, *P. shigelloides* is not commonly tested in medical settings, hence knowledge of this pathogen is limited [3, 15–21]. Because of the possibility for disease transmission, bacterial

pollution of surface waterways (including rivers) has remained a serious problem in quality of water [22]. Several bacteria found in surface water are well-known highly infectious that exhibit the difficult trait of drug resistance. *P. shigelloides*, *Vibrio* spp., and *Aeromonas* spp. are examples of pathogens that have been identified as the cause of a variety of opportunistic diseases. But while antibiotic resistance is a well-known nosocomial problem, the environment's role in the replication and dispersion of antibiotics, as well as the resistance factor, is becoming more widely recognized [23]. Releases from municipal wastewater systems, pharmacological sewage, and agricultural waste material are known to contain a high numbers of bacteria which have been revealed to some drug concentrations, resulting in resistant bacteria and their affiliated resistance genes being loaded into the external field [24, 25]. The development of organisms with numerous antibiotic resistances poses a serious public health danger, resulting in longer hospital stays, longer treatment times, and even failure in viral disease treatment.

As a result, *Plesiomonas* pathogenicity should be characterized using rapid specific culture-independent techniques. A large number of bacteria that have been exposed to some drug concentrations are found in municipal wastewater system releases, pharmacological effluent, and agricultural waste material, and these bacteria and their associated resistance genes are then released into the external field.

The development of organisms with numerous a severe threat to public health is posed by antibiotic resistance, resulting in longer hospital stays, longer treatment times, and even failure in viral disease treatment. It wouldn't only make clear previously unknown virulence characteristics related with certain strains, but it would also reveal previously unrecognized virulence characteristics connected with specific strains, but it would also contribute in the construction of quick and moderate methods for strain screening and diagnostic *Plesiomonas* characterization [26–29]. Immunodeficiency, cirrhosis, sickle cell disease and are the disorders that underlie extra pulmonary infections. To date, 10 examples of *P. shigelloides* sepsis and meningitis in newly born have been recorded, with 70% of these cases resulting in the patient's death. A child has been diagnosed with *P. shigelloides* meningoencephalitis and sepsis; this is the 2nd report in Europe, however the first

to survive [30]. To get a greater understanding of *P. shigelloides* genetic diversity and pathogenic ability, we produced complete sequences of genome with 12 of strains represent distinct serogroups which O-antigen genes clusters have already published. *P. shigelloides* phylogenetic position and species-specific genetic features were investigated utilizing phylogeny for a wider group of *Enterobacterales* organisms. *P. shigelloides* evolutionary and genetic diversity dynamics were studied using a pan-genome approach [13, 31].
Vaccination is now the most effective method of boosting the immune system and combating infectious diseases. Vaccine subunits could support fragments of antigenic proteins which can emulate current pathogen effects and lead to an immunological response to the infection [32]. The vaccine's components contain antigenic protein particles that can replicate the presence of a natural infection and trigger an immune response against with the infection being targeted.

Multi-epitope vaccines reduce undesirable parts of pathological immune reactions or adverse reactions when compared to traditional vaccine [33]. Epitope-based vaccines can be cost-effective, can focus immune responses on conserved epitopes, increase safety, and rationally develop epitopes for increased breadth and potency. Scientists have worked over many years to reduce the vaccine development side effects, time and costs. Currently, various immunoinformatic-based strategies for developing and designing competent and efficient new generation epitope-based vaccines are readily available [34].

Our first aim was to create a vaccine which would trigger a powerful immune response following immunisation because we were aware of all the advantages of a multi-epitope vaccine. In studies, a proteome of *P. shigelloides* was examined to identify the highest antigenic proteins and to predict the various T-cell and B-cell epitopes with their corresponding important histocompatibility complexes (MHC) alleles. Finally, using the most potent epitopes, a multi-epitope-based subunit vaccine (MEBSV) with appropriate adjuvants and linkers was designed. The MEBSV sequence was used for the physiochemical and immunological profiling followed by tertiary structure prediction. Disulfide engineering was employed in order to increase stability of the three-dimensional (3D) structure. Simulation and Molecular

docking were used to examine the vaccine-receptor complex's binding interaction and stability. Furthermore, immunological responses caused by the MEBSV antigen were simulated to see how effective it is in real life. Finally, the MEBSV codon was optimized for the E. coli system, followed by in *silico* cloning [3]. Reverse vaccinology approach was applied to proteome to identify three essential and antigenic proteins: LPS-assembly protein LptD, outer-membrane lipoprotein carrier protein, and lipopolysaccharide export system protein. The epitope mapping and prioritizing of these proteins resulted in selecting nine T cells and five B cell epitopes, which were attached together with the use of suitable linkers.

An adjuvant (Cholera Toxin B) was added to the vaccine construct's N-terminus to increase immunogenicity, and a stable MEBSV protein structure was generated. The designed MEBSV exhibited strong bonding and long-lasting interactions with Toll-like receptor4 and MHC molecules. Its high expression in the E. coli K12 strain was also confirmed by codon optimization and in-*silico* reverse translation. The multi-epitope vaccine suggested in this research can stimulate specific immune responses towards *P. shigelloides*, according to computer-aided tests done in this research. To assess this putative vaccine candidate's safety, wet lab verification is necessary.

Carbapenem were used as the last option for the treatment of infections caused by multidrug-resistant bacteria. These also showed strong activity against *AmPC* β-lactamass. Serious issues were faced by clinical management, so the *Carbapenem*s were introduced for the treatment of gram-negative [35]. As the world health organization started its efforts to produce novel drugs, CRE was listed as the most critical pathogen. No research work was done specifically for *P. shigelloides*, therefore the resistance against drugs could not be measured exactly. *Carbapenem* resistance in *P. shigelloides* has previously been linked to over expression of AMPC or ESBLs enzymes, furthermore alterations which influence membrane permeability from isolates of multiple countries, *Carbapenem*s such as NDM, KPC, and OXA-48 were predicted [36]. Factors and clinical outcomes of CRKA infection, as well as to describe the mechanism of *Carbapenem*- resistance, different works were performed. As there is no vaccine reported against *P. shigelloides*, the first

method that suggests an in-*silico*-evaluated completely functioning multi-peptide-based subunit vaccine (MEBSV). This study's MEBSV has advantages over standard vaccines. It has CTL, HTL, and B-cell epitopes, and may induce cellular and humoral host resistance. It consists of these epitopes which target distinct HLAs and identifies new, efficient T cell receptors [37]. Single vaccine can protect numbers of different proteins because it comprises immunogenic protein sequences that fuse into peptides. Vaccine effectiveness improves. This is a likely candidate for a *P. shigelloides* vaccination.
The latest research analyses the whole proteome of Mycobacterium tuberculosis F11 to look for new therapeutic and vaccine treatments against tuberculosis using subtractive genomics, alternative pathway analysis, a database of essential genes (DEG), and sub-cellular localization predictions. Tuberculosis (TB) is a widespread and often fatal infectious illness infectious disease Caused by mycobacterium (MTB) and other Mycobacterium strains. M. tuberculosis is believed to affect one-third of the worldwide people, with new infections occurring at a pace of one every second. Around 1.45 million individuals died in 2010, the majority of whom were from underdeveloped countries (WHO, 2011).

Furthermore, a couple of low - and - middle nations account for almost 80% of the world's active tuberculosis infection [38]. In the last decade, numerous microorganisms' entire genome sequencing has been discovered. The relative genomes and subtractive genomics methodologies can provide important knowledge on discovering treatments for various diseases caused by bacteria [39].

The subtractive genomics approach can be used to search out essential genes that are essential for pathogen survival but are lacking in the host. By discovering non-homologous proteins which are absent in humans, the risks of cross-reactivity and negative impacts can be reduced. By examining these genes with the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database, it is possible to find genes and their products that could be exploited as prospective therapeutic targets [40]. The genomes data is being used to find new medication targets. Non-homologous genes coding for proteins that are expressed in the pathogen but not in the recipient can be identified using a comparative genomics technique. Basic

Local Alignment Search Tool (BLAST) to against human utilizing the BLASTP tool can be used to find these genes. This would get rid of any homologous genes that exist in humans. The important genes needed for the pathogen's existence are then identified using the Database of Essential Genes (DEG) [41]. This method ensures that the drug target is exclusively found in the pathogen and not in people. New targets for variety of diseases have been successfully found using this method [42].

High throughput screening, excretion–toxicity screening, Combinatorial Chemistry, in *silico* absorption, distribution, metabolism, virtual screening, and de novo and structure-based drug design all contribute to speeding up and reducing the cost of current drug discovery. Structure-based computational drug design approaches primarily concentrate on molecules with known 3D structures for target sites, followed by a measurement of their attraction for the target, from which hits are acquired. Knowing about these targets and the medications that go with them, especially those in clinical trials and use, is extremely helpful in accelerating drug development [43].

Immuno-informatics-based strategies for identifying and designing vaccines based on essential components of immunity found in the genome sequence database. This data underscores the necessity for experimental, computational biology, and computation-driven experimental techniques for understanding host–pathogen connections and vaccine development. This research will also highlight the need for innovative treatments or vaccines to tackle antigenic variation [44]. Computational methods can be used to generate and develop new vaccine elements using an in *silico* approach. Bioinformatics and immunoinformatics are two excellent in *silico* technologies that have greatly aided vaccine development.

Bioinformatics is the interdisciplinary study of how to organize and preserve massive numbers of biological data generated by experiments in genetics, biotechnology, and molecular biology. Proteomics, genomes, and vaccinomics are among the investigations included. The identification and characterization of novel antigens has been speed considerably because to the information gained from these fields

[45]. Immunoinformatics is a branch of biology that uses bioinformatics software and apps to assist develop substantial immunological data.

The identification of a range of unique epitopes for B cell and T cell recognition using MHC class I and II molecules is among the most important applications of immunoinformatics. Immunoinformatics looks to be an effective method for identifying new antigenic epitopes which can be utilized in the development of new vaccines for a range of infectious diseases bacteria that causes viruses, fungi, and parasites. The main immunoinformatics software tools used in the last few years for presenting discrete and continual bacterial, viral, and tumor-specific epitopes for B and T cells are discussed [46].

1.2 Problem Statement

P. shigelloides is a newly discovered gram-negative rod-shaped, facultative and anaerobic bacteria which is only oxidase-positive member of the Enterobacteriaceae family, as well as the cause of very serious infections disease such as diarrhea. There is no immune-informatics or vaccine design study for *P. shigelloides* using subtractive genomics.

1.2.1 Aim

The purpose of this research is to find the targets for *P. shigelloides* by employing subtractive genomics. As a result of this work, we were able to identify T and B cell epitopes, which allowed us to create a *P. shigelloides* multi-epitope-based subunit vaccine.

1.2.2 Objectives

- To construct multiepitope subunit vaccines against *Plesiomonas shigelloides* infections.

- To build a MEBSV (Multiepitopes based subunit vaccine) is to control the humoral and cellular immune responses using immuno-informatics and subtractive genomics methods.
- To check humoral response of vaccine by cloning in different hosts.

Chapter 2

Literature Review

2.1 General Overview

Multiplication of bacterial strain within or on the causing harmful effects called as bacterial infection. Decompensation, organ failures, and death may occur due to bacterial infection. For public health, bacterial infections have a crucial impact. So prevention of these infections is very important. The main cause of bacterial infections is intestinal bacteria. About 10 million of the world's population is facing a bacterial infection. Infectious conditions become more critical, due to bacterial resistance against different drugs. Pathogenic bacteria are known for causing serious disease and infection. The bacterial infection develops when bacteria enter and increase their population in the body and defeat the good bacteria and healthy tissues that are predominantly sterile. Any part of the body is susceptible to infection or sickness, which may be brought on either by the organism itself or by the host's defenses.

Humans can be exposed to bacteria through the environment, via food or water, or by biological vectors. Basic modes for the transmission of bacteria to humans are factors, airborne, contact, vehicular, and droplet. For morbidity and mortality, prevention has dramatic effects [47]. For the last three decades, *Enterobacter aerogenes* have been considered as important opportunistic and multi resistant

bacterial pathogens for humans. Gram-negative bacteria were extensively studied and documented all throughout the various epidemics of hospital-acquired illnesses that occurred in Europe and France. The permeability of the membrane is controlled by redundant regulatory cascades, which also ensures that bacteria are protected and that antibiotics are broken down properly with the assistance of detoxifying enzymes. Strong antibiotic resistance of this bacterial species is due to its ability to adopt numerous genetic mobile elements. Furthermore, their fitness allows them to colonize the numbers of hosts and environments with efficient adaptability of their metabolism and physiology for environmental stress and external conditions. In colonized patients, it is able to resist promptly to antibiotic treatment.

Enterobacter aerogenes and *E. cloacae* are two important species of *Enterobacter*, that have clinical importance as opportunistic bacteria [48]. It belongs to *Enterobacteriaceae* family and genus of Gram-negative, facultative anaerobic, non-spore-forming, and rod-shaped bacteria. Before 1960 *Enterobacter aerogenes* was named *Agrobacteria aerogenes*, but after that, it was added to the genus *Enterobacter*. This species was renamed *Plesiomonas Shigelloides* in 1971 due to its mobility and genetic relatedness to *P. shigelloides* genus.

Enterobacteriaceae family lies within the *Gamma protea* bacteria according to phylogenetic analysis. These are distributed worldwide. They are found everywhere and in animals from insects to man. They are considered to be very important due to their quick generation time, facility to grow on specific media, economic importance, and pathogenicity for man and animals [49]. *Enterobacter aerogenes* is widely found in the gastrointestinal tract and environment. This opportunistic pathogen is considered to be very important, as it is resistant to multiple antibiotics that are used for its treatment [50]. Meningitis, surgical site infection, urinary tract infection, bacteremia, and pneumonia are the hospital-acquired infections caused by *Enterobacter aerogenes*, for antibacterial therapy phages are considered to be very essential. Despite the fact therapeutic reasons have generated, concerns about immunogenicity, restriction/ modification, fast toxin release via lytic action, the development of the bacterium resistance, and other issues,

phage therapy is still being studied for pneumonia, *siphoviridae* phage showed great results against *P. shigelloides* [51].

Plesiomonas was previously distributed like an element of the *Vibrionaceae* family, but it is now a member of the *Enterobacteriaceae* family. *Plesiomonas shigelloides* is a facultative anaerobic, rod-shaped, Gram-negative, flagellated, bacterium in the genus *Plesiomonas*. Humans, sheep, fish, cows, dogs, pigs and turkey vultures are some of the hosts, and it can also be found in aquatic areas. A serotyping approach based on antigenic variation in the lipopolysaccharide (O-antigen), of which there are currently 102 flagella and 51H-antigen was used for categorization, epidemiological investigation, and identification of *P. shigelloides* [52, 53].

Diarrheal disorders affected a large numbers of individuals in worldwide or are create severe medical issues. such as *Enterobacteriaceae*, Gram-negative bacteria and *Vibrionaceae* are commonly responsible for this disease [54, 55]. *Plesiomonas shigelloides*, a member of the *Enterobacteriaceae* family, has been identified as a major source of gastrointestinal illnesses.

It was 3rd in ranked Japan as a cause of diarrhea in tourists [56]. Human feces, several domestic and wild animals, and the environment, especially water, have all been found to have these germs. Drinking unfiltered water, eating uncooked shellfish, and travelling to underdeveloped countries have all been linked to *P. shigelloides* infections. 2 and 3 Different virulence factors, including a lipopolysaccharide, have been reported in an attempt to discover the harmful process of *P. shigelloides* [26, 57, 58].

2.2 Cellular structure

Microscopic representation of *Pl. shigelloides*, a facultative, anaerobic, gram - ve rod-shaped bacterium that was recently identified as the first oxidase-positive member of the *Enterobacteriaceae* family. *Pl. shigelloides* was found to be the first oxidase +ve member of the *Enterobacteriaceae* family as in Figure 2.1 [54].

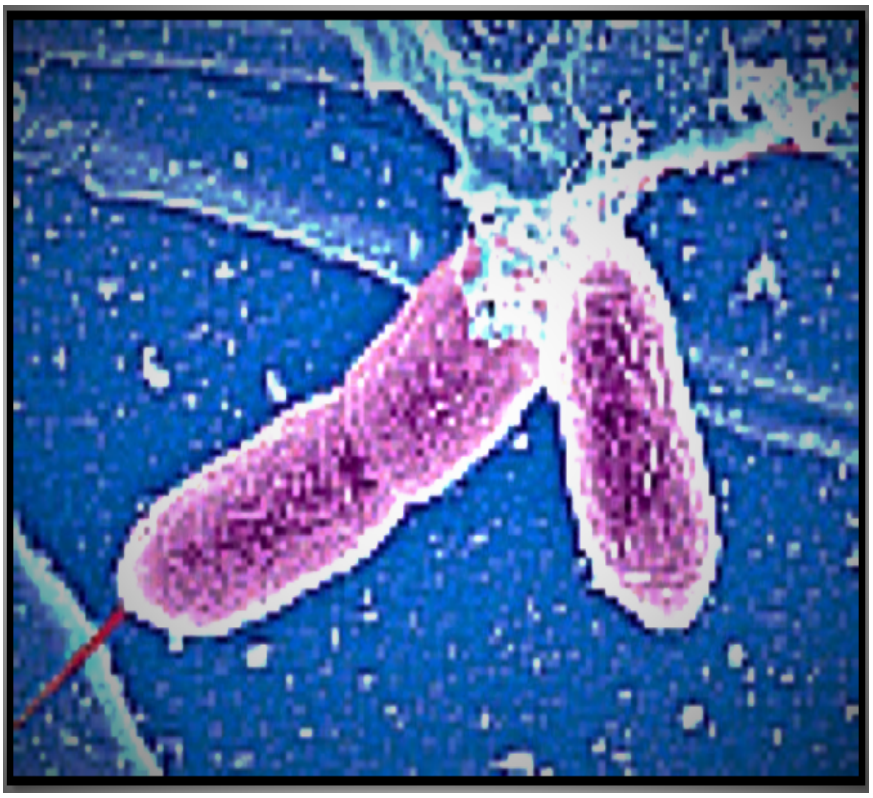

FIGURE 2.1: Microscopic representation of *P. Shigelloides*.

LPS is a pathogen-associated molecular sequence and innate immune stimulant. It's Gram-negative bacteria's main virulence factor, especially *Plesiomonas shigelloides*. Smooth bacteria's *LPS* molecules have a *core oligosaccharide, O-specific polysaccharide*, and *hydrophobic lipid A(LA)*. By modulating *LPS's* interaction with TLR4, the LA area affects endotoxin biological activity in macrophages and monocytes [59–61]. Signaling pathways are engaged, regulating transcriptional elements (e.g., NF-B) and creating proinflammatory cytokines (e.g., IL-6, IL-1, TNF) which influence the patient's immunological reaction to the infection. When Gram-negative bacteria produce bacteremia, a hypersensitivity to LPS can result in septic shock and sepsis [62]. The majority of bacteria are mono chromosome bacteria, meaning that their core genome is carried on a single bacterial chromosome. A subset of bacteria (known as multi chromosome bacteria) has a synthetic genome with the fundamental broken into two or even more replicons. The replication starting setting (replication origin and initiator protein) in the bigger replicon is similar to that found in mono chromosome bacteria, but the other (typically smaller) replicons are plasmid-derived. They're known as *chromids*. The fact that

plasmids are irregularly dispersed across bacteria, while *chromids* are consistently represented in all variants of a particular group, which could be a species, a genus, or an order, such as the *Vibrionales*, distinguishes *chromids* from plasmids. Other criteria for determining whether a replicon is a *chromid* have been suggested, including the transfer of core-genome genes onto the *chromid*, some of which may be critical [63].

LA has examined *P. shigelloides* LPS structures extensively. *P. shigelloides* LA molecules have varied acyl chain lengths, saturations, and PEtn replacement. Three *P. shigelloides O-serotypes* have defined with reference to LA, and structural examinations show that this area is largely conserved across all O-serotypes [64, 65]. *P. shigelloides* LA has a -GlcpN4P-(1 6)—GlcpN1P backbone with these fatty acids: 12 : 0(3-OH) at O-3 and O-3′, 14 : 0[3-(R)-OH] at N-2 and N-2′. In *P. shigelloides* LA, the acyl groups at N-2′ and O-3′ are replaced by secondary acyls: (1) cis-9-hexadecenoic acid ($9c - 16 : 1$) or *hexadodecanoic* acid (16 : 0) at N-2′ and 12 : 0 at O-3′ (strain dependent); (2)14 : 0 at N-2′ and 12 : 0 at O-3′; (3)12 : 0 at N-2′ and PEtn *P. shigelloides'* fatty acid chain and concentration differ from E.coli's. [66].

2.3 Ecology

Plesiomonas is a genus with such an only one homogenous specie (*P. shigelloides*). This bacterium first has been reported in 1947, it was established to share some traits among *shigella*. First strain had S. sonnei's major somatic antigen but was physiologically distinct. This organism, designated C27, and related individuals were used to show how *Pseudomonas, Vibrio* and *Aeromonas* evolved in chronological order. *Plesiomonas* was first recognized as a genus in the *Vibrionaceae* family by Schubert and Habs in 1962, based on a number of distinguishing characteristics, its resemblance to the genus of *Aeromonas ("plesio," neighbor, "monas," Aeromonas)*. It is portly flagellated, cytochrome oxidase positive and facultative anaerobic, just like the other species of the *Vibrionaceae* genus. Furthermore, isolates share a sensitivity to O/129 with species of the genus Vibrio. 5S rDNA

sequencing are found that the bacteria is very closely linked to the genus Proteus [67–69]. *Plesiomonas* present categorization is based on Schubert and Habs premise that the bacteria had a number of distinct characteristics that necessitated the creation of a new genus. Recent DNA genetic similarity and phenotypic analysis research have showed that within the species (independent of isolate source), a single hybridization group exists, indicating a very similar number of bacteria [59, 61].

P. shigelloides is typically transmitted and through environment, particularly freshwaters. Ingestion of polluted water, infected aquatic foods, or poisoned vegetables has been documented as a source of faecal-oral transmission. Additionally, *Plesiomonas* infections and illnesses have been linked to interaction with reptiles [2], amphibians, and recreational or occupational activities that enhance exposure to freshwaters [4]. Despite the fact that, *P. shigelloides* has been identified from streams, rivers, and lakes, no conclusive research on its occurrence in freshwater, particularly in South Africa, has been reported. The link among pathogen populations and physicochemical factors is usually a high priority for epidemiologists. Environmental conditions have an important role in pathogen survival and spreading in the environment (disease ecology). These variables are required for forecasting and preventing infectious disease outbreaks caused by aquatic pathogens [70, 71]. While ecological determinants can affect bacterial ecology, dispersion, and density, their effects on freshwater microorganisms are mainly examined at the community level. These microbial communities have been molded by environmental circumstances and spreading procedures [72] at regional and local levels [73].

In addition, environmental factors that encourage species-specific alterations in the bacterial population are typically ignored while attempting to comprehend the destinies of pathogens, the transmission of diseases, and outbreaks that occur in freshwater environments. In addition, the immune-regulatory power of climate change, particularly on weather variables (flooding, precipitation, temperature, and environmental deterioration), physical and chemical factors in aquatic environments, and its consequent effect on the time series of waterborne diseases, is

poorly observed in a few species. This is due to the fact that climate change is primarily affecting weather variables (flooding, precipitation, temperature, and environmental deterioration) [74]. Because infectious diseases are rapidly being linked to changes in lifestyle, climate change, and global warming, studies relating to above aspects are critical to understanding microbial destiny and transmission [75].

P. shigelloides' presence and distribution in freshwaters has gotten little study. However, at both municipal and industrial levels, freshwaters are continually and environmentally abused, and most of them act as actual destination for inadequately treated wastewater effluents. *P. shigelloides* is a bacterium that causes gastroenteritis and extra intestinal infections. *P. shigelloides* is a bacterium that causes gastroenteritis and extra intestinal infections. It is a well-known and growing waterborne and foodborne pathogen [76, 77].

2.4 Pathogenesis

Various bacterial diseases, in combination with environmental conditions, produce large mortality rates and poor growth, reducing yield and commercial viability of this fish. This bacterium, which is usually found in tropical and subtropical countries, has its principal habitat in fresh and estuary waters. This bacterium can be found in a variety of ability to stay and homeotherms, aquatic creatures, moisture reptiles, and birds.

P. shigelloides has been linked to high fatality in trout, both alone and in combination with other bacteria such as *Flavobacterium spp*, and *aeromonas hydrophilla* have both been linked to high death in trout. This pathogen has been discovered as a major pathogen in Beijing-area cultured sturgeons. *P. shigelloides* has just been identified from clinical studies of fish with mass mortality of *Ctenopharyngodo nidellus* and *Orechromis nilotica*, and has proven to be extremely harmful for such cultured fishes [76]. The drastic reductions in *E. aerogenes* pathogenicity were

reported among the resistant strains with drug-resistant membrane permeability changes. This variation is apparent even though this species is only moderately virulent in this model, even though the tested strains harbored the HPI pathogenic factor-encoding genes. Changes in the permeability of outer membrane (OM), such as a deficiency of porins that are a major nutrient entry channel and have a significant effect on bacterial fitness. Antibiotic stress promotes the establishment of resistant bacteria that lack porin and have lipopolysaccharide (LPS) alterations that result in an abnormal membrane state. This results in an adverse fitness cost, which changes the capacity for bacterium colonization.

P. shigelloides' taxonomic placement has been debated for quite some time. The primary antigen, Shigella sonnei phase I, and this bacterium were first placed in the family C27. Misidentification of *P. shigelloides* into genus of *Aeromonas* or the group of *Vibrionaceae* occurred because of its similarities to other bacterial taxa [5, 6]. Phylogenetic analysis using 5S or 16S data, in addition to data collected from the multilocus sequence typing (MLST), has shown been this branch is more closely related to the Enterobacteriaceae clade than it is to the Proteus taxon [78]. Verification of the bacterium's genetic variety has been accomplished by the use of pulsed-field gel electrophoresis (PFGE), randomly amplified polymorphic DNA (RAPD), and multilocus sequence typing (MLST), among other methodologies [79]. On the other hand, researchers that used approaches that were previously accessible found that it was impossible to collect taxonomy with greater resolution. Bacterial evolution, ecology, and toxicity can all be better understood with correct taxonomy. From a whole-genome viewpoint, therefore, *P. shigelloides* evolutionary position and genetic variation are yet to be fully characterized. Furthermore, *P. shigelloides* pervasive living condition and the population concentration of cohabiting bacteria in the human niche make genetic transmission between donor and recipient strains possible. HGT is a method of DNA transmission that results in unique traits such as niche adaptability and pathogenicity. Furthermore, the function of HGT in *P. shigelloides* pan-genome is likewise unclear [80]. Human vomiting and diarrhea, including severe secretory dysentery, aggressive *shigellosis*-like disorders, and cholera-like diseases, have all been linked

to P. shigelloides. The ingestion of seafood or fish, as well as polluted water, are commonly linked to outbreaks. It has also been linked to bacteremia, pneumonia, peritonitis, meningitis, hepatobiliary illness, septicemia, and pseudo appendicitis, among other extra intestinal conditions [81]. In vitro, *P. shigelloides* may attach to and penetrate host cells, and it has cytotoxicity or enters toxicity. This bacterium is sensitive to a number of antibiotic, with the exception of ampicillin ; although, multidrug-resistant characteristics have been described [82]. However epidemiological and clinical evidence suggests that *P. shigelloides* plays an important function in human infections, the actual mechanism of pathogenicity and the bacterium's genomic history have yet to be fully explored.

We created draught genome sequences for 12 of strains that represent diverse serogroups whose O-antigen gene clusters have already been published. This was done so that we could acquire a better knowledge of the genetic diversity of *P. shigelloides'* as well as its potential for causing harm [13]. *P. shigelloides'* phylogeny and genetic characteristics were explored using enterobacterales phylogeny. Pan-genome analysis revealed *P. shigelloides'* genetic diversity and evolution patterns. We found horizontally transmitted genes, donor taxa, gene of expansions and mobile genetic elements (MGEs), and quantity. Functional genomics was used to determine virulence-related genes. Genes of resistance were discovered for 18 of antibiotics, including aminoglycosides, beta-lactams, fluoroquinolones, tetra cyclines, polymyxin, and sulfonamides.

Furthermore, relevant models for researching *Plesiomonas* pathogenesis are still to be developed, and standard models have proven to be ineffective. *Plesiomonas* VFs must be detected using rapid low-cost and other polymerase chain reaction techniques, which have already been developed for the study of other, enter pathogens. There are currently no particular primer- and molecular-based methods to detecting for the majority, if not all, anticipated *Plesiomonas* VFs. Furthermore, there are no strain-specific or quick procedures for distinguishing between pathogenic and non-pathogenic *Plesiomonas* strains [83]. It is not completely out of the question that numerous occurrences of this kind took place in *Plesiomonas* due to the fact that the emergence of new diseases (including *Plesiomonas*) is connected

to a complex network of interconnected mechanisms. Among these mechanisms, horizontal and mutation gene transfer the primary forces at play. For most circumstances, gene transfer or loss will have an impact on genome evolution, turning a non-pathogenic strain pathogenic [84]. Despite the fact that these occurrences may have aided in the establishment of toxicity in *Plesiomonas*, no studies on the subject have been published. PAIs, resistant islands, IS elements, tRNA genes metabolic islands, phage-related genes, mobility genes, and directly repeated were all be studied in order to characterize *Plesiomonas* pathogenicity and genome evolution.

Plesiomonas pathogenicity should be characterised utilising culture-independent methods. Before wet-lab research to identify virulence marker candidates in Plesiomonas, a comparison of the accessible genomes could provide information and address concerns about the organism's rising pathogenicity. It would help provide quick and low-cost methods for strain testing and *Plesiomonas* diagnostic classification. In this study, we used in silico screening to discover the *Plesiomonas* full genome sequence components that were publically available with the type III, IV, VI, and VII secretion systems (T3SS, T4SS, T6SS, and T7SS); ICE elements; prophages; VFs; and T3SS, T4SS, T6SS, and T7SS.

P. shigelloides was found in the faces of a semi-wild YFP in China. Strain EE2's genome was sequenced. The whole genome was compared to eleven animal, human, and environmental *P. shigelloides* strains. This work uses the Path systems Resource Connection Center (PATRIC) pipeline to types known virulence genes and uncover new possible causal factors [14].

2.5 Genome Organization

Plesiomonas shigelloides is a facultative and Gram negative anaerobic bacteria which has been reported in surface water, freshwater and a range of domestic animals and wild [9, 85, 86]. *Lipopolysaccharide* (LPS) is an immunodominant and structural component found on the outer surface of Gram negative bacteria.

The three elements that make up this protein are core *oligosaccharide*, lipid A and O side chain or O-specific antigen. The outer constituent of LPS is the O-antigen, which is a polymer of oligosaccharide protein subunits. Another intriguing feature is the enormous chemical diversity of O-antigen LPS, which corresponds to genetically identical variance in the genes involved in their production, described as *wb* cluster. The genomics of O-antigen production in *Enterobacteriaceae* has been extensively researched, revealing that *wb* clusters comprise genes involved in beginning sugar biosynthesis, O-antigen polymerases, glycosyl transferases, and O antigen export [87].

The O antigen is a component of the facultative bacterial outer surface *LPS (lipopolysaccharide)*. It consists of several of an *oligosaccharide* unit (O unit). It provides O-serotype specificity and delivers significant antigen diversity to the surface of the cell. In the *E. coli* classification system, 166 known O antigens, and *Shigella* strains have 34 of them [88].

Shigella strains and Traditional *E. coli* of *E. coli* have a lot of similarities in O antigens, and 190 distinct O antigens present in *E. coli*. The host immune system selectively targets the surface O antigen, which could explain why *E. coli* maintains a variety of O-antigen types. *Sonnei* contains an O antigen that is similar that of *Plesiomonas shigelloides* serotype 17 and is not seen in other E. coli strains. Sonnei's O-antigen genes are unique in that they are present on the Pin v (invasion plasmid), this is over 180 kb in size or necessary for recipient epithelial cell penetration. *P. shigelloides* serotype O17 chromosomal O-antigen genes have been shown to produce offspring with this O-antigen gene cluster. According with majority of the plasmid-borne O-antigen gene clusters, the Sonnei O-antigen gene cluster is equivalent to that of *P. shigelloides* O-antigen for just a few hundred base pairs [89–91]. We acquired the whole *P. shigelloides* O17 O-antigen genome cluster, as well as a part of the Sonnei plasmid borne O-gene cluster, to obtain the complete genome. With exception of *wbgZ* and *wzz* genes, comparison between both sequences revealed that they are extremely similar. The close resemblance indicates that the O-antigen gene cluster was just passed through one species to the next. I believe the O-antigen genes of *Sonnei* were transported by

P. shigelloides to *E. coli* since they are on a plasmid; the significant concentration of homology for the most genes indicates that *Sonnei* very recently obtained. The difference between the two *wbgZ* and *wzz* genes is greater than the difference in the middle of the further gene in the two genes clusters, and there's indications that the plasmid's *wzz* gene has been selected for alteration in the new host [29, 92].

2.6 Metabolism

In the gamma family of *Proteobacteria*, the genus *Plesiomonas* seems to be situated between the *Vibrionaceae* and *Enterobacteriaceae* families [55]. *P. shigelloides* is an anaerobic facultative and gram-negative bacterium which is and chemoorganotrophic. The specialized PTS (phosphoenolpyruvate-dependent phosphotransferase system) in Gram-negative bacteria facilitates active transport of *hexitols* and concomitant phosphorylation of sugars [93, 94]. *P. shigelloides* is a very common pathogen that causes sepsis, newborn meningoencephalitis, and meningitis. Twelve infants have so far been reported to have the virus, and seven (58.3%) of them have perished. The very first case of *P. shigelloides* meningoencephalitis and sepsis in China is described in our example study of a newborn. For articles published between January 1970 and June 2014, we searched MEDLINE [95]. *Meningoencephalitis, meningitis*, and *Plesiomonas shigelloides or Aeromonas shigelloides* were the search phrases utilized. *Plesiomonas shigelloides/Aeromonas shigelloides* were implicated in 13 studies describing neonates with meningitis/meningoencephalitis, and two of those papers documented the same case. *Plesiomonas shigelloides* infections in newborn child can result in meningitis or meningoencephalitis and have an effect on the circulatory, gastrointestinal, respiratory, coagulation, hematologic, and other systems. Initial *Plesiomonas shigelloides*-induced sepsis and meningoencephalitis symptoms are comparable to those associated with a typical infection. Furthermore, this illness spreads quickly and kills a number of organs. If a physician waits too long to make a diagnosis or administers the incorrect antibiotic therapy, fatality rates will be high. Despite early use of sensitive

antibiotics, it could be challenging to stop the substantial central nervous system damage. The PTS uses phosphoenolpyruvate as a phosphoryl resource in a data management that involves Enzyme I, Enzyme II, and HPr. *P. shigelloides'* sugar metabolism helped us understand Gram-negative bacteria's diversity. Mannose is carried into cells via a PTS that creates mannose 6-phosphate. *P. shigelloides* grew less when cultured in broth culture with added mannose. In vivo, 13C-NMR was employed in resting cells, and in vitro, 31P-NMR was used in cell-free samples to assess mannose and glucose metabolism [96]. *P. shigelloides* has the ability to develop in both anaerobic and aerobic environments [97]. Because the enzymes needed for metabolic processes under aerobic and anaerobic conditions differ significantly, it is necessary to control the impression of target gene associated with cell responsibility like as excretion system biosynthetic pathways, macromolecule synthesis and nutrient absorption at same time to reply to the access to oxygen [29].

The *Arc* 2 different signal transduction system, which comprises of the kinase detector *ArcB* and its associated response regulator *ArcA*, is one of the system that allows. *E. coli* is able to adapt to changes in oxygen availability. In the existence of hypoxia and the energy supplied by *ATP, ArcB* is activated as a simple electron carrier. It has three cytoplasmic domains, and one of its tasks is to phosphorylate *ArcA* [95], the phosphate group is transported from His292 in the H1 domain to His717 in the H2 domain, Asp576 in the D1 domain, and lastly Asp54 in *ArcA* after auto phosphorylation of His292 in the H1 domain [98].

2.7 Clinical Manifestation

Plesiomonas shigelloides is a new bacterium that can cause extra intestinal and gastroenteritis infections in humans. We conducted a bibliometric survey to look at publishing trends in *Plesiomonas*-related research throughout place and time, international collaborative identify gaps, collaborations, and suggest future research options. The search phrase "*Plesiomonas shigelloides*" has been used to

find publications published in the Web of Science database between 1990 and 2017 [99]. The *Plesiomonas* group still has only one species, *P. shigelloides. P. shigelloides* is currently better recognized as a human intestinal infection than species of the *Aeromonas* group, thanks to oyster- as well as other water-related epidemics of human diarrheal disorders in humans. For all of these newcomers among increasing range of entero-pathogens, no human volunteer experiments with *A. hydrophilia or P. shigelloides* have been effectively carried out to fulfil Koch's postulate [100]. Along with *Plesiomonas shigelloides* ($n = 82, 52.9$ percent), nuclear magnetic resonance ($n = 13, 8.4$ percent) and lipopolysaccharide were among the top Authors Keywords and research concentration. Diarrhea ($n = 43, 27.7\%$), infections ($n = 31, 20.0\%$) and *Aeromonas* species ($n = 41, 26.5\%$) were all heavily represented. Two mega-clusters of author cooperation and collaboration and bonding networks were established, with nodes discussed exclusively by writers from greater countries [99]. Limited data about the rate of *P. shigelloides* infection in China was included in a recent comprehensive assessment of diseases caused by *P. shigelloides. P. shigelloides* has been linked to infections that cause gastroenteritis and extra-intestinal diseases, according to new research. We present a local viewpoint on *P. shigelloides*-related disorders in southeast China [101]. Bibliometric method is a measurement way of evaluating the quantitative and qualitative coverage and efficiency of scientific research in a certain field [102]. It can be used to predict development research goals, sources of funding, and multidisciplinary collaborations, as well as to define international and national research concentration and analyze research quality [103].

2.8 Antibiotic resistance

Antimicrobial resistance is one of the biggest challenges in medicine and veterinary medicine today. More concerning is the fact that despite their inherent antibiotic sensitivity, freshwater bacteria have developed considerable antibiotic resistance. Nowadays, it is believed that bacteria with antibiotic resistance and

the genes underlying that resistance can be found in marine ecosystem. Basically, manmade pollutants can enter aquatic habitats. from a variety of sources, such as storm water, wastewater treatment plants, agricultural wastes, hospital effluents, or surface runoffs, industrial wastes, landfill effluents, agricultural runoffs from slaughterhouses, flooding, and recreational areas [104]. Most of these are extensively contaminated with antibiotic residues, herbicides, pesticides, medicines, biocides, textile dyes and massive xenogenetic elements, organic compounds, genes and antibiotic-resistant bacteria [105].

The use of antibiotics is still a key component of infection treatment and prevention. Antibiotic resistance, therefore, is known to have a negative effect on the efficiency of treating *Plesiomonas* infections. Due to antibiotic medication error and tolerance, *Plesiomonas* infections have resulted in antimicrobial therapy and a number of fatalities. According to studies, treatment with at most one or a mixture of antibiotics from the aminoglycoside, cephalosporin, carbapenem, and quinolone antibiotic classes was necessary to ensure survival in neonates with extra intestinal infections caused by *shigelloides* [106].

E. aerogenes strains exhibit a wide range of antibiotic resistance [107]. They have a low level of naturally occurring *AmpC* chromosomal cephalosporins (Bush group 1), which confers resistance to cephalosporins (first-generation antibiotics). Hyper production of chromosomal *AmpC* in diagnostic and therapeutic strains results in resistance to all β-lactam antibiotics except *cefpirome, cefepime* and *carbapenems* following induction with the third-generation *cephalosporins* associated mutation. The *AmpC* cephalosporinase gene (blaCMY-10) carried on a plasmid exhibits similar phenotypes, by Extended - Spectrum Beta strains were identified in 1993. Pitout et al. separated strains capable of generating ESBLs that were immune to gentamicin and co-trimoxazole [108].

Numerous ESBLs have been listed as belonging to the TEM, SHV, and CTX β-lactamases families, but TEM-24 continues to be correlated with this species favorable conjugative plasmid. The response to carbapenems is retained in these producer strains. Simultaneously, several strains conferring resistance to imipenem

have been characterized. The paucity of antibiotic absorption in these isolated strains is primarily due to a change in the expression of porin: a modification in the *Omp35/Omp36* ratio is identified, preceded by an absence of porins in strains collected during the diagnosis [109].

Intriguingly, an entirely novel mechanism of impermeability has been described, involving the occurrence of a mutation in *Omp36* that significantly changes the permeability of the channel. Finally, *carbapenemases* of the NDM, IPM, or KPC types have been implicated in *carbapenem* resistance since 2008. Additionally, roughly 40% of multidrug-resistant (MDR) pathogenic isolates possess membrane transport mechanisms. The resistance against Quinolone occurs as a result of target alterations or as a result of resistance conferred by other species via plasmids (*qepA, or qnrS* codes for an efflux pump). Eventually, the complete resistance is not an uncommon phenotype in *E. aerogenes*, because a strain conferring resistance against all antibiotics, such as colistin, has been separated and investigated due to a mutation in *pmrA* [110].

2.9 Treatment and preventions

Antibiotic resistance is becoming an increasing issue when it comes to treating *Enterobacter* infections. β-lactams, carbapenems, β-lactamase inhibitors, aminoglycosides, fluoroquinolones, and trimethoprim are all possible therapies.

Cephalosporins of the 1st and 2nd generations are often ineffective against *Enterobacter* infections. While treatment with a 3rd- generation cephalosporin could be beneficial against some *Enterobacter* strains, but the treatment with the 3rd generation *cephalosporins* may result in multidrug resistance illness. The 3rd generation *cephalosporins* are essential to promote de-repressed *Enterobacter AmpC*β-lactamase variations, resulting in the enzyme's overproduction that causes the development of resistance. The 3rd generation cephalosporins are not advised for severe infections due to an escalated risk of resistance, notably among *Enterobacter aerogenes* and *Enterobacter cloacae* the two most clinically significant.

Other resistance strategies also exist like the insertion of a plasmid transferable *AmpC* gene and the occurrence of a mutation in the *AmpR* repressor. The 4th generation cephalosporins are stable against AmpCβ-lactamases and are thus regarded as an appropriate therapeutic choice in the absence of enzyme i.e., ESBL (Extended-Spectrum β-lactamase).This enzyme is capable of hydrolyzing oxyimino cephalosporins, rendering them ineffective against both 3rd and 4th generation cephalosporins [111].

Carbapenems have been found to be the most effective treatment for *Enterobacter aerogenes* that are multidrug-resistant. Impanel and has been demonstrated to be proficient against *Enterobacter aerogenes*. Historically, *carbapenems* were not frequently affected by ESBL but the resistance against them, on the other hand, has been growing in past few years. *Enterobacter* develops resistance to carbapenems via several plasmid-encoding enzymes, allowing for horizontal transfer and the propagation of resistance. Polymyxins, Fosfomycin, tigecycline, and carbapenems are all possible treatments for carbapenem-resistant *Enterobacter* (CRE) [112], A double carbapenem regimen can seem contradictory because one carbapenem drug (a "sacrificial" prescription) works best with a carbapenemase enzyme, allowing the less-favored concomitant carbapenem to stay at a higher concentration. In the cases of CRE that have more severe clinical symptoms, such as heart failure and quickly advancing disease, a mixture of antimicrobial drugs has been demonstrated to be more beneficial [111–113].

Using a combination of colistin with a backbone of ceftazidime has been found to be successful. Carbapenems, Fosfomycin and tigecycline are other antibiotics that can be taken with colistin. Colistin resistance has, however, been on the rise. There is an ongoing debate on whether Carbapenem-resistant strains of CRE should be treated with tigecycline or a mixture of two distinct carbapenems. Other antibiotics such as Fosfomycin, colistin, and aminoglycosides may be administered in combination with tigecycline. People with the renal disease must take care while taking drugs such as aminoglycosides, and colistin as these medications might cause more kidney damage [114]. Fosfomycin and aminoglycosides are commonly used as monotherapy in CRE UTIs when the patient is not critically ill. For

pyelonephritis, Fosfomycin should not be used due to limited oral absorption. CRE UTI patients should not take tigecycline or colistin due to the short half-life of these antibiotics in the urine. The use of additional therapy with a carbapenem, in conjunction with aminoglycoside or colistin, may be an effective therapeutic option for a patient with a CRE UTI if they are in a vulnerable situation [115].

There is a wide degree of diversity in the treatment for CRE, with many different combinations of drugs required. While the resistance remains, new medications are continuously being investigated as a possible solution.

2.10 Subtractive genomics

Modern advances in bioinformatics and computational biology have resulted in the development of numerous *in-silico* analyses and drug design methodologies, thereby avoiding the time and cost associated with trial-and-error testing during drug development [40]. These strategies are used to narrow down the list of prospective drug targets that will be subjected to lab testing.

Subtractive genomics is an *in-silico* approach that is utilized to identify therapeutic targets by identifying both the essential and non-homologous proteins of the respective pathogen [116]. The Database of Essential Genes (DEG) service can be used to identify proteins implicated in critical metabolic pathways necessary for the pathogen's persistence. Additionally, the identification of proteins that are identical to human proteins can be used to screen for potential adverse drug reactions during the computer-assisted drug discovery process.

The subtractive genomics technique identifies the novel therapeutic targets within the pathogen by identifying critical proteins required for pathogen proliferation and survival. The workflow opted during the subtractive genomics to develop the drug against any economically significant pathogen is shown in Figure 2.2.

The different steps were retrieval of whole genome, exclusion of paralogous sequences, identification by blast, draggability and conservancy analysis.

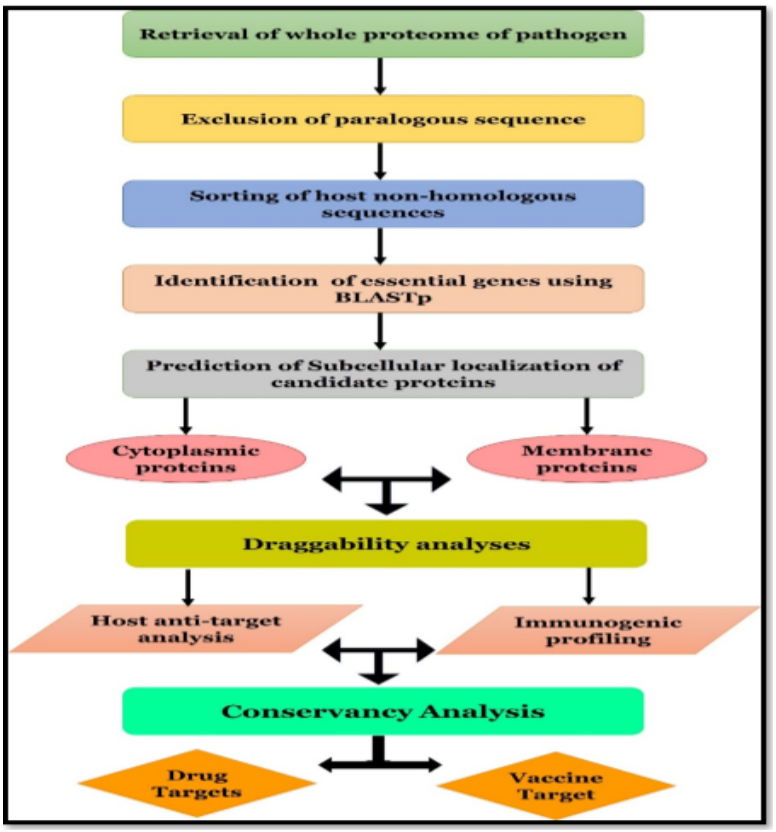

FIGURE 2.2: The steps followed during subtractive genomics in designing drug against pathogen.

2.11 Immuno-informatics

The desired immune response can be engineered by adjusting the various components of the immune system, which can be classed as cellular or humoral. An antigen can be recognized by the MHC molecules found in T lymphocytes if a vaccination that causes a cellular response is used for example, a tuberculosis vaccine , or a parasite vaccine against leishmaniosis [117]. The database used during the analysis is, IEDB. For locating MHC molecules, one can predict T-cell epitopes using numerous options, such as for a certain species, Type I or II, or even the alleles. On the other hand, a humoral response is also necessary for effective immune response, therefore the program must recognize B cell antigens, as is the case with influenza or HIV [40]. ABCpred is online server that search for sequential epitopes for B lymphocytes. Additionally, using several online tools a protein's three-dimensional structure can be used to anticipate conformational epitopes for B lymphocytes, such as the, Disco Tope. This online software programs is based on

computer programming on previously known epitopes and non-epitopes in order to assign values to new peptides and to forecast whether they are epitopes or not. Support Vector Machines (SVMs), Position-Specific Scoring Matrices (PSSMs), Artificial Neural Networks (ANNs), and Hidden Markov Models are just some of the methods that can be used to achieve this goal (HMMs). Many benefits and varying degrees of accuracy can be gained by using each technique [118].

To conduct an analysis, an individual's "immunome" must be accessed; it consists of all the genes and proteins expressed by the cells that participate in the individual's immune system. "Immunomics" is the study of all the various immune reactions that take place in the response, and each is unique to a particular organism. That is why it is vital to conduct studies using relevant information about the host organism. The exploration of immunomics, employing molecular biology and other molecular techniques, has led to significant advancements in understanding how the immune system functions [119].

Immunoinformatics is the application of computational methods to the study and design of algorithms to map B- and T-cell epitopes, which shortens the time and expense for testing pathogen gene products in a laboratory. By learning more about these areas, an immunologist will be able to examine possible novel vaccine targets. The whole approach is termed as "reverse vaccinology" which analyses the pathogen genome in order to find possible antigenic proteins [120], the best reason for favoring this approach is that the traditional procedures need cultivating of pathogens and then extracting their pathogenic proteins. Although infections develop quickly, it is difficult and time-consuming to extract their proteins, which are subsequently tested on a wide scale. The capability of detecting virulence and membrane-bound proteins via an immunoinformatic approach is not disputed. After identifying an antigen that elicits the desired reaction, through immunoinformatics it can be predicted either a portion of antigenic protein can create the optimal stimulus on its own or not. If a protein contains only one epitope, it can be used in a subunit vaccine and coupled with other epitopes from different organisms to create a multivalent vaccine, which considerably reduces the cost of its formulation. Epitopes can be generated synthetically or identified using molecular

biology methods. It makes a vaccine safer during the manufacturing process, as there is no possibility of passaging and escaping of infectious organisms [121].

To determine epitopes, proteins are examined for hydrophilic areas. The protein's three-dimensional structure is determined by the interactions of the amino acids with the media, and with the region exposed outside the membrane containing hydrophilic amino acids. The hydrophobic amino acids, on the other hand, are positioned in the structure's center. Hydrophobic amino acids are found in the middle of a structure. It is more likely that if this protein interfaces with immune cells, it will communicate with the hydrophilic area of the epitope. The stability of peptide complex bound with MHC can also be predicted using various tools like SOPMA and PSIPRED workbench, because some epitopes can relate to greater force and affinity, increasing the likelihood of immune activation. For instance, in case of vaccination against cancer the antigens found in B cells have been produced to aidin the eradication of cancerous cells.

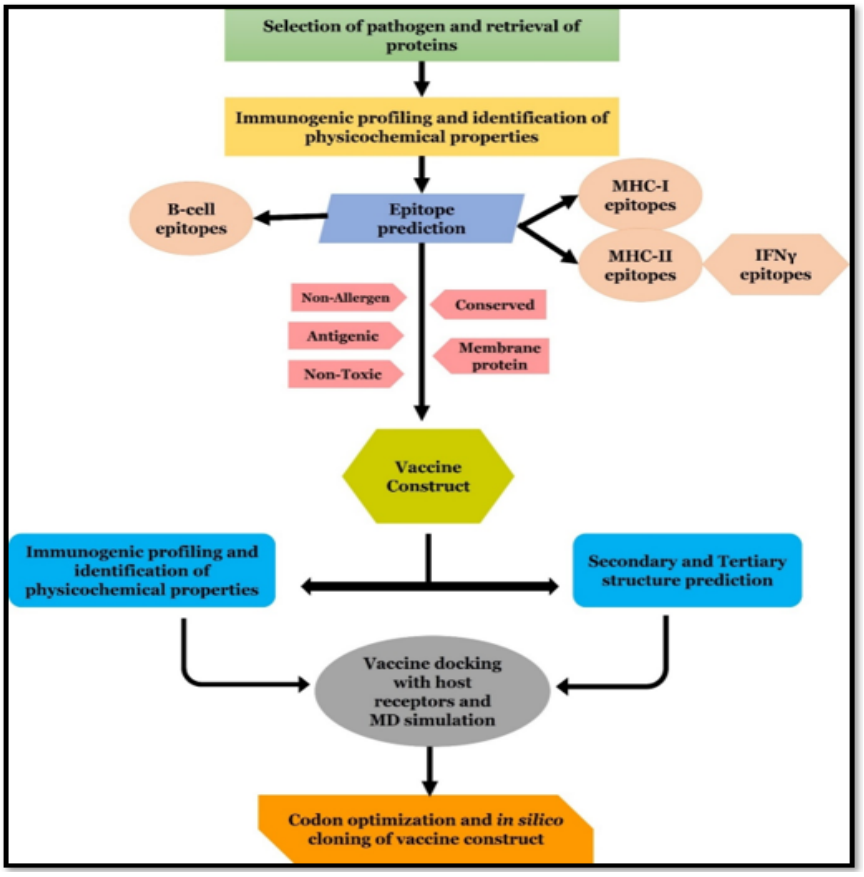

FIGURE 2.3: The workflow representation of the Immunoinformatic study for multiepitope vaccine development.

Moreover, antibodies directed against regulatory T cells have been observed to promote the tumor's shrinkage [122]. The latter paves the door for the identification of epitopes that could be employed in vaccinations, allowing for more effective and rapid disease elimination. Other online servers, such as ToxicPred, and AllergenFP can be used to identify the potential allergenicity and toxicity of the proteins in an allergy vaccine. The workflow followed during the study of immunoinformatic is illustrated in Figure 2.3.

2.12 Structural analyses

Proteins are macromolecules that consist of amino acid chains. A functional domain is a group of amino acid residues that execute a specific biological function. In multi-functional proteins, the size ranges from a few hundred to several thousand residues. Proteins are made up of thousands of amino acids arranged in a particular. The 3D structure is unique to each protein. A protein folds into a complex 3-dimensional shape during formation. Each protein has a specific conformation and folds itself into that specific orientation within a second. There are four levels of protein organization: the primary structure is simply the amino acid sequence of the protein. Each protein has its own distinct amino acid sequence. Secondary structure: the spatial arrangement of a protein's structure in the absence of its side chains. The tertiary structure of a protein is a 3D structure. Quaternary structure is a term that refers to the three-dimensional structures of proteins that are made up of two or more polypeptide chains referred as subunits. The proteins' primary structure is analyzed to determine the amino acid numbers, composition, half-life, molecular weight, polarity, thermal stability, isoelectric point (pI), hydrophobicity, hydrophobicity. For identifying these, Prot Param tool is an online available server that is mostly used by the researchers [123]. If the protein sequence is available, its phylodynamic analysis is easy to conduct after doing multiple sequence alignment (MSA) through different online software like clustalW [124], the alignment of the sequences provides us the data about the similarities and dissimilarities present at the specific sites of the sequences while the

conserved sequences among the multiple sequences can also be determined. After MSA, the phylogenetic tree can also be constructed to identify the phylogeny of the protein. The secondary structure information of the protein can be examined using the protein primary sequence through different softwares: SOPMA and PSIPRED etc. It describes the presence of alpha helices, beta turns, coils, sheets etc.

When it comes to figuring out what the protein does in the body, its three-dimensional structure is of the utmost importance. Several methodologies, including homology modelling, threading or fold recognition, and ab initio structure prediction, have been used in the process of structure prediction. I-TASSER was applied in order to perform structure prediction [125]. Following 3D structure prediction, it is necessary to evaluate the structure using various analyses tools.

Ramachandran plot analysis using PROCHECK is the most favorite analysis for verifying 3D structure. The Ramachandran plot depicts the summary statistics of the backbone dihedral angle pairings and. In theory, the Ramachandran plot's allowed regions indicate the range of feasible Phi/Psi angles for an amino acid, X, in an ala-X-alatripeptide [126]. In practice, the distribution of observed Phi/Psi values in a protein structure can be used to validate the structure. The Ramachandran plot depicts the dihedral angles' energetically permitted and prohibited areas. Numerous dihedral angles are discovered in the Ramachandran plot's banned zones for low-quality homology models. Such variations typically imply structural issues.

2.13 Molecular Docking and Simulation

Molecular docking is an appealing scaffold in designing drugs that interact with DNA and proteins, as well as for research into the mechanisms of pharmacokinetics by immobilizing a ligand molecule in the binding domain of the receptor so that a stable complex is formed that offers the potential for both efficacy and greater specificity [127]. By docking, we can determine the free energy, binding affinity, and stability of complexes. To predict the possible binding parameters in advance,

molecular docking is currently being used. Its major goal is to create a complex with the desired shape and a smaller free binding energy. Other parameters are also identified through it like, electrostatic (ΔG_{elec}), hydrogen bond (ΔG_{hbond}), desolvation energy (ΔG_{desolv}), and torsional free energy (ΔG_{tor}). Since an awareness of professional rules that regulate predicted free binding energy (ΔG_{bind}) and can provide additional clues about the nature of numerous types of interaction that results in molecular docking. It indicates that a thorough understanding of this ethical framework should be relevant to ΔG_{bind} knowledge [128].

2.14 Types of Molecular Docking

Numerous applications utilize docking technologies, all of which use search techniques such as genetic algorithms, Monte Carlo algorithms, fragment-based algorithms, and molecular dynamics algorithms. HADDOCK is among the primary applications for docking with high efficiency. The different docking approaches used during the analysis according to the objectives of the study are: Between the ligand and the target molecule either will be kept rigid or flexible, for example, if the ligand is kept flexible then the target will be treated as rigid. During rigid body interaction, both the molecules are kept rigid whereas during flexible interaction, both the molecules (ligand as well as target) are treated as flexible molecules [129].

Chapter 3

Material and Methods

Complete flow chart of whole procedure of research project is given in Figure 3.1.

FIGURE 3.1: Complete flow chart of whole procedure of research project

3.1 Retrieval of Proteome Analysis

The full proteome of the strain was extracted in FASTA format using UniProt [130]. UniProt is a vast collection of databases that helps the researchers/scientists by providing functional information and sequences of proteins. Over $550,000$ sequences are available in this database. Protein information is very easily accessible for researchers due to organized information, collected through literatures and experiments. The Geptop 2.0 server was used to find essential proteins. Identification off essential genes for prokaryotes are identified through Geptop. Based on orthology and phylogeny, it estimates the essential of any specie. It is available in both online and downloadable application. Human proteins were not being seen as vaccine candidates to prevent causing an autoimmune response. Non-homologous proteins were predicted using BLASTp [131].

PSORTb 3.0.3 was used to choose cellular membranes proteins for cellular localization. To understand the function of protein, genome annotation, detection of possible pharmaceutical targets, vaccine ingredients, or cell surface diagnostic indicators, protein subcellular localization is very important, and PSORTb is used for this purpose. Antigenicity relates to a person's able to respond quickly to an antigen and be immune to it, a protein high antigenicity was used for peptide based vaccine creation. The antigenicity of all proteins from the VaxiJen 2.0 server was also used to test *P. shigellodies* with a 0.5 threshold [132]. Based on physiochemical properties of proteins, VaxiJen allows entire classification of antigens. It is an easily available online server. VaxiJen has a $70 - 89$ percent accuracy when employing the autocross covariance transformation algorithm [132]. Because of their propensity to create difficulties during the purifying process, proteins that include transmembrane helixes were not required for the selection of the target protein.

Transmembrane helices are predicted with 97/98% accuracy with the help of TMHMM. With this server, proteins are evaluated by 99% of sensitivity and specificity, and differentiation between membrane and soluble proteins is done To predict transmembrane helix, the TMHMM v-2.0 server was performed [133].

3.2 CTL Epitope Selection and Evaluation

k Cytotoxic T-cells are important in the selection of specific antigens and the correct outline of CTL epitopes, which are essential to development of vaccine, more specifically, it saves money and time when compared to laboratory experiments. A method known (http://tools.immuneepitope.org/mhci) MHC-I binding were predicted using this method, and 40 MHC-I epitopes were predicted. After submitting a protein sequence in FASTA format, the consensus approach was used to predict CTL epitopes. Human species were chosen an origin, and all cast epitopes were picked. Because of their potential to bind, epitopes with a consensus score of 2 were chosen [134]. The immunogenicity tool was employed by an online database (IEDB MHC-I) to determine immunogenicity of CTL epitopes [134]. The immune epitope database is free, publicly available online database, which has over 1.6 million experiments; which represent the adaptive immune response to epitopes. The results of this method are quite significant due to the large number of HLA alleles included in the computation. All alleles were picked for prediction after the sequence was submitted in FASTA format. The antigenicity of epitopes was calculated using the VaxiJen-v-2.0 server to see if they may activate an immune response [135]. During the development of the vaccine, allergic or toxic reactions were avoided. To do so, Aller-TOP 2.0 an online server was used to assess the allergenicity of epitopes. Non-toxic CTL epitopes were found using the Toxin-Pred server. Toxin-Pred is a unique *in-silco* approach to predict toxicity of a protein. It helps in peptide-based drug discovery by predicting the toxic regions in proteins. It is an online freely available server [136].

3.3 Analysis and Selection of HTL Epitopes

Helper T-cells play a vital role in the immune system's cellular and humoral immunological responses to antigens from the outside world. As a result, HTL epitopes related to MHC-II alleles are critical in the creation of vaccines [137]. T-cells are important for activating B cells to produce antibodies towards effective viruses,

as well as for cells to react to cytotoxic T-cells and infections [138]. On selected proteins, the 27 HTL (Human Leukocyte Antigen) epitopes were predicted by server the IEDB (Immune Epitope Database) MHC-II binding tool5 with identical binding alleles and a 2 percentile threshold [139]. IL-4, IL-10, and FN-gamma help activate cytotoxic T-cells and other immune cells. Vaccine design relies primarily on HTL-producing cytokines. IFN-Gamma epitopes will be analysed using the IFN epitope server, SVM, hybrid motif, and IFN-Gamma vs. Non-IFN-Gamma model [140]. IL-4Pred and IL-10Pred were utilised so that the inducing qualities of IL-4 and IL-10, respectively, could be compared and contrasted [141]. We used SVM cutoffs of -0.3 for IL10Pred and 0.2 for IL4Pred. Responses from helper T lymphocytes (HTLs) improve both humoral and cellular immune responses. Vaccines used for both prophylaxis and treatment were thus likely to have HTL epitopes.

3.4 LBL Epitopes Identification and Evaluation

Because of their role in the immune system for increasing the adaptive immune response, B-cell epitopes are considered important basic components of a vaccine [142]. ABCPred will be used to anticipate B-cell epitopes [143]. For linear epitopes, the minimum predicting value was set at 0.5. For antigenicity, toxicity, and allergenicity profiling, online servers including Toxin Pred, VaxiJen v-2.0, and AllergenFP-v.1.0 will be employed [135, 136, 143]. Epitopes that are preserved and able of unique properties B-cell and T-cell (CD8 and CD4) reactions should be included in a vaccine [144].

3.5 Population Coverage Analysis

MHC (Major histocompatibility complex) there is not one uniform pattern of allelic distribution across the globe because it depends on factors such as ethnic groups and geographic locations. Therefore, when it comes to the creation of an

efficient vaccination, population coverage needs to be taken into consideration as a crucial aspect [145]. The IEDB (Immune Epitope Database) population coverage tool was used to pick MHC-I and MHC-II epitopes along with their accompanying HLA (Human Leukocyte Antigen)-binding alleles in order to calculate population coverage [146]. This tool calculated population coverage of each epitope for different regions of the world based on the distribution of human MHC binding alleles. This software predicts the proportion of persons who will react to a given set of epitopes based on the limits imposed by MHC. Allele and genotype frequencies for HLA genes are used in this calculation, with the latter based on the presumption that these loci are not subject to linkage disequilibrium. By comparing the epitope's sequence to itself at varied degrees of sequence identity, this tool assesses the extent to which the epitope is conserved within a given protein sequence collection. This is a handy tool that may be used either online or off.

3.6 Vaccine's Mapping

To create the multi-epitope vaccination, suitable linkers was mixed with B-cell, T-cell epitopes, and adjuvant [146, 147]. Adjuvant is used with caution to enhance vaccination immunogenicity. If peptides are chosen on their own, they may be insufficiently immunogenic [148]. As an adjuvant, cholera enterotoxin subunit B (Accessionno:P01556) will be used) [149]. The EAAAK linkers were employed to join the CTL epitope with adjuvant and remove the domains of a bi-functional fusion protein [150]. A linker was utilized to tie two epitopes together for effective epitope functioning [151]. GPGPG and AAY-linkers were utilized to combine CTL and HTL-epitopes as well as detect epitopes [150].

3.7 Structural Analysis

Blastp analysis was employed to corroborate the MEBSV sequence, which was created in a non-homologous fashion [152]. Physiochemical properties such as

theoretical pI, MW (molecular weight), instability index (II), GRAVY (Grand Average of Hydropathicity) AI (aliphatic index), in-vivo and in-vitro half-life was analyzed using the ProtParam server. It is a piece of software that can compute a number of different physical and chemical properties for a protein that has been stored in Swiss-Prot or for a protein sequence that the user has provided [152, 153]. The antigenic and immunogenic profiles were assessed using the IEDB immunogenicity tools and the VaxiJen-v 2.0 server. The Aller-TOP server was used to examine the allergic reactions to the vaccination [135]. Utilizing the PSIPRED workbench, we were able to determine the secondary structure of the vaccine. This test was used to estimate other properties such as random-coil, alpha-helices degree of beta-turns, and extended-chain. The Self Optimized Prediction method with alignment, also known as SOPMA, is a software program that can be accessed online and is used to predict the secondary structure of proteins [135].

3.8 Refinement, Confirmation and Prediction of Tertiary Structure

The lower energy structure of a protein is its three-dimensional structure, in which bends and twists exactly to provide maximum stability. I-TASSER server was used to develop our protein, which employed connection information to improve protein structural accuracy. Structure based annotation and structure prediction for protein is done through I-TASSER. After locating suitable PDB structural templates with the help of the LOMETS multithreaded technique, it then employs a template-based fragment assembly simulation to construct complete atomic models [154]. The Galaxy Refine server was used to refine and optimize 3D structured proteins. It uses molecular dynamic simulation to achieve recurrent structure perturbation and overall structural relaxation [155]. General systemic relaxation destroyed the structure in the dynamic simulation. The RAMPAGE server, which uses the Ramachandran plot, was used to validate our improved structure. For the conformation of a protein it shows simple view. Sci, Phi angles of a protein

backbone are plotted in 2d form [156]. The ProSA-webserver was used for structural verifications in later stages. A non-bonded relationship in vaccine structural data was analyzed using the ERRAT server. In this program crystallography is carried out to confirm protein structure. In given structure errors may occur on basics of interaction between non bonded atoms [82, 157].

3.9 B-cell Epitopes Screening

Using an online service like IEDB-AR v.2.22, linear B cell epitopes and conformational were discovered. The vaccine sequence was utilized as an input for the ABCPred server, with a value of 14 and a threshold of 0. All Ellipro tool values were set to standard, and the vaccine structure was given as input. For visualization of discontinuities in the final construction of vaccine, PYMOL-v.1.3,was employed [157].

3.10 MEBSV Stability through Disulfide Engineering

The MEBSV (Multi Epitope-based Subunit Vaccine) refined model's stability should be improved before the next step. Covalent disulfide bonds were used to ensure that the protein construct will properly stabilized, simulating the stability of a molecular interaction by verifying accurate geometric conformations. On the target protein's structure, covalently linkages were constructed by using the disulfide engineering method. With design 2.0, disulfide engineering will be done for an improved vaccine model.

An improved protein model was transferred for residue pair concealing, and the pairs were then used in disulfide engineering. For modification with cysteines residues, the mutated server function were utilized, and three pairs were chosen [158]. It is a tool that allows the users to create new disulfide connections protein.

When the appropriate amino acids are changed to cysteines, proteins structure files in PDB format are evaluated to find residue pairs that are likely to form a disulfide bond. Initially, the MEBSV refined model was uploaded and residue pairs were screened. The pair of residues were chosen for mutation with cysteines residues using the mutated function of the server [159].

3.11 Docking of MHC Molecules and TLR4 Receptor with Constructed Vaccine Disulphide

After the vaccination reacts with the host's immune system, an efficient immune system response is generated. The ability of vaccines to bind to receptors in the human immune system was investigated using molecular docking. The MHC-I, MHC-II and TLR4 receptors were used to study the stimulating and antimicrobial immune responses. Structures of TLR4 (ID: 4G8A), MHC-I receptor (ID: 1I1Y), and MHC-II receptor (ID: 1KG0) were retrieved from protein data bank. For the protein-protein interaction, the HADDOCK-v-2.4 server was used [160]. HADDOCK was performed to build model the bimolecular complexes. High Ambiguity Driven Protein-Protein Docking. That uses mutagenesis data or chemical shift perturbation data related to biochemical or biophysical interaction data.

Docking process is done on basis of Ambiguous Interaction Restraints (AIRs). AIR is an ambiguous distance that is found in all residues that are involved in interaction. The structures are selected on lowest intermolecular energies [161]. PyMOL-v.1.3 was used for visualizing of docked complexes [160]. For three-dimensional (3D) visualization of proteins, small molecules, nucleic acids, and surface, a molecular graphics tool is used called as PyMOL. There are many other plug-in tools available in this python-based software. Ray tracing and molecular editing is also done with this tool [162]. For evaluating interactions between docked complexes, PDB sum, an online server was used [163]. A macromolecule with 3D structure, molecular components, sequence and with many other features is available freely in this database. It is freely available database. The database helps in many ways.

3.12 Molecular Dynamic Simulation

Molecular dynamics plays a very basic role in determining the solidity of proteins and complexes in every *in-silico* study. By compared primary protein movements to traditional nodes, protein stability might be established [164, 165]. The Internal coordinates normal mode analysis service (iMOD was used to describe collective protein mobility in internal coordinates [166].

Even with huge macromolecules, iMODS permits the exploration of such modes and produces feasible transition routes between two homologous structures. This server, which describes the deformability, eigenvalues, covariance, and B-factors, were used to calculate the internal motions of the complexes. The value of standard mode was determined by the stiffness of motion. When the eigenvalue is low, it aids in the deformation of structures that have a certain energy.

3.13 Immune Simulation

An in-silico immunological simulation was performed using the internet server C-Imm Sim-10.1 to evaluate the immune reaction of the simulated vaccine. The thymus lymph nodes, and bone marrow three fundamental components of a working mammalian system, were reconstructed on the C-Imm Sim11 server. In presence of antigens in mammalian immune system, both humoral and cellular responses are determined at cellular scale through C-Imm-sim model. It is an online approach which is easily available for researchers [167].

The MEBSV was tested for its capacity to mimic a variety of immune cells and molecules, including cytotoxic T lymphocytes (CTLs), natural killer (NK) cells, B cytokines, and immunoglobulin. Two vaccine doses should be spaced by at least four weeks in clinical practice. Therefore, the simulation was running for 1000 steps, and two doses were given four weeks apart. Random-seed (12345), number of injection sets (1), volume (10), B-0702, B-0702, and HLA are all inputs used in the immunological simulation (A-0101, A-0101, D-RB1 0101, and DRB1 0101).

3.14 In-silico cloning and Optimization of Codons

Codon usage varies by species, and non - adapted codons might lead to low expression levels in the receptor. As conclusion, it should be created in line with the translational machinery of the host in order to boost expression of genes. This study was use the Java Codon Adaptation Tool (JCAT) to modify codon vaccine to *E.coli* strain K12, a common prokaryotic model [168]. To avoid the ribosome binding site in prokaryotes, restriction enzyme cleavage sites, rho-independent transcription termination and three more options were chosen at the bottom. The stability of CAI and GC (Guanine and Cytosine) will be investigated. At C and N terminal of optimal nucleotides sequences, restriction sites XhoI and NcoI have been introduced. Finally, the modified DNA sequence of the intended vaccination was cloned using Snap Gene 4.3. Complete graphical abstract is shown in Figure 3.2.

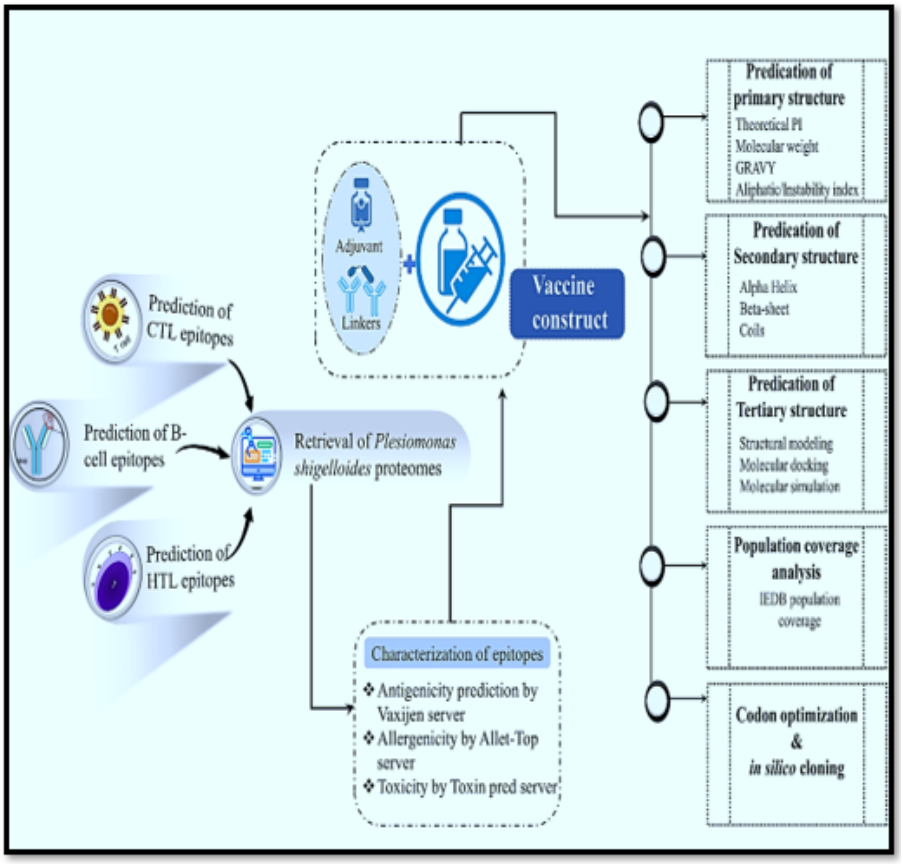

FIGURE 3.2: Complete graphical abstract

Chapter 4

Results and Discussion

4.1 Protein's Selection

Complete proteome of *P. shigelloides* (strain ATCC 13048 / DSM 30053) was retrieved from UniProt (Accession no: UP000008881). It is important to identify essential proteins because such proteins are continuously expressed by the surface cell and perform a very basic role in the environment in cellular crosstalk. Essential proteome has a broader treatment potential because it contains proteins that are required for the pathogen's cellular survival. Essential proteins play an important role in pathogen adhesion, infection, host persistence, growth, and survival.

Table 4.1: Top proteins with maximum antigenicity nature and extracellular locations.

Protein name	**Accession No**	**Antigenicity**	**Helices**	**Location**
Auto transporter domain-containing protein	>trR8ARB8	0.6154	0	Extracellular
Prophage tail fiber protein	>trR8APA5	0.7706	0	Extracellular

Table 4.1 continued from previous page

Protein name	Accession No	Antigenicity	Helices	Location
Tail fiber Protein	>trR8AQH0	0.8350	0	Extracellular
Flagellin	>trR8AS43	0.6989	0	Extracellular
Uncharacterized Protein	>trR8APD9	0.8919	0	Extracellular
Superoxide dismutase	>trR8APR1	0.5574	0	Extracellular

Plesiomonas complete proteome is made up of 3354 proteins. Geptop-0.5, an internet server, was used to determine 396 important proteins from all of these. By excluding human homologs from online BLASTp analysis, a total of 42 non-homologous proteins were discovered. These have been analysed using antigenicity values. Extracellular proteins were selected as the top six proteins with the maximum antigenicity values shows in Table 4.1.

4.2 Epitope CTL Evaluation and Selection

From the target protein of *Plesiomonas Shigelloides*, a total of 92 CTL epitopes consisting of 12 amino acids were generated. After being put through a series of tests to determine their immunogenicity, antigenicity, toxicity, and allergenicity, the top three antigenic, non-allergenic, immunogenic, and non-toxic epitopes were selected for the production of a vaccine. This was done by using a variety of procedures (Table 4.2).

There was a sum of ten basic HTL epitopes created. Using cytokines, the top only one epitope was selected for vaccine development (IL-10, IL-4, and IFN-Gamma) Table 4.3. Accordingly, 50 LBL, epitopes were chosen, with a total of two epitopes

being approved for vaccine manufacturing after being assessed for immunogenicity, allergenicity, and toxicity which were shown in the Table 4.4.

TABLE 4.2: Epitopes were chosen by CTL for vaccine development against *Plesiomonas Shigelloides.*

Epitope	**Protein**	**Alleles**	**Position**	**Anti-genicity**	**Immuno-genicity**
NRQVRGGYTFQP	Auto transporter domain-containing protein	HLA-B* 15:02 HLA-A* 24:02 HLA-A* 23:01	501-512	1.4726	0.1844
GSALGATYRIDY	Uncharacterized protein	HLA-A* 30:02 HLA-A* 01:01	133-144	1.2093	0.28084
YIDYRNVRPNYL	Superoxide dismutase	HLA-C* 06:02 HLA-C* 07:01 HLA-C* 08:02 HLA-C* 07:02 HLA-A* 01:01 HLA-C* 05:01	164-175	1.5493	0.12012

Table 4.3: HTL epitopes used in the development of a vaccination against *Plesiomonas Shigellodies*

Epitope	**Protein**	**Alleles**	**Position**	**Anti-genicity**	**IFY -N**	**IL-4**	**IL-10**
QPNS RYQ VTP FVA LNY	Auto transporter domain-containing protein	HLA-DRB1* 07:03	511-526	0.7055	+ve	Ind-ucer	Non-inducer

Table 4.4: Final selected B-Cell epitopes for vaccine construction against *Plesiomonas Shigelloides*

Epitope	**Protein**	**Score**	**Position**	**Anti-genicity**	**Immunogenicity**
GFGA GEVN GGIS YRIN	Auto transporter domain-containing protein	0.65	719	1.6757	0.36723
KGGT VGGV DYGV GDTL	Prophage tail fiber protein	0.89	433	1.6406	0.36096

4.3 Population Coverage Analysis

Coverage of the population is an essential component in the research and development of vaccines due to the fact that the HLA alleles that are present in different racial and ethnic groups as well as geographic locations of the world are not evenly distributed. In this particular piece of research, the overall population coverage of

specific T cell epitopes together with their associated HLA alleles was computed. It was determined that 95.91% of the world's population is protected by the selected epitopes, according to the overall coverage. The population was covered the best in North America at 91.29 percent, while South America had the lowest coverage at 60.24%. According to the findings of our research, the selected epitopes, which are depicted in Figure 4.1, would make outstanding candidates for vaccines.

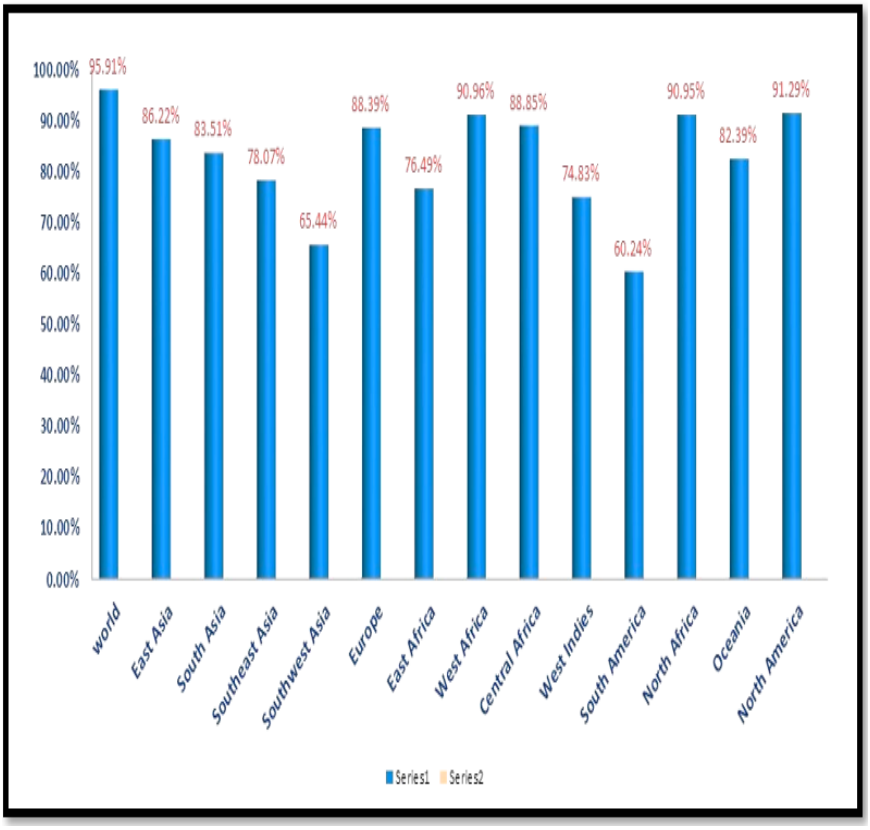

FIGURE 4.1: Population coverage

4.4 Construction of Vaccine

All of the epitopes which were selected were used to construct the vaccine. All LBL, MHC-I, and MHC-II epitopes were attached using KK, AAY and GPGPG Linkers Figure 4.2 . Such linkers were selected because they aid in immunization and epitope distribution and inhibiting junction epitope development [169, 170]. The

N-terminal of the final build vaccine was adjuvant with cholera enterotoxin subunit B (124 residues) in addition to the EAAAK linker. With efficient separation, the EAAAK-linker was also used to increase structural strength and diminish the interaction with the other protein regions are shown in Figure 4.3 [149].

FIGURE 4.2: Primary sequence of the MEBSV

The below figure 4.3 represent the schematic representation of MEBSV.

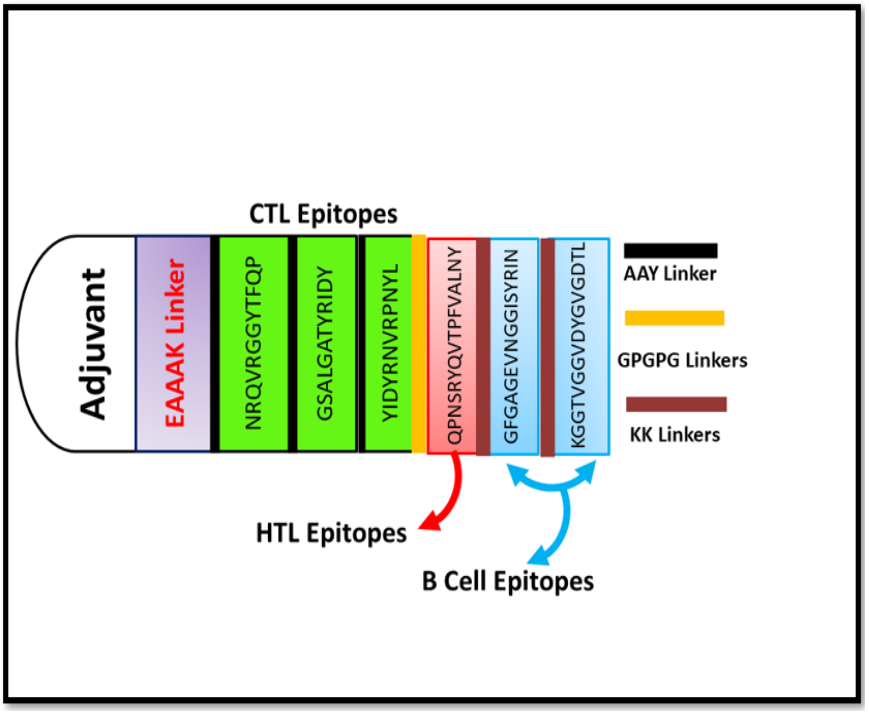

FIGURE 4.3: Schematic representation of the final MEBSV construct vaccine.

4.5 Physiochemical and Immunogenic Profiling

We dig deeper into the physiochemical and immunogenic features of the manufactured vaccine. When the homology of the generated vaccine was matched to the human proteome, the results revealed that no two human proteomes are identical. Afterwards, we conducted tests to determine the vaccine's antigenicity, allergenicity, and toxicity, and found that it is highly non-allergenic, antigenic, and safe. ProtParam was used to analyses physiochemical characteristics. The molecular weight of the vaccine was 30937.91 kDa, whereas the theoretical PI was 5.23 kDa. Our vaccine has a half-life in vitro of 30 hours, in vivo of ¿ 20 hours (Yeast), and in vivo of ¿ 10 hours (E.coli). GRAVY (grand average hydropathicity) was measured at -0.216. All of these characteristics pointed to *Plesiomonas Shigelloides* as a promising vaccine candidate.

4.6 Structural Evaluation

In order to investigate the secondary structure of the vaccine, the PSIPRED and SOPMA servers were utilized. There are 123 residues in the alpha-helix, which makes up 34.55 percent of the sequence. There are 102 residues in extended chains, which make up 28.65 percent of the structure, and there are 103 amino acids in coils, which make up 28.93 percent of the structure.

4.7 Prediction of tertiary Structure, Validation, and Refinement

The tertiary structure of the *Plesiomonas shigelloides* was constructed using I-TASSER, an internet server. In this server, the C-score was determined to be -1.66. To fine-tune the anticipated structure, the Galaxy-Refine server has been used. Galaxy Refine refined the structure. In the modified model study, 85.5% of

amino acids were in favored regions, 33% in tolerated regions and 38% in outlier regions in Figure 4.4. The Z-score was found to be −1.01. In ERRAT's quality check study, the optimized model received a score of 60. Figure 4.4 shows the results, which showed that the improved model was of good quality.

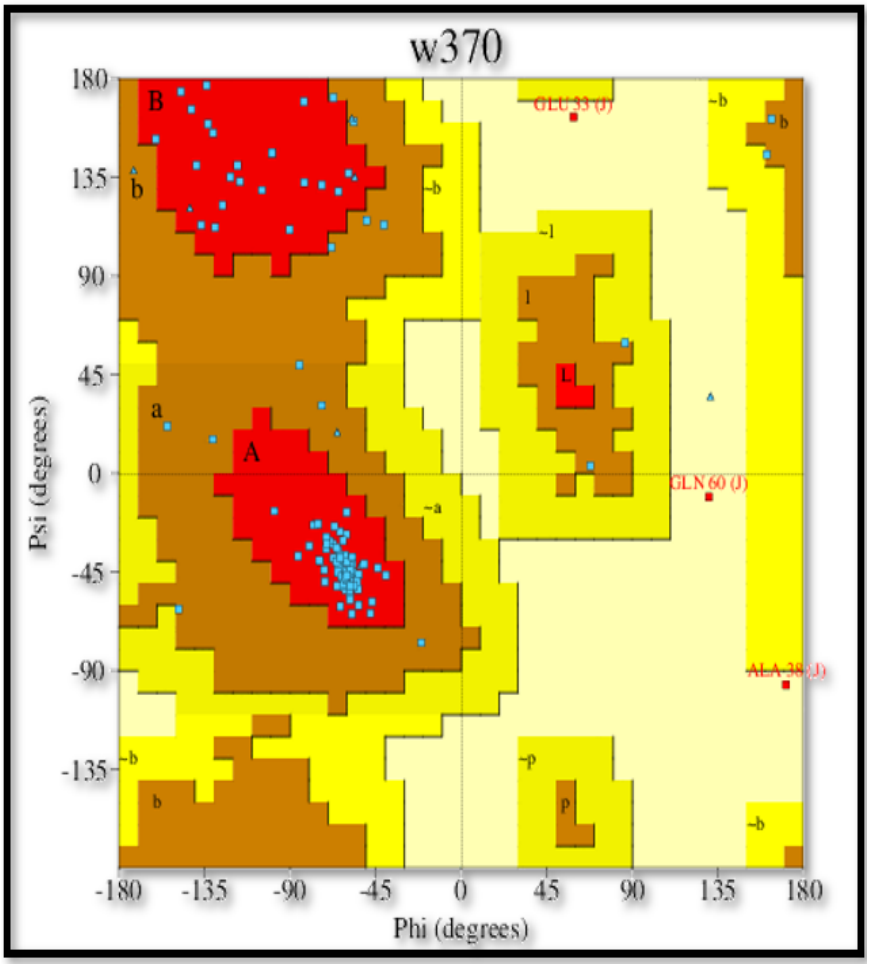

FIGURE 4.4: Schematic representation of the final MEBSV construct vaccine.

4.8 Selection of B-cell Epitopes

B-lymphocytes generate antibodies, which result in humoral immunity [171]. As an outcome, optimal domains of B-cell epitopes should be present in the vaccination. ABCPred 2.0 has been used to predict 3 conformational-discontinuous and 24 linear-continuous from the vaccine construct using default values. The conformation B cell epitopes were visualized in the vaccine construct using PyMOL v.1.3, selected B-cell for screening of vaccine are given in Table 4.5.

TABLE 4.5: Selected B cell for screening of vaccine

Epitope	**Protein**	**Score**	**Position**	**Antigenicity**	**Immuno-genicity**
GFGA GEVN GGIS YRIN	Auto transporter domain-containing protein	0.65	719	1.6757	0.36723
KGGT VGGV DYGV GDTL	Prophage tail fiber protein	0.89	433	1.6406	0.36096

4.9 Disulfide Engineering

The vaccine was improved using Disulfide by Design v-2.0, and disulfide manufacturing was employed to increase the vaccine's stability. For the disulfide synthesis, a total of 33 residues were employed. Three pairs of residues were chosen for disulfide synthesis because their energy and Chi3 scores were within the normal range. As a result, genetic variations in the residue pair emerged, involving PRO222-PRO251 with 1.67 kcal/mol of energy and -80.06 kcal/mol, and PRO222-PRO251 with -80.06 and 1.67 kcal/mol of energy, respectively Figure 4.5.

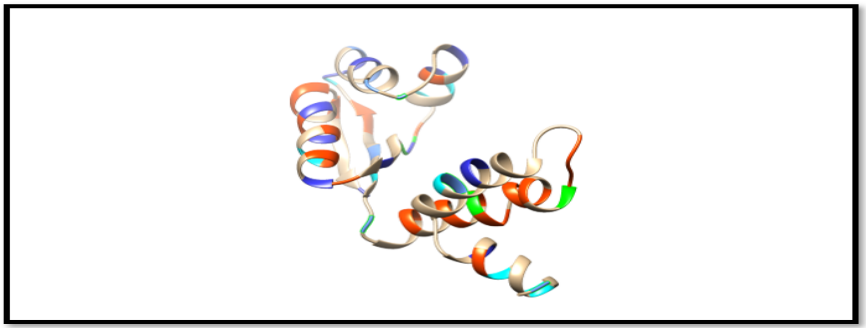

FIGURE 4.5: The green and red colors for two mutated pairs are chosen based on energy X3 and B-factor value.

4.10 Molecular Docking

A strong interaction among antigen and immune specific receptors is required for an active immune response. The docking of vaccination with TLR4, MHC-1, and MHC-II was done using HADDOCK v.2.4. (Immune receptor of human). TLR4 activates an efficient immune response in reaction to the bacterial identification. Vaccines and TLR4 have a substantial interaction, according to docking studies. TLR4 and vaccination had a binding score of 55.2 kcal/mol, which was computed. On the map, TLR4 was depicted as red, while Multi Epitopes Based Subunit Vaccine (MEBSV) was depicted as green (Figure 4.6). It was discovered both TLR4 and vaccination had ten hydrogen bond interaction within a 3.34 range. In addition, HADDOCK V.2.4 was used to dock the vaccine structure with the MHC1 and MHC II receptors. HADDOCK scores of $-23.9 + / - 10.6$ was found, with Desolvaion Energy of $-26.9 + / - 17.8$ (kcal/mol) respectively. Table (4.6) contains the docking results of TLR4.

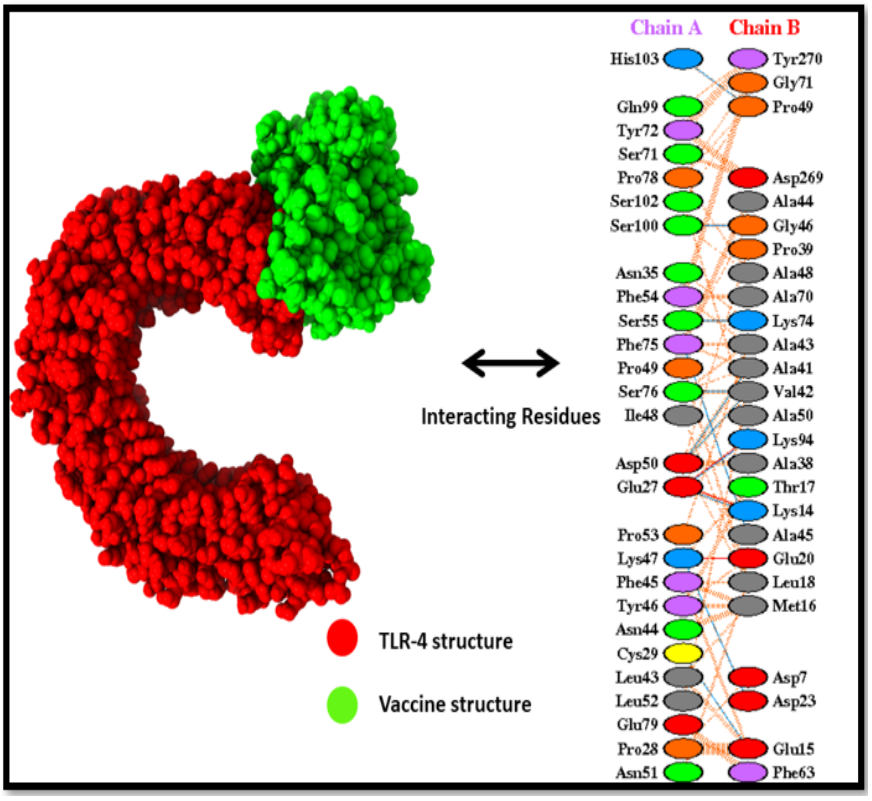

FIGURE 4.6: Docked complex of TLR4 receptor with the vaccine receptor indicating the interacting part along with the interacting residue.

TABLE 4.6: Docking table indicating the docking scores along different energy values of TlR-4 with vaccine construct

TLR4 (Parameters)	**Values**
HADDOCK-v.2.2 score	-23.9 +/-10.6
Cluster Size	38
RMSD from the overall lowest energy structure	4.9 +/ -0.5
Van-der-Waals energy	-84.6 +/ -4.0
Electrostatic Energy	-305.9 +/-49.3
Desolvaion Energy	-26.9 +/ -17.8
Restraints Violation Energy	1488.4 +/ -173.07
Buried Surface Area	2620.7 +/ -57.4
Z-Score	-0.9

In Figure 4.7 the docked complex of MHC-I with vaccine formation, seven hydrogen bond interactions were found with the HADDOCK scores of $-41.6 + / - 25.9$ with Desolvaion Energy of $29.7 + / - 11.7$.

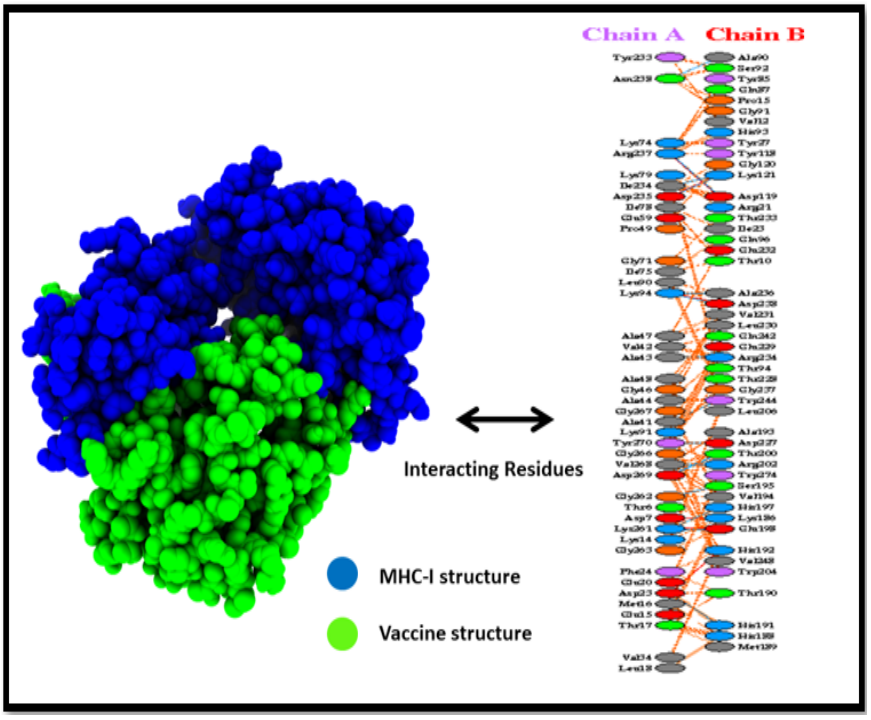

FIGURE 4.7: Intermolecular binding mode and residue level chemical interactions of MEBSV-MHCI complex. The MEBSV is shown green color while MHC I receptor is represented in cyan color.

Table 4.7 contains all results of docked complex of MHC-I. Even though the MHC-II complex has eleven connecting regions, as shown in Figures 4.8. With $-86.9 + / - 6.7$ of HADDOCK scores or Desolvaion Energy is $-73.0 + / - 14.6$ respectively. Table 4.8 shown all results of docked complex of MHC-II.

TABLE 4.7: Docking table indicating the docking scores along different energy values of MHC-I with vaccine construct

MHC1 (Parameters)	**Values**
HADDOCK-v.2.2 score	-41.6 +/ -25.9
Cluster Size	4
RMSD from the overall lowest energy structure	3.3 +/ -0.3
Van-der-Waals energy	-114.0 +/ -7.1
Electrostatic Energy	-467.0 +/ -74.2
Desolvaion Energy	-29.7 +/ -11.7
Restraints Violation Energy	1955.5 +/ -200.96
Buried Surface Area	4125.1 +/ -121.1
Z-Score	0.5

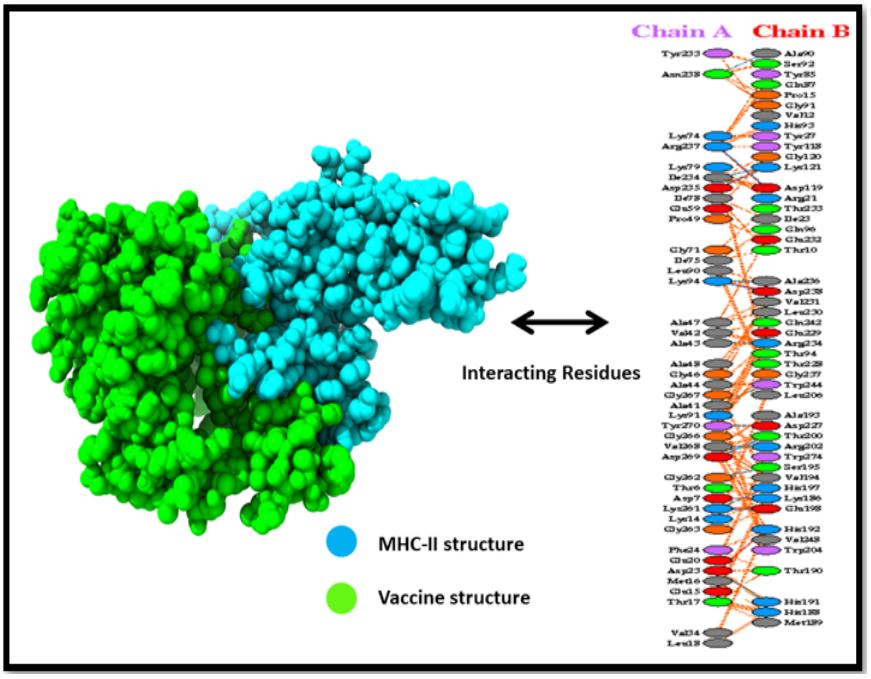

FIGURE 4.8: Intermolecular binding mode and residue level chemical interactions of MEBSV-MHCI complex. The MEBSV is shown in green color while the MHC II receptor is represented in grey color.

TABLE 4.8: Docking table indicating the docking scores along different energy values of MHC-II with vaccine construct

MHC2 (Parameters)	**Values**
HADDOCK-v.2.2 score	-86.9+/-6.7
Cluster Size	30
RMSD from the overall lowest energy structure	9.1+/-0.3
Van-der-Waals energy	-118+/-3.7
Electrostatic Energy	-325.5+/-32.8
Desolvaion Energy	-73.0 +/-14.6
Restraints Violation Energy	1697.7 +/ -217.49
Buried Surface Area	3807.7 +/ -77.8
Z-Score	-1.5

4.11 Molecular Dynamic Simulation

A large-scale normal mode analysis was used to study protein mobility and stability (NMA). I-MODS-server was employed for this analysis, which depends on inner coordinates of docked complexes.

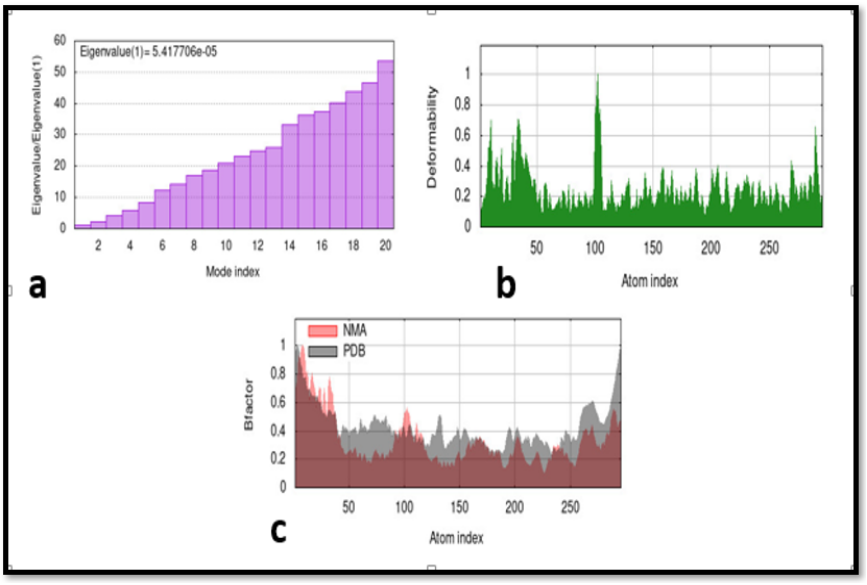

FIGURE 4.9: (a, b, c) Simulation of the vaccine-TLR4 complex, showing (a) eigenvalue, (b) deformability, (c) B-factor

Individual residue deformation influenced the complex's compressibility, as demonstrated by the chain's hinges (Figure 4.9). 2.788520e-05 was found to be the number for the complicated. By reversing the corresponding variance, the frequency of each standard operation was found (Figure 4.9) [172]. As a consequence of the normal mode analysis, the B-factor value was proportionate to the RMS (Figure 4.9). In a covariance matrix, which depicts the pairings of the residue pairs, several various sets of connected, disassociated, or randomly motions are depicted in various colors as reddish, blue, and white. These colors show the connectivity of the motions (Figure 4.10). The elastic map showed joint atoms connected by springs, with each point indicating one spring and grey colors denoting stiffer places, with intensity proportional to stiffness (Figure 4.10)

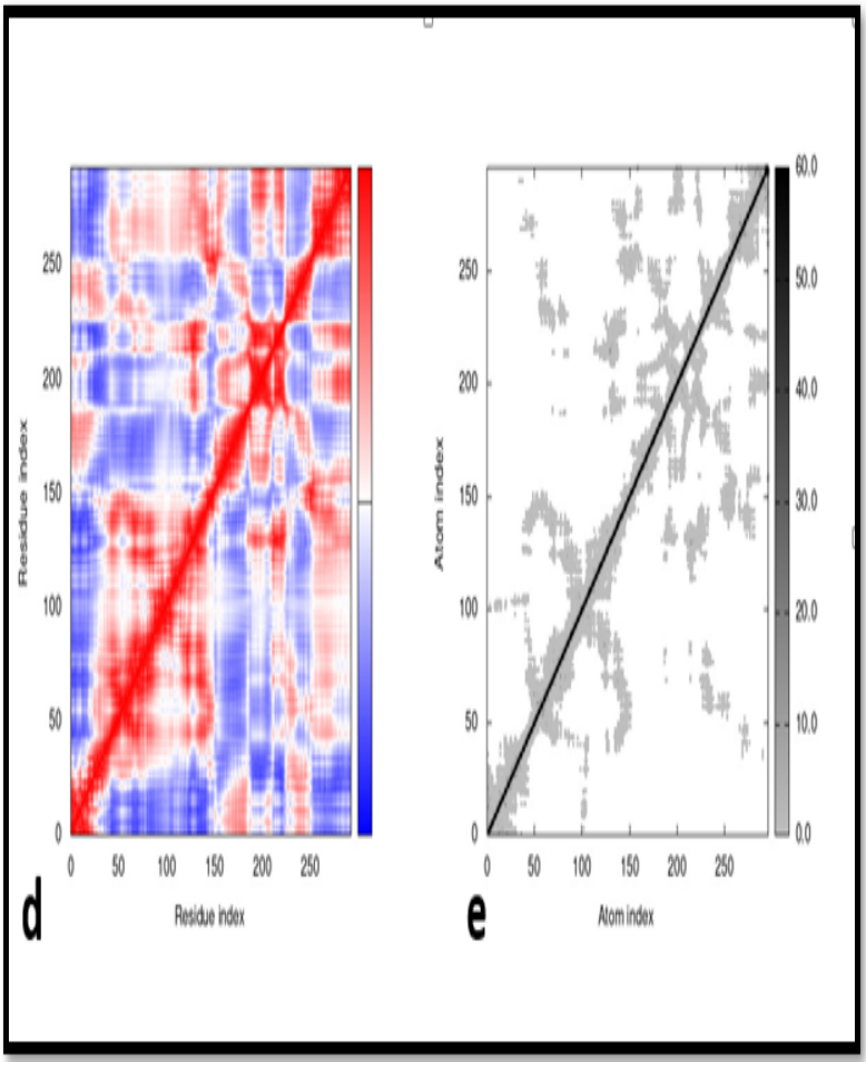

FIGURE 4.10: (d, e) Simulation of the vaccine-TLR4 complex, showing (d) covariance matrix, and (e) elastic network analysis.

4.12 Immune-Simulation

Secondary and primary immune response plays a major role in particular immune responses against with a disease. In a computer simulation, the immune system of the patient reacts to the antigen Figure 4.11. The primary response was generated as high IgG1 + IgG2 and IgM concentrations, followed by IgM, IgG1 + IgG2, and IgG1 in secondary and primary phases with synchronous antigen decrease. The effectiveness of interleukin and cytokine reactions has been discovered. Figure 4.11 shows a successful immune response to the vaccine, and also clearance after consecutive experiences.

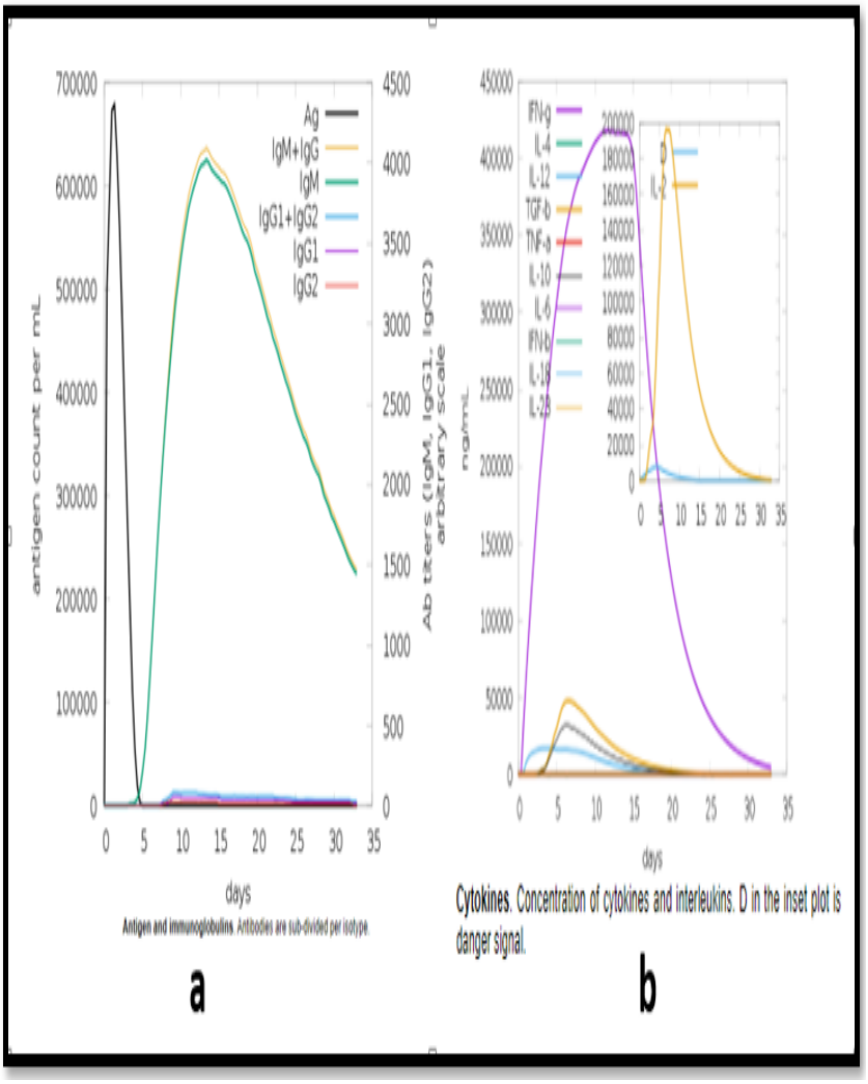

FIGURE 4.11: (a & b) C Imm Sim presentation of immune simulation: (a) Immunoglobulin production in different states and B-cell isotypes according to the Simpson index to the antigen. (b) Cytokine level.

4.13 *In-Silico* Cloning

The vaccine protein was produced successfully in the *E. coli* target host thanks to codon optimization and in-silco cloning. The vaccine's *P. Shigelloides* was modified based on the codon use of the *E.coli* K12 strain. In enhanced DNA, the GC content was 48.87%, and the CAI score was 0.9. As a reasonable compromise, CAI with a ratio of 1.0 was suggested. In the *E. coli* vector pET 30a (+, the synthesised codon was placed between the XhoI and NcoI conserved regions, as illustrated (Figure 4.12). The clone measured 5983bp in size.

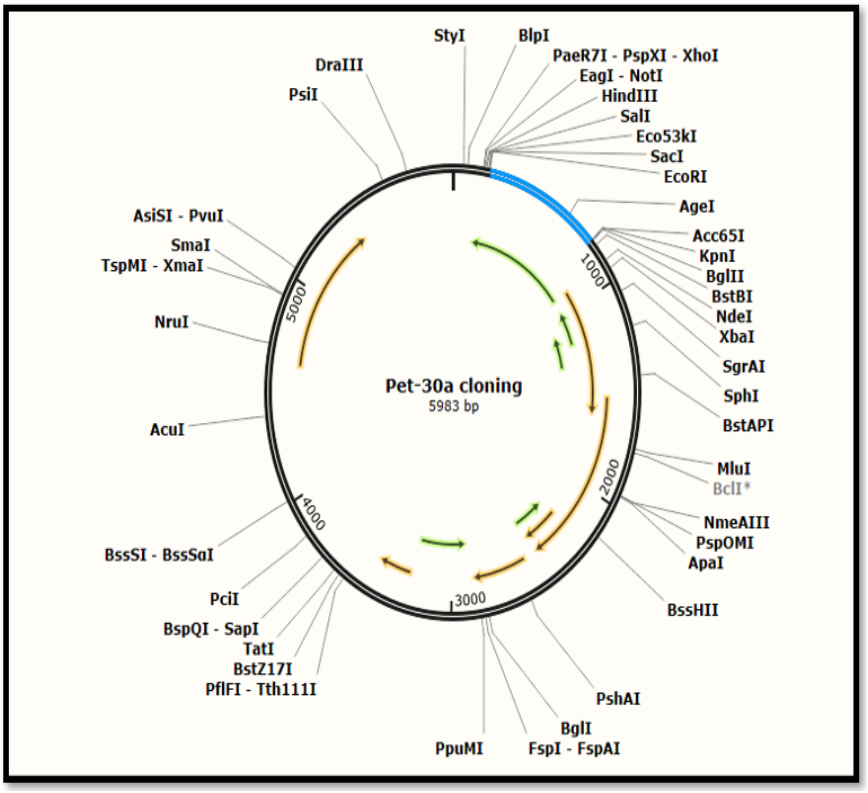

FIGURE 4.12: Cloned sequence of MEBSV (colored as red) into pET-30a (+) expression vector.

Plesiomonas was previously determined as a major element of the *Vibrionaceae* family, but it is now a member of the *Enterobacteriaceae* family. *Plesiomonas shigelloides* is a facultative anaerobic, rod-shaped, Gram-negative, flagellated, bacterium in the genus *Plesiomonas*. Humans, sheep, fish, cows, dogs, pigs and turkey vultures are some of the hosts, and it can also be found in aquatic areas. A serotyping approach based on antigenic variation in the lipopolysaccharide (O-antigen), of

which there are currently 102 flagella and 51H-antigen was used for categorization, epidemiological investigation, and identification of *P. shigelloides* [52, 53].

Diarrheal disorder affected a large number of individuals in worldwide and creates severe medical issues. such as *Enterobacteriaceae*, Gram-negative bacteria, and *Vibrionaceae* are commonly responsible for this disease [54, 55]. *Plesiomonas shigelloides*, a member of the *Enterobacteriaceae* family, has been identified as a major source of gastrointestinal illnesses. It was 3rd in ranked Japan as a cause of diarrhea in tourists [56].

Human faeces, several domestic and wild animals, and the environment, especially water, have all been found to have these germs. Drinking unfiltered water, eating uncooked shellfish, and travelling to underdeveloped countries have all been linked to *P. shigelloides* infections 2 and 3 Different virulence factors, including a lipopolysaccharide, have been reported in an attempt to discover the harmful process of *P. shigelloides* [26, 57, 58]. *Plesiomonas shigelloides* is a facultatively and gramme negative anaerobic bacteria which has been reported in-surface water, freshwater and a range of domestic animals and wild [9, 85, 86]. The creation and production of an effective live and attenuated vaccination is both costly and time-consuming [173]. Aside from that, typical attenuated vaccines' usage is restricted due to a number of difficulties, including their poor ability to elicit immune responses and a variety of side effects. Various methods for creating and generating effective epitope-based vaccines are already available [174]. In order to uncover semi-virulent, non-homologous, antigen, and non-allergenic vaccination candidates, a subtractive proteomics technique was applied to the primary proteome. Another essential need for excluding proteins is the demonstration of the presence of transmembrane helices. To be good targets, the chosen proteins have to be cutting-edge and capable of being detected by the immune system. Six proteins Auto transporter domain-containing protein, Prophage tail fiber protein, Tail fiber protein, Flagellin, Uncharacterized protein and Superoxide dismutase were elected as vaccine targets after extracting extracellular proteins. After the analysis was completed, the antigenicity of those proteins was evaluated using the VaxiJen server. Although Allertop was used to determine the allergenicity of those

proteins. If a long-lasting, substantial immune response is necessary, all B and T cell epitopes must join to establish either humoral or cell-mediated immunity. As a result, T and B cell epitopes from vaccine candidates have been predicted and thoroughly examined. The epitopes chosen accounted for 95.91 percent of the global population coverage. To make the vaccine, HTL, CTL, and B cell epitopes were bound to GPGPG, AAY, and KK linkers, respectively. The output, stabilization, and fold of MEBV can all be improved by incorporating linkers into the evolution process. The EAAAK linker is a stiff linker that has been employed in numerous vaccine development studies, including bacterial and viral illnesses, especially when different epitopes and adjuvants are required. In structural and biochemical studies, the availability of abundantly produced protein within an *E. coli* host is crucial [17]. A MEV protein's solubility was identified, as well as its ease of access to the host. The GRAVY score and the aliphatic index were employed to express thermos ability and hydrophobicity, respectively. To ensure a new vaccine's basic character, the theoretical pI estimate was used. Furthermore, the total stability index derived in this study confirms the protein's lifespan after expression, showing that its utility has grown.

The 3D structure not only helps researchers better understand protein behavior and function as well as protein-protein interactions by providing information about the spatial assembly of essential protein components. *E. coli* expression systems were created to generate recombinant proteins. In *E. coli* K12, codon optimization was used to achieve a high degree of synthetic vaccine protein production. Such vaccines may elicit mucosal immunological responses when administered intranasal or sublingually, inhibiting pathogen entry into the host's body by encouraging the production of host-defensive T and B lymphocytes in the mucosal and systemic regions. As a result, multi-epitope immunizations may become an important tool in the future fight against diseases. Additional laboratory testing is required to demonstrate the vaccine's efficacy and safety, as the current study was based only on an integrated computation process. As there is no vaccine reported against *P. Shigelloides*, this is the first technique that recommends a full functional multi-peptide-based subunit vaccine (MEBSV) that has been evaluated using in *silico*

approaches. MEBSV designed in this study has very important properties that give it an advantage of traditional vaccines. It contains B-cell, CTL, and HTL epitopes, and thus it may be capable of inducing cellular and humoral resistance in the host. It consists of epitopes that target different HLAs and allows the identification of new T cell receptors which have proved effective in a large population. As it contains many immunogenic protein regions, which seem to fuse into a fragment of peptide, a single vaccine can have many targeted proteins. This increases the vaccine's efficacy. Hence, it is an outstanding element for vaccine development against *P. Shigelloides* infection.

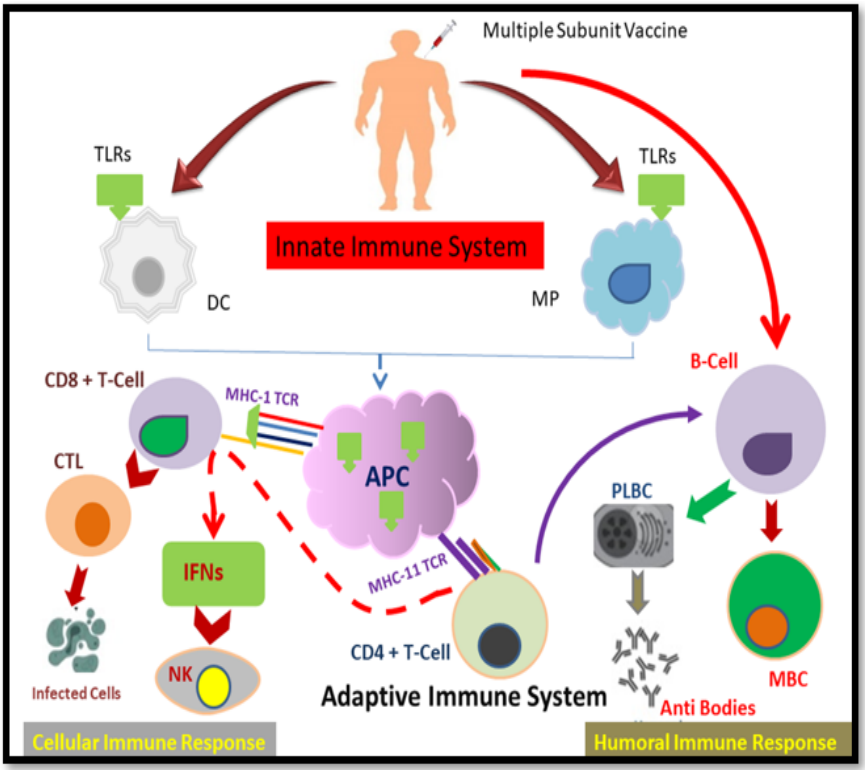

FIGURE 4.13: Adaptive Immune system

Figure 4.13 explain toll like receptors (TLRs) helps vaccine to attach with dendritic cells (DC) and macrophages (MP), to enhance immune response when it enters in the body. Vaccine's epitopes were processed by (APC) antigen presenting cells and presented to T-cell which enhances the immune response. T-cells can either kill infected cells on their own or assist other immune cells in doing so. On the other hand, plasma B-cells (PLBC) manufacture antibodies to deactivate viruses, and memory B-cells (MBCs) retain information in order to activate a robust immune

response in the event of re-infection. Cluster of Differentiation abbreviated as CD. The acronym MHC stands for the major histocompatibility complex. T-cell receptors, or TCR for short. Cytotoxic T cells are abbreviated as CTL. IFN stands for "interferons." NK: natural killer cells.

Chapter 5

Conclusion and Future Aspects

Bacterial infections can cause dangerous illnesses, organ failure, decomposition, and death. Bacterial infections afflict 10 million people worldwide. Foreign bacteria that enter the body and cause diseases and infections kill healthy bacteria and tissues. *P. Shigelloides* causes skin, tissue, respiratory, and urinary tract infections in people. *P. Shigelloides* causes high ICU mortality. *P. Shigelloides* is a rod-shaped, Gram-negative, motile enterbacteriaceae. Misuse of medications and healthcare equipment disseminated this microorganism. This pathogen is drug- and antibiotic-resistant. *P. Shigelloides* needs medications or a vaccination.

This work aimed to design a *P. Shigelloides* multi-epitope vaccine. Subtractive genomics was utilised to uncover bacterial therapeutic proteins lacking from the host. UniProt found 3354 proteins in *P. Shigelloides'* proteome. Geptop 2.0 identified 396 of 3354 proteins as essential. Essential proteins were submitted to BlastP to prevent drug interactions with human proteins. Out of 398 proteins, only 193 were non-homologous to humans. BlastP analysis was carried out to choose best proteins. Only three proteins were found to be extracellular that were selected for further process. Cytotoxic T-cells were predicted through IEDB database as these play essential role in selection for particular antigens for vaccine designing. MHC-I and MHC-II epitopes were also predicted through this database. Integrity of theses epitopes was calculated through different tools like, VaxiJen, Aller-TOP and Toxin-Pred. Selection for HTL epitopes was done through IL4Prd and IL10Pred.

B-cell epitopes play a fundamental role as building blocks of vaccine. ABCPred tool was utilized to predict the B-cell epitopes, and validated through VaxiJen; Aller-TOP and Toxin-Pred.

Population coverage is very important part to construct effective vaccine. Population coverage was calculated through IEDB population coverage tool. Vaccine was constructed by combining of epitopes with adjuvants and linkers. In order to fuse the selected epitopes in a rationally immunogenic fashion suitable adjuvants and linkers were used. Structural analysis was done by using different tools like ProtParam and Blastp. Tertiary structure was predicted by I-TASSER, evaluated and refined through RAMPAGE and Galaxy Refine server. Screening for B-cell epitopes was done through *IEDB-AR*. To stabilize the protein construct, disulfide engineering was done through design 2.0 tool. Binding ability of vaccine was verified by using HADDOCK and visualized through *PyMOL*. Molecular dynamics and integrity was measured by molecular dynamic simulation by using *iMODS* server. Immune response was validated by *in-silico* approach for simulation by using C-Imm Sim server. At the end in-silico cloning was performed through JCAT server to check the expression of vaccine in the host. *P. Shigelloides* infection is considered a major epidemic and a global health issue as well. Medical authorities are failed to control this infection with different medicines as well as vaccines. Different antibiotics were tested, but none of them showed to be effective against this infection. Different techniques related to immunoinformatics as well as subtractive genomics are concentrated to synthesize the vaccine subunit. The proposed vaccine could contribute towards the development of the *P. Shigelloides* infection vaccine in combination with computational analysis and immune data. Our proposed vaccine model will have a positive effect on *P. Shigelloides* infection treatment study. Although, the additional laboratory research is needed to confirm the MEBS efficacy and safety.

Bibliography

[1] J. Janda, "Family I. Enterobacteriaceae, genus XXVII. Plesiomonas Habs and Schubert 1962," *Bergey's Manual of Systematic Bacteriology*, vol. 2, pp. 740-744, 2005.

[2] I. Ciznar, C. González-Rey, K. Krovacek, and A. Hostacka, "Plesiomonas shigelloides," *In Food-Borne Pathogens, ed: Springer*, 2006, pp. 73-80.

[3] J. M. Janda, S. L. Abbott, and C. J. McIver, "Plesiomonas shigelloides revisited," *Clinical Microbiology Reviews*, vol. 29, pp. 349-374, 2016.

[4] J. P. Keating, "Chronic diarrhea," Pediatrics in Review, vol. 26, pp. 5-14, 2005.

[5] W. Ewing, R. Hugh, and J. G. Johnson, "Studies on the aeromonas group," *Studies on the Aeromonas Group*, 1961.

[6] M. Veron, "The taxonomic position of Vibrio and certain comparable bacteria," *Comptes Rendus Hebdomadaires des Seances de L'academie des Sciences. Serie D: Sciences Naturelles*, vol. 261, pp. 5243-5246, 1965.

[7] G. M. Garrity and J. G. Holt, "The road map to the manual," *In Bergey's Manual® of Systematic Bacteriology, ed: Springer*, 2001, pp. 119-166.

[8] J. C. Escobar, D. Bhavnani, G. Trueba, K. Ponce, W. Cevallos, and J. Eisenberg, "Plesiomonas shigelloides infection, Ecuador, 2004–2008," *Emerging infectious diseases*, vol. 18, p. 322, 2012.

[9] R. Brenden, M. A. Miller, and J. M. Janda, "Clinical disease spectrum and pathogenic factors associated with Plesiomonas shigelloides infections in humans," *Clinical Infectious Diseases*, vol. 10, pp. 303-316, 1988.

[10] Å. Ljungh and T. Wadström, "Aeromonas and Plesiomonas as possible causes of diarrhoea," *Infection*, vol. 13, pp. 169-173, 1985.

[11] M. L. Miller and J. A. Koburger, "Plesiomonas shigelloides: an opportunistic food and waterborne pathogen," *Journal of Food Protection*, vol. 48, pp. 449-457, 1985.

[12] D. MCNEELEY, P. IVY, J. C. CRAFT, and I. COHEN, "Plesiomonas: biology of the organism and diseases in children," *The Pediatric Infectious Disease Journal*, vol. 3, pp. 176-181, 1984.

[13] D. Xi, X. Wang, K. Ning, Q. Liu, F. Jing, X. Guo, and B. Cao, "O-antigen gene clusters of Plesiomonas shigelloides serogroups and its application in development of a molecular serotyping scheme," *Frontiers in Microbiology*, vol. 10, p. 741, 2019.

[14] M. S. Edwards, R. W. McLaughlin, J. Li, X. Wan, Y. Liu, H. Xie, Y. Hao, and J. Zheng, "Putative virulence factors of Plesiomonas shigelloides," *Antonie Van Leeuwenhoek*, vol. 112, pp. 1815-1826, 2019.

[15] D. A. Eschenbach, "Tratamiento del aborto séptico espontaneo e inducido," *Obstet Gynecol*, vol. 125, pp. 1042-8, 2015.

[16] S. Ahmed, "Plesiomonas shigelloides periprosthetic knee infection after consumption of raw oysters," *Am J Orthop*, vol. 46, pp. E32-E34, 2017.

[17] M. A. Pence, "The brief case: wound infection with Plesiomonas shigelloides following a freshwater injury," *ed: Am Soc Microbiol*, 2016.

[18] H. Long, B. Xu, Y. Luo, and K. Luo, "Artemisinin protects mice against burn sepsis through inhibiting NLRP3 inflammasome activation," *The American Journal of Emergency Medicine*, vol. 34, pp. 772-777, 2016.

[19] Y. Chen, X. Chen, F. Yu, M. Wu, R. Wang, S. Zheng, D. Han, Q. Yang, H. Kong, and F. Zhou, "Serology, virulence, antimicrobial susceptibility and molecular characteristics of clinical Vibrio parahaemolyticus strains circulating in southeastern China from 2009 to 2013," *Clinical Microbiology and Infection*, vol. 22, pp. 258. e9-258. e16, 2016.

[20] D. Bhavnani, "Diarrheal Disease in Northwestern Ecuador: Prevalence, Pathogenicity, and Transmission of Enteric Pathogens Across the Region," *University of Michigan*, 2012.

[21] R. B. Clark, G. R. Westby, H. Spector, R. R. Soricelli, and C. L. Young, "Fatal Plesiomonas shigelloides septicaemia in a splenectomised patient," *Journal of Infection*, vol. 23, pp. 89-92, 1991.

[22] Y. Titilawo, T. Sibanda, L. Obi, and A. Okoh, "Multiple antibiotic resistance indexing of Escherichia coli to identify high-risk sources of faecal contamination of water," *Environmental Science and Pollution Research*, vol. 22, pp. 10969-10980, 2015.

[23] X.-X. Zhang, T. Zhang, and H. H. Fang, "Antibiotic resistance genes in water environment," *Applied Microbiology and Biotechnology*, vol. 82, pp. 397-414, 2009.

[24] G. A. Khan, B. Berglund, K. M. Khan, P.-E. Lindgren, and J. Fick, "Occurrence and abundance of antibiotics and resistance genes in rivers, canal and near drug formulation facilities–a study in Pakistan," *PLoS One*, vol. 8, p. e62712, 2013.

[25] T. U. Berendonk, C. M. Manaia, C. Merlin, D. Fatta-Kassinos, E. Cytryn, F. Walsh, H. Bürgmann, H. Sørum, M. Norström, and M.-N. Pons, "Tackling antibiotic resistance: the environmental framework," *Nature Reviews Microbiology*, vol. 13, pp. 310-317, 2015.

[26] Y. Okawa, Y. Ohtomo, H. Tsugawa, Y. Matsuda, H. Kobayashi, and T. Tsukamoto, "Isolation and characterization of a cytotoxin produced by Plesiomonas shigelloides P-1 strain," *FEMS Microbiology Letters*, vol. 239, pp. 125-130, 2004.

[27] M. Kaszowska, K. Stojkovic, T. Niedziela, and C. Lugowski, "The O-antigen of Plesiomonas shigelloides serotype O36 containing pseudaminic acid," *Carbohydrate research*, vol. 434, pp. 1-5, 2016.

[28] C. Alexander and E. T. Rietschel, "Invited review: bacterial lipopolysaccharides and innate immunity," *Journal of endotoxin research*, vol. 7, pp. 167-202, 2001.

[29] E. Aquilini, S. Merino, and J. M. Tomßs, "The Plesiomonas shigelloides wbO1 gene cluster and the role of O1-antigen LPS in pathogenicity," *Microbial Pathogenesi*s, vol. 63, pp. 1-7, 2013.

[30] O. Ozdemir, S. Sari, S. Terzioglu, and A. Zenciroglu, "Plesiomonas shigelloides sepsis and meningoencephalitis in a surviving neonate," *Journal of Microbiology, Immunology and Infection*, vol. 43, pp. 344-346, 2010.

[31] Z. Yin, S. Zhang, Y. Wei, M. Wang, S. Ma, S. Yang, J. Wang, C. Yuan, L. Jiang, and Y. Du, "Horizontal Gene Transfer Clarifies Taxonomic Confusion and Promotes the Genetic Diversity and Pathogenicity of Plesiomonas shigelloides," *Msystems*, vol. 5, pp. e00448-20, 2020.

[32] M. J. Giddins, N. Macesic, M. K. Annavajhala, S. Stump, S. Khan, T. H. McConville, M. Mehta, A. Gomez-Simmonds, and A.-C. Uhlemann, "Successive emergence of ceftazidime-avibactam resistance through distinct genomic adaptations in bla KPC-2-harboring Klebsiella pneumoniae sequence type 307 isolates," *Antimicrobial Agents and Chemotherapy*, vol. 62, pp. e02101-17, 2018.

[33] S. Yuan, L. Feng, K. Wang, J. Pang, M. Bosch, C. Lollar, Y. Sun, J. Qin, X. Yang, and P. Zhang, "Stable metal–organic frameworks: design, synthesis, and applications," *Advanced Materials*, vol. 30, p. 1704303, 2018.

[34] N. Martín-María, M. Miret, F. F. Caballero, L. A. Rico-Uribe, A. Steptoe, S. Chatterji, and J. L. Ayuso-Mateos, "The impact of subjective well-being on mortality: a meta-analysis of longitudinal studies in the general population," *Psychosomatic medicine*, vol. 79, pp. 565-575, 2017.

[35] F. An, G. An, Q. An, V. Antonelli, E. Baussan, J. Beacom, L. Bezrukov, S. Blyth, R. Brugnera, and M. B. Avanzini, "Neutrino physics with JUNO," *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, p. 030401, 2016.

[36] E. Papaemmanuil, M. Gerstung, L. Bullinger, V. I. Gaidzik, P. Paschka, N. D. Roberts, N. E. Potter, M. Heuser, F. Thol, and N. Bolli, "Genomic classification and prognosis in acute myeloid leukemia," *New England Journal of Medicine*, vol. 374, pp. 2209-2221, 2016.

[37] A. Khajuria and D. Joshi, "Effects of vibrotactile feedback on postural sway in trans-femoral amputees: A wavelet analysis," *Journal of Biomechanics*, vol. 115, p. 110145, 2021.

[38] S.-S. Kim, S. I. Lee, H.-S. Jin, and S. Park, "Tuberculosis risk is associated with genetic polymorphisms in the LRP2, CUBN, and VDR genes," *Genes & Genomics*, vol. 42, pp. 1189-1196, 2020.

[39] A. S. De Groot, H. Sbai, C. S. Aubin, J. McMurry, and W. Martin, "Immunoinformatics: Mining genomes for vaccine components," *Immunology and Cell Biology*, vol. 80, pp. 255-269, 2002.

[40] D. Barh, S. Tiwari, N. Jain, A. Ali, A. R. Santos, A. N. Misra, V. Azevedo, and A. Kumar, "In silico subtractive genomics for target identification in human bacterial pathogens," *Drug Development Research*, vol. 72, pp. 162-177, 2011.

[41] R. Zhang and Y. Lin, "DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes," Nucleic acids research, vol. 37, pp. D455-D458, 2009.

[42] S. Anishetty, M. Pulimi, and G. Pennathur, "Potential drug targets in Mycobacterium tuberculosis through metabolic pathway analysis," *Computational Biology and Chemistry*, vol. 29, pp. 368-378, 2005.

[43] J. J. May, R. Finking, F. Wiegeshoff, T. T. Weber, N. Bandur, U. Koert, and M. A. Marahiel, "Inhibition of the D-alanine: D-alanyl carrier protein ligase from Bacillus subtilis increases the bacterium's susceptibility to antibiotics that target the cell wall," *The FEBS journal*, vol. 272, pp. 2993-3003, 2005.

[44] A. N. Oli, W. O. Obialor, M. O. Ifeanyichukwu, D. C. Odimegwu, J. N. Okoyeh, G. O. Emechebe, S. A. Adejumo, and G. C. Ibeanu, "Immunoinformatics and vaccine development: an overview," *ImmunoTargets and Therapy*, vol. 9, p. 13, 2020.

[45] A. Kazi, C. Chuah, A. B. A. Majeed, C. H. Leow, B. H. Lim, and C. Y. Leow, "Current progress of immunoinformatics approach harnessed for cellular-and antibody-dependent vaccine design," *Pathogens and Global Health*, vol. 112, pp. 123-131, 2018.

[46] E. Raoufi, M. Hemmati, S. Eftekhari, K. Khaksaran, Z. Mahmodi, M. M. Farajollahi, and M. Mohsenzadegan, "Epitope prediction by novel immunoinformatics approach: a state-of-the-art review," *International Journal of Peptide Research and Therapeutics*, vol. 26, pp. 1155-1163, 2020.

[47] S. I. Doron, P. L. Hibberd, and S. L. Gorbach, "Probiotics for prevention of antibiotic-associated diarrhea," *Journal of Clinical Gastroenterology*, vol. 42, pp. S58-S63, 2008.

[48] M. L. Mezzatesta, F. Gona, and S. Stefani, "Enterobacter cloacae complex: clinical impact and emerging antibiotic resistance," *Future Microbiology*, vol. 7, pp. 887-902, 2012.

[49] L. Galluzzi, B.-S. Pedro, I. Vitale, S. Aaronson, J. Abrams, D. Adam, E. Alnemri, L. Altucci, D. Andrews, and M. Annicchiarico-Petruzzelli, "Essential versus accessory aspects of cell death: recommendations of the NCCD 2015," *Cell Death & Differentiation*, vol. 22, pp. 58-73, 2015.

[50] A. Davin-Regli and J.-M. Pagès, "Enterobacter aerogenes and Enterobacter cloacae; versatile bacterial pathogens confronting antibiotic treatment," *Frontiers in Microbiology*, vol. 6, p. 392, 2015.

[51] Q. Cao, D. Yu, M. Georgescu, Z. Han, and J. Wu, "Impacts of land use and land cover change on regional climate: a case study in the agro-pastoral transitional zone of China," *Environmental Research Letters*, vol. 10, p. 124025, 2015.

[52] R. Schubert, "Genus IV. Plesiomonas," *Bergey's Manual of Systematic Bacteriology ed. Krieg, NR and Holt, JG*, pp. 548-550, 1984.

[53] I. Huber, B. Spanggaard, K. Appel, L. Rossen, T. Nielsen, and L. Gram, "Phylogenetic analysis and in situ identification of the intestinal microbial community of rainbow trout (Oncorhynchus mykiss, Walbaum)," *Journal of Applied Microbiology*, vol. 96, pp. 117-132, 2004.

[54] G. Garrity, J. Bell, and T. Lilburn, *Bergey's Taxonomic Outline: Springer*, 2003.

[55] I. Stock, "Plesiomonas shigelloides: an emerging pathogen with unusual properties," *Reviews in Medical Microbiology*, vol. 15, pp. 129-139, 2004.

[56] C. González-Rey, S. B. Svenson, L. Bravo, A. Siitonen, V. Pasquale, S. Dumontet, I. Ciznar, and K. Krovacek, "Serotypes and anti-microbial susceptibility of Plesiomonas shigelloides isolates from humans, animals and aquatic environments in different countries," *Comparative immunology, microbiology and infectious diseases*, vol. 27, pp. 129-139, 2004.

[57] W. Gu and R. E. Levin, "Quantification of viable Plesiomonas shigelloides in a mixture of viable and dead cells using ethidium bromide monoazide and conventional PCR," *Food Biotechnology*, vol. 21, pp. 145-159, 2007.

[58] W. Gu and R. E. Levin, "Innovative methods for removal of PCR inhibitors for quantitative detection of Plesiomonas shigelloides in oysters by real-time PCR," *Food Biotechnology*, vol. 22, pp. 98-113, 2008.

[59] O. Takeuchi and S. Akira, "Pattern recognition receptors and inflammation," *Cell*, vol. 140, pp. 805-820, 2010.

[60] J. Y. Khalil, T. Langlois, J. Andreani, J.-M. Sorraing, D. Raoult, L. Camoin, and B. La Scola, "Flow cytometry sorting to separate viable giant viruses from amoeba co-culture supernatants," *Frontiers in Cellular and Infection Microbiology*, vol. 6, p. 202, 2017.

[61] D. Artner, A. Oblak, S. Ittig, J. A. Garate, S. Horvat, C. c. Arrieumerlou, A. Hofinger, C. Oostenbrink, R. Jerala, and P. Kosma, "Conformationally constrained lipid A mimetics for exploration of structural basis of TLR4/MD-2 activation by lipopolysaccharide," *ACS Chemical Biology*, vol. 8, pp. 2423-2432, 2013.

[62] K. Honda, A. Takaoka, and T. Taniguchi, "Type I inteferon gene induction by the interferon regulatory factor family of transcription factors," *Immunity*, vol. 25, pp. 349-360, 2006.

[63] G. C. Dicenzo and T. M. Finan, "The divided bacterial genome: structure, function, and evolution," *Microbiology and Molecular Biology Reviews*, vol. 81, pp. e00019-17, 2017.

[64] J. Lukasiewicz, T. Niedziela, W. Jachymek, L. Kenne, and C. Lugowski, "Structure of the lipid A–inner core region and biological activity of Plesiomonas shigelloides O54 (strain CNCTC 113/92) lipopolysaccharide," *Glycobiology*, vol. 16, pp. 538-550, 2006.

[65] M. Kaszowska, W. Jachymek, J. Lukasiewicz, T. Niedziela, L. Kenne, and C. Lugowski, "The unique structure of complete lipopolysaccharide isolated from semi-rough Plesiomonas shigelloides O37 (strain CNCTC 39/89) containing (2S)-O-(4-oxopentanoic acid)-α-D-Glcp (α-D-Lenose)," *Carbohydrate research*, vol. 378, pp. 98-107, 2013.

[66] J. Lukasiewicz, M. Dzieciatkowska, T. Niedziela, W. Jachymek, A. Augustyniuk, L. Kenne, and C. Lugowski, "Complete lipopolysaccharide of Plesiomonas shigelloides O74: H5 (strain CNCTC 144/92). 2. Lipid A, its structural variability, the linkage to the core oligosaccharide, and the biological

activity of the lipopolysaccharide," *Biochemistry*, vol. 45, pp. 10434-10447, 2006.

[67] W. Ferguson and N. Henderson, "Description of strain C27: a motile organism with the major antigen of Shigella sonnei phase I," *Journal of Bacteriology*, vol. 54, pp. 179-181, 1947.

[68] J. Kramer, R. Gilbert, and M. Doyle, "Foodborne bacterial pathogens," *By MP Doyle, Marcel Dekker, Inc., New York and Basel*, pp. 22-70, 1989.

[69] M. P. Doyle, F. Diez-Gonzalez, and C. Hill, *Food microbiology: fundamentals and frontiers: John Wiley & Sons*, 2020.

[70] L. Vezzulli, R. R. Colwell, and C. Pruzzo, "Ocean warming and spread of pathogenic vibrios in the aquatic environment," *Microbial ecology*, vol. 65, pp. 817-825, 2013.

[71] G. J. Gutiérrez-Salazar, Z. J. Molina-Garza, M. Hernández-Acosta, J. A. García-Salas, R. Mercado-Hernández, and L. Galaviz-Silva, "Pathogens in Pacific white shrimp (Litopenaeus vannamei Boone, 1931) and their relationship with physicochemical parameters in three different culture systems in Tamaulipas, Mexico," *Aquaculture*, vol. 321, pp. 34-40, 2011.

[72] E. S. Lindström and S. Langenheder, "Local and regional factors influencing bacterial community assembly," *Environmental Microbiology Reports*, vol. 4, pp. 1-9, 2012.

[73] A. C. Yannarell and E. W. Triplett, "Geographic and environmental sources of variation in lake bacterial community composition," *Applied and Environmental Microbiology*, vol. 71, pp. 227-239, 2005.

[74] R. Shope, "Global climate change and infectious diseases," *Environmental Health Perspectives*, vol. 96, pp. 171-174, 1991.

[75] X. Wu, H. Tian, S. Zhou, L. Chen, and B. Xu, "Impact of global change on transmission of human infectious diseases," *Science China Earth Sciences*, vol. 57, pp. 189-203, 2014.

[76] B. Behera, A. Bera, P. Paria, A. Das, P. Parida, S. Kumari, S. Bhowmick, and B. Das, "Identification and pathogenicity of Plesiomonas shigelloides in Silver Carp," *Aquaculture*, vol. 493, pp. 314-318, 2018.

[77] J. K. Bowman, X. C. Zhang, and J. B. Hack, "Plesiomonas shigelloides meningitis in an adult in the ED," *The American Journal of Emergency Medicine*, vol. 34, pp. 1329. e1-1329. e2, 2016.

[78] R. Ruimy, V. Breittmayer, P. Elbaze, B. Lafay, O. Boussemart, M. Gauthier, and R. Christen, "Phylogenetic analysis and assessment of the genera Vibrio, Photobacterium, Aeromonas, and Plesiomonas deduced from small-subunit rRNA sequences," *International Journal of Systematic and Evolutionary Microbiology*, vol. 44, pp. 416-426, 1994.

[79] M. Shigematsu, M. Kaufmann, A. Charlett, Y. Niho, and T. Pitt, "An epidemiological study of Plesiomonas shigelloides diarrhoea among Japanese travellers," *Epidemiology & Infection*, vol. 125, pp. 523-530, 2000.

[80] C. Gyles and P. Boerlin, "Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease," *Veterinary pathology*, vol. 51, pp. 328-340, 2014.

[81] T. C. Ekundayo and A. I. Okoh, "Pathogenomics of virulence traits of Plesiomonas shigelloides that were deemed inconclusive by traditional experimental approaches," *Frontiers in Microbiology*, vol. 9, p. 3077, 2018.

[82] H. Abdelhamed, O. Ozdemir, H. C. Tekedar, M. A. Arick, C.-Y. Hsu, A. Karsi, and M. L. Lawrence, "Complete genome sequence of multidrug-resistant Plesiomonas shigelloides strain MS-17-188," *Genome Announcements*, vol. 6, pp. e00387-18, 2018.

[83] D. Che, M. S. Hasan, and B. Chen, "Identifying pathogenicity islands in bacterial pathogenomics using computational approaches," *Pathogens*, vol. 3, pp. 36-56, 2014.

[84] T. K. McDaniel and J. B. Kaper, "A cloned pathogenicity island from enteropathogenic Escherichia coli confers the attaching and effacing phenotype on E. coli K-12," *Molecular Microbiology*, vol. 23, pp. 399-407, 1997.

[85] B. Mandal, K. Whale, and B. Morson, "Acute colitis due to Plesiomonas shigelloides," *British medical journal (Clinical Research ed.)*, vol. 285, p. 1539, 1982.

[86] T. Wong, H. Tsui, M. So, J. Lai, S. Lai, C. Tse, and T. Ng, "Plesiomonas shigelloides infection in Hong Kong: retrospective study of 167 laboratory-confirmed cases," *Hong Kong Med J*, vol. 6, pp. 375-80, 2000.

[87] C. R. Raetz and C. Whitfield, "Lipopolysaccharide endotoxins," *Annual Review of Biochemistry*, vol. 71, pp. 635-700, 2002.

[88] C. L. Gyles, *Escherichia coli in domestic animals and humans: CAB international*, 1994.

[89] D. Anderson, O. Dixon, and E. Lizzio, "Immunization and culture of rainbow trout organ sections in vitro," *Veterinary Immunology and Immunopathology*, vol. 12, pp. 203-211, 1986.

[90] F. Fenner, "The effects of changing social organisation on the infectious diseases of man," *The impact of Civilization on the Biology of Man*, pp. 48-76, 1970.

[91] R. Lan and P. R. Reeves, "Recombination between rRNA operons created most of the ribotype variation observed in the seventh pandemic clone of Vibrio cholerae," *Microbiology*, vol. 144, pp. 1213-1221, 1998.

[92] G. M. Garrity, M. Winters, and D. B. Searles, "Taxonomic outline of the procaryotic genera," *Bergey's Manual of Systematic Bacteriology*, vol. 2, 2001.

[93] M. Kaszowska, W. Jachymek, T. Niedziela, S. Koj, L. Kenne, and C. Lugowski, "The novel structure of the core oligosaccharide backbone of the lipopolysaccharide from the Plesiomonas shigelloides strain CNCTC 80/89 (serotype O13)," *Carbohydrate Research*, vol. 380, pp. 45-50, 2013.

[94] D. Henderson, E. Wyckoff, C. Rashidi, H. Verlei, and A. Oldham, "Characterization of the Plesiomonas shigelloides genes encoding the heme iron utilization system," *Journal of bacteriology*, vol. 183, pp. 2715-2723, 2001.

[95] F.-Q. Xia, P.-N. Liu, and Y.-H. Zhou, "Meningoencephalitis caused by Plesiomonas shigelloides in a Chinese neonate: case report and literature review," *Italian Journal of Pediatrics*, vol. 41, pp. 1-5, 2015.

[96] M. N. Rager, M. R. Binet, G. Ionescu, and O. M. Bouvet, "31P-NMR and 13C-NMR studies of mannose metabolism in Plesiomonas shigelloides: Toxic effect of mannose on growth," *European Journal of Biochemistry*, vol. 267, pp. 5136-5141, 2000.

[97] F. C. Herrera, J. A. Santos, A. Otero, and M.-L. García-López, "Occurrence of Plesiomonas shigelloides in displayed portions of saltwater fish determined by a PCR assay based on the hugA gene," *International Journal of Food Microbiology*, vol. 108, pp. 233-238, 2006.

[98] J. Yan, Y. Li, X. Guo, X. Wang, F. Liu, A. Li, and B. Cao, "The effect of ArcA on the growth, motility, biofilm formation, and virulence of Plesiomonas shigelloides," *BMC Microbiology*, vol. 21, pp. 1-16, 2021.

[99] T. C. Ekundayo and A. I. Okoh, "A global bibliometric analysis of Plesiomonas-related research (1990–2017)," *PLoS One*, vol. 13, p. e0207655, 2018.

[100] T. Wadström and Å. Ljungh, "Aeromonas and Plesiomonas as food-and waterborne pathogens," I*nternational Journal of Food Microbiology*, vol. 12, pp. 303-311, 1991.

[101] X. Chen, Y. Chen, Q. Yang, H. Kong, F. Yu, D. Han, S. Zheng, D. Cui, and L. Li, "Plesiomonas shigelloides infection in Southeast China," *PLoS One*, vol. 8, p. e77877, 2013.

[102] O. Ellegaard and J. A. Wallin, "The bibliometric analysis of scholarly production: How great is the impact," *Scientometrics*, vol. 105, pp. 1809-1831, 2015.

[103] F. Geaney, C. Scutaru, C. Kelly, R. W. Glynn, and I. J. Perry, "Type 2 diabetes research yield, 1951-2012: bibliometrics analysis and density-equalizing mapping," *PLoS One*, vol. 10, p. e0133009, 2015.

[104] E. Korzeniewska, A. Korzeniewska, and M. Harnisz, "Antibiotic resistant Escherichia coli in hospital and municipal sewage and their emission to the environment," *Ecotoxicology and environmental safety*, vol. 91, pp. 96-102, 2013.

[105] Y.-N. Jiao, H. Chen, R.-X. Gao, Y.-G. Zhu, and C. Rensing, "Organic compounds stimulate horizontal transfer of antibiotic resistance genes in mixed wastewater treatment systems," *Chemosphere*, vol. 184, pp. 53-61, 2017.

[106] J. D. Eason and D. Peacock, "Plesiomonas shigelloides septicemia and meningitis in a neonate," *Canadian Journal of Infectious Diseases*, vol. 7, pp. 380-382, 1996.

[107] G. Prats, B. Mirelis, T. Llovet, C. Muñoz, E. Miró, and F. Navarro, "Antibiotic resistance trends in enteropathogenic bacteria isolated in 1985–1987 and 1995–1998 in Barcelona," *Antimicrobial Agents and Chemotherapy*, vol. 44, pp. 1140-1145, 2000.

[108] J. Pitout, K. Thomson, N. Hanson, A. Ehrhardt, E. Moland, and C. Sanders, "β-Lactamases responsible for resistance to expanded-spectrum cephalosporins in Klebsiella pneumoniae, Escherichia coli, and Proteus mirabilis isolates recovered in South Africa," *Antimicrobial Agents and Chemotherapy*, vol. 42, pp. 1350-1354, 1998.

[109] J.-P. Pignon, H. Tribodet, G. V. Scagliotti, J.-Y. Douillard, F. A. Shepherd, R. J. Stephens, A. Dunant, V. Torri, R. Rosell, and L. Seymour, "Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative

Group," *In Database of Abstracts of Reviews of Effects (DARE): Quality-Assessed Reviews [Internet], ed: Centre for Reviews and Dissemination (UK)*, 2008.

[110] A. Thiolas, C. Bollet, B. La Scola, D. Raoult, and J.-M. Pagès, "Successive emergence of Enterobacter aerogenes strains resistant to imipenem and colistin in a patient," *Antimicrobial Agents and Chemotherapy*, vol. 49, pp. 1354-1358, 2005.

[111] D. Rawat and D. Nair, "Extended-spectrum β-lactamases in Gram Negative Bacteria," *Journal of Global Infectious Diseases*, vol. 2, p. 263, 2010.

[112] B. Suay-García and M. T. Pérez-Gracia, "Present and future of carbapenem-resistant Enterobacteriaceae (CRE) infections," *Antibiotics*, vol. 8, p. 122, 2019.

[113] L. Tzouvelekis, A. Markogiannakis, M. Psichogiou, P. Tassios, and G. Daikos, "Carbapenemases in Klebsiella pneumoniae and other Enterobacteriaceae: an evolving crisis of global dimensions," *Clinical Microbiology Reviews*, vol. 25, pp. 682-707, 2012.

[114] A. Cancelli, C. Cottone, A. Giordani, G. Asta, D. Lupoi, V. Pizzella, and F. Tecchio, "MRI-guided regional personalized electrical stimulation in multisession and home treatments," *Frontiers in Neuroscience*, vol. 12, p. 284, 2018.

[115] U. Arumugam, N. Stalin, and G. P. Rebecca, "Isolation, molecular identification and antibiotic resistance of Enterococcus faecalis from diseased Tilapia," *Int. J. Curr. Microbiol. Appl. Sci*, vol. 6, pp. 136-146, 2017.

[116] G. Subramanian, U. Vetrivel, and M. I. Mohamedyousuff, "Deciphering novel potential antibacterial targets in tomato pathogen Ralstonia solanacearum GMI1000 through integration of in silico subtractive genomics, codon usage and protein–protein interaction analyses," *Australasian Plant Pathology*, pp. 1-11, 2022.

[117] M. A. Seyed, I. Jantan, S. N. A. Bukhari, and K. Vijayaraghavan, "A comprehensive review on the chemotherapeutic potential of piceatannol for cancer treatment, with mechanistic insights," *Journal of Agricultural and Food Chemistry*, vol. 64, pp. 725-737, 2016.

[118] L. Backert and O. Kohlbacher, "Immunoinformatics and epitope prediction in the age of genomic medicine," *Genome Medicine*, vol. 7, pp. 1-12, 2015.

[119] G. B. Tomar, R. K. Srivastava, N. Gupta, A. P. Barhanpurkar, S. T. Pote, H. M. Jhaveri, G. C. Mishra, and M. R. Wani, "Human gingiva-derived mesenchymal stem cells are superior to bone marrow-derived mesenchymal stem cells for cell therapy in regenerative medicine," *Biochemical and Biophysical Research Communications*, vol. 393, pp. 377-383, 2010.

[120] C. F. Dormann, J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, and W. Daniel Kissling, "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review," *Ecography*, vol. 30, pp. 609-628, 2007.

[121] J. Ou, F. Carbonero, E. G. Zoetendal, J. P. DeLany, M. Wang, K. Newton, H. R. Gaskins, and S. J. O'Keefe, "Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans," *The American Journal of Clinical Nutrition*, vol. 98, pp. 111-120, 2013.

[122] I. Delany, R. Rappuoli, and E. De Gregorio, "Vaccines for the 21st century," *EMBO Molecular Medicine*, vol. 6, pp. 708-720, 2014.

[123] E. Gasteiger, C. Hoogland, A. Gattiker, M. R. Wilkins, R. D. Appel, and A. Bairoch, "Protein identification and analysis tools on the ExPASy server," *The Proteomics Protocols Handbook*, pp. 571-607, 2005.

[124] T. Barrett and R. Edgar, "Gene Expression Omnibus: microarray data storage, submission, retrieval, and analysis," *Methods in Enzymology*, vol. 411, pp. 352-369, 2006.

[125] S. C. Biddie, S. John, P. J. Sabo, R. E. Thurman, T. A. Johnson, R. L. Schiltz, T. B. Miranda, M.-H. Sung, S. Trump, and S. L. Lightman, "Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding," *Molecular Cell*, vol. 43, pp. 145-155, 2011.

[126] G. Ramachandran, "Protein structure and crystallography," *Science*, vol. 141, pp. 288-291, 1963.

[127] I. A. Guedes, C. S. de Magalhães, and L. E. Dardenne, "Receptor–ligand molecular docking," *Biophysical Reviews*, vol. 6, pp. 75-87, 2014.

[128] V. Agarwal, G. W. Bell, J.-W. Nam, and D. P. Bartel, "Predicting effective microRNA target sites in mammalian mRNAs," *Elife*, vol. 4, p. e05005, 2015.

[129] M. L. Lamb and W. L. Jorgensen, "Computational approaches to molecular recognition," *Current Opinion in Chemical Biology*, vol. 1, pp. 449-457, 1997.

[130] P. Kalita, D. L. Lyngdoh, A. K. Padhi, H. Shukla, and T. Tripathi, "Development of multi-epitope driven subunit vaccine against Fasciola gigantica using immunoinformatics approach," *Int J Biol Macromol*, vol. 138, pp. 224-233, 2019.

[131] K. Azhagesan, B. Ravindran, and K. Raman, "Network-based features enable prediction of essential genes across diverse organisms," *PloS One*, vol. 13, 2018.

[132] I. A. Doytchinova and D. R. Flower, "VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines," *BMC Bioinformatics*, vol. 8, p. 4, 2007.

[133] A. Krogh, B. Larsson, G. Von Heijne, and E. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes," *Journal of Molecular Biology*, vol. 305, pp. 567-580, 2001.

[134] M. Moutaftsi, B. Peters, V. Pasquetto, D. C. Tscharke, J. Sidney, H.-H. Bui, H. Grey, and A. Sette, "A consensus epitope prediction approach identifies the breadth of murine T CD8+-cell responses to vaccinia virus," *Nature Biotechnology*, vol. 24, pp. 817-819, 2006.

[135] I. A. Doytchinova and D. R. Flower, "VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines," *BMC Bioinformatics*, vol. 8, pp. 1-7, 2007.

[136] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G. P. Raghava, and O. S. D. D. Consortium, "In silico approach for predicting toxicity of peptides and proteins," *PloS One*, vol. 8, p. e73957, 2013.

[137] M. Nielsen, C. Lundegaard, and O. Lund, "Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method," *BMC Bioinformatics*, vol. 8, pp. 1-12, 2007.

[138] J. Zhu and W. E. Paul, "CD4 T cells: fates, functions, and faults," *Blood, The Journal of the American Society of Hematology*, vol. 112, pp. 1557-1569, 2008.

[139] P. Wang, J. Sidney, Y. Kim, A. Sette, O. Lund, M. Nielsen, and B. Peters, "Peptide binding predictions for HLA DR, DP and DQ molecules," *BMC Bioinformatics*, vol. 11, pp. 1-12, 2010.

[140] S. K. Dhanda, P. Vir, and G. P. Raghava, "Designing of interferon-gamma inducing MHC class-II binders," *Biology Direct*, vol. 8, pp. 1-15, 2013.

[141] G. Nagpal, S. S. Usmani, S. K. Dhanda, H. Kaur, S. Singh, M. Sharma, and G. P. Raghava, "Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential," *Scientific Reports*, vol. 7, pp. 1-10, 2017.

[142] M. D. Cooper, "The early history of B cells," *Nature Reviews Immunology*, vol. 15, pp. 191-197, 2015.

[143] S. Saha and G. P. S. Raghava, "Prediction of continuous B-cell epitopes in an antigen using recurrent neural network," *Proteins: Structure, Function, and Bioinformatics*, vol. 65, pp. 40-48, 2006.

[144] H. A. Dar, T. Zaheer, M. Shehroz, N. Ullah, K. Naz, S. A. Muhammad, T. Zhang, and A. Ali, "Immunoinformatics-aided design and evaluation of a potential multi-epitope vaccine against Klebsiella pneumoniae," *Vaccines*, vol. 7, p. 88, 2019.

[145] U. K. Adhikari, M. Tayebi, and M. M. Rahman, "Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus," *Journal of Immunology Research*, vol. 2018, 2018.

[146] H.-H. Bui, J. Sidney, K. Dinh, S. Southwood, M. J. Newman, and A. Sette, "Predicting population coverage of T-cell epitope-based diagnostics and vaccines," *BMC Bioinformatics*, vol. 7, pp. 1-5, 2006.

[147] Z. Nain, F. Abdulla, M. M. Rahman, M. M. Karim, M. S. A. Khan, S. B. Sayed, S. Mahmud, S. R. Rahman, M. M. Sheam, and Z. Haque, "Proteome-wide screening for designing a multi-epitope vaccine against emerging pathogen Elizabethkingia anophelis using immunoinformatic approaches," *Journal of Biomolecular Structure and Dynamics*, vol. 38, pp. 4850-4867, 2020.

[148] W. Li, M. D. Joshi, S. Singhania, K. H. Ramsey, and A. K. Murthy, "Peptide vaccine: progress and challenges," *Vaccines*, vol. 2, pp. 515-536, 2014.

[149] R. Arai, H. Ueda, A. Kitayama, N. Kamiya, and T. Nagamune, "Design of the linkers which effectively separate domains of a bifunctional fusion protein," *Protein Engineering*, vol. 14, pp. 529-532, 2001.

[150] R. K. Pandey, T. K. Bhatt, and V. K. Prajapati, "Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein," *Scientific Reports*, vol. 8, pp. 1-11, 2018.

[151] N. Nezafat, Y. Ghasemi, G. Javadi, M. J. Khoshnoud, and E. Omidinia, "A novel multi-epitope peptide vaccine against cancer: an in silico approach," *Journal of Theoretical Biology*, vol. 349, pp. 121-134, 2014.

[152] A. Mahram and M. C. Herbordt, "Fast and accurate NCBI BLASTP: acceleration with multiphase FPGA-based prefiltering," *In Proceedings of the 24th ACM International Conference on Supercomputing*, 2010, pp. 73-82.

[153] J. M. Walker, *The proteomics protocols handbook: Springer*, 2005.

[154] W. Zheng, Y. Li, C. Zhang, R. Pearce, S. Mortuza, and Y. Zhang, "Deep-learning contact-map guided protein structure prediction in CASP13," *Proteins: Structure, Function, and Bioinformatics*, vol. 87, pp. 1149-1164, 2019.

[155] L. Heo, H. Park, and C. Seok, "GalaxyRefine: Protein structure refinement driven by side-chain repacking," *Nucleic Acids Res*, vol. 41, pp. W384-W388, 2013.

[156] S. C. Lovell, I. W. Davis, W. B. Arendall III, P. I. De Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, Structure validation by Cα geometry: and C deviation," *Proteins: Structure, Function, and Bioinformatics*, vol. 50, pp. 437-450, 2003.

[157] J. Ponomarenko, H.-H. Bui, W. Li, N. Fusseder, P. E. Bourne, A. Sette, and B. Peters, "ElliPro: a new structure-based tool for the prediction of antibody epitopes," *BMC Bioinformatics*, vol. 9, pp. 1-8, 2008.

[158] D. B. Craig and A. A. Dombkowski, "Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins," *BMC Bioinformatics*, vol. 14, pp. 1-7, 2013.

[159] P. Craig and M. Petticrew, "Developing and evaluating complex interventions: reflections on the 2008 MRC guidance," *International Journal of Nursing Studies*, vol. 50, pp. 585-587, 2013.

[160] G. Van Zundert, J. Rodrigues, M. Trellet, C. Schmitz, P. Kastritis, E. Karaca, A. Melquiond, M. van Dijk, S. De Vries, and A. Bonvin, "The HADDOCK2. 2 web server: user-friendly integrative modeling of biomolecular complexes," *Journal of Molecular Biology*, vol. 428, pp. 720-725, 2016.

[161] C. Dominguez, R. Boelens, and A. M. Bonvin, HADDOCK a protein protein docking approach based on biochemical or biophysical information," *Journal of the American Chemical Society*, vol. 125, pp. 1731-1737, 2003.

[162] Y. Yuan, D. Cao, Y. Zhang, J. Ma, J. Qi, Q. Wang, G. Lu, Y. Wu, J. Yan, and Y. Shi, "Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains," *Nature Communications*, vol. 8, pp. 1-9, 2017.

[163] R. A. Laskowski, "PDBsum new things," *Nucleic Acids Res*, vol. 37, pp. D355-D359, 2009.

[164] D. M. Van Aalten, B. L. De Groot, J. B. Findlay, H. J. Berendsen, and A. Amadei, "A comparison of techniques for calculating protein essential dynamics," *Journal of Computational Chemistry*, vol. 18, pp. 169-181, 1997.

[165] K. Wüthrich, G. Wagner, R. Richarz, and W. Braun, "Correlations between internal mobility and stability of globular proteins," *Biophysical Journal*, vol. 32, pp. 549-560, 1980.

[166] J. R. López-Blanco, J. I. Aliaga, E. S. Quintana-Ortí, and P. Chacón, "iMODS: internal coordinates normal mode analysis server," *Nucleic Acids Res*, vol. 42, pp. W271-W276, 2014.

[167] N. Rapin, O. Lund, M. Bernaschi, and F. Castiglione, "Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system," *PloS One*, vol. 5, p. e9862, 2010.

[168] A. Grote, K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D. C. Hempel, and D. Jahn, "JCat: a novel tool to adapt codon usage of a target gene to its potential expression host," *Nucleic Acids Res*, vol. 33, pp. W526-W531, 2005.

[169] B. Livingston, C. Crimi, M. Newman, Y. Higashimoto, E. Appella, J. Sidney, and A. Sette, "A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes," *The Journal of Immunology*, vol. 168, pp. 5499-5506, 2002.

[170] N. Nezafat, Z. Karimi, M. Eslami, M. Mohkam, S. Zandian, and Y. Ghasemi, "Designing an efficient multi-epitope peptide vaccine against Vibrio cholerae via combined immunoinformatics and protein interaction based approaches," *Computational Biology and Chemistry*, vol. 62, pp. 82-95, 2016.

[171] F. E. Lund, "Cytokine-producing B lymphocytes—key regulators of immunity," *Current Opinion in Immunology*, vol. 20, pp. 332-338, 2008.

[172] J. A. Kovacs, P. Chacón, and R. Abagyan, "Predictions of protein flexibility: first-order measures," *Proteins: Structure, Function, and Bioinformatics*, vol. 56, pp. 661-668, 2004.

[173] A. Umar, A. Haque, Y. S. Alghamdi, M. M. Mashraqi, A. Rehman, F. Shahid, M. Khurshid, and U. A. Ashfaq, "Development of a Candidate Multi-Epitope Subunit Vaccine against Klebsiella aerogenes: Subtractive Proteomics and Immuno-Informatics Approach," *Vaccines*, vol. 9, p. 1373, 2021.

[174] S. Aslam, S. Ahmad, F. Noor, U. A. Ashfaq, F. Shahid, A. Rehman, M. Tahir ul Qamar, E. A. Alatawi, F. M. Alshabrmi, and K. S. Allemailem, "Designing a Multi-Epitope Vaccine against Chlamydia trachomatis by Employing Integrated Core Proteomics, Immuno-Informatics and In Silico Approaches," *Biology*, vol. 10, p. 997, 2021.