Phenotypic and Genotypic Analysis of Clinical MDR *E. coli*

By

Zainab Mazhar

(00000205976)

Atta-ur-Rahman School of Applied Biosciences,

National University of Sciences and Technology,

Islamabad, Pakistan.

Phenotypic and Genotypic analysis of Clinical MDR *E. coli*

A thesis submitted as a final year project as a requirement of MS

in

Industrial Biotechnology

By

Zainab Mazhar

(00000205976)

Atta-ur-Rahman School of Applied Biosciences,

National University of Sciences and Technology,

Islamabad, Pakistan

Annex A to NUST Letter No.

0972/102/Exams/Thesis-Cert

Thesis Acceptance Certificate

Certified that final copy of MS/MPhil thesis written by Ms. **Zainab Mazhar** Registration No. **00000205976** of **ASAB** has been vetted by undersigned, found complete in all aspects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfilment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature with stamp: _____

Name of Supervisor: ___Dr. Saadia Andleeb_____

Date: _____

Signature of HoD with stamp: _____

Date: _____

Countersign by

Signature (Dean/Principal): _____

Date: _____

Declaration

I certify that this research work titled “Phenotypic and Genotypic Analysis of Clinical MDR *E. coli*” is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

_____

Zainab Mazhar

Certificate for Plagiarism

It is to confirm that MS thesis entitled 'Phenotypic and Genotypic Analysis of Clinical MDR E. coli' by Ms. Zainab Mazhar, Registration No. 00000205976 has been examined by me. I undertake that,

1. Thesis has significant new work/ knowledge as compared to already elsewhere. No sentence, table, equation, diagram, photograph or section has been copied verbatim from previous work except when placed under quotation marks and duly referenced.
2. The work presented is original and own work if the author i.e. there is no plagiarism. No idea, results or words or others have been presented as author's own work.
3. There is no fabrication of data or results such that the research is not accurately represented in the records. The thesis has been checked using Turnitin (a copy of the originality report attached and found within the limits as per HEC plagiarism policy and instructions issued from time to time).

Supervisor
Dr. Saadia Andleeb
Associate Professor and
HOD Research
Industrial Biotechnology
ASAB, NUST

Dedicated to

Parents & Siblings

Acknowledgments

First, I thank ALLAH Almighty to bless me with strength to fulfil all the requirements for completion of this dissertation.

Words are inadequate to express my gratitude to my supervisor Dr. Saadia Andleeb. She was a constant source of support during this period. She is extremely dedicated towards her work. She ensures that her research student does quality work and knows the impact of his contribution to the science community. Not only does she focus on developing scientific skills in her students but also on carrying out the research work ethically.

I would also like to acknowledge Dr. Sadaf Zaidi for friendly environment during this stressful task. And Dr. Amjad Ali for his guidance during this strenuous computational biology research work. I'm also grateful to Dr. Abdur Rehman for his kind suggestions.

I would also like to acknowledge Robert F. Potter for his help in genome sequencing that was a vital part of this research project. I'm also thankful to Dr. Hussnain Janjua, Principal Aasb, for his relentless efforts in ensuring the smooth running of this institute.

I'm extremely thankful to seniors Dr. Ibrahim Rashid, Sidra Rehmat and Naimat Ullah for their guidance whenever needed. I could not have completed this work without the assistance of Tehreem Zaheer, Shehroz and Mahnoor Majid. I also thank my juniors Minahil, Hassan, Kashaf and Zomia for all the queries they asked related to their projects through which I learnt a lot.

This act of acknowledgement wouldn't complete without paying regards to my parents and siblings for their constant support during this tenure. Thanks to Papa and Mama for being the best and providing me an opportunity to study at an expensive institute. I still remember when I stumbled during this journey and my family stepped forward to support me. This made me appreciate and realize the importance of the support of loved ones in my life.

Abstract

World Health Organization listed *Escherichia coli* as the number one priority pathogen for the development of new antibiotic due to the fast evolution of its pathogenicity and multidrug resistance. Understanding the mechanism behind increasing bacterial antibiotic resistance is as significant as development of new antibiotics. Often, bacterial resistance and virulence attributes are harboured on highly mobile pieces of DNA that can transfer to other bacteria. Prophages, plasmids, transposons and insertion sequences all are mediators of this form of genetic transfer. A case of UTI was reported in PIMS hospital Islamabad contributed by an extended spectrum β-lactamase (ESBL) *Escherichia coli* exhibiting resistance to Ceftazidime (3rd generation cephalosporin). Phenotypic analysis by antimicrobial susceptibility assay by using Kirby Bayer Disk Diffusion revealed that isolate was resistant against commonly prescribed antibiotics for treatment of urinary tract infection. Isolate was susceptible to carbapenem but resistant against more than three classes of antibiotics suggesting the isolate to be classified as Multidrug Resistant (MDR). Genotypic analysis was done by identification of β-lactamases by PCR revealing the presence of *bla*CTX-M. Whole Genome Sequencing and downstream analysis were employed for extensive genotypic analysis. Whole Genome Annotation revealed the genome to be ~4.85Mbp having 5095 CDSs, 62 tRNA and 5 rRNA genes. Further downstream analysis exhibited the presence of multiple insertion and mobile elements present along with resistance genes that provide resistance to Sulphonamides, Aminoglycosides, Trimethoprim and Beta-lactams. For long term of infection control it is significant to develop a vaccine against Uropathogenic *Escherichia coli* which shows susceptibility to few antibiotics. There is no vaccine available against this pathogen up to date. The 3D structure of proposed multi-epitope vaccine was refined and validated using bioinformatics tool. The proposed vaccine is antigenic and stably interacts with TLR-2 and TLR-5.

Table of Contents

List of Figures

List of Tables

Introduction

Escherichia coli belongs to family Enterobacteriaceae, which are anaerobically rod shaped bacteria. *Escherichia coli* are single straight rods with width 1.1–1.5μm and length 2–6 μm. They can be either nonmotile or motile. For its motility *Escherichia coli* produces lateral flagella. Other than flagella, *Escherichia coli* produces proteinaceous appendages (pili or fimbriae) projecting outward from surface of bacteria and have role in the attachment to the surface including host tissue and other cells (Desmarchelier and Fegan, 2016).

Antibiotic resistance requires urgent global attention due to worldwide public health emergency. Intensive usage of antimicrobials causes selection of resistance in bacterial pathogens present among animals and human, which lead problems in infection treatment (Laxminarayan, 2013). *Escherichia coli* and K*lebsiella pneumoniae* are commensals bacteria, they are involved in sepsis and urinary tract infections. Exposure of antibiotics to such bacteria results in increased mortality, morbidity and longer duration of hospitalization. According to CDC one report at least 23,000 people in United states face death annually due to multidrug resistant infections and this rate is even three times higher in lower income countries (Allegranzi, 2013). Global spread of multi-drug resistant CTX-M type ESBL producing Enterobacteriaceae strains have become extensively untreatable especially in low income countries. According to infection society for infectious disease the frequency of hospital acquired urinary tract infection is 12.9% in the USA, 19.6% in the Europe and 24% in the developing countries like Pakistan, India.

After the birth of human infant, the gastrointestinal tract is colonized by *Escherichia coli*. Commensal *E. coli* colonizes mucous layer of the mammalian colon. *E. coli* is the best competitor among facultative anaerobes of the human intestinal microflora. As compared to the gut other residents the *Escherichia coli* utilizes gluconate more efficiently which suggests its high metabolic activities. To cause a wide spectrum of diseases multiple virulence factors are acquired by *E. coli*.

Antimicrobial resistance has increased in bacteria *Escherichia coli*, most prevailing source of urinary tract infections. There is an extensive increase in population of extended spectrum beta lactamases MDR *E. coli*. These enzymes provide resistance to third generation

cephalosporins which are prescribed in urinary tract infections (Chen *et al*., 2013). These strains are not only present among hospitals wards but also among general population. *E. coli* clonal type O25:H4-ST131 has emerged as MDR bacterial pathogen worldwide. *E. coli* isolates have become resistant to first line antibiotics like amoxicillin, trimethoprim-sulfamethoxazole and clavulanic acid which are used to treat nosocomial infections worlwide. The release of extended spectrum cephalosporins and fluoroquinolones in 1980s increased the treatment efficacy but resistance against fluroquinolones and extended spectrum cephalosporins have been developed which diminished hopes (Gupta, 1999).

To develop a specific PATHOTYPE of *E. coli* the most effective combination of virulence factors must be persisted. Three most common clinical syndromes i.e. UTI, enteric/diarrhoeal and sepsis/meningitis diseases are resulted from infection with one of the *Escherichia coli* pathotypes. Among intestinal disease-causing pathogens six are well characterized: enterohaemorrhagic *E. coli*, enteropathogenic *E. coli*, enteroaggregative *E. coli*, enterotoxigenic *E. coli*, enteroinvasive *E. coli* and diffusely adherent *E. coli*. Meningitis and sepsis are contributed by extraintestinal pathotype called meningitis-associated *Escherichia coli*. Uropathogenic *E. coli* are also called as ExPEC (extraintestinal *E. coli*) (Kaper *et al*., 2004).

The metabolic fitness of ST 131 and its potential to be transmitted efficiently was investigated by multiple studies. About 300 UPEC were collected between 2007 and 2009 and investigated by Gibreel and colleagues. Among them 37 belong to ST131. The capability of these strains to utilize carbon source and enzymatic activity were tested 47 biochemical tests added in Vitek 2 compact Automated Expert System (AES). Non ST131 showed fewer positive results than ST131. By UPGMA cluster analysis it was revealed that ST131 clones were closely related to the antibiotic resistance strains. Comparative study performed between commensal *E. coli* strains and ST131 in murine models it was revealed that latter strain out-compete the former strain and had more potential to colonize the gut without evoking potential immune response (Vimont *et al*., 2012).

According to last few decades whole genome sequencing is identified as important technique in clinical microbiology. In 1995 the first bacterial genome was sequenced. WGS now become a most promising tool in clinical laboratories as a result of high throughput sequencing technologies. Isolate characterization, finding sources of infection, antimicrobial resistance profiling and between patient transmission are common applications of WGS in

diagnostics laboratories (Pallen, 2010). WGS is used for accurately distinguishing lineages compared to pulse field gel electrophoresis, MALDI-TOF and variable number tandem repeat (Neville, 2011).

Limited information is available among low- and middle-income countries about clonality of circulating UPEC strains. As resistance against frequently prescribed antibiotics (fluroquinolones, Beta-lactamases and Ampicillin) is increasing. Resistance against available antibiotics is increasing (Fluroquinolones and Cephalosporins) which limit the available treatment options to cure urinary tract infections. The current study focuses first on the determination of the genetic diversity in *Escherichia coli* genome isolated from PIMS hospital Islamabad, Pakistan and then a prolonged infection control strategy was developed by using reverse vaccinology approach.

Objectives

The objectives of the current research are

1. Whole genome sequencing to determine the genomic diversity, antibiotic resistance and virulence attributes in *E. coli* genome
2. Reverse vaccinology-based vaccine designing for Uropathogenic *Escherichia coli*

Literature Review

2.1 Basic Structure of Bacterial Cell

In gram positive bacteria the inner membrane is bounded by a rigid and tough mesh like structure called cell wall. While cell wall is a thin layer in gram negative bacteria. In gram negative bacteria cell wall is bounded by another membrane called outer membrane. Periplasm is a space present between cytoplasmic membrane and outer membrane (Figure 1). Outer membrane present in gram negative bacteria provides shield and inhibits entry of many substances into the bacteria. This outer membrane contains many channels called porins, which permit entry of molecules like drugs inside the cell. Cell wall is a tough layer that gives bacterium a defined shape and refrains it from mechanical and osmotic stress (Hauser, 2015).

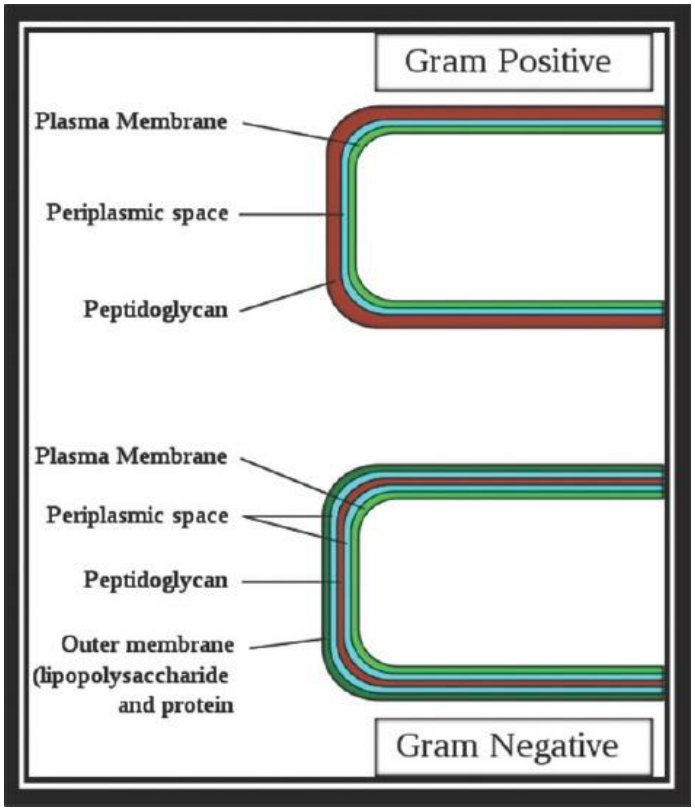


Figure 1 Bacterial cell envelope structure

2.2 Classification of Antibiotics on basis of Mechanism of Action

Antibiotics are classified into various classes on basis of action mechanism as described in figure 2.

2.2.1 Antibiotics with Cell Wall as Target

Long sugar polymers made up of peptidoglycan are present in bacterial cell wall. The peptidoglycan undergoes glycan strands cross linking through action of transglycosidases. Sugars present in polymeric form contain peptide chains (Reynolds, 1989). These peptide chains cross link each other through D alanyl-alanine portions. Cross linking of the peptide chains involved glycine residues of D-alanyl portion of peptides in presence of penicillin binding proteins. Peptide chains are cross linked to strengthen the cell wall. Cell wall synthesis is inhibited by the glycopeptides and β-lactams.

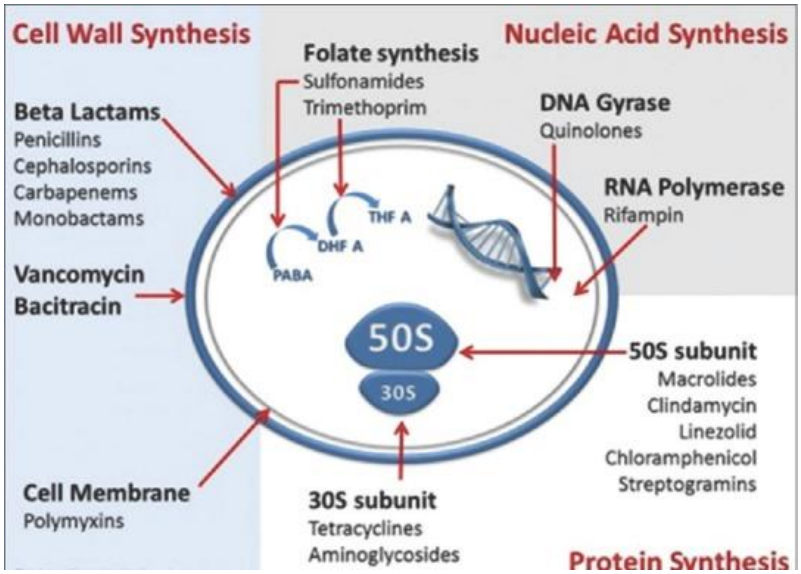


Figure 2 Mechanism of action of antibiotics

2.2.1.1 Beta-lactam Antibiotics

PBPs are major target of β-lactams. PBPs bind to beta-lactam rings similar in structure to D-alanyl D alanine of peptide. PBPs interacting with Beta lactams are not available for peptidoglycan synthesis. Bacterial cell is lysed through disruption of peptidoglycan.

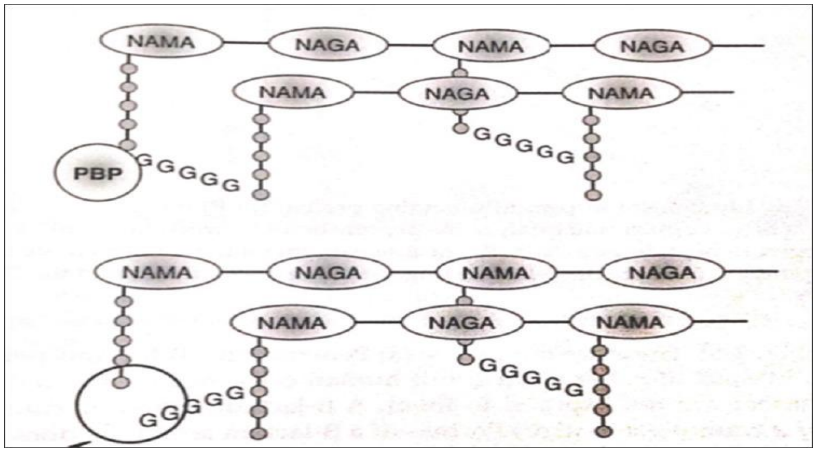


Figure 3 Mechanism of action of β-lactam antibiotics

2.2.3 Inhibitors of Protein Biosynthesis

Transcription is a process in which a molecule called m-RNA is synthesized from DNA (figure 4). Then, translation occurs. Translation is a process in which macromolecular structure called ribosomes synthesized proteins present in form of information in m-RNA. Process of proteins synthesis is controlled by ribosomes and multiple cytoplasmic components. Ribonucleoprotein subunits such as 50S and 30S are the major components of 70S bacterial ribosome (Yoneyama and Katsumata, 2006). Antimicrobial agents inhibit protein synthesis by targeting bacterial ribosomes subunits 50S or 30S (Vannuffel and Cocito, 1996).

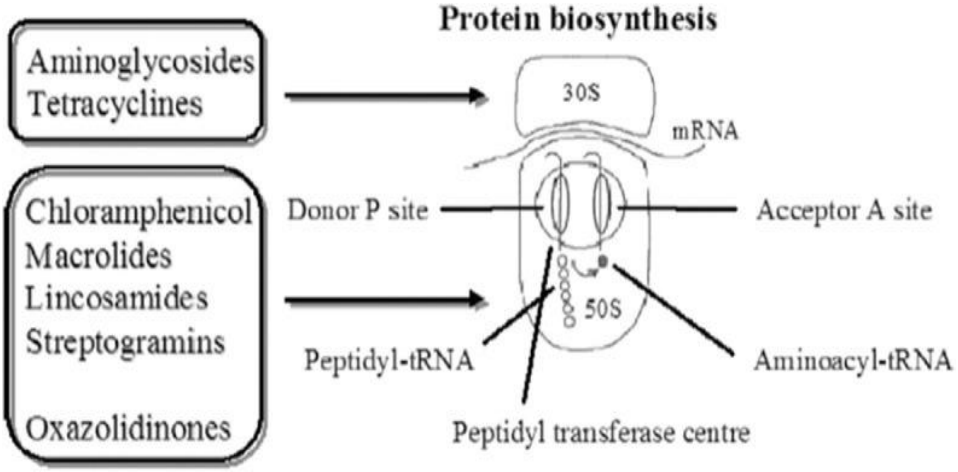


Figure 4 Site of action of protein biosynthesis inhibitors

2.2.3.1 **Subunit 30S Inhibitors**

2.2.3.1.1 Aminoglycosides

Aminoglycosides have positively charged nature through which these are attracted towards outer membrane which is negatively charged in nature. Attraction of positively charged compounds towards negatively charged membrane allows large pores formation in outer membrane and introduces entry of antibiotic within bacterial cell. To reach on its target ribosomes, antibiotic molecules must have to pass through cytoplasmic membrane which is energy dependent step and requires an active proton movement. For such reasons' AGs require oxygen and work in aerobic condition rather anaerobic one. AGs work along with those antibiotics that inhibit cell wall synthesis. Through hydrogen bonding AGs act near a site A of the 30S ribosomal subunit. AGs cause premature termination of translation process.

2.2.3.1.2 Tetracyclines

Tetracyclines, such as doxycycline, chlortetracycline, or minocycline, refrain t-RNA binding to the site A by acting on highly conserved sequence of 30S ribosomal subunit (Yoneyama and Katsumata, 2006).

2.2.3.4 50S subunit Inhibitors

2.2.4.1 Chloramphenicol

It interacts on 50S ribosomal subunit 23S r-RNA and binds with the peptidyl transferase cavity. To inhibit protein synthesis, it binds to ribosome site A and refrains t-RNA binding on its site in ribosome.

2.2.4.2 Macrolides

Macrolides affect the translocation activity of the 50S subunit. It targets the 23S r-RNA conserved sequence of peptidyl transferase. This process results in non-functional peptide chains from m-RNA (Wise, 1999).

2.2.5 Inhibitors of DNA replication

2.2.5.1 Quinolones

Bacterial DNA gyrase is inhibited by fluoroquinolones (FQs), which introduces nicks in DNA double strands. It decreases super coiling in DNA strand during replication or transcription and rebinds the nicked ends. DNA gyrase contains two subunits of A and B. Subunit A performs breakage of DNA double strands and subunit B decreases super coiling and at the end subunit

A re-joins the broken end. The FQs bind with higher affinity to DNA gyrase A and interferes with DNA strands breaking and rebinding functions (Higgins *et al*., 2003).

2.2.6 Inhibitors of Folic acid Metabolism

2.2.6.1 Sulfonamides and Trimethoprim

Both these drugs target different folic acid metabolism steps. Trimethoprim and sulpha drugs show synergy through action on different stages of similar biosynthesis pathway. Sulfonamides show higher affinity for inhibition of dihydropteroate synthetase than natural substrate p-amino benzoic acid. Sulfonamides act in a competitive manner. Trimethoprim inhibit dihydrofolate reductase at last stage of folic acid synthesis (Yoneyama and Katsumata, 2006).

2.3 Mechanisms Involved in Antimicrobial Resistance

Inhibition in accumulation of antimicrobials either by increasing efflux or by decreasing uptake from the cell via changes in outer membrane permeability

Drug molecules transfer to the cell through different mechanisms such as diffusion through self-uptake, porins and diffusion through bilayer. In outer membrane of the bacteria porins channels are present. The outer membrane facilitates the movement of small hydrophilic molecules through porins. The decrease entry of beta lactams antibiotics and FQs occur due to the decrease in number of porin channels.

2.3.1 Efflux Pumps

These pumps export antibiotics out of bacterial cell and maintain low intracellular concentrations. Before antimicrobial agents reach to their targets they are exported out of the cell at the same speed (Wise, 1999). In the cytoplasmic membrane efflux pumps are present while porins are present in outer membrane. These pumps are specific for antibiotics. Most of these pumps are multi drug transporter that allow efflux of range of antibiotics i.e. fluroquinolones, tetracycline and macrolides. Thus, contribute mainly towards emergence of multidrug resistance (Lambert, 2002).

2.3.2 Modification of Target Molecule

Binding of antimicrobial agents to the target site is prevented by various changes in it. Any mutations in bacterial genes present on the chromosome often results target site changes. Minor

alteration in drug target molecule have effect on antibiotic binding and specific interaction with target molecule.

2.3.3 Antibiotic Inactivation

There are three enzymes that are involved in inactivation of antibiotics such as Chloramphenicol acetyltransferases (AACs), Aminoglycoside-modifying enzymes and Beta-lactamases (Dockrell, 2004).

2.3.3.1 Beta-lactamases

Beta-lactamases hydrolyze those antibiotics that have amide and ester bond e.g., monobactam, penicillins, carbapenems and cephalosporins. Nearly about 300 Beta lactamases are known till yet. Beta lactamases are prevailing enzymes that are further classified into two systems: Functional (Bush–Jacoby–Medeiros) and Structural (Ambler).

2.3.3.2 Aminoglycoside Modifying Enzymes (AGE's)

Specific enzymes neutralize AG such as Nucleotidyl-transferases, Phosphoryl-transferases or Adenylyl-transferases. These modifying enzymes lessen affinity of modified molecule, delay binding to 30S ribosomal subunit and provide resistance to FQ and AG's (Maurice, 2008).

2.3.3.3 Chloramphenicol-Acetyltransferases

Few bacterial pathogens and some *Haemophilus influenzae* isolates are present resistant against chloramphenicol. These bacteria contain an enzyme chloramphenicol transacetylase that acetylates hydroxyl group of chloramphenicol. Binding of modified chloramphenicol is not able to bind to the ribosome 50S subunit properly (Tolmasky, 2000).

2.4 Bacterial Genome

It is necessary for bacteria to maintain genome stability and integrity for its survival (Venttone *et al*., 2014). A genome must undergo variable degree of modifications, in order to adopt to the changing environmental conditions. Bacteria have developed offense, defence and robust repair apparatus to overcome environmental factors that introduce damaging effects on the genome (Denamur and Matic, 2006). To avoid extreme leaning in any direction, bacteria use genomic features to maintain balance between plasticity and genome stability.

2.4.1 Horizontal Transfer of Gene

The transfer of genetic material between different genomes is known as horizontal gene transfer. Through horizontal transfer a huge portion of bacterial DNA is shared between bacteria. First gene transfer was observed in publication of japan in 1959, which reported phenomena of sharing of antibiotic resistant genes through horizontal transfer of gene (Ochiai *et al*., 1959). The primary reason which present resistance to various antibiotics classes and worldwide epidemics of community and hospital acquired infections is horizontal transfer of gene (Dorman, 2014). Genes of antibiotic resistance are linked to resistance islands, transposons, integrons and are often present on transferable plasmid. Various mechanisms of recombination are used by these mobile elements to get integrated into recipient bacterial genome.

2.4.1.1 Mobile Genetic Elements

These genetic elements consist of diverse members i.e. plasmids and viruses of which some are non-canonical (genomic islands), some canonical which together form mobilome. These all are driving forces for bacterial evolution, adaptation, evolution, speciation and diversification. Mobile genetic elements can carry beneficial, parasitic and cooperative genes, probability and distribution of which is influenced by evolutionary and ecological conditions (Rankin *et al*., 2011).

2.4.1.2 Bacteriophages Mediating Genome Plasticity

Temperate bacteriophages are vehicle for genome variation and horizontal transfer of genes. DNA encapsulated within bacteriophages is protected against attack of many solvents and nucleases. Phages have an ability to survive in environment suitable for bacteria even without loss of its biological activity. There are two main types by which bacteriophage transfer genetic material, specialized and generalized transduction. When phage genome sized DNA molecule packed inside phage head and replaces phage DNA, this condition is known as generalized transduction. Any part of host genome can be packaged inside phage head and transferred via bacteriophage. Examples of generalized transduction are coliphages T4, T1and P1 and P22 of *Salmonella* (Masters, 1996). In contrast, both viral DNA and host DNA sequences are present in specialized transducing phages. Specialized transducing phages via aberrant excision from host chromosome. Example of transducing phage is coliphage lamda. Those phage markers that present near prophage insertion site are transduced through specialized transduction (Weisberg, 1996).

Lysogenic phages play role in spread of genes of antibiotic resistance (O'Shea and Boyd, 2002). Most of the studies suggested that the major reservoirs of ARGs are viruses. Genes for *ermB* (macrolides resistance), *qnrS* (fluoroquinolones resistance), *blaTEM* (β-lactams resistance), *tetW* (tetracyclines resistance) and *sul* (sulphonamides resistance) have been found in the viruses isolated from environmental and activated sludge water samples (Parsley *et al*., 2010), and present on the DNA of phage extracted from human faecal samples and mucus layer of intestine of wild freshwater fishes (Marti *et al*., 2018).

One of the studies assess the contribution of prophages in evolution of *A. baumannii* pathogenicity. Multiple virulence factors are encoded by *A. baumannii* prophages that maybe involved in bacterium's capacity to evade the host immune system, tolerate antibiotics, subsist in unfavourable environments and colonize host niches. These results suggest the contribution of prophages towards evolution and dissemination of pathogenicity in *A. baumannii* (Wang *et al*., 2010).

Prophages that become inactive to form cell lysis, plaque formation and production of phage like particles may become entrapped in the chromosomes as result of gradual decay and genome rearrangements these elements are called as cryptic prophages. Since its divergence 100 million years ago from *Salmonella* sp. *Escherichia coli* K-12 comprise 3.6% of genome that harboured nine cryptic prophage elements and 1600 kbp of novel DNA (18%) (Wang *et al*., 2010). For example, Bor protein of phage lambda results in increased *E. coli* survival by significantly increasing its serum mediated killing resistance.

2.4.2 Conjugation

Bacteria have ability to share their genetic content with other bacteria by tube-like structure called pilus from donor cell to a recipient cell this process is called conjugation. Conjugative transposons and double stranded plasmids could be exchanged (Juhas *et al*., 2009). Through conjugation antibiotic resistant genes transferred in gram negative bacteria and numerous studies have reported conjugation mediated transfer of antibiotic resistant genes. Studies have reported successful transfer of antibiotic resistant genes through conjugation via integrons from clinical isolate *Acinetobacter* sp. to environmental isolates (Fu *et al*., 2012).

2.4.2.1 Transposons

Transposition is a process of jumping of transposons within a genome. These elements are also called as jumping genes. Inverted repeats are present at both ends of transposons. Transposons are different from insertion sequences. Transposons carry genes while no genes are present on insertion sequences. Phenotype of the cell can be changed by transposons for example, genes of antibiotic resistance present in transposons that confer resistance to particular antibiotics. Transposons are capable of transferring genes between plasmid and chromosomal DNA (Bennett, 2008).

2.4.2.2 Integrons

These genetic elements can transfer and capture genes, like antibiotic resistance, by a mechanism of site-specific recombination across bacterial genomes. Two characteristics make transposons different from integrons. First, repeat sequences are present on both ends of genes present on transponsons, while integrons have no such repetitive sequences and like bacteriophage site specific integrase gene are present on integrons but they don't own enzymes that are involved in process of transposition (Boucher *et al*., 2007). Genes for virulence factors and antibiotic resistance are carried on genes cassettes (Roe and Pillai, 2003). Genes for several antibiotics' classes (aminoglycoside, trimethoprim, quinolones, chloramphenicol and beta lactams) are reported to be present on ARGs cassettes.

2.4.2.3 Plasmids

This is extrachromosomal DNA present in the cell which is capable of independent replication. Transfer of antibiotic resistance genes occur through plasmids against several classes of antibiotics (Svara and Rankin, 2010). Conjugation is most common mechanism involved in the dissemination of genes of antibiotic resistance via plasmid. Specialized membrane structure pilus are actively involved in transfer of plasmid between conjugating bacteria by a process called conjugation (Huddleston, 2014). The transfer of plasmid is linked to the bacterial genome plasticity due to exchange of gene content occur during this mechanism (Treangen and Rocha, 2010). Emergence of super antibiotic resistant bugs takes place via plasmid transfer through conjugation process (Johnson *et al*., 2010).

2.5 Genes Conferring Resistance to Quinolones and Fluoroquinolones

To treat various infections, occur in humans' fluoroquinolones and quinolones are important antimicrobial agents for them. Against all bacteria fluoroquinolones are used as bactericidal agents. Reduced permeability of the outer membrane, mutations in drug targets may also occur mainly in the genes for topoisomerase IV and DNA gyrase but other mechanism such as upregulated expression of efflux pumps may also play an important role to confer resistance against these antimicrobial agents.

Fluoroquinolones have two targets primary and secondary. Primary target of fluoroquinolones is gyrase which has two subunits of GyrA and GyrB. Topoisomerase IV is another target of fluoroquinolones which consists of two subunits of ParE and ParC. Major mutations are observed in quinolones determining region present between Gln107 and Ala67. Quinolones determining region mutations occur at 87 and 83 codons (Hopkins *et al*., 2005). Resistance to quinolones occur by single mutation in gyrA but mutations within parC and gyrA gene are needed is needed for fluoroquinolones resistance. Mutation occur at codon 80 and 84 of ParC.

2.5.1 Plasmid-Borne Resistance Mechanisms Confers Resistance to Fluoroquinolones

Since 1997 global dissemination of PMQR determinants is a serious alarm when plasmid mediated quinolone resistance determinants (PMQR) qnrA1 was first identified (Rodríguez-Martínez *et al*., 2016). Some plasmid related resistances have been identified such as (i) active efflux pumps (ii) Qnr-like proteins protect DNA from quinolone binding such as (QnrA, QnrD QnrB) and (iii) the AAC (6')-Ib-cr acetyltransferase alters fluoroquinolones i.e. enrofloxacin and ciprofloxacin. Overall these resistance determinants confer reduced resistance to quinolones. However, these determinants might play a role in selection of those clinical isolates conferring high level of resistance through chromosomally encoded mechanisms (Rodríguez-Martínez *et al*., 2016).

Methodology

3.1 Sample Collection

This study included one hundred and fifty-seven bacterial strains obtained from PIMS hospital, Islamabad during period of January 2012 and June 2013. Bacterial cultures were obtained from pus, sputum, high vaginal swab, stools and various infections of blood. Collected samples were identified through standard biochemical techniques and Gram staining as described previously (Cheesbrough, 2006).

3.2 Antibiotic Resistance Phenotypes

Each isolate resistance profile was established using Kirby Bauer Disk diffusion method. Phenotypic test for AST was performed according to (CLSI) guidelines (CLSI, 2013). Bacterial strains were tested against susceptibility to 12 antibiotics.

3.3 DNA Extraction

Bacteria was grown in 5mL of LB for overnight. 1.5mL of culture in a sterile microfuge was centrifuged at 13,000 rpm for 2 minutes. After centrifugation supernatant was removed and pellet was re-dissolved in 576 µL of TE buffer. 3 µL (20mg/mL) proteinase K and 30 µL (10 % SDS) was added and rested at 37 ºC for 1 hour. 100 µL (5 M NaCl) was added in it. 80 µL CTAB/NaCl solutions was added followed by incubation for 10 minutes at 65ºC. 700 µL (1:24) isoamyl alcohol/chloroform was added and centrifuged for 5 at 13,000 rpm minutes. Supernatant was separated in other microfuge tube and phenol/chloroform/isoamyl alcohol about 1 volume was added and followed a 5 minutes centrifuge. In a microfuge tube supernatant was separated and DNA was precipitated by adding isopropanol into it. DNA pellet was dissolved in 70% ethanol. Ethanol was evaporated and in 50 µL TE buffer pellet was re-suspended. At -20°C suspended DNA was stored.

3.3.1 Amplification of *bla*CTX-M Gene

Reported primers were used to amplify *bla*CTX-M group (Monstein *et al.,* 2007). Amplicon size and primer sequences are provided in table 1.

Table 1 Primers for PCR amplification

Group	Primers	Sequence	Amplicon size (base pairs)
CTX-M	CTX-M-F1	5'-ATGTGCAGTACCAGTAAAGTTATGGC-3'	593
	CTX-M-F2	5'-TGGGTGAAGTAAGTGACCAGAATCAGCGG-3'	

PCR profile for primer sequences was optimized. PCR tubes (Axygen^{R,} California, USA) were used for reaction mixture preparation. 25 µl of reaction mixture was prepared containing 10x PCR buffer, 50 mM of $MgCl_{2,}$ each primer 50 pmol, 1 unit of Taq polymerase, 2 mM of dNTPs and 3 µl of DNA sample (Fermentas, USA). Then microcentrifuge was used to vortex the prepared reaction mixture (Edison, New Jersey, USA), and followed by PCR setup by using thermal cycler Swift^{TM} MaxPro (Applied Biosystem, Foster city, USA).

3.3.2 PCR Profile of *bla*CTX-M15

Initial step of denaturation was performed at 94 ºC for 4 minutes followed by final denaturation of 35 PCR cycles at 94 ºC for 45 seconds, annealing at 56 ºC for 45 seconds and extension was performed at 72 ºC for 2 minutes. The reaction was ended at 72 ºC for 10 minutes with a final extension and hold at 4 ºC until tubes were carried out of the cycler.

3.3.3 Agarose Gel Electrophoresis

For analysis of PCR product, in 1X TAE buffer 2% of agarose gel was prepared (400 µL of 0.5 EDTA pH 8.0, 1 M Tris HCl (10mL) and 990 mL distilled water). For gel preparation, 1X TAE buffer (40 mL) was used to add agarose powder (0.8 g) and heated until gets boil in microwave.

Agarose mixture was cooled to room temperature and to stain the gel 5µL of ethidium bromide (0.5µg/mL) was added upon cooling. For analysis, 1X loading dye (3 µL) was mixed with (7 µL) PCR product. To resolve the PCR product into DNA bands gel electrophoresis was

performed (Wealtec, Sparks, USA) at voltage 60mA and for 30 minutes. To completely resolve DNA product current was increased to 80 mA. UV trans-illuminator (Biometra, Goettingen, Germany) was used to view a gel. Gel Doc Setup (Wealtec Dolphin Doc Sparks, USA) was used to photograph a gel.

3.4 Whole Genome Sequencing

Single colony of selected isolate was cultured in LB broth. Total DNA was extracted (Invitrogen Charge Switch gDNA Mini Bacteria kit) as per the manufacturer's protocol than followed by library preparation using Illumina Nextera XT Kit (Baym *et al*., 2015). Prepared libraries were sequenced by Illumina NextSeq 2500, at Washington University School of Medicine. Deconseq and trimmomatic were used to remove potentially contaminating human reads and Illumina adaptors (Schmieder & Edwards, 2011). SPAdes 3.10.0 was used to assemble paired end reads. QUAST was used to analyse contigs (Vyahii & Tesler, 2013). These generated contigs were used for further analysis.

3.5 Creating single FASTA file using Artemis

The generated .FNA sequence file, containing reordered contigs, was retrieved from respective folder and opened on Artemis tool. Go to 'find' option and. Select 'Write' then select 'all bases' to generate a single sequence FASTA file for its use in further analysis (Carver *et al*., 2005).

3.6 Genome Annotation

3.6.1 Genome Annotation using Parokka

The assembled draft genome was annotated using Parokka by using command line prokka/home/asab/filename. Fna. Generated output files were saved for further analysis (Seemann, 2014).

3.6.2 Genome Annotation using PATRIC ver 3.5.43

Obtained draft genome was annotated using Patric ver. 3.3.43. Patric employs RASTtk for genome annotation which is update RAST server which is still under trial. The Patric server identifies CDS, rRNA and tRNA, and analyzes different subsystems present in genome (Wattam *et al*., 2017; Brettin *et al*., 2015).

3.7 Genotypic Characterization using Center for Genomic Epidemiology

The Center for Genomic Epidemiology offers various tools for typing, phenotyping and phylogeny such as PlasmidFinder, pMLST, SeroTypeFinder, Resfinder and CSI phylogeny using available at http://www.genomicepidemiology.org/.

3.7.1 Isolate Typing using Multi Locus Sequence Typing from CGE

Multilocus Sequence Typing was also performed by MLST 2.0 tool available at (https://cge.cbs.dtu.dk/services/MLST/). MLST configuration used was Escherichia coli# 1 and data type was Assembled Genome/Contigs in .FASTA format (Larsen *et al*., 2012).

3.7.2 PlasmidFinder-CGE

Plasmidfinder was used to identify plasmid sequences present in assembled genome. PlasmidFinder will be available at (https://cge.cbs.dtu.dk/services/PlasmidFinder/). Enterobacteriaceae was selected as a database with threshold value was 95% and least coverage selected was 60% (Carattoli *et al*., 2014).

3.7.3 SNP Determination

Single Nucleotide Polymorphism (SNP) were detected against three genomes. The genome available at NCBI *E. coli* (str. K-12 substr. MG1655111) was a reference genome and two other strains, *Escherichia coli* UK and *Escherichia coli* PK and *Escherichia coli* INDIA Escherichia coli USA were used along with the sequenced isolate to identify SNPs present within the genome. CSI Phylogeny hosted by Center of Genomic Epidemiology was used to identify the SNP presents within genome (Kaas, Leekitcharoenphon, Aarestrup, & Lund, 2014). Following Input Parameters were used for computation of SNP.

Select min. depth at SNP positions:10

Relative depth at SNP positions: 10

Minimum distance between SNPs: 10

Minimum SNP quality:30

Minimum read mapping quality: 25

Minimum Z-score: 1.96

SNP analysis was also done using RealPhy with default settings. The input included *E. coli* (str. K-12 substr. MG1655111), *Escherichia coli US* and *Escherichia coli* India and *Escherichia coli* from Pakistan. The output was obtained in the form of phylogenetic tree.

3.8 Identification of Antibiotic Resistance Genes

3.8.1 Identification of ARGs by ResFinder

ResFinder 2.1 was used for identifying acquired antimicrobial resistance genes as well as chromosomal mutation present in ARGs genes. 90% threshold was used for identification of resistance genes with least length of 60%. Reads type used was Assembled Genome/Contigs (Zankari *et al*., 2012).

3.8.2 Identification of ARGs by CARD

Genes of antibiotic resistance were identified by Resistance Gene Identifier (RGI) tool available at online server. The input file is in .FASTA format, data type was DNA sequence and 'Discovery-Perfect, Strict and Loose hits' was selection criteria ((Jia *et al*.,2017). Comprehensive resistance database is offered at (https://card.mcmaster.ca/).

3.9 Identification of Genetic Determinants

3.9.1 Determination of Insertion Sequences

Insertion sequences were predicted in draft genome by IS finder. Input for analysis is FASTA format and blastn is used to perform analysis. Genome sequence was pasted in query sequence and default settings were used.

3.9.2 Determination of Prophage Sequences

PHASTER is a webserver used for prophage sequences determination and annotation present in bacterial plasmids and genomes. Select an input type 'upload file' and upload nucleotide sequence file in FASTA format and submit it.

3.9.3 Determination of Virulence Factors

A web-based virulence factors database is used for determination of virulence factors within bacterial genome. A blast-based search is performed, and virulence factors is selected as option in query field.

3.10 Vaccine Development

3.10.1 Proteins Sequence Search

In first stage of vaccine development genome of *Escherichia coli* was analysed. Through literature two virulent and conserved proteins of Uropathogenic *Escherichia coli* were selected. Amino acid sequences of proteins (AGA03821.1 and ANK05584.1) were retrieved from NCBI database in FASTA format. Both, T and B-cell antigens were analyzed to generate both cell-mediated and humoral immune response.

3.10.2 B cell Epitope Prediction

B lymphocytes bind to the antigen specific fragments called B cell epitope. BepiPred-2.0 web server was used to predict B cell epitope (http://www.cbs.dtu.dk/services/BepiPred/). BepiPred-2.0 is based upon algorithm learned to predict B cell epitopes annotated from antigen-antibody protein structures. Moreover, two online servers were employed to predict linear epitopes. SVMTrip, is Support Vector Machine method for linear epitopes prediction which fuses propensity and tri-peptide similarity results, for epitope prediction. IEDB server (http://tools.iedb.org/mhcii/) was also used to discover antigenic epitopes.

3.10.3 Cytotoxic T lymphocytes (CTL) Epitope Prediction

Cytotoxic T lymphocytes (CTL) epitopes were analysed by web-server NetCTL 1.2 (http://www.cbs.dtu.dk/services/NetCTL/). This method predicts proteasomal C-terminal cleavage, TAP (Transporter Associated with Antigen Processing) transport efficiency and MHC class I binding peptides. For predictions of CTL epitopes only A1 supertype was used. By employing artificial neural networks proteasomal cleavage and MHC I binding were performed. Using weight matrix TAP transporting efficiency is predicted. For CTL epitope prediction of the CTL epitope default setting (0.75) was used.

3.10.4 Helper T-cell (HTL) Epitope Prediction

NetMHCII 2.2 (http://www.cbs.dtu.dk/services/NetMHCII/) was employed to analyze 15 k mer length HTL epitopes for MHC II human alleles. This server predicts epitopes to HLA-DR, HLA-DP and HLA-DQ various alleles by employing artificial neural networks. Epitopes of MHC II were predicted on basis of receptor affinity deduced from percentile ranks and IC_{50} values given to each predicted epitope. High affinity peptides have IC_{50} values <50 nM,

intermediate affinity binding peptides have IC_{50} value <500 nM and <5000 nM value recommends low affinity. The percentile rank can be directly related to the IC_{50} and inversely related to the affinity of the epitope.

3.10.5 Construction of Multi-Subunit Vaccine

Predicted high binding affinity HTLs and B-cell epitopes and high scoring CTLs were used to create vaccine candidate. Various epitopes were linked by GPGPG and AAY linkers. Moreover, a Toxin B subunit was added as an adjuvant **(**Accession id. Q7X2D2) at N terminus using EAAAK linker. The adjuvant sequence was obtained from UniProt. Adjuvant was attached to improve immunogenicity.

3.10.6 Antigenicity and Allergenicity Prediction

VaxiJen v2.0 and ANTIGENpro were utilized to examine antigenic nature of constructed vaccine. ANTIGENpro (http://scratch.proteomics.ics.uci.edu/) generates an antigenicity index. Based upon cross-validation this webserver has 76% of accuracy in its results. The freely accessible VaxiJen 2.0 server (http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html) is works on cross and auto-variance transformation of proteins into uniform vectors to focus on the antigenicity of multi epitope construct.

3.10.7 Physiochemical Properties and Solubility Prediction

By webserver ProtParam (http://web.expasy.org/protparam/) various physiochemical properties were assessed. These parameters included theoretical pI, included amino acid composition, *in-vivo* half-life, molecular weight, aliphatic index and grand average of hydropathicity (GRAVY). SOLpro (http://scratch.proteomics.ics.uci.edu/) was employed to evaluate solubility of vaccine construct.

3.10.8 Vaccine-Structure Modelling

I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) based on homology modelling predicts tertiary structure of the multi-epitope construct. I TASSER builds three-dimensional structure from convergence of structural assembly simulations and multiple threading alignments (Roy, Kucukural & Zhang, 2010). In the last five CASP experiments I-

TASSER is ranked the best among different servers for protein structure prediction (Yang *et al*., 2015).

3.10.9 Vaccine Structure Refinement

Obtained 3D model was refined in two stages, first through ModRefiner (https://zhanglab.ccmb.med.umich.edu/ModRefiner/) and later further refined by the GalaxyRefine (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE). Protein structure is constructed and refined from Cα traces through energy minimization at atomic level by ModRefiner. Such step performs in improvement of both local and global structures with better hydrogen bonding systems, less atomic level overlaps and more accurate side chains arrangements (Xu & Zhang, 2011).

Refinement of tertiary model was done webserver GalaxyRefine (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE). This method improves the quality of local and global structures. Whole protein structure was refined by both aggressive and mild relaxation method and a refined model was generated with no more structural deviation. Refinement of the structure was obtained by repacking and rebuilding of side chains. The output gives five model structures in which first model was obtained through structural perturbation applied on the clusters of side chains and 2 to 5 models were proposed by vigrous perturbation of secondary structural elements and loops (Heo, Park, & Seok 2013).

3.10.10 Tertiary Structure Model Validation

During model validation the critical steps are performed to detect potential errors in generated protein 3D model (Khatoon, Pandey & Prajapati, 2017). Qualitative analysis of protein model was carried out by ProSA web (https://prosa.services.came. sbg.ac.at). This server is employed for structure authentication through NMR spectroscopy, X-ray study and hypothetical estimations. ProSA displays the overall quality score of input structure with respect to the all known protein structures. In case, the analysed score falls outside the native protein structures score, it mostly likely contains an error (Wiederstein & Sippl, 2007). Rampage webserver (http://mordred.bioc.cam.ac. uk/rapper/rampage.php) measures protein model quality by Ramachandran plot. This algorithm is used improvement in high resolution structures. It predicts psi and phi angels of amino acids present for allowed and disallowed regions on the van der Waal radius of the side chains of amino acids (Lovell, Davis & Arendall, 2003).

3.10.11 Molecular Docking between Designed Multi-epitope Construct and Immune Receptors

Interaction between specific immune receptor and antigenic molecule plays a significant role in production of specific immune response. Molecular docking was performed by HADDOCK server between TLR-2 (PDB ID 2Z7X) receptor and multi-epitope vaccine was performed to find binding between receptor and ligand and generation of respective immune response. TLR-2 is reported to play significant role in urinary tract infections (Hayes & Abraham, 2016). To predict such binding, data-driven docking of constructed chimeric protein-TLR2 was performed. Data-driven docking predicts overall shape of complexes and interacting subunits involved in the interaction. To find out the interface residues of TLR2 and constructed vaccine online server CPORT was used (https://milou.science.uu.nl/services/CPORT/) (de Vries & Bonvin, 2011).

Docking simulations between structural complexes was performed by HADDOCK webserver (http://haddock.science.uu.nl/services/HADDOCK2.2) (van Zundert *et al*., 2016). HADDOCK (High Ambiguity Driven DOCKing) contains large python scripts written on basis of NMR and crystallography procedures for structure findings. After evaluation of proper docking calculation and complex orientation HADDOCK passes resultant complex to final MD simulation stage for slight refinement in an 8 Å interface of TIP3P water molecules. Intra- and intermolecular forces are measured by employing van der Waals and electrostatic energy systems and by employing the pairwise backbone root mean square distance at the active surface residues final structures are obtained. To visualize the refined structure of vaccine-TLR chimera was used. PDBsum was used to map the interacting residues between both of them.

Results

4.1 Genome Annotation

Obtained sequence reads were assembled using Patric tool, which resulted in 4638970 bp (strain EC-562.48468) sized draft genome. This genome contained 404 contigs. The GC content of the genome was 51.5 %. Evaluation of the genome completeness and presence of any contamination showed it was >99.9% complete genome and minimal contamination was observed. *E. coli* EC-562.48468 genome consisted of 5095 coding sequences (CDS), tRNA (62) and rRNA (5). The circular representation of this genome is depicted in Figure 1. The order of contigs was corrected in draft genome by using *Escherichia coli* str. K-12 substr. as reference genome.

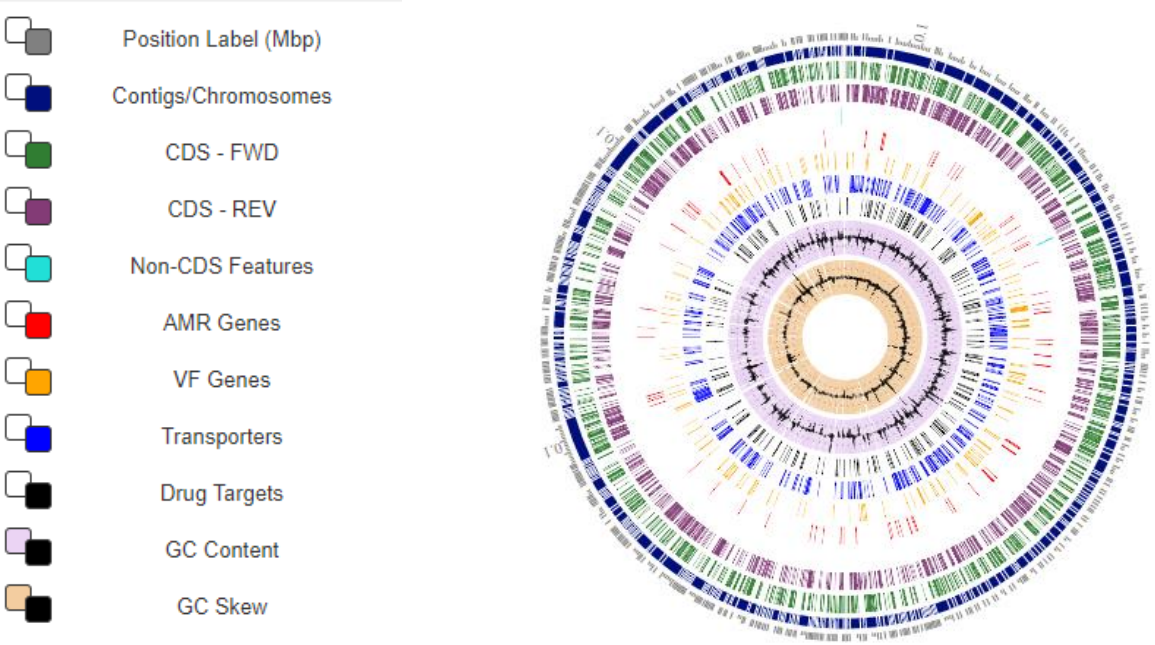


Figure 5 Circular plot of Escherichia coli EC-562.48468 depicting annotation using PATRIC, v.5.43 Track Order outwards toward the center of the circular genome depiction is: Forward Strand CD

4.2 Genotypic Characterization using CGE

4.2.1 Clonal Typing

In silico (MLST) analysis of seven house-keeping genes classified EC-562.48468 as ST-131 according to Achtman scheme.

Table 2 MLST done by MLST 2.0 for determination of strain sequence type

Locus	**Identity**	**Coverage**	**Alignment length**	**Gaps**	**Alleles**
adk	100	100	536	0	Adk 53
fumC	100	100	469	0	fumC 40
gyrB	100	100	460	0	gyrB 47
icd	100	100	518	0	Icd 13
mdh	100	100	452	0	mdh 36
purA	100	100	478	0	purA 28
recA	100	100	510	0	recA 29

4.2.2 PlasmidFinder from CGE

PlasmidFinder identified four plasmids. Three plasmids (IncFIA, IncFIB, and IncFII) are conjugative plasmids while ColpVC is non conjugative plasmid. It has a role in virulence factors of the pathogenic strain.

Table 3 Putative plasmid sequences identified by PlasmidFinder v1.3

Plasmid	**Identity**	**Query / Template length**	**Position in contig**	**Accession number**
ColpVC	98.96	193 / 193	1107..1299	JX133088
IncFIA	99.74	388 / 388	2551992..2552379	AP001918
IncFIB(AP001918)	99.71	682 / 682	1566595..1567276	AP001918
IncFII(pRSB107)	100	261 / 261	2182504..2182764	AJ851089

4.2.3 SNP Determination

The SNP analysis revealed that 80.36% of reference genome *E. coli* (str. K-12 substr. MG1655111) was covered by all three genomes including *Escherichia coli* USA, *Escherichia coli* UK and *Escherichia coli* India and clinical Isolate. Size of reference genome was 4641652 and clinical isolate had 3827124 valid positions as compared to reference sequence. The SNP Matrix index is shown in table 3.

Table 4 SNP score index obtained from CSI Phylogeny tool available at CGE

	***Escherichia coli* Pakistan**	***Escherichia coli* USA**	***Escherichia coli* UK**	***Escherichia coli* India**	**Reference**
***Escherichia coli* Pakistan**	0	24796	531	440	42265
***Escherichia coli* USA**	24796	0	24667	24743	42548
***Escherichia coli* UK**	531	24667	0	335	42355
***Escherichia coli* India**	440	24743	335	0	42383
Reference	42265	42548	42355	42383	0

4.3 Identification of Antibiotic Resistance Determinants

Different databases were used for identification of Antibiotic Resistance Determinants present within the whole genome sequence. The results of the antibiotic determinants identified by different databases are given below.

4.3.1 Antibiotic Resistance Genes Identified by CARD

Antibiotic resistance determinants present within assembled and annotated genome were identified by Resistance Gene Identifier. Increased in antibiotic resistance is resulted due

to various antibiotic efflux pumps present within the annotated sequence (Jia *et al*., 2017). Table 2 summarizes the antibiotic resistant genes present within the annotated sequence. Identified antibiotic resistance genes were shown in the form of RGI wheel (figure 6) and tabular form (table 2).

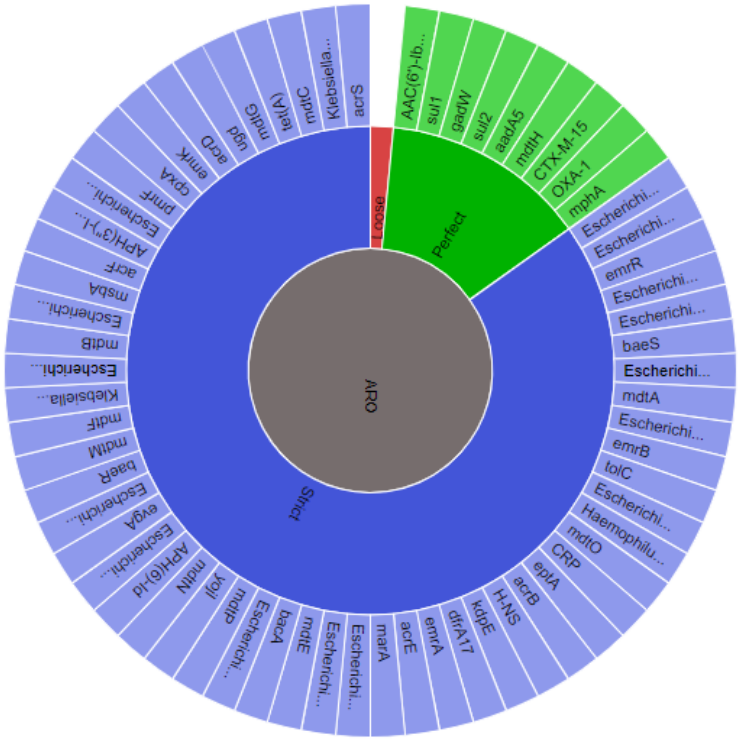


Figure 6 RGI Wheel showing Antibiotic Resistance genes identified in the sequence with Perfect and strict Hits

Table 5 Antibiotic resistance genes identified by giving DNA sequence as input on RGI with default perfect and strict hits only

Perfect hits					
ARO_name	**Model_type**	**ARO_category**	**bit_score**	**Best_Hit _e_value**	**Percentage_ Identity**
mphA	Protein homolog model	antibiotic inactivation, Macrolid Resistance protein	594.7	2.4e-171	100
AAC(6')-Ib-cr	Protein homolog model	antibiotic inactivation, fluoroquinolone antibiotic, aminoglycoside antibiotic	405.2	1.8e-114	100
gadW	Protein homology model	antibiotic efflux	481.5	2.4e-137	100
sul1	Protein homology model	antibiotic target replacement	549.7	8.2e-158	100
OXA-1	Protein homology model	antibiotic inactivation	546.6	6.9e-157	100
mdtH	Protein homology model	antibiotic efflux	784.6	2.2e-228	100
sul2	Protein homology model	antibiotic target replacement	517.3	4.4e-148	100
CTX-M-15	Protein homology model	antibiotic inactivation	557.4	4.1e-160	100

Strict hits					
ARO_name	**Model_type**	**ARO_category**	**bit_score**	**Best_Hit _e_value**	**Percentage Identity**
eptA	Protein homology model	antibiotic target alteration	1062	9.6e-312	96.3
tet(A)	Pprotein homology model	antibiotic efflux	752.7	9.1e-219	99.74
emrR	Protein homology model	antibiotic efflux	346.7	6.7e-97	99.43
emrB	Protein homology model	antibiotic efflux	995.3	995.3	99.8
Escherichia coli EF-Tu mutants conferring resistance to Pulvomycin	Protein homology model	antibiotic target alteration	785.8	9.6e-229	99.75
emrK	Protein homology model	antibiotic efflux	297.7	3.1e-82	99.34
acrS	Protein homology model	antibiotic efflux	435.3	1.8e-123	98.64
kdpE	Protein homology model	antibiotic efflux	446	1e-126	99.55
Escherichia coli ampH	Protein homology model	antibiotic inactivation	446	1e-222	99.22
mdtA	Protein homology model	antibiotic efflux	766.1	8.3e-223	98.55
mdtP	Protein homology model	antibiotic efflux	911.8	1.4e-266	98.16
Escherichia coli mdfA	Protein homology model	antibiotic efflux	761.9	1.5e-221	97.07

Escherichia coli ampC	Protein homology model	antibiotic inactivation	682.9	7.4e-198	97.9
yojI	Protein homology model	antibiotic efflux	1056.6	4e-310	99.09
APH(3'')-Ib	Protein homology model	antibiotic inactivation	553.9	4.2e-159	99.63
cpxA	Protein homology model	antibiotic efflux	530.4	5.6e-152	100
evgA	Protein homology model	antibiotic efflux	398.7	1.7e-112	99.51
mdtC	Protein homology model	antibiotic efflux	1903.3	1.7e-112	98.73
tolC	Protein homology model	antibiotic efflux	923.3	4.8e-270	99.8
mdtE	Protein homology model	antibiotic efflux	723.4	5.7e-210	99.74
ugd	Protein homology model	antibiotic target alteration	756.1	8e-220	99.23
Escherichia coli soxS with mutation conferring antibiotic resistance	Protein homology model	antibiotic target alteration, reduced permeability to antibiotic, antibiotic efflux	219.2	9.8e-59	100
acrB	Protein homology model	antibiotic efflux	1986.5	1.2e-345	99.81
acrF	Protein homology model	antibiotic efflux	523.1	8.4e-150	99.28
mdtF	Protein homology model	antibiotic target alteration, antibiotic efflux	1934.1	1.1e-543	99.42

Escherichia coli acrR with mutation conferring multidrug antibiotic resistance	Protein homology model	antibiotic target alteration, antibiotic efflux	426.8	1.2e-567	100
acrD	Protein homology model	antibiotic efflux	1979.9	1.2e-643	99.9
*dfr*A17	Protein homology model	antibiotic target replacement	316.6	8.4e-88	99.36
Escherichia coli gyrA conferring resistance to fluoroquinolones	Protein homology model	antibiotic target alteration	1674.1	-	99.66
baeR	Protein homology model	antibiotic efflux	483.8	4.8e-138	99.58
*Escherichia coli par*C conferring resistance to fluoroquinolone	Protein homology model	antibiotic target alteration	1459.9	-	99.2
*bac*A	Protein homology model	antibiotic target alteration	523.1	-	99.27
mdtM	Protein homology model	antibiotic efflux	764.2	3.1e-222	97.07
Escherichia coli GlpT with mutation conferring resistance to fosfomycin	Protein homology model	antibiotic target alteration	920.2	3.7e-269	99.78

CRP	Protein homology model	antibiotic efflux	419.5	9.7e-119	99.52
msbA	Protein homology model	antibiotic efflux	1094.7	-	99.66
APH(6)-Id	Protein homology model	antibiotic inactivation			
Klebsiella pneumoniae KpnE	Protein homology model	antibiotic efflux	176.4	8.2e-46	82.2
acrE	Protein homology model	antibiotic efflux	730.7	3.6e-212	99.48
Haemophilus influenzae PBP3 conferring resistance to beta-lactam antibiotics	Protein homology model	antibiotic target alteration	574.7	5e-165	53.29
mdtN		antibiotic efflux	636	98.83	1.1e-183
H-NS	Protein homology model	antibiotic efflux	256.5	7.1e-70	99.27
mdtG	Protein homology model	antibiotic efflux	776.2	7.9e-226	99.75
Klebsiella pneumoniae KpnF	Protein homology model	antibiotic efflux	189.5	8.5e-50	84.4
Escherichia coli UhpT with mutation conferring resistance to fosfomycin	Protein homology model	antibiotic target alteration	893.6	3.8e-261	99.78
Escherichia coli emrE	Protein homology model	antibiotic efflux	221.1	2.7e-59	98.18

pmrF	Protein homology model	antibiotic target alteration	626.3	8e-181	99.07
baeS	Protein homology model	antibiotic efflux	902.9	6.4e-264	95.93
Escherichia coli marR mutant conferring antibiotic resistance	Protein Homology model	antibiotic target alteration, antibiotic efflux	287	5.1e-79	98.61
mdtO	Protein homology model	antibiotic efflux	1320.8	3.1e-97	98.39
emrA	Protein homology model	antibiotic efflux	739.6	7.8e-215	99.74
Escherichia coli acrA	Protein homology model	antibiotic efflux	748.4	1.7e-217	99.75
marA	Protein homology model	reduced permeability to antibiotic, antibiotic efflux	260	5.9e-71	99.21
Escherichia coli PtsI	Protein homology model	antibiotic target alteration	1089.7	2.1e-56	99.65

4.3.2 Acquired Antibiotic Resistance Genes

In clinical isolate few genes were plasmid borne shown in the table 6.

Table 6 Antibiotic resistance genes identified by Resfinder using DNA sequence as input and acquired resistance genes as default setting

Antibiotic Class	Resistance gene	Identity	Query Template length	Contig position	Predicted phenotype	Accession number
Quinolone	*aac(6')-Ib-cr*	100	600 / 600	2751744..2752343	Fluoroquinolone and Aminoglycoside resistance	DQ303918
Sulphonamide	*sul1*	100	840 / 840	1951755..1952594	Sulphonamide resistance	U12338
	sul2	100	816 / 816	1559405..1560220	Sulphonamide resistance	AY034138
Beta-lactam	*blaCTX-M-15*	100	876 / 876	2538515..2539390	Beta-lactam resistance Alternate name; UOE-1	AY044436
	blaOXA-1	100	831 / 831	2698327..2699157	Beta-lactam resistance	HQ170510

Aminoglycoside	*aac(6')-Ib-cr*	100	600 / 600	2751744..2752343	Fluoroquinolone and aminoglycoside resistance	DQ303918
	aadA5	100	789 / 789	1950420..1951208	Aminoglycoside resistance	AF137361
	aph(3'')-Ib	100	804 / 804	1560281..1561084	Aminoglycoside resistance	AF321551
	aph(6)-Id	100	831 / 831	1561090..1561920	Aminoglycoside resistance A	CP000971
Trimethoprim	*dfrA17*	100	474 / 474	1949816..1950289	Trimethoprim resistance	FJ460238
Tetracycline	*tet(A*	100	1200 / 1200	969008..970207	Tetracycline resistance	AJ517790
Macrolide	*mdf(A)*	97.81	1233 / 1233	2466973..2468205	Macrolide resistance	D16251

4.4 Determination of Insertion Sequences

Multiple prophage sequences were discovered in clinical Escherichia coli isolate as shown in table 7. Only one insertion sequence IS6 was recognized for its association with antibiotic resistance gene.

Table 7 Insertion sequences identified by IS Finder using DNA sequence as input and by using blast default setting

IS Family	**Bp Length**	**Polypeptide Length**	**Antibiotic Resistant Genes**	**Origin**
IS4-ssgr-IS4	1426 bp	442 aa	N/A	*E. coli*
ISL3	1722bp/ 1885bp/ 1316 bp	539 aa/ 1221 aa/ 406aa	N/A	*E. coli/ E. coli/ K. oxytoca*
IS4_ssgr_IS50	1534 bp	476 aa	N/A	*E. coli*
IS3_ssgr_IS3	1258 bp	387 aa	N/A	*E. coli*
IS6	820 bp/880 bp/ 891 bp	234 aa/254 aa/ 250 aa	*bla*CTX-M15	*S. panama/ M. fortuitum/ E. coli*
IS200_IS605_ssgr_IS200	709 bp	151 aa	N/A	*S. typhimurium*
IS481	1020 bp/1229 bp	323 aa/384 aa	N/A	*E. coli*
Tn3	3854 bp/4950 bp	213 aa/286 aa/991 aa	N/A	*E. coli/ X. campestris*

4.5 Determination of Virulence Factors

Multiple virulence factors genes were present in clinical isolate. These genes are involved in pathogenesis of *Escherichia coli.*

Table 8 Virulence Factors identified by Virulence Factors Database using DNA sequence as input and by using blast default setting

Virulence Factors	**Genes**	**Function**
Fim (type 1)	*fimABCDEFGHI*	Colonization
Yersiniabactin	*fyuA, irp1, irp2, ybtA, ybtE, ybtP, ybtQ, ybtS, ybtT, ybtU, ybtX*	Iron Uptake
Factor adherence *E. coli*	*fdeC*	Aggregation, Colonization
Aerobactin	*iucA, iucB, iucD, iutA*	Iron Uptake
Hemin uptake	*chuA, chuS, chuT, chuU, chuV, chuW, chuX, chuY*	Heme Uptake
OmpA	*OmpA*	Adherence, Invasion

4.6 Determination of Prophage Sequences

Multiple prophage sequences were present in clinical isolate shown in table 9.

Table 9 Prophage Sequences identified by PHASTER using DNA sequence as input

Name	Start Position	End Position	Size
PHAGE_Entero_P88_NC_026014	313524	340053	26.5Kb
PHAGE_Burkho_BcepMu_NC_005882	951530	998474	46.9Kb
PHAGE_Entero_DE3_NC_042057	690799	712424	21.6Kb
PHAGE_Entero_BP_4795_NC_004813	1566220	1578685	12.4Kb
PHAGE_Cronob_vB_CsaM_GAP32_NC_019401	2503337	2511795	8.4Kb
PHAGE_Entero_P88_NC_026014	3289332	3320235	30.9Kb
PHAGE_Entero_HK97_NC_002167	3887767	3899707	11.9Kb

5 .1 Preliminary Evaluations and Retrieval of Protein Sequences

Sequence of both proteins (AGA03821.1 and ANK05584.1) were obtained from NCBI database and used for designing a multi-subunit vaccine against Uropathogenic *E. coli*. Localization analyses predicted these proteins as extracellular. Both these proteins showed no significant similarity with Non-Host homologous proteins. Functional sequences for these proteins were subjected to T-cell and B-cell epitope prediction.

Table 10 Characteristics of Potential Vaccine Candidates

Protein Accession	Protein Note	Localization	Transmembrane Helices	Non-Host Homologous Proteins	Antigenicity	Molecular Weight
AGA03820.1	FimH	Extracellular	0	No significant similarity	0.65	31.44 kDa
ANK05584.1	fliC	Extracellular	0	No significant similarity	0.79	57.54 kDa

6.1 B-cell Epitope Prediction

Antigenic determinants of varying lengths were obtained from multiple servers (BepiPred 2.0, SVMTrip, and IEDB). Some of resultant B-cell determinants were also predicted to be T-cell determinants and successfully utilized in chimeric structure construction (Table 11).

6.2 Cytotoxic T Lymphocytes Epitopes

By using default threshold settings, total 25 CTL (9-mer) antigenic determinants were obtained for these two proteins. From these, 4 epitopes were selected based on their high scores. These antigenic determinants were used in constructing chimera (Table 11).

6.3 Helper T Lymphocytes Epitopes

With NetMHCII 2.2 web server high binding HTL determinants for human alleles HLA-DQ, -DP and -DR were obtained based on their IC_{50} scores. For final vaccine peptide a total of 4 high binding HTL antigenic determinants were selected. (Table 11).

Table 11 List of Selected T and B Cell Epitopes

Table 11 List of Selected Antigenic Determinants

Cytotoxic T lymphocytes epitopes	Helper T-cell epitopes	B cell epitopes
TSETPRVVY	ANVYVNLAPAVNVGQ	LLMGWSVNAWSFACKTANGT
VTLQRGSAY	NVYVNLAPAVNVGQN	
TTDNTTKLY	KAQIIQQAGNSVLAK	KALDEAISQIDKFRSSLGAI
SSALTSNDY	KAQIIQQAGNSVLAK	

7.1 Multi-subunit Vaccine Construct

The total number of predicted antigenic determinants used in chimera designing were 2 linear B-cell, 4 CTL and 4 HTL antigenic determinants. Resulted peptide sequences containing T-cell and B-cell determinates were joined by AAY and GPGPG linkers. The CTB (adjuvant) was added to the N terminus of constructed vaccine through EAAAK linker to enhance the immune response against specific antigen. Furthermore, 6xHis tag was added to peptide C-terminus to assist in protein identification and purification. Final vaccine construct had 290 amino acids which was obtained by merging 10 peptide sequences.

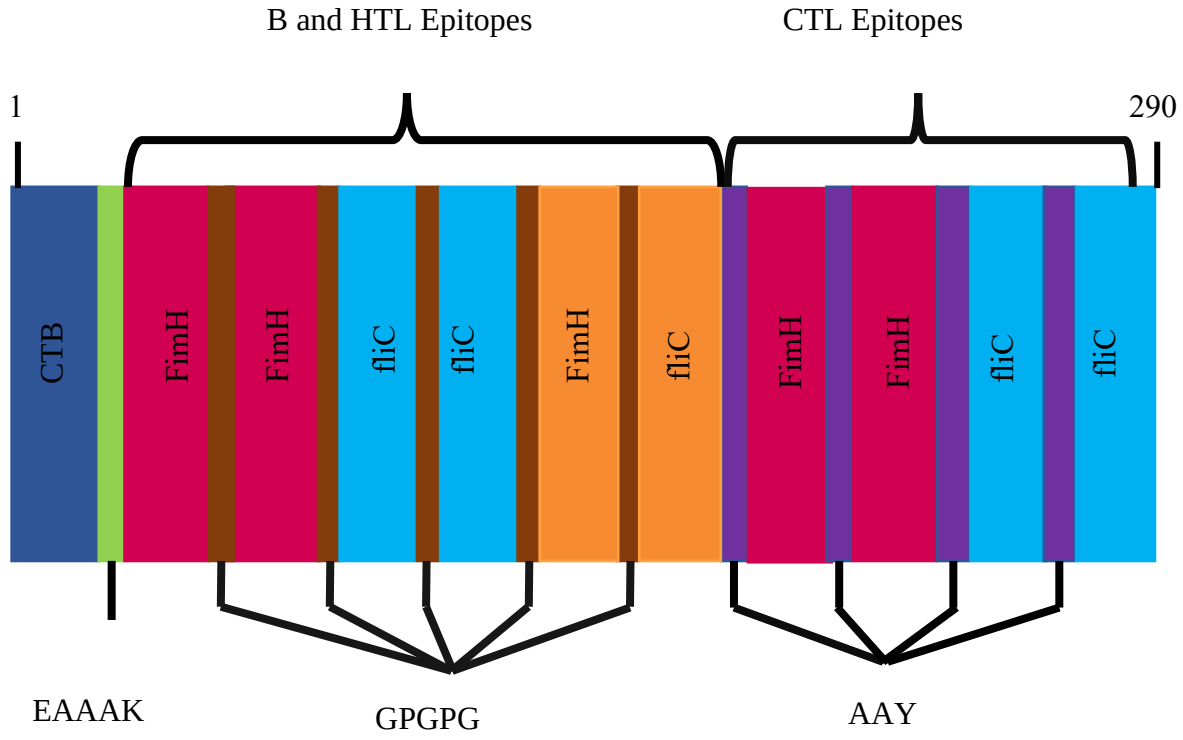


Figure 7 Multi-epitope vaccine construct

7.1.1 Antigenicity, Allergenicity and Solubility Prediction

ANTIGENpro and VaxiJen 2.0 were employed for prediction of vaccine construct antigenicity. The antigenicity score predicted from ANTIGENpro was 0.94 which showed vaccine construct has antigenic nature. By using VaxiJen 0.57 score was obtained which indicates vaccine construct has good antigenic nature. Threshold score of VaxiJen is 0.4. Both server results showed that peptide construct is a good antigen. AllerTOP v. 2.0 and AllergenFP v.1.0 server found vaccine construct has non-allergenic behavior.

7.1.2 Sequence Characteristics of Design Vaccine Construct

Chemical and Physical properties of the selected proteins were analysed. Parameters such as a grand average of hydropathicity which indicates solubility of the proteins (negative GRAVY (hydrophilic), (positive GRAVY (hydrophobic)), pI (Isoelectric point), aliphatic index and instability index are present below (Table 12).

Table 12 Physiochemical Properties of Multi-epitope vaccine

Number of amino acids	314		
Molecular weight	32.54		
Theoretical pI	8.96		
Amino acid composition	Ala (A)	40	13.8%
	Arg (R)	6	2.1%
	Asn (N)	19	6.7%
	Asp (D)	8	2.8%
	Cys (C)	3	1.1%
	Gln (Q)	13	4.6%
	Glu (E)	10	3.4%
	Gly (G)	29	10.2%
	His (H)	9	3.2%
	Ile (I)	17	5.9%
	Leu (L)	19	6.7%
	Lys (K)	17	5.6%
	Met (M)	5	1.8%
	Phe (F)	5	1.8%
	Pro (P)	16	5.6%
	Ser (S)	18	6.3%
	Thr (T)	22	7.7%
	Trp (W)	3	1.1%
	Tyr (Y)	14	4.9%
	Val (V)	18	6.3%

	Pyl (O)	0	0.0%
	Sec (U)	0	0.0%
Total number of negatively charged residues (Asp +Glu)	18		
Total number of positively charged residues (Arg + Lys)	23		
Atomic composition	Carbon C		1353
	Hydrogen H		2120
	Nitrogen N		378
	Sulfur S		8
	Oxygen O		413
Total atoms number	$C_{1353}H_{2120}N_{378}O_{413}S_8$		
Extinction coefficients	37485		
Estimated half-life	30 hours in mammalian cells >20 hours in yeast cell >10 hours in *Escherichia coli*		
Instability index	This classifies the protein as stable		
Aliphatic index	78.86		
Grand average of hydropathicity (GRAVY)	-0.214		

Solubility	0.87

8.1 Tertiary Structure Modelling

I-TASSER server evaluated five tertiary structure models of the designed chimeric protein based on 10 threading templates of which 4l6t, 1ltrA and 3chbD were the best. Based on the *Z-score* values (from 0.86 to 9.12) the selected 10 templates showed better alignment scores. Predicted models had *C-score* varies from -4.37 to -2.84. The *C-score* range is generally between -5 to 2. Higher C values indicate more confidence. Model with highest C-score value was submitted for refinement (Fig.8).

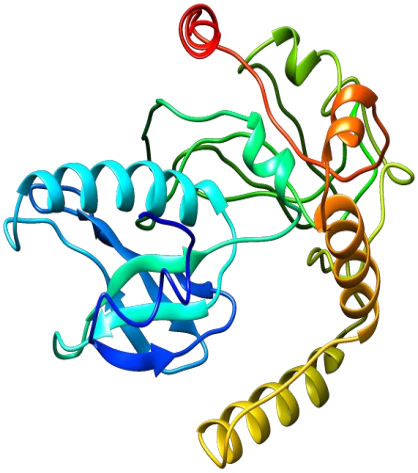
8.2 Tertiary Structure Refinement

Initial vaccine structure was refined by ModRefiner server and further refined by GalaxyRefine server which resulted in five structures. Based on the quality scores the best model was selected. Model 2 was the best based on various properties i.e. RMSD (2.203) and MolProbity (1.687). The clash score was 3.6, poor rotamers score was 0.0 and Ramachandran plot was 90.3%. This model was selected as final vaccine construct used in further analysis (Fig. 8).

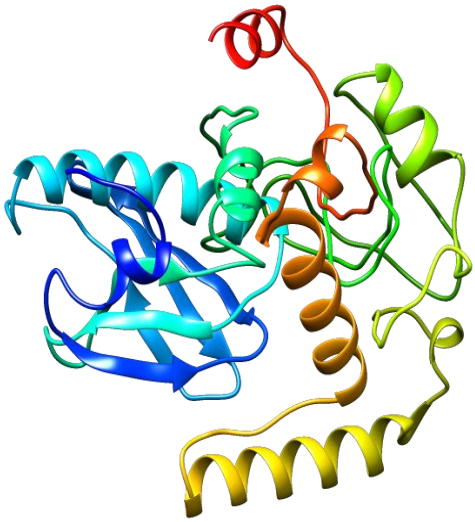
8.3 Tertiary Structure Validation

Ramachandran plot results showed that 90.3% of residues of proteins are present in favoured regions. This is consistent with the 90.3% score obtained from GalaxyRefine 2. Moreover, 6.6% residues were present in allowed regions and 1.1% were present in outlier region (Fig. 8). Significant errors present in the model structure were verified by ERRAT and ProSA web. ProSA web gave Z score about -6.1 for constructed model which was lying within the score range of the native protein models (Fig. 8) while the refinement model had an overall quality factor of 80.85 (not shown). Above findings concluded the predicted tertiary model was highly stable and of good quality.

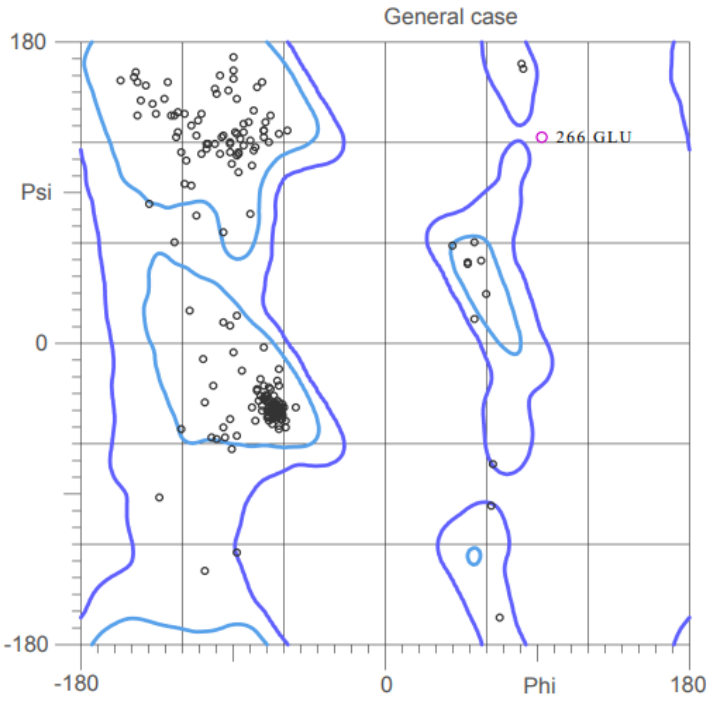
(A)



(B)



(C)



(D)

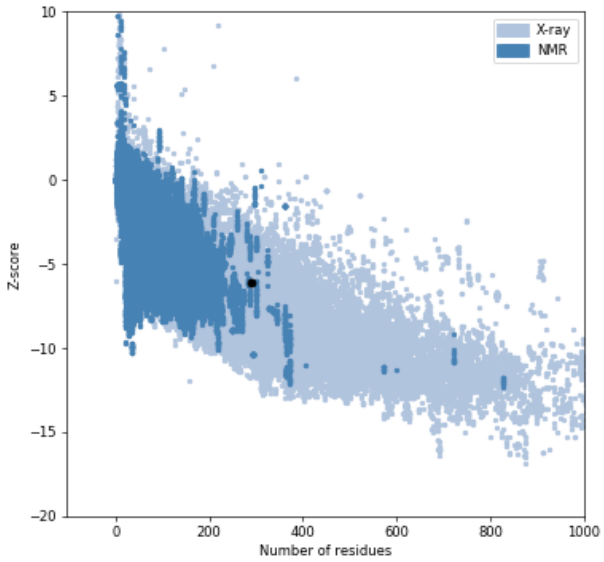


Figure 8 (A) Final 3D structure of multi-subunit vaccine based upon homology modelling by I-TASSER. (B) Refined 3D model by GalaxyRefine (C) Ramachandran plot analysis showing 90.3% of protein residues in favoured region and (D) ProSA-web plot showing -6.1 score.

9.1 Docking of Multi-epitope Vaccine with Immune Receptor (TLR-2)

HADDOCK was employed for docking of multi-epitope construct with TLR2. Given active interface amino acid residues were provided by CPORT: I541, R540, K539, P575, I573 and D543 from TLR-2 and S157, N159, A160, F163, A164, V158, P128, Q126, and C165 from chimeric protein were predicted and used further to drive molecular docking. In case of TLR-2 docking, 144 structural complexes grouped in 21 clusters by HADDOCK, representing 36.0 % of the water-refined models generated by HADDOCK. Top cluster is the lowest HADDOCK scored, -136.5 +/- 7.3 was the most reliable cluster. This representative model was forwarded towards refinement. The HADDOCK refinement server grouped 20 structures in 1 cluster that showed 100.0 % of the water-refined models were generated by HADDOCK.

Table 13 Statistics observed from refined cluster is given below

Parameters	**Values**
HADDOCK score	-140.5 +/- 2.9
Cluster size	20
RMSD from the overall lowest-energy structure	0.3 +/- 0.2
Electrostatic energy	-82.4 +/- 3.4
Desolvation energy	-12.3 +/- 4.9
Restraints violation energy	0.0 +/- 0.00
Buried Surface Area	2206.9 +/- 52.3
Z-Score	0.0

Further detailed analysis of docking results showed good interaction between TLR-2 and vaccine construct (Fig.). To get complete insights into intermolecular forces and interface residues between bounded complex, the PDBsum was used. A total of 19 residues of TLR-2 were associated with the total 20 interface residues of the multi-epitope vaccine. The interface area of the TLR-2 was found to be 1092 Å2 while that of vaccine was found to be 1120 Å2.

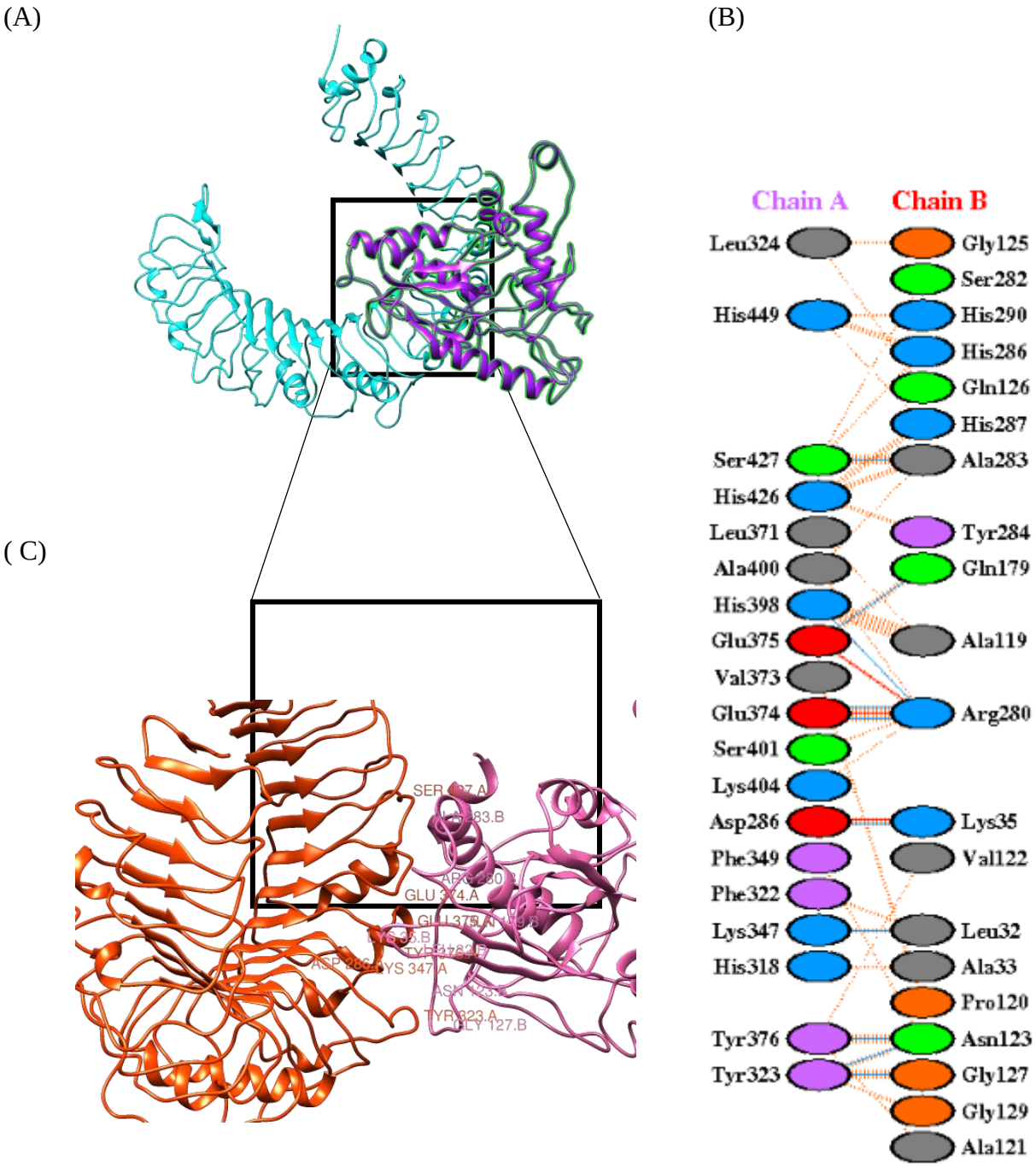


Figure 9 Molecular docking exhibiting vaccine construct-TLR2 docked structure. Vaccine construct in purple colour while TLR2 in blue colour; (B) Residues present between interacting TLR2 and vaccine construct; (C) few prominent hydrogen bonds

Our predicted vaccine formed a total of 10 hydrogen bonds [Chain A (TLR2)-B(Vaccine); 427-283, 374-280, 398-280, 286-35, 347-32, 376-123 and 323-127]. Structural analysis revealed that ASP286-LYS35 formed a hydrogen bond at distance of 2.74 Å; similarly, TYR323-THR167 formed a hydrogen bond at a distance 3.21Å. Likewise, TYR323-GLY127 forms hydrogen bond at 2.90 Å, LYS347-LEU32 at 2.86 Å, GLU374-ARG280 at 2.78 Å, GLU374-ARG280 at 2.74 Å, GLU375-GLN179, TYR376-ASN127 at 2.87 Å, HIS398-ARG280 at 2.88 Å and SER427-ALA283 at 3.30 Å.

9.2 Docking of Multi-epitope Vaccine with Immune Receptor (TLR-5)

HADDOCK was employed for analysis of multi-epitope construct with TLR5. Given active interface amino acid residues were provided by CPORT: 27D, 43Q, 92R and 70E and S157, N159, A160, F163, A164, V158, P128, Q126, and C165 from chimeric protein were predicted and used further to drive molecular docking. In case of TLR-5 docking,173 structural complexes were grouped in 27 clusters by HADDOCK, representing 43.25 % of water-refined models generated by HADDOCK. Top cluster is the lowest HADDOCK scored, -136.5 +/- 7.3 was the most reliable cluster. This representative model was forwarded towards refinement. The HADDOCK refinement server grouped 20 structures in 1 cluster that shows 100.0 % water-refined models are generated by HADDOCK.

Table 14 Statistics observed from refined cluster is given below

Parameters	**Values**
HADDOCK score	-151.5 +/- 4.2
Cluster size	20
RMSD from the overall lowest-energy structure	0.3 +/- 0.2
Electrostatic energy	-94.5 +/- 2.4
Desolvation energy	-24.1 +/- 5.3
Restraints violation energy	0.3 +/- 0.21
Buried Surface Area	2713.8 +/- 27.8
Z-Score	0.0

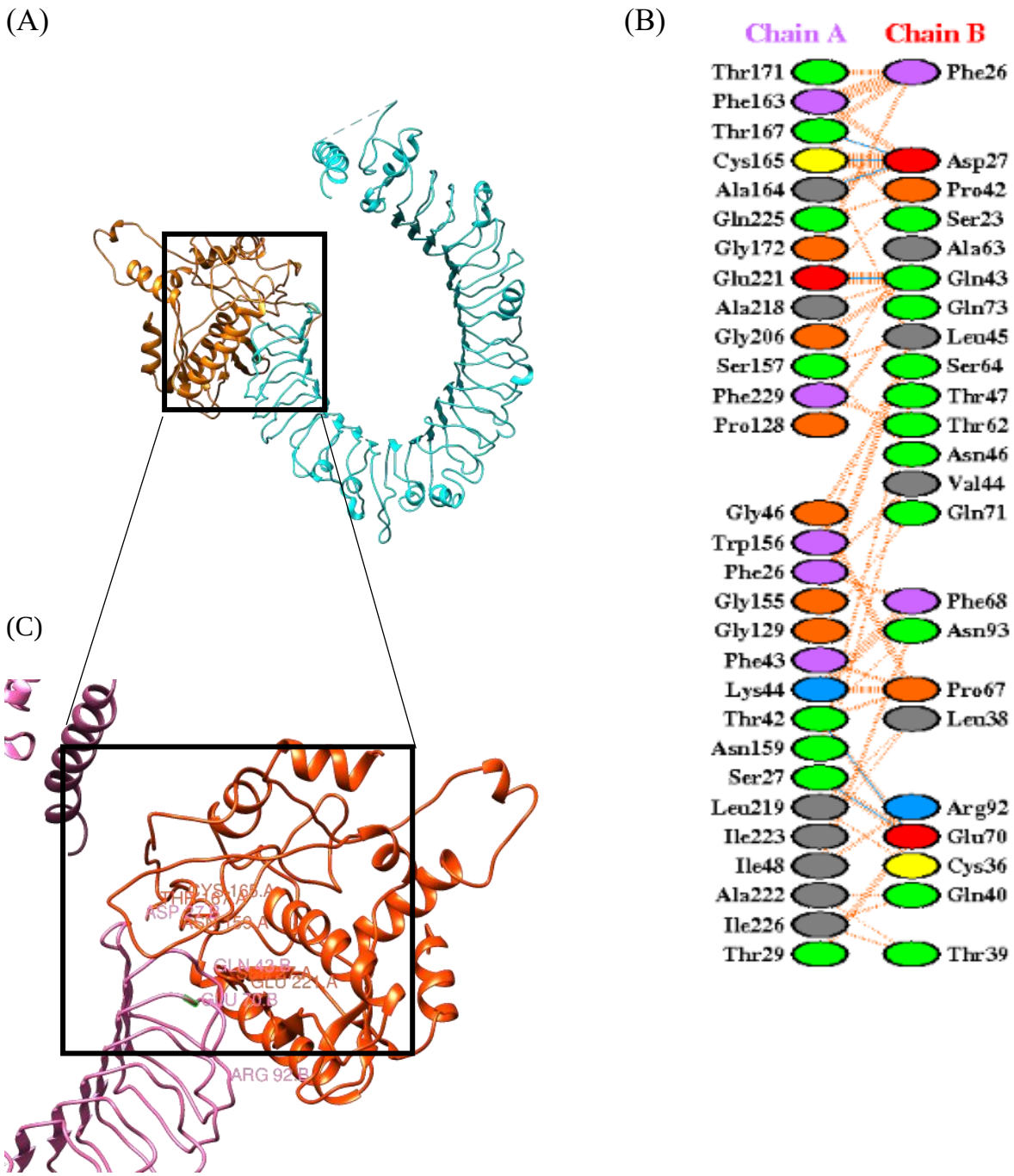


Figure 10 Molecular docking exhibiting vaccine construct-TLR2 docked structure. Vaccine construct in orange colour while TLR2 in blue colour; (B) Residues present between interacting TLR2 and vaccine construct; (C) few prominent hydrogen bonds

Our predicted vaccine formed 6 hydrogen bonds [Chain A (TLR2) – B (Vaccine); 167-27, 165-27, 165-27, 221-43, 159-92, 27-70]. Structural analysis revealed that SER27-GLU70 at 2.63 Å, THR42-GLU70 at 2.59 Å, ALA164-ASP27 at 2.71 Å, CYS165-ASP27 at 2.78Å, THR167-ASP27 at 3.05Å formed a hydrogen bond similarly, GLU221-GLN43 at 2.73 Å.

Discussion

The present work aimed to identify genetic determinants and genotypes which are present in the genome, associated with pathogenicity and success of clinical *Escherichia coli*.
E. coli 101-2 was isolated in June 2013 from a patient who visited Pakistan Institute of Medical Sciences, Islamabad, Pakistan. Antimicrobial susceptibility assay is used to classify a pathogen as either MDR, XDR or PDR depending upon the number of antibiotics to which an organism is resistant (Magiorakos *et al*., 2012). The selected isolate exhibited resistance to more than three antibiotic classes that were tested including Penicillin, Aminoglycosides, Cephalosporins, Fluroquinolones and Sulphonamides while it was susceptible to the Carbapenem class of antibiotics. This highlights the emergence of a multidrug resistant *Escherichia coli*.

ESBL are plasmid borne enzymes that have ability to hydrolyze oxyimino-*β* lactams i.e. 3rd generation cephalosporins. Such plasmids also contain genes against other antibiotics including chloramphenicol, sulphonamides, tetracycline, trimethoprim and aminoglycosides. ESBL-producing *Escherichia coli* isolates can be susceptible and resistant to ceftazidime. Subsequent study also revealed that *Escherichia coli* carrying ESBL along with other resistance genes present on plasmid sequences. Most of the resistance acquired genes are present on the plasmid. Resistance and virulence determinants are mostly acquired through three ways i.e. conjugation or transformation and transduction. One of the studies exhibited that resistance determinants shared between the closely related resistant species such as *S. enterica* and *E. coli* are somewhat like the transconjugant having resistance similar to the donor group which established through the conjugation experiments (Morita *et al*., 2010).

To identify the resistance determinants, Beta-lactamase (*bla$_{CTX\text{-}M}$*, *bla$_{TEM}$* and *bla$_{SHV}$*) resistance determinants were detected via PCR. Only *blaCTX-M* showed positive PCR results. According to one study *bla*$_{\text{CTX-M-15}}$ was highly prevalent genotype in *E. coli* isolates causing urinary tract infections. Due to the increased usage of cephalosporins ESBL have been emerged worldwide (Kumar *et al*., 2014). The increasing *bla*$_{\text{CTX-M-15}}$ allele has ability to hydrolyze aztreonam, cefotaxime and ceftazmidime. *bla*$_{\text{CTX-M}}$ containing clinical isolates are resistant against wide range of antibiotics (beta-lactams) against those not containing a *bla*$_{\text{CTX-M}}$ gene.

This gave an indication that presence of *bla*$_{\text{CTX-M}}$ gene can act as a good biomarker for implementation in routine susceptibility testing protocols.

To better understand the resistance determinants whole genome sequencing was performed by Illumina which could help to explain the extended antimicrobial resistance of clinical isolates. We obtained 404 contigs that were annotated by PATRIC and Prokka. *Escherchia coli* is among one of the highly studied pathogens. Its genome size varies from 4.6 Mbp to 5.9 Mbp with 4200 to 5500 genes and GC content of 50.6%. *Escherichia coli* are considered to have a large variable genome with high horizontal gene transfer and homologous recombination, which is often increased by host's immunocompromised conditions (Stecher *et al*., 2012). These findings confirmed our strain shows similarity to *Escherichia coli*. Extraintestinal *E. coli* genomes have larger genome than commensal or K12 isolates because they have genes involved in survival in sites other than intestinal tract. Genomes of UPEC isolates CFT536, UTI89 and CFT073 contains 8-22% ORFs and these genomes are 6-13% larger than K-12 reference strain (Brzuszkiewicz *et al*., 2006).

Each *E. coli* pathotype has unique characteristic virulence and phylogenetic pattern which allow them to invade and colonize their host (Bielaszewska *et al*., 2014) . *E. coli* sequence type 131 is a pandemic and multidrug resistant UPEC clone which causes community and hospital acquired antimicrobial resistant infection. In one study EC958 was extracted from patient urine sample suffering with community acquired UTI reported as a representative strain with ST131 and its produced *CTX-M-15* lineage that was resistant to ciprofloxacin and susceptible to gentamicin (Woodford, N *et al*., 2007). For the treatment of UTIs frequently prescribed antibiotics are fluroquinolones and sulphamethoxazole-trimethoprim but resistance to these antibiotics is frequently increasing. Alternative therapeutic agents such as β-lactam inhibitors, fosfomycin and nitrofurantoin can be prescribed. Resistance of fluoroquinolones in ST131 strains is high (Ali, I *et al*., 2019).

PlasmidFinder server found multiple plasmid sequences within the WGS data provided. Plasmid sequences for IncFIA, IncFIB, IncFII and ColpVC were obtained in analysis. No antimicrobial gene was identified on the contigs carrying these replicons which can be explained by the lack of such genes on plasmids or inability of the available sequencing techniques or methods used to assemble complete genome (Carattoli *et al*., 2014). Plasmids that belong to incompatibility group F are conjugative plasmids with size ranging from 45 to 200 kb and low copy number. These plasmids encode for several replicons including FIA, FIB

and FII (Rozwandowicz *et al*., 2018). However, the selected *E. coli* strain harbour *CTX-M-15* gene. On study exhibited that *E. coli* ST 131 strains harbour ESBL encoding IncF plasmid (Nicolas-Chanoine *et al*., 2014).

Uropathogenic *E. coli* causes 150 million UTI infections annually worldwide. Community-acquired UTIs are mainly caused by UPEC (80–90%) (Flores-Mireles *et al*., 2015). Many virulence factors i.e. surface polysaccharides, iron quenching system, flagella, adhesins and toxins are identified in UPEC (Bien *et al*., 2012). Multiple virulent factors are source of UTI infections through UPEC (Hannan *et al*., 2012). Other than Uropathogenic *E. coli* urinary tract infections is also caused by *K. pneumoniae* (7%), *P. mirabilis* (5%), *Streppococcus bovis*, *Pseudomonas aeruginosa*, *Enterococcus faecalis*, and *the fungus Candida albicans* (Parish and Holliday, 2012). There are present fimbrial projections on the surface of bacteria that allows attachment to urinary tract mucosa, gastrointestinal and respiratory cells. Additional virulence factors of *Escherichia coli* have an ability to capture iron from surrounding medium through yersiniabactin and enterobactin. Both are less molecular weight compounds which take up iron present in host cell proteins. Recent study has shown that enterobactin has higher iron binding capacity than aerobactin (Watts *et al*., 2012). Yersiniabactin plays effective role in sequestrating host-derived copper, biofilm formation and protection against intracellular killing (Chaturvedi *et al*., 2012). All these factors are potential components for research and development of new therapeutic (Klemm *et al*., 2010).

Many prophage sequences were observed in clinical isolate some of them were complete or intact prophage sequences while remaining were incomplete phages. Incomplete phages suggest that they can be cryptic phages which no longer encode an integrase gene (Colavecchio *et al*., 2017). Prophage sequences as result of genomic rearrangements and gradual deterioration become trapped to the chromosomes and become inactive in phage particle production, cell destruction and plaque formation. While active prophage shows prophage capable of forming plaques. On study exhibited that poorly studied cryptic prophages increase bacterial growth in medium, provide enhanced protection against β-lactam and quinolone antibiotics, protection against acids, oxidative stress and biofilm formation. Such prophage genes provide affect on cell physiology (Wang *et al*., 2010).

For studying bacterial epidemiological typing whole genome sequencing can detect a single base change present between two genome and concentrate on exploitation and identification of SNPs to differentiate one isolate from other (Pallen *et al*., 2010). Least SNPs were detected

between *Escherichia coli* isolate from Pakistan and India. This might be due to the similar ST observed for both the strains.

A novel strategy for generation of specific immune response and evading response against unlikable antigenic determinants (epitopes that might drive immune modulating or immunopathogenic response) involve epitope-based vaccines. Possible advantages of multiepitope vaccine include an ability to evoke immune response against conserved epitopes and provide more safety (Zhou *et al.,* 2009). Subsequent study focused development of potential multi-epitope vaccine design against Uropathogenic *Escherichia coli* by using two highly conserved proteins involved in UTI infection and pathogenesis (fimH and fliC) (Brien, V.P.O 2016). In immunomic studies the selected proteins displayed potential to be vaccine candidates. One study suggested 99.9% reduction in murine bladder colonization upon inoculation of FimH, type 1 fimbrial adhesin, bounded to periplasmic chaperone, FimC. FimH gene observes the phenomena of phase variation. The prevalence of this gene is present among both commensal *E. coli* and uropathogenic *E. coli* but the expression is different which is low in case of commensal strains (Qin, X *et al*., 2013).

It was established immunity to urinary tract infections depends upon T- and B- cells response (humoral and cell mediated response). To evoke immune response earlier sensing of infection is a critical step. It has been declared by experimental bladder infection that deficiencies in the expression of TLR2, 4, 5 and 11 are detrimental to the host (Abraham *et al*., 2015, Ingersoll *et al*., 2013).

The common TLRs emerged in the urinary tract infection include TLR5 (recognizes flagellin) and TLR2 (recognizes lipoprotein). UPEC contains peritichous flagella which enhances the expression of TLR5 proteins. From evolutionary point of view bacterial flagellum is one of the virulence factors which has a contribution to several types of UTIs (Behzadi, E. 2016). TLR2 and TLR5 are considered for its role in immune modulation and cytokine production in cases related to urinary tract infections.

Predicted T and B cell epitopes from two virulent proteins were fused with suitable linkers for formation of multi-epitope vaccine. Specialized spacers were used to improve the vaccine design (Meza *et al*., 2017). GPGPG and AAY linkers were introduced between the predicted multi-epitope vaccine to formulate a peptide sequence with optimal antigenicity therefore, suggesting a rational multi-epitope vaccine design. EAAAK linker was added in vaccine design for joining adjuvant (CTB) along first B-cell epitope. Role of this linker is

reported in enhancing expression of fused protein present in designed bifunctional proteins (Zhang, L *et al*., 2018).

In order to obtain enhanced bioactivity and increased expression of fusion proteins the EAAAK linker formerly reported for bifunctional protein (Arai, R *et al*., 2001) was added between adjuvant and fused epitopes. Designed multi-epitope vaccine lacked allergenic properties which supports its potential as vaccine candidate. Multi-epitope vaccines are poor immunogens and coupling with adjuvants is required to enhance its antigenicity. The proposed protein structure exhibited high antigenicity score with or without adjuvant sequence against selected adjuvant CTB. Such results comprehend it may be suggestible to express multi-epitope vaccine without adjuvant employed here and evaluate the response with other adjuvants in various tests.

Molecular weight of proposed vaccine is 32.05 kDa and it is soluble upon its expression in *E. coli* host. It is necessary to investigate various functional and biochemical properties for its expression in *E. coli* and assists in purification of expressed multi-subunit vaccine (Khatoon *et al*., 2017). The isoelectric point of vaccine predicted to be 8.90 indicating its basic nature. Instability index showed that protein will be stable which strengthened its potential use. Potential hydrophobicity exhibited by aliphatic index confirming its hydrophobic nature. Designed protein model was thermally stable at various temperatures. Ramachandran plot verified that large number of protein amino acids were present on favourable region exhibiting no steric clashes. Thus, confirming its satisfactory quality.

To predict interaction between the chimeric vaccine peptide and TLR2 and TLR5 molecular docking was employed since TLR5 agonist was used as adjuvant in the proposed chimeric protein. Binding energies calculated from chimeric protein interaction with immune receptor showed that proposed vaccine will enhance the effective immune response.

Conclusion

Present study focused on characterizing a clinical isolate which was highly resistant Escherichia coli isolated from urine sample. The isolate was phenotypically characterized by antibiotic susceptibility assay by using antibiotics which showed the isolate to be highly resistant. Susceptibility assay performed against commonly used antibiotics showed that our strain was resistant against fluroquinolones and 3rd generation cephalosporins.

Genotypic characterization was done by identification of ESBL genes through PCR and whole genome sequencing. Whole genome sequence was annotated by PATRIC and RAST. *Insilico* MLST and SNP analysis was performed by web-based tools available at CGE. Resistance determinants were identified by CARD and ResFinder. Various resistance genes including quinolones and β-lactamase were also present in clinical isolate.

To curb the urinary tract infection introduction of novel control method is necessary. By applying immunoinformatics tools a rational multiepitope construct was designed that will evoke different types of immune responses (innate and adaptive). Proteins selected in this study express in specifically pathogenic *Escherichia coli*. This work warrants more experimental validations to signify this study. It is expected that suggested vaccine construct will depict successful results against control of UTI.

Future Prospect

Escherichia coli, the etiological agent of Urinary tract Infection is a highly infectious is a highly infectious pathogen that is globally prevalent. The current study exhibits the multidrug resistant *Escherichia coli* that exhibited resistance to all clinically prescribed antibiotics.

The present study can be furthered advanced in various ways. Firstly, the selected clinical isolate showed resistance to antibiotics prescribed for urinary tract infection. This calls for detailed surveillance of antimicrobial resistance against fluroquinolone and higher generation of cephalosporins drugs in *Escherichia coli* present within endemic regions. Increased resistance against these antimicrobial agents could be helpful to indicate the presence of novel antimicrobial genes present within the genome of *Escherichia coli* along with novel virulence determinants.

Secondly, strategy can be also be devise that includes the use of UPEC-specific bacteriophages that could be helpful to reduce burden to these pathogenic strains without affecting normal microbiota which could be another effective way to combat antibiotic resistance.

References

Abraham, S. N. & Song, J. 2008. Innate and adaptive immune responses in the urinary tract. *Eur J Clin Invest*, 38 (S2), 21–28.

Abraham, S. N., & Miao, Y. (2015). The nature of immune responses to urinary tract infections. *Nat*, 15(10), 655–663.

Adamus-Białek, W., Baraniak, A., Głuszek, S., Gad, B., Wróbel, K., Bator, P., Majchrzak, M. and Wawszczak, M. (2018). *Mol Biol Rep*, 45(5), 1054-1055.

Ali, I., Rafaque, Z., Ahmed, I., Tariq, F., Graham, S. E., Salzman, E., … Dasti, J. I. (2019). Phylogeny, sequence-typing and virulence profile of uropathogenic *Escherichia coli* (UPEC) strains from Pakistan. *BMC infectious diseases*, 19(1), 620.

Allegranzi, B. N.S., Combescure, C., Graafmans, W., Attar, H., Donaldson, L. & Pittet, D. (2011). Burden of endemic health-care-associated infection in developing countries: systematic review and meta-analysis. *Lancet*, 377,228–41.

Anjum, M. F. (2015). Screening methods for the detection of antimicrobial resistance genes present in bacterial isolates and the microbiota. *Future Microbiol*, 10,317–320.

Arai, R., Ueda, H., Kitayama, A., Kamiya, N. & Nagamune, T. (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng*, 14, 529–32.

Behzadi, E., & Behzadi, P. (2016). The role of toll-like receptors (TLRs) in urinary tract infections (UTIs). *Cent European J Urol*, 69(4), 404–410.

Bennett, P.M. (2008). Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Br. J. Pharmacol*, 153-347.

Bennuru, S., Cotton, J.A., Jose M. C. Ribeiro, Alexandra Grote, A., Harsha, B., Nancy Holroyd, N., Mhashilkar, A., Molina, D. M., Arlo Z. Randall, A.Z., Adam D. Shandling, A.D., Unnasch, T. R., Ghedin, E., Berriman, M., Lustigman, S. and Nutman, T.B. (2016). Stage-Specific Transcriptome and Proteome Analyses of the Filarial Parasite Onchocerca volvulus and Its Wolbachia Endosymbiont, *mBio*, 7 (6).

Binkowski, T. A., Naghibzadeh, S. & Liang, J. CASTp. (2003). Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res*, 31, 3352–3355.

Brzuszkiewicz, E., Brüggemann, H., Liesegang, H., Emmerth, M., Olschläger, T., Nagy, G., Albermann, K., Wagner, C., Buchrieser, C., Emody, L., Gottschalk, G., Hacker, J. and Dobrindt. (2006). *UProc Natl Acad Sci USA*, 103(34), 12879-84.

Carattoli, A., Zankari, E., García-Fernández, A., Voldby, Larsen, M., Lund, O., Villa, L.… and Hasman, H. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *AAC*, 58(7), 3895–3903.

Chaturvedi, K.S., Hung, C.S., Crowley, J.R., Stapleton, A.E. and Henderson, J.P. (2012) The siderophore *yersiniabactin* binds copper to protect pathogens during infection. *Nat Chem Biol*, 8, 731–736.

Chen, Y.H., Ko, W.C. and Hsueh, P.R. (2013). Review Emerging resistance problems and future perspectives in pharmacotherapy for complicated urinary tract infections. *Expert Opin Pharmacother*, 14(5),587-96.

Colavecchio, A., D'Souza, Y., Tompkins, E., Jeukens, J., Freschi, L., Emond-Rheault, J. G. and Goodridge, L. D. (2017). Prophage Integrase Typing Is a Useful Indicator of Genomic Diversity in Salmonella enterica. *Front Microbiol*, 8, 1283.

Darmon, E. & Leach, D.R.F. (2014). Bacterial genome instability. *Microbiol Mol Biol R*, 78, 1–39.

De Vries, S. J. & Bonvin, A. M. CPORT. (2011): a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PLoS One*, 6, e17695.

Dorman, C.J. (2014). H-NS-like nucleoid-associated proteins, mobile genetic elements and horizontal gene transfer in bacteria. *Plasmid*, 75,1-11

Flores-Mireles, A.L., Walker, J.N., Caparon, M. and Hultgren, S.J. (2015) Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*, 13, 269–284.

Foxman, B. (2010). The epidemiology of urinary tract infection. *Nat*. 7: 653–660.

Fu, Y., Du, X., Ji, J., Chen, Y., Jiang, Y. and Yu, Y. (2012). Epidemiological characteristics and genetic structure of *bla*$_{\text{NDM-1}}$ in non-*baumannii Acinetobacter* spp. in China. *J Antimicrob*, 67, 2114–2122.

Garcia, E.C., Brumbaugh, A.R. and Mobley, H.L. (2011) Redundancy and specificity of *Escherichia coli* iron acquisition systems during urinary tract infection. *Infect Immun*, 79, 1225–1235.

Gupta, K., Scholes, D. & Stamm, W.E. (1999). Increasing prevalence of antimicrobial resistance among uropathogens causing acute uncomplicated cystitis in women. *JAMA,* 281,736–738.

Hayes, W, B., & Abraham, S, N. (2016). Innate Immune Responses to Bladder Infection. *MicrobiolSpectr,* 4(6).

Heo, L., Park, H. & Seok, C. (2013). GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res*, 41, W384–W388.

Higgins, P.G., Fluit, A.C. and Schmitz, F.J. (2003). Fluoroquinolones: Structure and target sites. *Curr Drug Targets*, 4,181–90.

Hopkins, K.L., Davies, R.H. and Threlfall, E.J. (2005). Mechanisms of quinolone resistance in *Escherichia coli* and *Salmonella*: recent developments. *Int*, 25,358–373.

Ingersoll, M., & Albert, M. (2013). From infection to immunotherapy: host immune responses to bacteria at the bladder mucosa. *Mucosal Immunol,* 6, 1041–1053.

Juhas, M., van der Meer, J.R., Gaillard, M., Harding, R.M., Hood, D.W. and Crook, D.W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev*, 33(2), 376-93.

Kaper, J., Nataro, J. & Mobley, H. (2004). Pathogenic *Escherichia coli*. *Nat Rev Microbiol*, 2, 123–140.

Khatoon, N., Pandey, R. K. & Prajapati, V. K. (2017). Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci Rep* 7, 8285.

Khatoon, N., Pandey, R. K. & Prajapati, V. K. (2017). Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci Rep* 7, 8285.

Klemm, P., Vejborg, R. M., and Hancock, V. (2010). Prevention of bacterial adhesion. *Appl Microbiol Biotechnol*, 88, 451–459.

Kumar, D., Singh, A.K., Ali, M.R. and Chander, Y. (2014). Antimicrobial Susceptibility Profile of Extended Spectrum Beta-Lactamase (ESBL) Producing *Escherichia coli* from Various Clinical Samples. *Infect Dis (Auckl)*, 7,1–8.

Lambert, P.A. (2002). Mechanisms of antibiotic resistance in *Pseudomonas aeruginosa*. *J Roy Soc Med*, 95, 22–6.

Lane, M.C., Alteri, C.J., Smith, S.N. and Mobley, H.L. (2007) Expression of flagella is coincident with uropathogenic *Escherichia coli* ascension to the upper urinary tract. *PNAS*, 104, 16669–16674.

Laxminarayan, R. D.A., Wattal, C., Zaidi, A.K., Wertheim, H.F. & Sumpradit, N. (2013). Antibiotic resistance, the need for global solutions. *Lancet Infect Dis*, 13(12),1057–98.

Lloyd, A.L., Rasko, D.A. and Mobley, H.L. (2007). Defining genomic islands and uropathogen-specific genes in uropathogenic *Escherichia coli*. *J Bacteriol Res*, 189, 3532–3546.

Lovell, S.C., Davis, I.W, Arendall, W.B., de Bakker, P.I., Word, J.M. and Prisant, M.G. (2003). Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. *Proteins*, 50, 437–450.

Magiorakos, A. P., Srinivasan, A., Carey, R. B., Carmeli, Y., Falagas, M. E., Giske, C. G., Monnet, D. L. (2012). Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect*, 18(3), 268-281.

Marraffini, L.A. & Sontheimer, E.J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet*, 11,181–190.

Marti, E., Huerta, B., Rodríguez-Mozaz, S., Barceló, D., Marcé, R. & Balcázar, J. L. (2018). Abundance of antibiotic resistance genes and bacterial community composition in wild freshwater fish species. *Chemosphere.* 196,115–119.

Maurice, F., Broutin, I., Podglajen, I., Benas, P., Collatz, E. & Dardel, F. (2008). Enzyme structural plasticity and the emergence of broad-spectrum antibiotic resistance. *EMBO Reports*, 9,344–9.

Meza, B., Ascencio, F., Sierra-Beltrán, A. P., Torres, J. & Angulo, C. (2017). A novel design of a multi-antigenic, multistage and multi-epitope vaccine against *Helicobacter pylori*: An *in silico* approach. *Infect Genet Evol*, 49, 309–317.

Moreno E, Andreu A, Pigrau C, Kuskowski MA, Johnson JR, Prats G. (2008). Relationship between *Escherichia coli* strains causing acute cystitis in women and the fecal *E. coli* population of the host. *JMC*, 46, 2529–2534.

Neville, S.A. (2011). Utility of matrix-assisted laser desorption ionization-time of flight mass spectrometry following introduction for routine laboratory bacterial identification. *JCM*, 49, 2980–2984.

Nicastri, E. & Leone, S. (2018). Hospital-acquired urinary tract infection. *Infect Control Hosp Epidemiol*, 35(9),1147–55.

Nicolas-Chanoine, M. H., Bertrand, X., &Madec, J. Y. (2014). *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Newsl*, 27(3), 543–574.

Pallen, M.J. (2010). High-throughput sequencing and clinical microbiology: progress, opportunities and challenges. *Curr Opin Microbiol*, 13, 625–631.

Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. (2018). Mobile genetic elements associated with antimicrobial resistance. *Clin Microbiol Newsl*, 31,88-17.

Qin, X., Hu, F., Wu, S., Ye, X., Zhu, D., Zhang, Y. and Wang, M. 2013. Comparison of Adhesin Genes and Antimicrobial Susceptibilities between Uropathogenic and Intestinal Commensal *Escherichia coli* Strains. *Plos One*, 8(4): e61169.

Rahman, M.M., Haq, J.A., Hossain, M.A., Sultana, R., Islam, F. and Islam, A.H. (2004). Prevalence of extended-spectrum beta-lactamase-producing *Escherichia coli* and *Klebsiella pneumoniae* in an urban hospital in Dhaka, Bangladesh. *Int J Antimicrob Agents*, (5),508-10.

Rankin, D.J., Rocha, E.P.C. & Brown, S.P. (2011). What traits are carried on mobile genetic elements, and why? *Heredity* (Edinb), 106,1-10.

Reynolds, P.E. (1989). Structure, biochemistry and mechanism of action of glycopeptide antibiotics. *Eur J Clin Microbiol Infect Dis*, 8,943–50.

Richardson, J.S., & Richardson, D.C. (2003). Structure validation by Calpha geometry: phi, psi and C beta deviation. *Proteins*, 50(3),437-50.

Rogers, B. A., Sidjabat, H. E. & Paterson, D. L. (2011). *Escherichia coli* O25b-ST131: a pandemic, multiresistant, community-associated strain. *J Antimicrob Chemother*, 66, 1–14.

Roy, A., Kucukural, A. & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoco*, 5, 724-725.

Schaechter, M. (2001). *Escherichia coli* and *Salmonella* 2000: the view from here. *MMBR*, 65(1), 119–130.

Stecher, B., Denzler, R., Maier, L., Bernet, F., Sanders, M.J., Pickard, D.J., Barthel, M., Westendorf, A.M., Krogfelt, K.A., Walker, A.W., Ackermann, M., Dobrindt, U., Thomson, N.R. and Hardt, W.D. (2012). Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae. *Proc Natl Acad Sci USA*, 109(4),1269-74.

Taslima, T., Lina, T.T., Khajanchi, B.K., Azmi, I. J., Islam, M.A., Mahmood, B., Akter, M., Banik, A., Alim, R., Navarro, A., Perez, G., Cravioto, A. and Talukder, K. A. (2014). Phenotypic and Molecular Characterization of Extended-Spectrum Beta-Lactamase-Producing *Escherichia coli* in Bangladesh. *Plos one*, 9(10), e108735.

Temkin, E., Adler, A., Lerner, A. and Carmeli, Y. (2014). Carbapenem-resistant Enterobacteriaceae: biology, epidemiology, and management. *Ann N Y Acad Sci*, 1323, 22-42.

Tenover, F.C. (2006). Mechanisms of antimicrobial resistance in bacteria. *Am. J.* 119:3–10.

Tolmasky, M.E. (2000). Bacterial resistance to aminoglycosides and beta-lactams: The Tn1331 transposon paradigm. *Front Biosci*, 5,20–9.

Van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M, Schmitz, C., Kastritis, P.L., Karaca. E, Melquiond, A.S.J., van Dijk, M., de Vries, S.J., Bonvin, A.M.J.J. (2016). The HADDOCK2.2 WebServer: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol*, 428(4), 720-725.

Vannuffel, P. & Cocito, C. (1996). Mechanism of action of streptogramins and macrolides. *Drugs*, 51, 20–30.

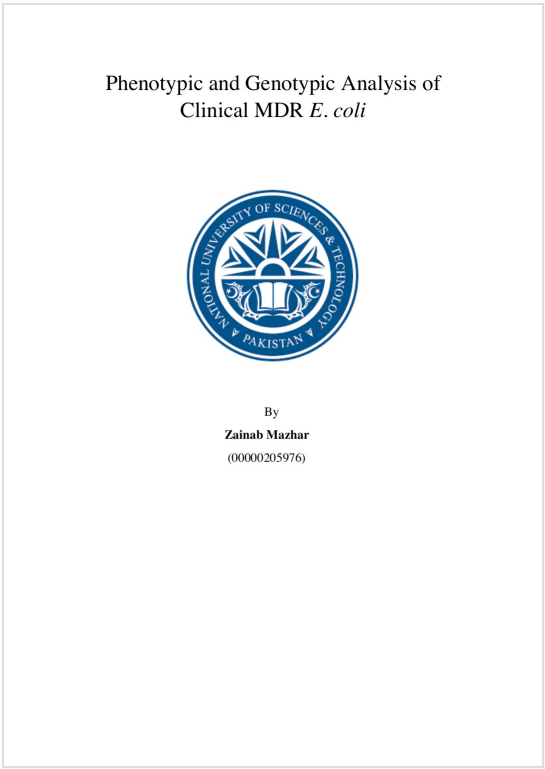
Vettone, A., Perugino, G., Rossi, M., Valenti, A. & Ciaramella, M. (2014). Genome stability: recent insights in the topoisomerase reverse gyrase and thermophilic DNA alkyltransferase. *Extremophiles*, 18, 895–904.

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Zainab Mazhar
Assignment title:	Plagiarism Detection 2019 Part 4 (M…
Submission title:	Thesis
File name:	96446_Zainab_Mazhar_Thesis_510…
File size:	1.8M
Page count:	62
Word count:	11,634
Character count:	70,317
Submission date:	27-Feb-2020 05:45PM (UTC+0500)
Submission ID:	1265207430



Phenotypic and Genotypic Analysis of
Clinical MDR *E. coli*

By

Zainab Mazhar

(00000205976)

Thesis

ORIGINALITY REPORT

12%	6%	9%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	2%
2	www.nature.com Internet Source	2%
3	Mahnoor Majid, Saadia Andleeb. "Designing a multi-epitopic vaccine against the enterotoxigenic Bacteroides fragilis based on immunoinformatics approach", Scientific Reports, 2019 Publication	1%
4	Submitted to Cardiff University Student Paper	1%
5	Rajan Kumar Pandey, Mudassar Ali, Rupal Ojha, Tarun Kumar Bhatt, Vijay Kumar Prajapati. "Development of multi-epitope driven subunit vaccine in secretory and membrane protein of Plasmodium falciparum to convey protection against malaria infection", Vaccine, 2018	<1%

Publication

6	Udomluk Leungtongkam, Rapee Thummeepak, Kannipa Tasanapak, Sutthirat Sitthisak. "Acquisition and transfer of antibiotic resistance genes in association with conjugative plasmid or class 1 integrons of Acinetobacter baumannii", PLOS ONE, 2018 Publication	<1%
7	www.frontiersin.org Internet Source	<1%
8	Submitted to University of Sheffield Student Paper	<1%
9	www.ncbi.nlm.nih.gov Internet Source	<1%
10	Utpal Kumar Adhikari, Mourad Tayebi, M. Mizanur Rahman. "Immunoinformatics Approach for Epitope-Based Peptide Vaccine Design and Active Site Prediction against Polyprotein of Emerging Oropouche Virus", Journal of Immunology Research, 2018 Publication	<1%
11	Submitted to RMIT University Student Paper	<1%
12	Aruna Narula, Rajan Kumar Pandey, Nazia Khatoon, Amit Mishra, Vijay Kumar Prajapati. "Excavating chikungunya genome to design B	<1%

and T cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection", Infection, Genetics and Evolution, 2018
Publication

---

13 Robert Adamu Shey, Stephen Mbigha Ghogomu, Kevin Kum Esoh, Neba Derrick Nebangwa et al. "In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases", Scientific Reports, 2019 <1%
Publication

---

14 Babita Sharma. "EGF domain II of protein Pb28 from Plasmodium berghei interacts with monoclonal transmission blocking antibody 13.1", Journal of Molecular Modeling, 04/2009 <1%
Publication

---

15 Mudassar Ali, Rajan Kumar Pandey, Nazia Khatoon, Aruna Narula, Amit Mishra, Vijay Kumar Prajapati. "Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection", Scientific Reports, 2017 <1%
Publication

---

16 Submitted to Dokuz Eylul Universitesi <1%
Student Paper

17	Manica Negahdaripour, Mahboobeh Eslami, Navid Nezafat, Nasim Hajighahramani et al. "A novel HPV prophylactic peptide vaccine, designed by immunoinformatics and structural vaccinology approaches", Infection, Genetics and Evolution, 2017 Publication	<1%
18	repositorium.sdum.uminho.pt Internet Source	<1%
19	www.physiology.org Internet Source	<1%
20	"Abstracts cont.", Clinical Microbiology and Infection, 4/2005 Publication	<1%
21	Submitted to Birla Institute of Technology Student Paper	<1%
22	libres.uncg.edu Internet Source	<1%
23	elib.suub.uni-bremen.de Internet Source	<1%
24	academic.oup.com Internet Source	<1%
25	Stephane Fraga de Oliveira Tosta, Mariana Santana Passos, Rodrigo Kato, Álvaro Salgado et al. "Multi-epitope based vaccine against	<1%

yellow fever virus applying immunoinformatics approaches", Journal of Biomolecular Structure and Dynamics, 2020
Publication

26	Nasim Hajighahramani, Mahboobeh Eslami, Manica Negahdaripour, Mohammad Bagher Ghoshoon et al. "Computational design of a chimeric epitope-based vaccine to protect against Staphylococcus aureus infections", Molecular and Cellular Probes, 2019 Publication	<1%
27	Mette Voldby Larsen, Katrine G. Joensen, Ea Zankari, Johanne Ahrenfeldt et al. "Chapter 5 The CGE Tool Box", Springer Science and Business Media LLC, 2017 Publication	<1%
28	Submitted to University of KwaZulu-Natal Student Paper	<1%
29	Submitted to Instituto Gulbenkian de Ciencia Student Paper	<1%
30	Submitted to October University for Modern Sciences and Arts (MSA) Student Paper	<1%
31	Sajad Abdollahi, Iraj Rasooli, Seyed Latif Mousavi Gargari. "An in silico structural and physicochemical characterization of TonB-dependent copper receptor in A. baumannii",	<1%

Microbial Pathogenesis, 2018
Publication

32	Aarti Rana, Shweta Thakur, Nupur Bhardwaj, Devender Kumar, Yusuf Akhter. " Excavating the surface-associated and secretory proteome of for identifying vaccines and diagnostic markers relevant immunodominant epitopes ", Pathogens and Disease, 2016 Publication	<1%
33	oup.silverchair-cdn.com Internet Source	<1%
34	Ali Dalir Ghaffari, Abdolhossein Dalimi, Fatemeh Ghaffarifar, Majid Pirestani. "Structural predication and antigenic analysis of ROP16 protein utilizing immunoinformatics methods in order to identification of a vaccine against Toxoplasma gondii: An in silico approach", Microbial Pathogenesis, 2020 Publication	<1%
35	Submitted to University of Massachusetts, Lowell Student Paper	<1%
36	Rajan Kumar Pandey, Rupal Ojha, Amit Mishra, Vijay Kumar Prajapati. "Designing B- and T-cell multi-epitope based subunit vaccine using immunoinformatics approach to control Zika virus infection", Journal of Cellular Biochemistry,	<1%

2018
Publication

37	Submitted to University of Birmingham Student Paper	<1%
38	Nina Chatterjee, Rupal Ojha, Nazia Khatoon, Vijay Kumar Prajapati. "Scrutinizing Mycobacterium tuberculosis membrane and secretory proteins to formulate multiepitope subunit vaccine against pulmonary tuberculosis by utilizing immunoinformatic approaches", International Journal of Biological Macromolecules, 2018 Publication	<1%
39	Submitted to Silpakorn University Student Paper	<1%
40	"Immunoinformatics", Springer Science and Business Media LLC, 2014 Publication	<1%
41	Submitted to University of Edinburgh Student Paper	<1%
42	Cecilia M. Abe, FÃ¡bia A. Salvador, Ivan N. Falsetti, MÃ´nica A.M. Vieira et al. " Uropathogenic (UPEC) strains may carry virulence properties of diarrhoeagenic ", FEMS Immunology & Medical Microbiology, 2008 Publication	<1%