Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2026.241

RESEARCH ARTICLE

Exploration of Vaccine Candidates against Avian Pathogenic *E. coli* Using Proteome Mining

Amara Saif[1], Ayesha Siddiqa[1], Muhammad Imran[2], Moazur Rahman[3], Muhammad Muddassar[4], Sakandar Rauf[5], Nayla Munawar[6]* and Muhammad Saleem[1*]

[1]School of Biological Sciences, University of the Punjab, Lahore; [2]KAM School of Life Sciences, Forman Christian College (A Chartered University), University Lahore; [3]Centre of Excellence in Molecular Biology, University of the Punjab, Lahore; [4]Department of Biosciences, COMSATS University, Islamabad; [5]School of Biochemistry and Biotechnology, University of the Punjab, Lahore; [6]Department of Chemistry, College of Science, United Arab Emirates University, Al Ain, 15551, UAE.
*Corresponding author: MSaleem.sbs@pu.edu.pk (MS);nmunawar@uaeu.ac.ae (NM)

ARTICLE HISTORY (26-964)

Received: August 01, 2026
Revised: September 18, 2026
Accepted: September 21, 2026
Published online: October 11, 2026

Key words:
Antigen
APEC
Epitope
Outer membrane proteins (OMPs)
Vaccines

ABSTRACT

Avian Pathogenic *Escherichia coli* (APEC) is a major causative agent of avian colibacillosis, causing significant financial losses to the poultry industry. Current vaccines offer limited protection and also pose logistical challenges related to storage and administration. Considering this fact, a subtractive proteome analysis along with a reverse vaccinology approach was performed for the screening of promising vaccine candidates that would offer broad protection against heterologous serotypes of APEC. O1, O2, and O78 are APEC's most prevalent and pathogenic serotypes, responsible for over 80% of infections in poultry. Therefore, the essential, highly virulent, and strongly antigenic outer membrane proteins (OMPs) shared among these pathogenic serotypes of APEC were screened through comprehensive *In Silico* analysis. Thus, six OMPs were identified as viable vaccine targets, including *OmpA, PapC, Wza, YehB, PhoE,* and *EtsC*. These OMPs belong to the four functional families: outer membrane protectin, transporters, capsular biogenesis, and fimbrial biogenesis. Afterwards, a novel chimeric vaccine was designed against APEC by using four B-cell, four CTL, and five HTL epitopes, which were highly conserved among the three highly pathogenic serotypes. Molecular docking and dynamic simulations revealed the stable and high-affinity interactions between the constructed chimeric vaccine and the mammalian TLR4 receptor. Nonetheless, *in vitro* and *in vivo* experimental validation will be required to confirm the immunogenicity and protective efficacy of the vaccine ensemble.

To Cite This Article: Saif A, Siddiqa A, Imran M, Rahman M, Muddassar M, Rauf S, Munawar N and Saleem M, 2026. Exploration of vaccine candidates against avian pathogenic *E. coli* using proteome mining. Pak Vet J, 46(10): 2407-2419. http://dx.doi.org/10.29261/pakvetj/2026.241

INTRODUCTION

Avian pathogenic *Escherichia coli* (APEC) is a small, cylindrical, gram-negative bacterium that belongs to the subclass of extraintestinal pathogenic *Escherichia coli* (ExPEC) (Nawaz *et al.,* 2024). It is a causative agent of the deadliest infectious diseases in the poultry industry, particularly colibacillosis, which causes major economic losses to the industry (Christensen *et al*., 2021). Colibacillosis is a syndromic disease frequently associated with various local and systemic infections, mainly cellulitis, perihepatitis, pericarditis, airsacculitis, and other respiratory disorders (Dziva and Stevens, 2008; Fancher *et al*., 2020). Chickens acquire the pathogen through contaminated water and feed, and the disease can further spread via aerosol or fecal-oral routes (Joseph *et al.,* 2023). Researchers have even termed APEC as an opportunistic pathogen because it invades and colonizes chickens' respiratory and gastrointestinal tracts without causing illness and only transfers to extra-intestinal sites under stressors (Kathayat *et al.,* 2021; Watts and Wigley, 2024). Moreover, in turkeys and chickens, APEC is responsible for osteomyelitis and swollen head syndrome, resulting in less meat and egg production and costing the poultry industry millions of dollars worldwide (Christensen *et al.,* 2021; Kathayat *et al.,* 2021).

The global incidence of APEC varies from approximately 9.52 to 37%, affecting broiler chickens during their egg production period. However, its prevalence is higher in developing countries due to

intensive poultry farming practices and poor disease management strategies (Lutful Kabir, 2010; Christensen *et al.,* 2021). Although there are several serotypes of APEC, some are more frequently linked to virulence and illness in poultry. APEC O1, O2, and O78 are the most prevalent and highly pathogenic serotypes, responsible for more than 80% of APEC infections (Ghunaim *et al.,* 2014). According to recent studies, the O145 serotype is prominent in Pakistan and China, causing serious human illnesses (Kathayat *et al.,* 2021; Nawaz *et al.,* 2024). APEC exhibits genomic resemblance to human ExPECs, particularly to uropathogenic *E. coli* (UPEC) and neonatal meningitis *E. coli* (NMEC), and possesses their virulent genes associated with urinary tract infections (UTI) and meningitis, respectively. Moreover, the K1 capsule, aerobactin iron transporter, and P and type 1 fimbriae have been reported as the common virulence factors between APEC and ExPEC (Mellata, 2013; Fancher *et al.,* 2020).

The immune response against APEC infection is mediated through innate and adaptive immunity. Macrophages and heterophages (homologous to mammalian neutrophils) are crucial in providing innate immunity against APEC in chickens by inducing a strong pro-inflammatory response that promotes bacterial clearance. Moreover, the humoral immune response is derived by IgY, IgA, and IgM antibodies that stimulate complement activation and opsonization to inhibit bacterial colonization and proliferation. However, both the cellular (Th1 or Th2) and humoral responses contribute to protective immunity against APEC (Alber *et al.,* 2021).

Primarily, various antibiotics are used to cure colibacillosis, but the growing resistance of APEC towards different classes of antibiotics presents challenges in controlling APEC infections (Monroy *et al.,* 2025). Various killed, subunit, autogenous, and live-attenuated bacterial vaccines were used to prevent APEC outbreaks. Autogenous vaccines only protect against the homologous serotypes, however, subunit vaccines can protect against heterologous serotypes, but the limitations include side effects from adjuvants and stress induction in birds (Kathayat *et al.,* 2021; Pokharel *et al.,* 2023). Currently, only two vaccines, including live-attenuated Poulvac® (aroA mutant) and inactivated Nobilis® *E. coli* vaccine (having FT flagellar and F11 fimbrial antigens), are able to trigger both cell-mediated and humoral protection against APEC. Still, they do not offer protection against the diverse pathogenic serotypes of APEC (Watts and Wigley, 2024). Therefore, identifying pathogenic and virulent factors conserved among APEC serotypes is crucial for developing an effective vaccine for chickens. APEC employs multiple virulence factors to promote colibacillosis in chickens, including invasins, adhesins, toxins, protectins, iron acquisition systems, transcriptional regulators, and quorum sensors. These pathogenic factors help the bacterium to attach, invade, colonize, and escape from the immune system, contributing to host tissue destruction (Mellata *et al.,* 2003; Hu *et al.,* 2022). Despite the discovery of these factors, no vaccine is currently available to protect chickens against the heterologous serotypes of APEC. Therefore, it is crucial to design a multi-epitope-based vaccine that can trigger strong mucosal and cellular immunity to confer long-term protection against the heterologous serotypes of APEC.

In this study, a suite of computational tools involving subtractive proteome analysis is employed to identify potential vaccine candidate OMPs and formulate a multi-epitope-based vaccine against the most prevalent and pathogenic APEC serotypes, as outlined Fig.1. The developed method overcomes the limitations of traditional antigen discovery methods in terms of time, cost, and precision. Moreover, the immunoinformatic techniques used in these studies help evaluate specific immunogenic effects, ultimately contribute to reducing the burden of *in vivo* and *in vitro* research.

Even though some potentially important information may have been excluded due to this stringent methodology, the chosen OMPs were confirmed to be highly promising candidates for further investigation. They can further be used to formulate multi-antigen-based, epitope-based, or peptide-based vaccines that may trigger the protective immune response against the APEC serotypes.

MATERIALS AND METHODS

The flow diagram illustrating the methodology used in this study is presented in Fig.1.

Upstream In Silico Analysis

Data Retrieval: Initially, the whole proteome of the APEC O1 serotype (UniProt ID CP000468.1) was retrieved by using the Universal Protein Knowledgebase (UniProtKB) (https://www.uniprot.org/) because the proteomes of the O2 and O78 serogroups were not fully annotated (Breuza *et al.,* 2016). Additionally, the OMPs were short-listed based on their subcellular localization mentioned in the database. They were further sorted using a subtractive proteomics pipeline to rank the potential vaccine candidates. Although several proteins can serve as vaccine targets, the OMPs have been reported as the preferred vaccine candidates because they play a critical role in the physiology and pathogenesis of the bacteria.

Identification of Essential OMPs: Essential genes are critically indispensable for bacterial growth and survival, making them ideal vaccine candidates. Here, essential genes were determined by aligning the retrieved OMPs against the Database of Essential Genes (DEG) using BLASTp analysis (Liang *et al.,* 2024). DEG (http://www.essentialgene.org/) is an online repository of experimentally validated essential genes present in the genomes of various organisms. The OMPs with a bit score of $\geq$100 and an E-value of $\leq 10^{-4}$ were marked as essential proteins of the APEC.

Identification of Virulent OMPs: Pathogens produce certain virulent proteins that contribute to their pathogenesis by promoting adhesion, colonization, and immune response evasion, which weakens the host immune system and results in illness. Here, the virulence of the selected OMPs was evaluated by using BLASTp against the VFDB (Virulence Factor Database) (http://www.mgc.ac.cn/VFs/) (Chen *et al.,* 2005). It holds data on virulence genes present in various bacterial

pathogens. The OMPs satisfying the following criteria were considered as virulent proteins: a bit score of ≥100 and an E-value of $\leq 10^{-4}$.

Additionally, by using the VirulentPred 2.0 server (https://bioinfo.icgeb.res.in/virulent2/predict.html), the previously filtered essential and virulent OMPs were further investigated for their virulence functions to improve the accuracy of the method (Sharma *et al.,* 2023). It classifies proteins as virulent or non-virulent, based on machine-learning algorithms trained on the experimentally verified virulent proteins. The OMPs marked as virulent were considered putative vaccine candidates.

Antigenicity and Allergenicity Assessment: Meanwhile, to improve the efficiency of the vaccine, the antigenicity and allergenicity of the potential targets were analyzed via VaxiJen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) and AllerTop v 2.0 (https://www.ddg-pharmfac.net/AllerTOP/), respectively (Dimitrov *et al.,* 2014; Zaharieva *et al.,* 2019). These servers are alignment-free software that predicts the protective antigens and allergens based on the proteins' physicochemical properties.

Assessment of Physicochemical Properties: The molecular weight and instability index of the selected OMPs were predicted via the ExPASy ProtParam tool (http://us.expasy.org/tools/protparam.html) (Gasteiger *et al.,* 2005). OMPs with a molecular weight of ≤110kDa were selected because small-sized proteins are easy to purify in their stable state and might be better vaccine targets. Moreover, the TMHMM server (https://services.healthtech.dtu.dk/services/DeepTMHMM-1.0/) was used to identify transmembrane helices that lie in the structure of OMPs. The proteins with <1 transmembrane helix were preferred for the current study.

Homology analysis of selected OMPs to other Serotypes: The vaccine candidates shared by all the pathogenic serotypes are highly desirable because they offer broad protection, enhanced herd immunity, and lower the chances of vaccine escape during mutations. Therefore, the proteins shared by all three serogroups were screened by performing BLASTp analysis of previously shortlisted OMPs against the annotated proteome of O2 (taxid: 1401689) and O78 (taxid: 1274814) serotypes via NCBI (Mahram and Herbordt, 2015). The obtained hits with a Percentage identity ≥90% and an E-value $\leq 10^{-4}$ were assessed for further analysis.

Downstream In Silico Analysis

Domains Search: Subsequently, the Conserved Domain Database (CDD) and Pfam were accessed to find out the functional domains of the shortlisted proteins (Yang *et al.,* 2020). CDD integrated within the NCBI's Entrez query system facilitates the annotation of protein sequences by identifying their functional sites and conserved domains. Pfam is an online database that categorizes proteins into families based on their functional sites and conserved domains via the Hidden Markov Models (HMMs) and multiple sequence alignment (Mistry *et al.,* 2021). Furthermore, the VICM server was used to study the structural and functional features of selected OMPs based on their sequences and involvement in various biological processes.

Epitopes Prediction and Analysis: The identification of signal peptides aids in determining the function and subcellular location of proteins. Thus, the SignalP-6.0 server (https://services.healthtech.dtu.dk/services/SignalP-6.0/) was used to predict the existence of signal peptides in OMPs (Nielsen *et al.,* 2024). It applies multiple artificial neural networks to detect the signal peptide sequence and the possible cleavage sites of eukaryotic, archaeal, Gram-positive, and Gram-negative bacterial proteins. Since the signal peptide sequence is a transient component of the mature protein, it is preferable to eliminate it before the structure analysis and epitope prediction.

Prediction of B-Cell Epitopes: The selected antigenic proteins were analyzed by BCPREDS (http://ailab-projects2.ist.psu.edu/bcpred/predict.html) for the prediction of linear B-cell epitopes with a residual length of 12 and 14 amino acids (Liu *et al.,* 2024). The criteria used to predict the potential epitopes involve a threshold of more than 0.9. Further, the antigenic and allergenic nature of the predicted epitopes were evaluated by using the VaxiJen and AllerTop servers, respectively.

Prediction of T-Cell Epitopes: The absence of chicken MHC alleles in computational immunology impedes the development of T-cell-based vaccines against poultry pathogens. Prior studies have suggested that the chicken B-F alleles can primarily trigger an immunological response similar to the human class I alleles, leading to the substitution of human alleles for the chicken MHC alleles. Therefore, the three alleles from the MHC class I (HLA*B 40:06, HLA*B 41:04, and HLA*B 41:03) and four alleles (DRB1:1482, DRB1:1366, DRB1:1310, and DRB1:1445) from the MHC class II were identified from the previous study as the best human MHC alleles for chicken (Valdivia-Olarte *et al.,* 2015; Fatoba *et al.,* 2022).

NetMHCpan–4.0 server (https://services.healthtech.dtu.dk/services/NetMHCpan-4.0/) and NetMHCIIpan-4.2 server (https://services.healthtech.dtu.dk/services/NetMHCIIpan-4.2/) were approached for the identification of T-cell antigenic determinants, which restricted well to the previously selected MHC-I and MHC-II alleles, respectively (Jurtz *et al.,* 2017; Nilsson *et al.,* 2023). These *in silico* tools predict the binding affinity of peptides to their respective alleles by applying artificial neural networks (ANNs). Only the 9mer Class-I and 15mer Class-II epitopes that met the criteria of ≤0.5% percentile rank and ≤100nM IC50 were chosen. Furthermore, the antigenicity and allergenicity of the significant epitopes were calculated via VaxiJen and AllerTop, respectively.

3D Structural Assessment: The transmembrane topology and the solvent accessibility prediction of the shortlisted OMPs were visualized via TMVisDB (https://tmvisdb.rostlab.org/). It serves as an online resource

for the exploration and structural visualization of transmembrane proteins (TMPs) derived via AlphaFold DB (Marquet *et al*., 2022). It was used to anticipate the per-residue transmembrane topology and the solvent accessible regions of selected proteins. Pred-TMBB server was also used to assess the surface-exposed regions of the nominated OMPs (http://bioinformatics.biol.uoa.gr/PRED-TMBB/). The web server uses a Hidden Markov Model to predict the transmembrane topology of the OMPs of Gram-negative bacteria (Bagos *et al.,* 2004).

Chimera Designing: The detailed analysis of targeted proteins through various bioinformatics tools and approaches resulted in the selection of potential B-cell and T-cell epitopes, to be included in the designing of a multi-epitope-based vaccine against APEC. Moreover, ClustalW was used to examine the conservation of predicted epitopes among O1, O2, and O78 serotypes. The selected epitopes were fused via short hydrophobic linker sequences. The epitopes with overlapping sequences were joined into a single epitope sequence. Additionally, a 6x His-tag was attached to the C-terminus of the protein for improved purification.

Immunological Profiling and Solubility Analysis: The antigenicity and allergenicity of the chimeric vaccine were assessed via VaxiJen and AllerTop v 2.0 servers. However, Protein-Sol (https://protein-sol.manchester.ac.uk/) was used to analyse the solubility of the constructed chimera (Hebditch *et al.,* 2017). The cut-off value of 0.45 was applied.

Assessment of Physicochemical Properties: The physicochemical properties like molecular weight, instability index, grand average of hydropathicity (GRAVY), isoelectric pH, and half-life, were predicted via the ExPASy ProtParam tool.

Structural Modelling, Refinement and Validation: The PSIPRED (https://bioinf.cs.ucl.ac.uk/psipred/) and AlphaFold server (https://alphafoldserver.com/) were used to predict the secondary and tertiary structure of the chimeric vaccine construct. It provides a broad range of tools for protein annotation and predictions with a primary focus on the 2D structural annotations (Buchan and Jones, 2019). The 3D structure attained by AlphaFold was further refined by using GalaxyRefine Server (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE). AlphaFold is an online server that uses neural networks to predict the 3D structure of proteins from their amino acid sequences. It uses multiple sequence alignment to examine the evolutionary relationships and generate accurate protein models (Jumper *et al.,* 2021). GalaxyRefine improves the predicted 3D structures by optimizing overall stereochemistry, structural relaxation, and side-chain modelling (Heo *et al.,* 2013). Finally, the accuracy of the 3D structure was evaluated by using Structure Analysis and Verification Server v 6.1- SAVES (https://saves.mbi.ucla.edu/). It is a meta-server that analyzes the stereochemistry of protein structures by verifying the structural geometry from individual amino acids to the overall protein structure.

Molecular Docking and Simulations: Molecular docking is an ideal method to predict the development of an immune response by the interaction between the chimeric vaccine and immune receptors. Here, the ClusPro server (https://cluspro.org) was accessed to evaluate the interactions and binding energies of the chimeric construct and the TLR4 immune receptors (2Z65). GROMACS (version: 2020.7) was used to study the dynamic behavior of the chimeric antigen and TLR4 complex. The CHARMM36 force field was utilized to accurately model the molecular interactions within the system. Topology files for the chimeric antigen and TLR4 were generated using the "pdb2gmx" tool. The system was solvated in a triclinic box using SPC water molecules (TIP3P model), and neutralization was achieved by adding Na+ and Cl- ions. The system underwent energy minimization using the steepest descent algorithm to remove steric clashes, followed by equilibration with NVT equilibration at 300K for 1ns and NPT equilibration at 1atm for an additional 1ns. These steps were performed to stabilize the system's temperature and pressure. Subsequently, a 200ns production molecular dynamics simulation with a time step of 2fs was carried out. The trajectory data were analyzed using GROMACS Toolkits and XMGRACE software to evaluate the stability and interactions of the protein-vaccine complex.

RESULTS

Upstream In Silico Analysis

Data Retrieval: APEC O1 is comprised of one chromosome and four plasmids that collectively produce 4870 proteins. The F-type plasmid pAPEC-O1-ColBM encodes a potential virulence cluster and produces the B and M colicins. pAPEC-O1-R plasmid contains eight antibiotic resistance genes, while the cryptic plasmids (pAPEC-O1-Cryptic1 and pAPEC-O1-Cryptic2) share some similarities with Yersinia-type plasmids.

Among the whole proteome of APEC O1, UniProt characterized 77 proteins as OMPs based on experimental data and retrieved them for the subtractive proteome analysis.

Identification of Essential and Virulent OMPs: Firstly, the BLASTp analysis of the selected OMPs against the experimentally validated proteins of DEG and VFDB identified 37 essential and 35 virulent proteins, respectively. Moreover, the virulence of the shortlisted OMPs was also evaluated via VirulentPred 2.0 to improve the efficacy of the selection criteria. Altogether, the manual comparison revealed that 18 OMPs showed both the essential and virulent characteristics, thus assessed for the downstream analysis. The details of putative vaccine targets are listed in Table 1.

Antigenicity and Allergenicity Assessment: The OMPs screened from the previous step were subjected to antigenicity and allergenicity assessment. VaxiJen analysis predicted 14 highly antigenic proteins having a prediction score of more than 0.5. Moreover, 13 non-allergenic proteins were identified by the AllerTop analysis, as mentioned in Table 1. However, to improve the vaccination effectiveness and reduce any adverse effects, the OMPs that were found to be highly antigenic and non-allergenic proteins were taken into account for further investigation.

Assessment of Physicochemical Properties: The instability index and molecular weight assessment by the ExPasy ProtParam tool revealed that 11 proteins had a molecular weight of less than 110kDa and an instability index score less than 40. Moreover, transmembrane helices were not detected in the chosen OMPs, according to TMHMM server analysis, which confirmed that they are suitable vaccine targets. The details are mentioned in Table 1.

Homology analysis of selected OMPs to other Serotypes: Finally, the homology analysis of selected antigenic OMPs of APEC O1 was performed to filter the conserved vaccine candidates among all three pathogenic serogroups (O1, O2, and O78) of APEC. The BLASTp analysis identified 6 OMPs that fall under the set criteria. The details are mentioned in Table 2.

The detailed subtractive proteome analysis resulted in the selection of 6 highly antigenic OMPs that are considered to be potential vaccine targets against APEC caused by multiple pathogenic serogroups. Even though some potentially important information may have been excluded due to this stringent methodology, however the chosen OMPs were confirmed to be highly promising targets for additional investigation.

Table 2: List of shortlisted vaccine candidate OMPs and their role in the pathogenesis of APEC.

Sr. No	Protein Names	Homology %	VICM	Functions
1	Outer membrane protein/porin A (OmpA)	>95%	Cellular Process	Protectin (Intracellular Survival)
2	Outer membrane porin (PhoE)	>98%	Virulent Factor	Phosphate uptake
3	Polysaccharide export protein (Wza)	~100%	Virulent Factor	Capsular Biogenesis
4	Outer membrane efflux protein (EtsC)	>99%	Virulent Factor	Auto-transporter
5	Outer membrane usher protein (YehB)	>97%	Cellular Process	Adhesion, motility, intracellular survival, colonization.
6	Outer membrane usher protein (PapC)	>99%	Cellular Process	Adhesion, motility, intracellular survival, colonization.

Table 1: The subsequent proteome mining to screen the highly potential Vaccine targets.

	OMPs of APEC O1: K1			Step 1		Step 2		Step 3		
Sr. No	Entry	Protein Name	Length (aa)	DGE	VirulentPred & VFDB	Antigenicity	AllerTOPv2.0	M.W (kDa)	TMHMM 2.0	Instability Index
1	A0A0H2YXT7	ompA	370	Essential	Virulent	Antigen	Non-allergen	40.2	0	24.5
2	A0A0H2XJN4	tsh	1377	Essential	Virulent	Antigen	Non-allergen	148.2	-	-
3	A0A0H2YVV9	phoE	353	Essential	Virulent	Antigen	Non-allergen	39.2	0	32.2
4	A0A0H2YWF4	vat	1376	Essential	Virulent	Antigen	Non-allergen	148	-	-
5	A0A0H2YZW3	nmpC	360	Essential	Virulent	Antigen	Non-allergen	40	0	13.9
6	A0A0H2Z0D8	wza	379	Essential	Virulent	Antigen	Non-allergen	42	0	37.9
7	A0A0H2Z1K8	yfcU	884	Essential	Virulent	Antigen	Non-allergen	97.4	0	31
8	A0A0H2XK18	etsC	456	Essential	Virulent	Antigen	Non-allergen	52	0	39.6
9	A0A0H2XLR1	silC	562	Essential	Virulent	Non-Antigen	-	-	-	-
10	A0A0H2YVK7	htrE	862	Essential	Virulent	Non-Antigen	-	-	-	-
11	A0A0H2YWY2	cusC	460	Essential	Virulent	Non-Antigen	-	-	-	-
12	A0A0H2YZ34	Outer membrane channel protein	457	Essential	Virulent	Non-Antigen	-	-	-	-
13	A0A0H2Z0D2	yehB	826	Essential	Virulent	Antigen	Non-allergen	92.3	0	39.2
14	A0A0H2Z328	yqiG	840	Essential	Virulent	Antigen	Non-allergen	93.5	0	31.2
15	A0A0H2Z3A7	papC	839	Essential	Virulent	Antigen	Non-allergen	92	0	33.6
16	A0A0H2Z3C0	hofQ	412	Essential	Virulent	Antigen	Allergen	-	-	-
17	A0A0H2Z3R1	Iron compound receptor	713	Essential	Virulent	Antigen	Non-allergen	79	0	35.6
18	A0A0H2Z498	hofD	654	Essential	Virulent	Antigen	Non-allergen	71.1	0	33.8

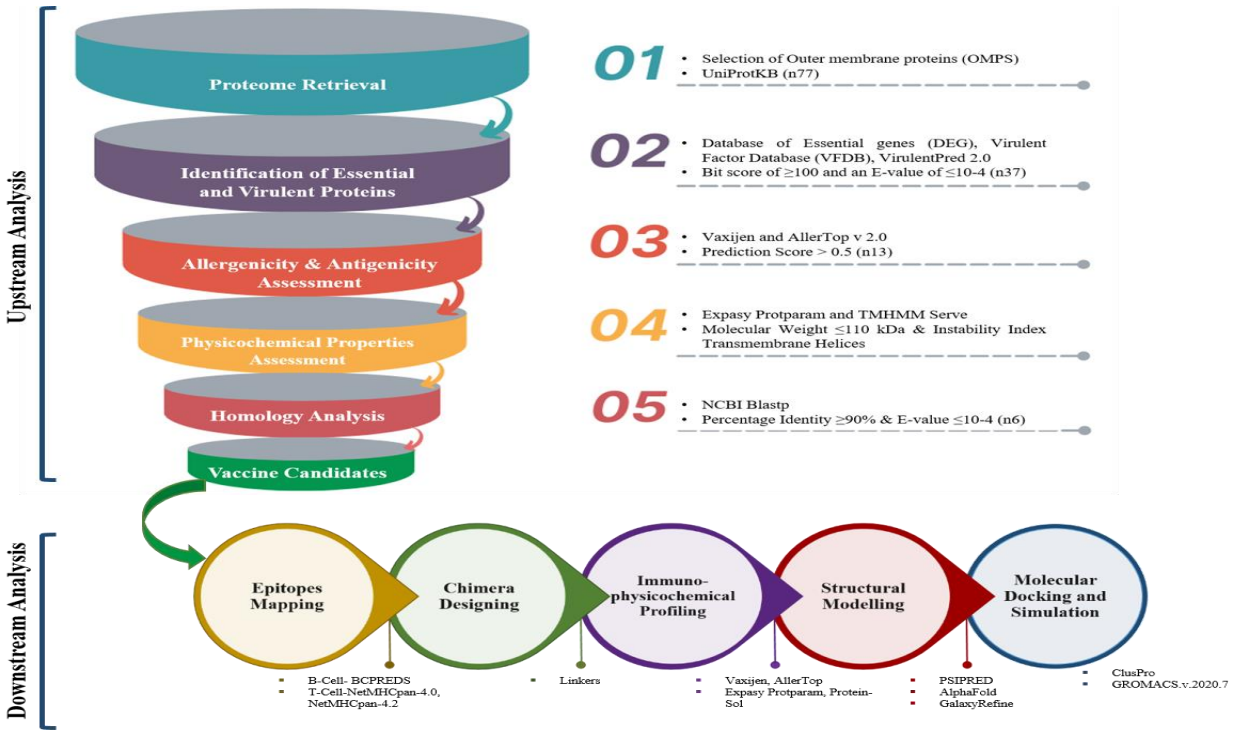


Fig. 1: The methodology used to screen the highly potential vaccine targets of APEC through proteome mining.

Downstream In Silico Analysis

Domain Search: CDD and Pfam were used to identify the functional units of the vaccine candidate proteins. OmpA is a porin responsible for the penetration of small nutrients. It also provides stability during bacterial conjugation and host-pathogen interactions, indicating its role in bacterial pathogenicity. PhoE belongs to the porin superfamily and is involved in the transport of negatively charged molecules like phosphorylated compounds and inorganic phosphate. Wza is responsible for the export of polysaccharides for capsular biosynthesis. EtsC is an outer membrane efflux protein of the TolC family, involved in the type 1 secretion system of proteins. YehB and PapC are the usher proteins, involved in the export and assembly of fimbrial subunits for the fimbrial biogenesis.

Prediction of Epitopes: Furthermore, the primary sequences of shortlisted OMPs were examined for the presence of signal peptide sequences via SignalP-6.0. Prior to epitope mapping, the predicted signal sequences were eliminated from the functional sequence of the mature OMPs to screen out the potential immunogenic regions.

After passing the whole screening process, the six highly efficient vaccine candidate OMPs were mapped for the epitope prediction using various computational approaches. The predicted B and T-cell Epitopes are listed in Table 3 and 4, respectively.

3D Structural Assessment: The potential B-cell and T-cell epitopes were analyzed via Pred-TMBB to annotate the surface-exposed regions, as marked in Fig. 2, respectively. Moreover, the three-dimensional structures of the respective proteins were retrieved via UniProt and further analyzed to annotate the transmembrane topology and solvent accessibility through TMVisDB. The per-residue structural aspects of the OMPs are coded in different colors for clear representation, as mentioned in Fig. 3.

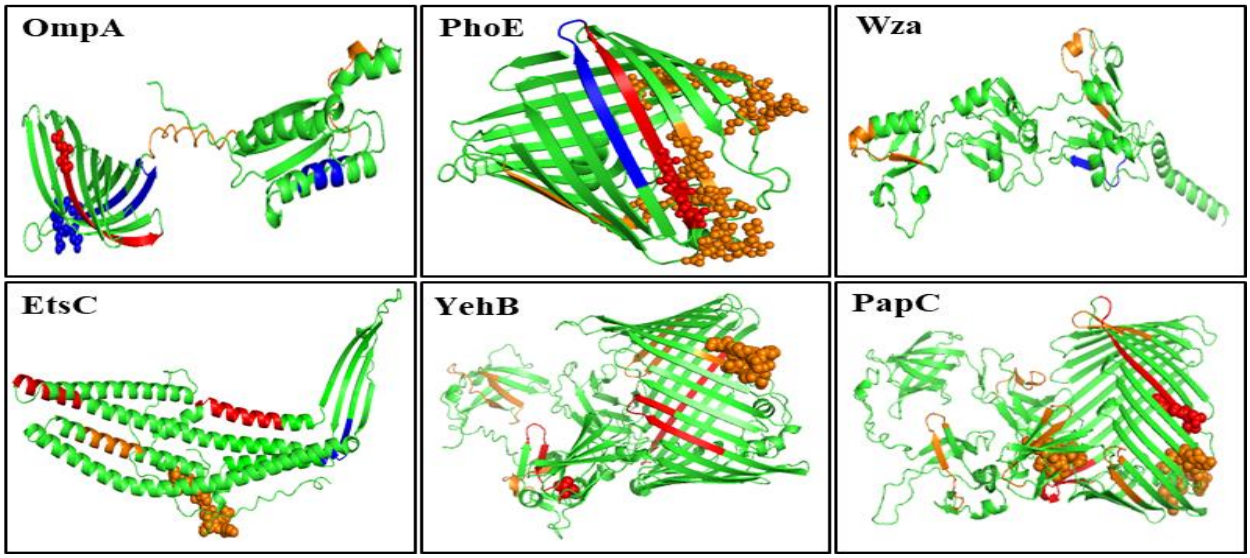


Fig. 2: 3D structural assessment. Three-dimensional structures of native proteins with highlighted regions showing the predicted B-cell and T-cell (Class I and Class 2) epitopes in orange, blue, and red. The spherical representation of residues shows the surface-exposed regions of the selected epitope.

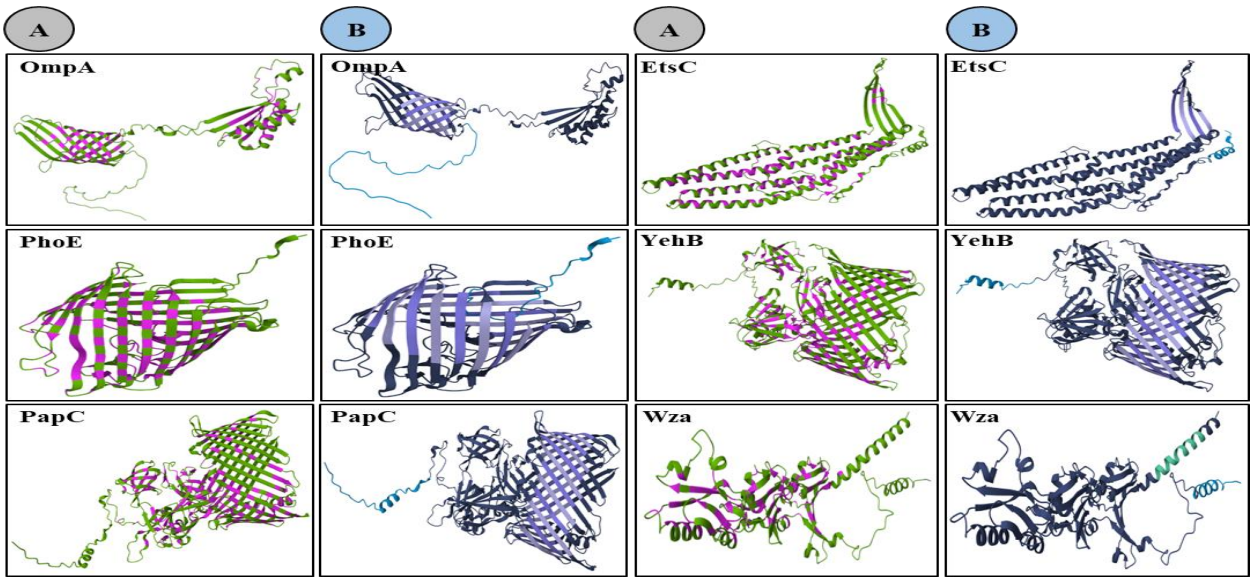


Fig. 3: Solvent accessible regions and Transmembrane topology of shortlisted proteins via TMVisDB. (A) The solvent accessible residues are marked as green, while the buried residues are marked as pink. (B) The signal peptide is marked as blue while the transmembrane region is marked as purple, containing the light and dark purple color scheme represents the outside-to-inside and inside-to-outside topology of the transmembrane region.

Table 3: List of shortlisted B-cell epitopes screened via BCPREDS analysis.

Sr. No	Protein (UniProt ID)	Epitope	Start-End	Score
1	ompA A0A0H2YXT7	PVVAPAPAPAPE	222-233	1
		EAAPVVAPAPAPAP	219-232	0.97
		ESNPVTGNTCDN	326-337	0.93
2	phoE A0A0H2YVV9	KKQNGDGFGTSL	184-195	0.98
		SDNDSKDGDQSY	47-58	0.963
		DIEGIGDEDLVNYV	295-308	0.967
		KKQNGDGFGTSLTY	184-197	0.96
		SDNDSKDGDQSYIR	47-60	0.924
3	Wza A0A0H2Z0D8	RLATYIADPQVDVN	151-164	0.98
		VIRPLKGEGGRNGK	300-313	0.95
4	etsC A0A0H2XK18	NWTKGDVDGNTSPF	37-50	0.98
		AQEKRRQAELRLDE	420-433	0.97
5	yehB A0A0H2Z0D2	ASHAYDENHHEEKR	538-551	1
		GANLTWNAPVATVN	621-634	0.96
		YEIIVKDNPQET	70-81	0.998
		RADGQPLTFGYE	757-768	0.96
		GHNIGVVGQGSQ	774-785	0.948
6	papC A0A0H2Z3A7	DGVGGVPVDGGQVV	688-701	1
		RMLPPRLRGYAPQI	281-294	0.991
		ETNARVVVSQQGRV	297-312	0.98
		GLPGVDIRPDTGAG	132-145	0.964
		AGAGWDLGVPGT	417-428	1
		RYRNDYSSREKE	488-499	1
		SVPLGTGTASYS	574-590	0.94

*Note: The blue-highlighted amino acids are exposed on the surface of the protein as predicted via the PRED-TMBB server.

Table 4: List of shortlisted APEC T-cell epitopes and their respective MHC-Alleles.

S. No	Protein (UniProt ID)	MHC-Class	MHC-Allele	Epitope	Start-End	IC50 nM	Percentile Rank %
1	ompA A0A0H2YXT7	Class- I	HLA-B*40:06	VEYAITPEI	172-180	68.8	0.02
			HLA-B*41:04	-		36.6	0.09
			HLA-B*41:03	-		35	0.09
			HLA-B*41:04	LEYQWTNNI	185-192	65.4	0.19
			HLA-B*41:04	SERRAQSVV	298-306	86.4	0.28
		Class- II	DRB1*1482	AQGVQLTAKLGYPIT	119-133	54.22	0.44
2	phoE A0A0H2YVV9	Class- I	HLA-B*41:03	FEAVAQYQF	270-278	63.2	0.19
		Class- II	DRB1*1482	DFGLRPSLGYVLSKG	279-293	53.17	0.42
				QFDFGLRPSLGYVLS	276-291	52.56	0.41
3	Wza A0A0H2Z0D8	Class- I	HLA-B*41:04	GEVKKQSTL	261-269	96.0	0.32
			HLA-B*41:03	-	261-269	90.0	0.3
4	etsC A0A0H2XK18	Class- I	HLA-B*41:04	SEYYPTFNL	299-307	36.9	0.11
			HLA-B*41:03	-	-	27.5	0.0604
		Class- II	DRB1*1310	NNDIAIAALRIHRAQ	71-85	12.20	0.2
				DIAIAALRIHRAQLD	73-87	11.13	0.15
				NDIAIAALRIHRAQL	72-86	9.97	0.11
				SNNDIAIAALRIHRA	70-84	15.47	0.36
				YAKETLRLAEVRYRA	189-203	17.76	0.5
				IAIAALRIHRAQLDA	74-88	11.57	0.17
5	yehB A0A0H2Z0D2	Class- II	DRB1*1366	DIGVFRLDFSVPQAW	119-133	79.0	0.44
				IGVFRLDFSVPQAWV	120-134	66.6	0.3
				QIAYNKFVSQTSTRF	422-436	74.1	0.4
				KSTYVRFNSGLNLLG	174-188	82.77	0.5
6	papC A0A0H2Z3A7	Class- II	DRB1*1366	YLFRAIPRWRANLTL	242-256	22.1	0.01
			DRB1*1310	-		8.03	0.05
			DRB1*1445	-		78.7	0.02
			DRB1*1366	TDYYTVSVNRYFNVF	531-545	81.6	0.47

*Note: The blue-highlighted amino acids are exposed on the surface of the protein as predicted via the PRED-TMBB server.

Chimera Designing: The chimeric vaccine was constructed by using four B-cell, four Cytotoxic T Lymphocytic (CTL), and five helper T Lymphocytic (HTL) antigenic determinants, from the six shortlisted OMPs. All the selected epitopes were proven to be 100% conserved across all the mentioned serotypes, except one T-cell epitope of the yehB protein that showed two substitutions in the sequence of the O2 serotype. The KK, AAY, and GGGS linkers were used between B-cell, CTL, and HTL epitopes. However, the GPGPG linker was used to join CTL with the HTL peptide sequence. Moreover, the EAAAK linker was added at the C-terminus to join the adjuvant and 6x-His tag at the N-terminus of the chimeric vaccine. The sequence of the constructed ensemble is given in Fig. 4.

Immunological Profiling and Solubility Analysis: The constructed vaccine was predicted to be highly antigenic by the VaxiJen server, with an antigenicity score >1.0 for the bacterial model. Moreover, the AllerTop server estimated that the chimeric protein was non-allergenic. Furthermore, the chimeric vaccine showed a solubility value of more than 0.6 as assessed via ProteinSol, predicted to be a highly soluble protein upon expression.

Assessment of Physicochemical Properties: The whole chimeric peptide consisted of 214 amino acids with a molecular weight of 23.33kDa and a theoretical isoelectric point (pI) of 9.3, indicating it to be slightly basic. According to the ExPASy ProtParam analysis, the chimeric peptide has a half-life of approximately 1hour in vitro (mammalian reticulocytes) and more than 10hours in vivo (in *E. coli*). Moreover, it was predicted as a stable protein with an instability index (II) value of 37, which is below 40. Additionally, it displayed a GRAVY value of -0.55, suggesting that the constructed chimera is predominantly hydrophilic. These characteristics correlate with the enhanced antigen expression and immune response generation in the host.

Structural Modelling, Refinement and Validation: The secondary structure prediction via PSIPRED revealed that the chimeric peptide consisted of 30% alpha helices, 16.3% beta sheets, and 53.7% coils. Moreover, the AlphaFold server predicted five models for the three-dimensional structure of the constructed chimera, as illustrated in Fig. 4. However, the model with higher confidence and containing epitopes in their native conformation was retrieved and refined via GalaxyRefiner. Additionally, the SAVES server validated the quality of the predicted 3D structure. According to the ERRAT server, the overall quality factor of the 3D model was 91.25, which was above the acceptable range (>50). The Ramachandran Plot by PROCHECK projected 97.8% residues (175aa) in the most favourable region and 2.2% residues (4aa) in the additionally allowed regions.

Molecular Docking and Simulation: The ClusPro 2.0 server was used to analyze the possible interactions and binding energies between the TLR4 receptor and the constructed multi-epitope vaccine. The model with the lowest binding energy score was retrieved showing a high binding affinity and stability of the complex. The results revealed several hydrophobic interactions and hydrogen bonds between (Ser194-Asp60, Lys200-His159, Tyr198-Glu135, Ser184-Asn156, Gln118-His258, Lys35-Gly291, Arg36-Cys289, Glu33-Lys288, Gly116-Lys254, Pro115-Lys254, Tyr81-Gln39, and Ser79-Gln39) of the chimeric vaccines and the TLR4 receptor, respectively, as shown in Fig. 5.

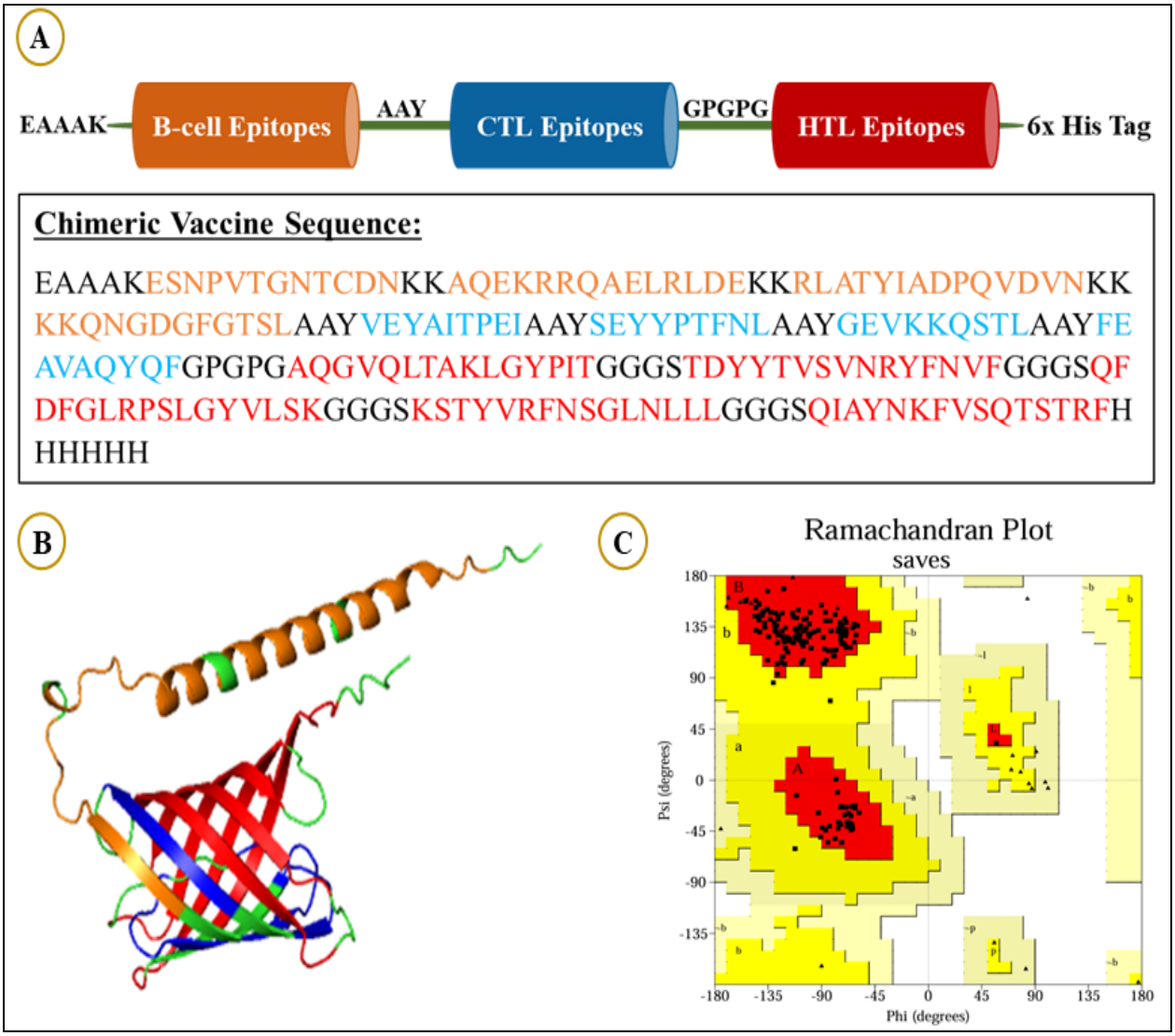


Fig. 4: Vaccine designing. A) Schematic diagram and sequence of the constructed chimera with B-cell, CTL, and HTL epitopes joined via KK, AAY, and GGGS linkers, respectively. The ensemble is a 214aa-long peptide with EAAAK adjuvant at the C-terminal and 6x His tag at the N-terminal. B) and C) The three-dimensional structure and validation of the chimeric peptide.

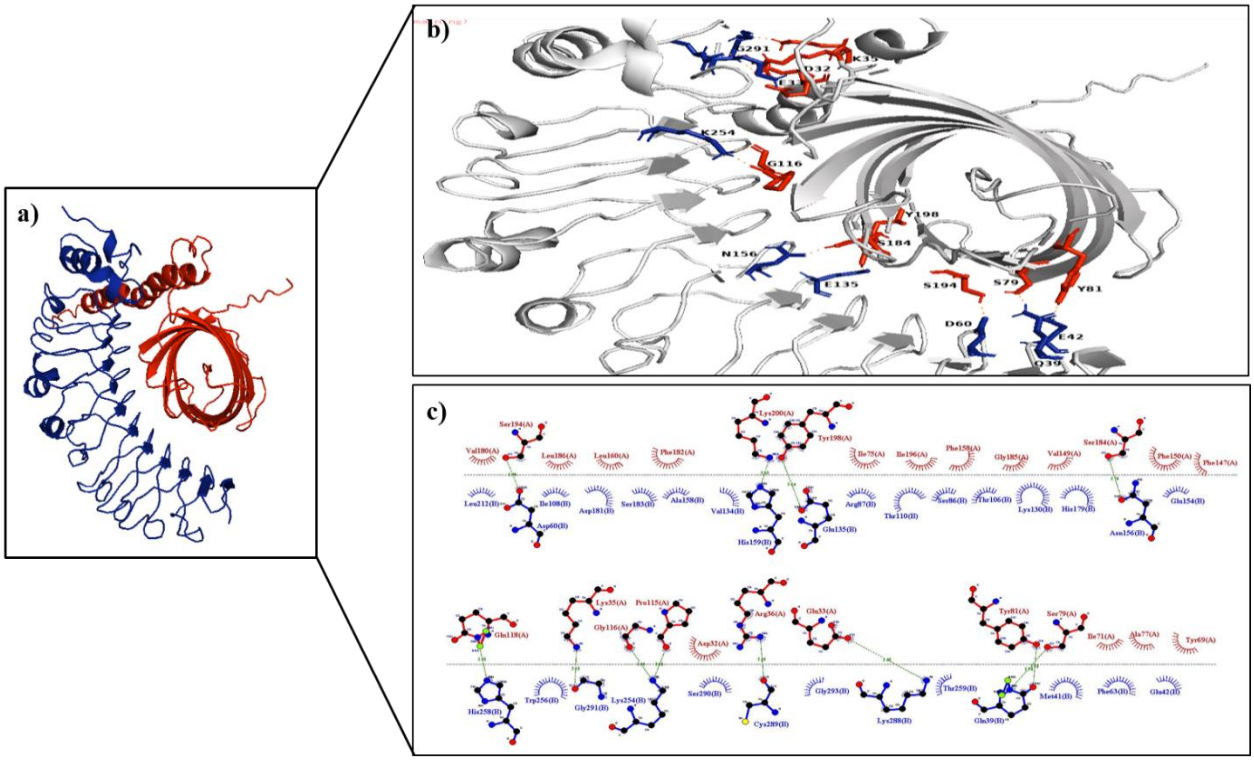


Fig. 5: Molecular docking between the constructed chimeric vaccine (shown in red) and the mammalian TLR4 receptor (shown in blue), a) three-dimensional structure of the docked complex, b) 2D interactions indicating the types of bonds and interacting residues of the docked complex are highlited in the form of stickes, c) the hydrophbic interaction are displayed in the form of eyelas structures and the hydrogen-bonds are displayed in green doted lines.

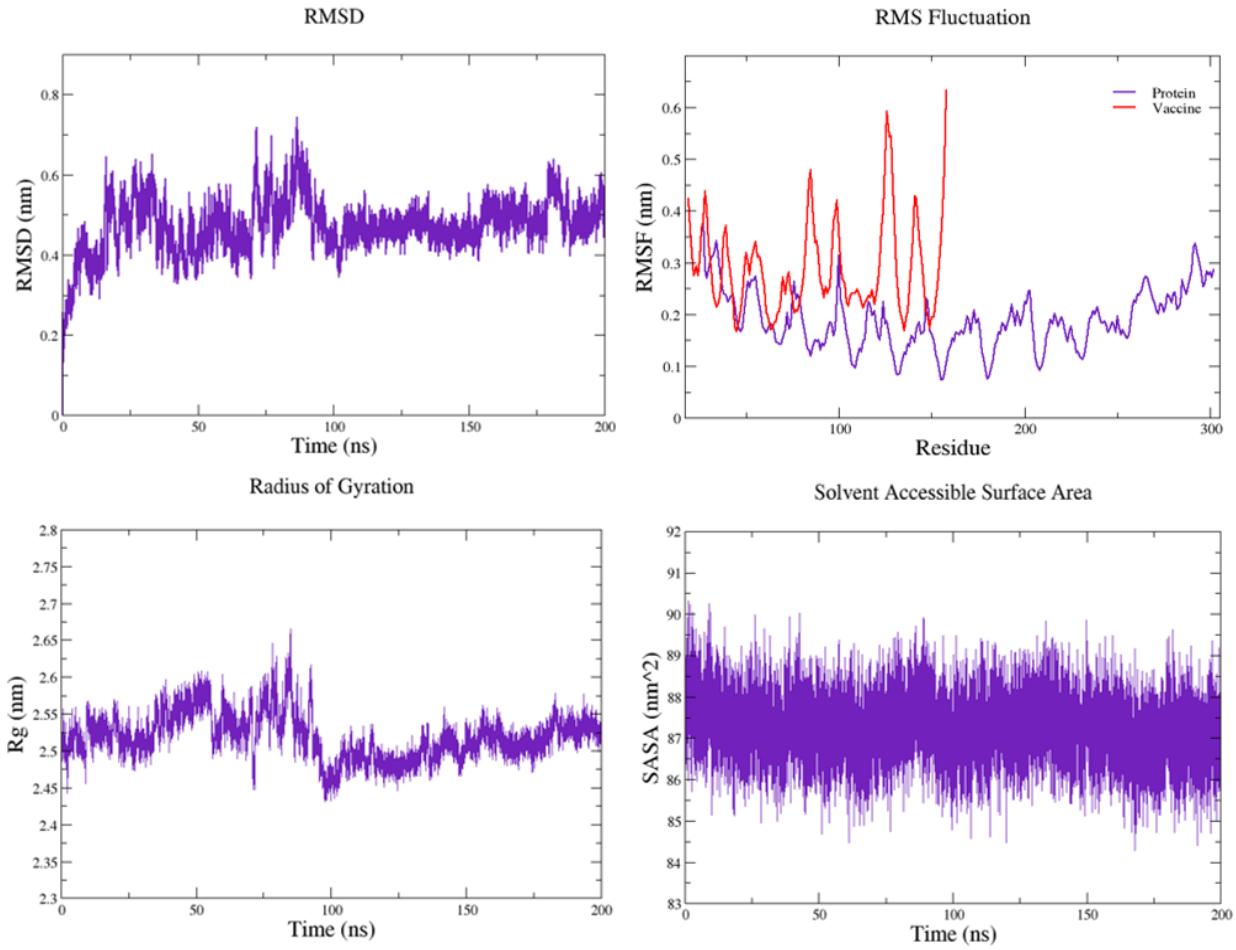


Fig. 6: Structural Dynamics of Protein and Vaccine: The RMSD, RMSF, Radius of Gyration (Rg), and SASA plots show the stability and flexibility of both the protein and vaccine during the simulation.

RMSD (Root Mean Square Deviation): The RMSD plot demonstrated a relatively stable trend throughout the 200ns simulation period. In the first 100ns, fluctuations were observed, where the RMSD values initially ranged from 0.2nm to 0.6nm, exhibiting some notable variation within this range. After the 100ns mark, the RMSD values stabilized, remaining consistently within the range of 0.4nm until the conclusion of the simulation. This steady behavior of the RMSD plot indicates that the overall tertiary structure of the antigen and TLR4 complex remained stable during the simulation, with fluctuations confined to a narrow range (0.2 to 0.6nm), as illustrated in Fig. 6. Such minor fluctuations are typical in molecular dynamics simulations, reflecting localized adjustments in the protein's structure, such as side-chain flexibility and loop movements.

RMSF (Root Mean Square Fluctuation): The RMSF plot revealed significant fluctuations in both the protein and vaccine across their respective amino acid sequences. The N-terminus of the protein exhibited high fluctuations, starting at approximately 0.4nm and gradually decreasing to a range of 0.1 to 0.2nm over the first 80aa. Between residues 81 and 200, the RMSF values remained relatively stable, fluctuating within the range of 0.1 to 0.2nm. Beyond residue 200, a notable increase in fluctuations was observed towards the C-terminus, with the plot concluding at around 0.3nm. These fluctuations suggest the presence of loop regions or regions undergoing conformational changes, which exhibit higher mobility during the simulation.

In contrast, the vaccine displayed a similar fluctuation pattern to the TLR4 protein in the initial 70 residues. However, a sharp rise in RMSF values was observed, reaching up to 0.6nm, with this peak pattern extending throughout the remaining sequence, culminating at 0.65nm at the C-terminus, as displayed in Fig.6. These sharp peaks indicate regions of the vaccine that are more dynamic, potentially undergoing substantial structural rearrangements. In summary, the overall trend suggests that the protein maintains structural rigidity, with localized flexibility observed primarily in the loop regions. Whereas, the vaccine demonstrates more pronounced fluctuations, highlighting a greater susceptibility to dynamic movements during the simulation run.

Radius of Gyration (Rg): The radius of gyration (Rg) was computed to assess the overall compactness of the protein throughout the simulation. The Rg plot initially started at 2.5nm and remained stable within this range for the first 25ns. Following this, notable fluctuations were observed between 25ns and 100ns, with the Rg values oscillating between 2.45nm and 2.65nm. After 100ns, the plot stabilized around 2.5nm, maintaining this value for the remainder of the simulation, as shown in Fig.6. These observations indicate that the protein maintained its overall compact conformation throughout the simulation period. This behavior reflects a stable tertiary structure during the simulation, with no major conformational changes affecting its overall compactness.

Solvent Accessible Surface Area (SASA): The Solvent Accessible Surface Area (SASA) was calculated to evaluate the protein's folding and stability throughout the simulation. The SASA plot exhibited minor fluctuations, remaining within the range of 86 to 89nm² over the course of the simulation, as dipicted in Fig.6. This consistent SASA behavior underscores the protein's structural stability, suggesting that its surface remained appropriately accessible for molecular interactions. Such stability is crucial for maintaining the protein's functional role in the vaccine, ensuring its continued biological activity.

DISCUSSION

This study was formulated to screen out the OMPs of APEC that have the potential to induce an effective immune response against the three most pathogenic serotypes of APEC, via proteome mining. Various APEC serotypes have been linked to avian colibacillosis; however, O1, O2, and O78 serotypes are responsible for more than 80% of the cases (Ghunaim *et al.,* 2014). The proteome mining led to the identification of 6 out of 77 OMPs of APEC as probable conserved and potential targets against the major APEC serotypes. However, the proteomes of O2 and O78 APEC serotypes are not fully annotated; therefore, the preliminary screening was performed on the complete proteome profile of the O1 serotype. Initially, the proteome was searched for OMPs because they provide the first point of interaction between the host and the pathogen, can be easily recognized, and are capable of eliciting strong immune responses owing to their surface-exposed epitopes and high conservation across the diverse serotypes. Bacterial pathogenesis and survival greatly depend upon these OMPs (Maiti *et al.,* 2020). Previous studies have utilized OMPs in vaccine design against *E. coli*, highlighting their potential as an important target in vaccine development (Hejair *et al*., 2017; Forsyth *et al.,* 2020; Nielsen *et al.,* 2020; Mohammed *et al.,* 2021). Furthermore, essential and virulent proteins, owing to their involvement in critical cellular processes, were predicted as vaccine targets (Abd El-Baky *et al.,* 2020; Chand and Singh, 2021; Khan *et al.,* 2022). These proteins are crucial for pathogen physiology, aiding in infection establishment and survival within the host. This study identified 18 out of 77 OMPs as both essential and virulent. These were further evaluated for their potential to elicit protective immunity. Therefore, 13 OMPs were predicted to possess strong antigenic and non-allergenic potentials when analysed using the AllerTOP and VaxiJen servers, emphasizing them as exciting vaccine candidates. This approach has been widely used in multiple studies (Shey *et al.,* 2019; Rawal *et al.,* 2021).

Subsequently, the physicochemical properties of potential targets, like molecular weight, instability index, and transmembrane helices, were studied to enhance the prediction accuracy of the process. Prior studies often emphasize selecting possible vaccine targets with a lower molecular weight to simplify the expression and purification process (Yan and Gao, 2020). It was also found that proteins with fewer transmembrane helices are easy to produce in large quantities through recombinant DNA technology (Yan and Gao, 2020). Consequently, OMPs exhibiting favorable physicochemical properties were selected for further analysis. This was followed by

BLASTp homology analysis of shortlisted OMPs to confirm their conservation across the three most pathogenic serotypes of APEC. As a result, 6 most highly immunogenic OMPs, OmpA, PhoE, Wza, EtsC, YehB, and PapC, were selected for the downstream analysis. Among these potential targets, OmpA, EtsC, and PapC have previously been identified as viable vaccine candidates against APEC, further validating the accuracy of this study (Kathayat *et al.,* 2021). Moreover, it emphasizes that the computational pipeline used various predefined biological parameters to gradually shortlist and prioritize the potential vaccine candidates instead of just providing random results.

In gram-negative bacteria, about 2-3% of the protein-coding genes express OMPs, crucial for their virulence and pathogenicity. The potential OMPs selected in this study belonged to four functional families: outer membrane protectin, capsular biogenesis, and fimbrial biogenesis.

OmpA is an outer membrane porin involved in the transfer of solutes across the membrane and promotes bacterial adhesion to the host epithelial cells, stimulating invasion and colonization into the host tissues. Besides, its protectin nature contributes to biofilm formation, immune evasion, serum, antibiotic resistance, and intracellular survival of the APEC (Confer and Ayalew, 2013; Nielsen *et al.,* 2020). Structurally, it comprises two domains: a surface-exposed β-barrel with a narrow pore at the N-terminal, while the periplasmic C-terminal domain residing in the periplasm (Reusch, 2012). Moreover, OmpA possesses four extracellular loops that protrude from the β-barrel and can be quickly recognized by the host immune system. These loops vary in sequence and structure, contributing to their strain-specific antigenicity (Confer and Ayalew, 2013; Nielsen *et al.,* 2020). OmpA has been widely reported as a vaccine candidate in multiple studies. Moreover, OmpA and EtsC were reported to induce a strong humoral immune response (with high IgY antibody titers) and cytokine response against the diverse serotypes of APEC. (Pore and Chakrabarti, 2013; Novinrooz *et al.,* 2017; Van Goor *et al.,* 2017).

PhoE is another outer membrane porin that facilitates nutrient uptake under phosphate-limiting conditions. Studies have also suggested its role as a carrier of antigenic peptides containing epitopes for vaccination that trigger a highly protective immune response (Janssen and Tommassen, 1994; Crépin *et al.,* 2011). Additionally, their involvement in bacterial survival and strong immunogenic properties make them a promising target for vaccine development. Structurally, it exists as a trimer, with each monomer forming a β-barrel composed of 16 antiparallel strands and the internal loop responsible for the narrowing of the channel and selective binding of ions (Cowan *et al.,* 1992).

Wza is an integral lipoprotein that is essential for the export of group 1 capsular polysaccharides (CPS) across the outer membrane. It has an octameric structure that consists of three ring-like domains which expand into the periplasm. However, an outer membrane channel made up of α-helices is present at the C-terminus of wza, and a polysaccharide export sequence is present at the N-terminus (Sande *et al.,* 2019; Sande and Whitfield, 2021). EtsC is part of the gram-negative bacteria's type 1 secretion system and functions as an auto-transporter. Its structure and functions are not well studied in *E. coli*, but it has been proven as a vaccine target based on its pathogenicity. Previously, a recombinant vaccine containing multiple antigens (OmpA, TraT, EtsC, and OmpT) was reported to elicit a strong humoral immune response in chickens against the multiple APEC serotypes (Kathayat *et al.,* 2021).

YehB and PapC are the virulence factors associated with adhesion to the host and are crucial for bacterial pathogenesis by promoting colonization, motility, and biofilm formation. YehB is an outer membrane usher protein that belongs to the yehDCBA chaperon-usher fimbrial operon (Ravan and Amandadi, 2015). PapC is an usher protein of P pili and has been reported as a remarkable virulence factor of UPEC, ExPEC, and APEC (Mitchell *et al.,* 2015; Meena *et al.,* 2021). They are associated with the assembly of fimbrial or pilus subunits and the transport of fimbriae or pilus fibers to the outer surface (Ravan and Amandadi, 2015). PapC comprises five distinct domains, including a transmembrane domain that forms a channel in the outer membrane, four globular domains comprising two C-terminal and one N-terminal domain found in the periplasmic region, and a plug domain (Volkan *et al.,* 2012). Among these proteins, PhoE represented significant T-cell epitopes in its surface-exposed region, suggesting that it would be a highly promising vaccine target.

Afterwards, the OMP focused reverse vaccinology approach was integrated to construct a chimeric vaccine by joining the promising B- and T-cell epitopes with their potential linkers. A similar technique has been used to formulate multiple vaccines against different pathogens like *Ebola*, *S. typhi H58* (XRD strain), and *Onchocerca volvulus* (Bazhan *et al.,* 2019; Shey *et al.,* 2019; Khan *et al.,* 2022). All the selected epitopes were proven to be more than 95% conserved across all the mentioned serotypes. Linkers ensure the proper folding and presentation of epitopes on antigen-presenting cells (APCs) by reducing steric hindrance between the epitopes. The KK, AAY, GPGPG, and GSSS linkers were incorporated between the potential epitopes to enhance the immunogenicity of the designed chimeric vaccine (Shahab *et al.,* 2022). Furthermore, the resulting chimera was predicted to be a highly antigenic protein with no allergenic properties, making it a potential vaccine candidate. Moreover, the 23.33kDa chimeric vaccine was evaluated to provide a soluble and stable protein upon expression. It is highly required for the recombinant protein to be expressed in a soluble and stable form for the immunization studies. Also, it showed GRAVY score of -0.55, which suggests the presence of hydrophilic epitopes on the outer surface of the designed chimera. Thus, the multi-epitope vaccine candidate has demonstrated strong ability to elicit a significant immune response in the host. According to the secondary structure assessment, the chimeric vaccine is composed of 16.3% beta sheets, 30% alpha helices, and 53.7% random coils. The 3D structure of the constructed chimera was predicted via AlphaFold and further reconstructed via GalaxyRefiner, which revealed the proper folding of multiple epitopes. Moreover, the native 3D structure of individual epitopes was highly conserved in the chimeric vaccine.

Additionally, the Ramachandran plot validated the 3D structure of the designed chimera, with 97.8% of amino acid residues lying in the most favourable region and a few in the additionally allowed region.

Toll-like receptors (TLRs) are expressed on the surfaces of several cell types, like endothelial cells and antigen-presenting cells (APCs). TLR4 plays a vital role in recognizing conserved bacterial components, leading to the induction of a strong proinflammatory cytokine response in the host. Moreover, a promising CTL response towards the chimeric vaccine requires its presentation to the CD8+ T-cells by the APCs (Kar *et al.,* 2022). Since the three-dimensional structure of avian TLR4 receptors is not available yet, the mammalian TLR4 receptor was accessed to predict the probable interactions with the constructed chimeric vaccine. It has been previously reported that the chicken TLR4 receptor has displayed 64% similarity and 46% identity to the human TLR4 protein (Ijaz *et al.,* 2025).

Furthermore, the molecular docking and simulations revealed that the vaccine-receptor complex would be stable enough to activate both innate and adaptive immune responses inside the host. Afterwards, molecular dynamics (MD) simulations were performed to evaluate the stability and dynamic behavior of the docked complex over a 200ns time period. The *in silico* results revealed that the chimeric vaccine and the TLR4 receptor maintained their compactness and structural stability throughout the MD simulations, with localized flexibility confined to the exposed loop regions. Altogether, it is anticipated that the constructed multi-epitope vaccine will make stable interactions with TLR4 receptors in vivo, presented by the APCs, inducing an effective immune response against the diverse serotypes of APEC. However, the in vitro analysis is crucial to confirm the direct molecular interactions between the recombinant vaccine and the receptor.

Conclusions: Immunoinformatics provides a faster and more streamlined alternative to traditional vaccine development methods. In this study, a computational pipeline was employed to identify six outer-membrane vaccine targets, which led to the designing of a novel chimeric vaccine. This research may address the limitation of current APEC vaccines in terms of providing broad-spectrum protection against the diverse serotypes of the pathogen. However, further experimental validation particularly through in vivo studies, is still necessary to validate the findings. Moreover, the current study provides an efficient framework for the development of future therapeutic interventions related to avian diseases.

Limitations: Nonetheless, the limitations posed by the computational research remained, such as dependence on the O1 reference proteome due to the incomplete annotation of O2 and O78 proteomes, which may have hampered the prediction of conserved targets. Moreover, the reliance on computational tools owing to the unavailability of experimental data regarding the antigen and immunogenic potential of OMPs, and the use of the human TLR4 receptor for docking and the human MHCs for epitope prediction, due to the lack of avian-specific data, may restrict the anticipated potential of the chimera against the avian pathogen and demands additional experimental validation.

Acknowledgment: The authors gratefully acknowledge the Higher Education Commission (HEC) of Pakistan for financial support through the HEC-NRPU Research Project grant (No. NRPU-15555), the Office of Research, Innovation and Commercialization (ORIC), University of the Punjab, Lahore and NM's research is supported by AUA-UAEU and NTU-UAEU grant codes G00004971 and G00005028, College of Science, United Arab Emirates University, AL-Ain, UAE for their support.

Conflict of interest: The authors declare that there is no conflict of interest.

AI-assisted writing and editing disclosure statement: While preparing this manuscript, the author(s) utilized OpenAI's ChatGPT to assist with writing, editing, and enhancing clarity, grammar, and overall structure. Following the use of this tool, the author(s) reviewed and revised the content to ensure accuracy and integrity. The final version of the manuscript reflects the author(s)' own interpretations and modifications, and they assume full responsibility for its content.

REFERENCES

Abd El-Baky RM, Ibrahim RA, Mohamed DS, *et al.,* 2020. Prevalence of virulence genes and their association with antimicrobial resistance among pathogenic *E. coli* isolated from Egyptian patients with different clinical infections. Infection and Drug Resistance 13:1221-1236.

Alber A, Stevens MP and Vervelde L, 2021. The bird's immune response to avian pathogenic *Escherichia coli*. Avian Pathology 50(5):382-391.

Bagos PG, Liakopoulos TD, Spyropoulos IC, *et al.,* 2004. PRED-TMBB: a web server for predicting the topology of β-barrel outer membrane proteins. Nucleic Acids Research 32:W400-W404.

Bazhan SI, Antonets DV, Karpenko LI, *et al.,* 2019. *In silico* designed Ebola virus T-cell multi-epitope DNA vaccine constructions are immunogenic in mice. Vaccines 7(2):34.

Breuza L, Poux S, Estreicher A, *et al.,* 2016. The UniProtKB guide to the human proteome. Database 2016:bav120.

Buchan DW and Jones DT, 2019. The PSIPRED protein analysis workbench: 20 years on. Nucleic Acids Research 47(W1):W402-W407.

Chand Y and Singh S, 2021. Prioritization of potential vaccine candidates and designing a multiepitope-based subunit vaccine against multidrug-resistant *Salmonella Typhi* str. CT18: a subtractive proteomics and immunoinformatics approach. Microbial Pathogenesis 159:105150.

Chen L, Yang J, Yu J, *et al.,* 2005. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Research 33:D325-D328.

Christensen H, Bachmeier J and Bisgaard M, 2021. New strategies to prevent and control avian pathogenic *Escherichia coli* (APEC). Avian Pathology 50(5):370-381.

Confer AW and Ayalew S, 2013. The OmpA family of proteins: roles in bacterial pathogenesis and immunity. Veterinary Microbiology 163(3-4):207-222.

Cowan S, Schirmer T, Rummel G, *et al.,* 1992. Crystal structures explain functional properties of two *E. coli* porins. Nature 358(6389):727-733.

Crépin S, Chekabab SM, Le Bihan G, *et al.,* 2011. The Pho regulon and the pathogenesis of *Escherichia coli.* Veterinary Microbiology 153(1-2):82-88.

Dimitrov I, Bangov I, Flower DR, *et al.,* 2014. AllerTOP v. 2—a server for *in silico* prediction of allergens. Journal of Molecular Modeling 20(6):2278.

Dziva F and Stevens MP, 2008. Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic *Escherichia coli* in their natural hosts. Avian Pathology 37(4):355-366.

Fancher CA, Zhang L, Kiess AS, *et al.,* 2020. Avian pathogenic *Escherichia coli* and *Clostridium perfringens*: challenges in no antibiotics ever broiler production and potential solutions. Microorganisms 8(10):1533.

Fatoba AJ, Adeleke VT, Maharaj L, *et al.,* 2022. Design of a multiepitope vaccine against chicken anemia virus disease. Viruses 14(7):1456.

Forsyth VS, Himpsl SD, Smith SN, *et al.,* 2020. Optimization of an experimental vaccine to prevent Escherichia coli urinary tract infection. mBio 11(2):10-28.

Gasteiger E, Hoogland C, Gattiker A, *et al.,* 2005. Protein identification and analysis tools on the ExPASy server. The Proteomics Protocols Handbook:571-607.

Ghunaim H, Abu-Madi MA and Kariyawasam S, 2014. Advances in vaccination against avian pathogenic *Escherichia coli* respiratory disease: potentials and limitations. Veterinary Microbiology 172(1-2):13-22.

Hebditch M, Carballo-Amador MA, Charonis S, *et al.,* 2017. Protein–Sol: a web tool for predicting protein solubility from sequence. Bioinformatics 33(19):3098-3100.

Hejair HM, Ma J, Zhu Y, *et al.,* 2017. Role of outer membrane protein T in pathogenicity of avian pathogenic *Escherichia coli.* Research in Veterinary Science 115:109-116.

Heo L, Park H and Seok C, 2013. GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Research 41(W1):W384-W388.

Hu J, Afayibo DJA, Zhang B, *et al.,* 2022. Characteristics, pathogenic mechanism, zoonotic potential, drug resistance, and prevention of avian pathogenic *Escherichia coli* (APEC). Frontiers in Microbiology 13:1049391.

Ijaz A, Putman M, Broere F, *et al.,* 2025. Activation of chicken macrophages by isomalto/malto-polysaccharide (IMMP) is facilitated by toll-like receptor 4 (TLR4). Poultry Science 104:105690.

Janssen R and Tommassen J, 1994. PhoE protein as a carrier for foreign epitopes. International Reviews of Immunology 11(2):113-121.

Joseph J, Jennings M, Barbieri N, *et al.*, 2023. Characterization of avian pathogenic *Escherichia coli* isolated from broiler breeders with colibacillosis in Mississippi. Poultry 2(1):24-39.

Jumper J, Evans R, Pritzel A, *et al.,* 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596(7873):583-589.

Jurtz V, Paul S, Andreatta M, *et al.,* 2017. NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. The Journal of Immunology 199(9):3360-3368.

Kar PP, Araveti PB, Kuriakose A, *et al.,* 2022. Design of a multi-epitope protein as a subunit vaccine against lumpy skin disease using an immunoinformatics approach. Scientific Reports 12(1):19411.

Kathayat D, Lokesh D, Ranjit S, *et al.,* 2021. Avian pathogenic *Escherichia coli* (APEC): an overview of virulence and pathogenesis factors, zoonotic potential, and control strategies. Pathogens 10(4):467.

Khan K, Jalal K and Uddin R, 2022. An integrated *in silico* based subtractive genomics and reverse vaccinology approach for the identification of novel vaccine candidate and chimeric vaccine against XDR *Salmonella typhi* H58. Genomics 114(2):110301.

Liang YT, Luo H, Lin Y, *et al.,* 2024. Recent advances in the characterization of essential genes and development of a database of essential genes. iMeta 3(1):e157.

Liu F, Yuan C, Chen H, *et al.,* 2024. Prediction of linear B-cell epitopes based on protein sequence features and BERT embeddings. Scientific Reports 14(1):2464.

Lutful Kabir S, 2010. Avian colibacillosis and salmonellosis: a closer look at epidemiology, pathogenesis, diagnosis, control and public health concerns. International Journal of Environmental Research and Public Health 7(1):89-114.

Mahram A and Herbordt MC, 2015. NCBI BLASTP on high-performance reconfigurable computing systems. ACM Transactions on Reconfigurable Technology and Systems (TRETS) 7(4):1-20.

Maiti B, Dubey S, Munang'Andu HM, *et al.,* 2020. Application of outer membrane protein-based vaccines against major bacterial fish pathogens in India. Frontiers in Immunology 11:1362.

Marquet C, Grekova A, Houri L, *et al.,* 2022. TMvisDB: resource for transmembrane protein annotation and 3D visualization. BioRxiv:2022.11.30.518551.

Meena PR, Yadav P, Hemlata H, *et al.,* 2021. Poultry-origin extraintestinal *Escherichia coli* strains carrying the traits associated with urinary tract infection, sepsis, meningitis and avian colibacillosis in India. Journal of Applied Microbiology 130(6):2087-2101.

Mellata M, 2013. Human and avian extraintestinal pathogenic *Escherichia coli*: infections, zoonotic risks, and antibiotic resistance trends. Foodborne Pathogens and Disease 10(11):916-932.

Mellata M, Dho-Moulin M, Dozois CM, *et al.,* 2003. Role of virulence factors in resistance of avian pathogenic *Escherichia coli* to serum and in pathogenicity. Infection and Immunity 71(1):536-540.

Mistry J, Chuguransky S, Williams L, *et al.,* 2021. Pfam: the protein families database in 2021. Nucleic Acids Research 49(D1):D412-D419.

Mitchell NM, Johnson JR, Johnston B, *et al.,* 2015. Zoonotic potential of *Escherichia coli* isolates from retail chicken meat products and eggs. Applied and Environmental Microbiology 81(3):1177-1187.

Mohammed GM, ElZorkany HE, Farroh KY, *et al.,* 2021. Potential improvement of the immune response of chickens against *E. coli* vaccine by using two forms of chitosan nanoparticles. International Journal of Biological Macromolecules 167:395-404.

Monroy I, Catalá-Gregori P and Sevilla-Navarro S, 2025. Assessment of antibiotic resistance and virulence in *Escherichia coli* strains isolated from poultry in Spain. Poultry Science 104(2):104838.

Nawaz S, Wang Z, Zhang Y, *et al*., 2024. Avian pathogenic *Escherichia coli* (APEC): current insights and future challenges. Poultry Science 103(12):104359.

Nielsen DW, Ricker N, Barbieri NL, *et al.,* 2020. Outer membrane protein A (OmpA) of extraintestinal pathogenic Escherichia coli. BMC Research Notes 13(1):51.

Nielsen H, Teufel F, Brunak S, *et al.,* 2024. SignalP: the evolution of a web server. Protein Bioinformatics. Springer:331-367.

Nilsson JB, Kaabinejadian S, Yari H, *et al.,* 2023. Accurate prediction of HLA class II antigen presentation across all loci using tailored data acquisition and refined machine learning. Science Advances 9(47):eadj6367.

Novinrooz A, Zahraei Salehi T, Firouzi R, *et al.,* 2017. In-silico design, expression, and purification of novel chimeric *Escherichia coli* O157:H7 OmpA fused to LTB protein in *Escherichia coli*. PLoS One 12(3):e0173761.

Pokharel P, Dhakal S and Dozois CM, 2023. The diversity of *Escherichia coli* pathotypes and vaccination strategies against this versatile bacterial pathogen. Microorganisms 11(2):344.

Pore D and Chakrabarti MK, 2013. Outer membrane protein A (OmpA) from *Shigella flexneri* 2a: a promising subunit vaccine candidate. Vaccine 31(36):3644-3650.

Ravan H and Amandadi M, 2015. Analysis of yeh fimbrial gene cluster in *Escherichia coli* O157:H7 in order to find a genetic marker for this serotype. Current Microbiology 71(2):274-282.

Rawal K, Sinha R, Abbasi BA, *et al.,* 2021. Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. Scientific Reports 11(1):17626.

Reusch RN, 2012. Insights into the structure and assembly of *Escherichia coli* outer membrane protein A. The FEBS Journal 279(6):894-909.

Sande C, Bouwman C, Kell E, *et al.,* 2019. Structural and functional variation in outer membrane polysaccharide export (OPX) proteins from the two major capsule assembly pathways present in *Escherichia coli.* Journal of Bacteriology 201(14):10-28.

Sande C and Whitfield C, 2021. Capsules and extracellular polysaccharides in *Escherichia coli* and *Salmonella*. EcoSal Plus 9(2):eESP-0033-2020.

Shahab M, Hayat C, Sikandar R, *et al*., 2022. *In silico* designing of a multi-epitope vaccine against *Burkholderia pseudomallei*: reverse vaccinology and immunoinformatics. Journal of Genetic Engineering and Biotechnology 20(1):100.

Sharma A, Garg A, Ramana J, *et al.,* 2023. VirulentPred 2.0: an improved method for prediction of virulent proteins in bacterial pathogens. Protein Science 32(12):e4808.

Shey RA, Ghogomu SM, Esoh KK, *et al.,* 2019. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. Scientific Reports 9(1):4409.

Valdivia-Olarte H, Requena D, Ramirez M, *et al.,* 2015. Design of a predicted MHC restricted short peptide immunodiagnostic and vaccine candidate for Fowl adenovirus C in chicken infection. Bioinformation 11(10):460.

Van Goor A, Stromberg ZR and Mellata M, 2017. A recombinant multi-antigen vaccine with broad protection potential against avian pathogenic *Escherichia coli*. PLoS One 12(8):e0183929.

Volkan E, Ford BA, Pinkner JS, *et al.,* 2012. Domain activities of PapC usher reveal the mechanism of action of an *Escherichia coli* molecular machine. Proceedings of the National Academy of Sciences 109(24):9563-9568.

Watts A and Wigley P, 2024. Avian pathogenic *Escherichia coli*: an overview of infection biology, antimicrobial resistance and vaccination. Antibiotics 13(9):809.

Yan F and Gao F, 2020. A systematic strategy for the investigation of vaccines and drugs targeting bacteria. Computational and Structural Biotechnology Journal 18:1525-1538.

Yang M, Derbyshire MK, Yamashita RA, *et al.,* 2020. NCBI's conserved domain database and tools for protein domain analysis. Current Protocols in Bioinformatics 69(1):e90.

Zaharieva N, Dimitrov I, Flower DR, *et al.,* 2019. VaxiJen dataset of bacterial immunogens: an update. Current Computer-Aided Drug Design 15(5):398-400.