scientific reports

Check for updates

OPEN

Bioinformatics analysis of the in silico engineered protein vaccine with and without *Escherichia coli* heat labile enterotoxin adjuvant on the model of *Klebsiella pneumoniae*

Kimia Jafari Ranjbar[1], Parisa Sarkoohi[2], Behzad Shahbazi[3,4], Maryam Babaei[1] & Khadijeh Ahmadi[5]✉

***Klebsiella pneumoniae* (*K. pneumoniae*) has been identified as a major cause of nosocomial infections with multidrug-resistant phenotypes. Vaccination is one of the most effective methods to prevent infectious diseases. We aim to design a vaccine candidate based on the epitope-rich domains of the OmpA, OMPK17, and fimb proteins of *K. pneumoniae* that could protect against this infection. A vaccine structure was constructed by selecting five epitope-rich domains from three proteins. We decided to add the heat-labile toxin (LT) of *Escherichia coli* as an adjuvant to the designed protein structure. The evaluation of the vaccine candidates' interaction with the immune system's receptors showed an appropriate interaction of the specially adjuvated protein with TLR2 and TLR4. The stability of the interactions was also studied by molecular dynamics (MD) for to 100 ns. All parameters showed that the structure of the candidate proteins alone and in complex with TLR2 and TLR4 are stable, especially the adjuvanted protein. Immune response simulations showed that both candidates induce acceptable protective immune responses. Overall, the LT-adjuvanted design protein may have the potential to induce more favorable protective immune responses. However, further in vitro and in vivo studies are required to obtain more definitive results.**

Keywords *Klebsiella pneumoniae*, Epitope, Vaccine, Docking, Molecular dynamics

Klebsiella pneumoniae, a member of the *Enterobacteriaceae* family, is characterized as an encapsulated, nonmotile, gram-negative bacterium. *Klebsiella pneumoniae* has been prioritized by the World Health Organization (WHO) as one of the top three pathogens of international concern in 2017[1]. There are emerging issues of global health importance today, and one of them is the rapid spread of multidrug-resistant (MDR) bacteria[2]. The antimicrobial resistance of *K. pneumoniae* in severe healthcare-associated infections is approximately 50% worldwide, according to a global WHO report[3]. Thus, there is an urgent need for effective strategies to prevent *K. pneumoniae* infections[4]. Many investigators have explored routes for the development of vaccines against *K. pneumoniae*, including inactivated whole cell *K. pneumoniae*, capsular polysaccharide-based vaccine, ribosomal-based vaccines and etc[4–6]. However, to date, no vaccine has been licensed to provide complete protective immunity against this bacterium.

Vaccination is widely acknowledged as one of the most efficacious methods for preventing infectious diseases.

Immunoinformatics, a specialized branch of bioinformatics that uses computational tools and algorithms to predict, analyze, and validate immune responses, plays a key role in the design of vaccine epitopes. Epitope-based vaccines aim to induce specific immune responses by targeting key antigenic components of pathogens, making immunoinformatics an essential tool in this process[7,8]. Immunoinformatics tools facilitate the identification of B and T cell epitopes from pathogen genomes and proteomes. These epitopes are critical regions on antigens

[1]Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran. [2]Department of Pharmacology and Toxicology, Faculty of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran. [3]School of Pharmacy, Semnan University of Medical Sciences, Semnan, Iran. [4]Nervous System Stem Cells Research Center, Semnan University of Medical Sciences, Semnan, Iran. [5]Department of Medical Biotechnology, School of Paramedicine, Bushehr University of Medical Sciences, Bushehr, Iran. ✉email: khadijeh_6482@yahoo.com

recognized by the immune system. Advanced algorithms predict MHC-binding peptides (both class I and class II) that are essential for T cell-mediated immunity and surface-exposed or structural epitopes for B cell activation[9,10]. Antigenic and immunogenicity computational models evaluate predicted candidate epitopes, ensuring that the selected epitopes are effective and can elicit robust immune responses[11]. The tools also assess allergenic potential to ensure the safety of vaccine candidates[12]. Immunoinformatics helps design multi-epitope vaccines by combining T- and B-cell epitopes in a single construct. This approach ensures broader protection by targeting diverse immune pathways and pathogen strains[13]. Immunoinformatics tools analyze pathogen diversity to identify epitopes that are conserved across multiple strains or serotypes[14]. This is particularly useful for designing universal or broad-spectrum vaccines[15]. By reducing reliance on traditional experimental methods, immunoinformatics accelerates the vaccine design process. Computational predictions significantly reduce cost and time while increasing the accuracy of selecting promising candidates for experimental validation[16]. Multi-epitope vaccine strategy also offers several advantages over whole organism vaccines, including lack of infectivity, immunogenicity, ease of production, reduced allergic and reactive responses. In the case of peptide vaccines, there is evidence that some peptides alone are capable of eliciting the necessary immune responses[17]. Recent advancements in bioinformatics, immunoinformatics, and vaccinology have ushered in a new era of vaccine development, offering a more contemporary, robust, and realistic approach to designing potent immunogens. Computational tools and techniques provide a rational and efficient framework for the development of safe, effective, and specific vaccines, especially against challenging pathogens[18]. One of the most fundamental steps in the design of a multi-epitopic vaccine is the selection of appropriate antigens. The selected antigens should play an essential role in the function or survival of the pathogen. In this study, we aim to implement a principled and rational approach to the immunoinformatics design of a multiepitopic vaccine against *K. pneumoniae*. Since both types of B cells and T cells have an important role in the fight against *K. pneumoniae* infection, the prediction of B cell and T cell epitopes may be important in the design of a vaccine against *K. pneumoniae* in order to stimulate both arms of the immune system[19]. *Klebsiella pneumoniae*-associated diseases express through various virulence factors, including capsules, mucoviscosity-associated exopolysaccharides, lipopolysaccharides (LPSs), adhesins, and iron uptake systems[20,21]. Outer membrane proteins (OMPs) serve as crucial immunogenic entities in bacteria, imparting characteristics such as immune evasion, stress tolerance, and resistance to antibiotics/antimicrobials, as well as protecting bacteria against lung collectins[22–24]. While several factors may be involved in the pathogenic potential of *K. pneumoniae*, OmpA stands out significantly, contributing to the pathogen's virulence, complement resistance, and biofilm formation[23]. Another key member of the OMP family implicated in *K. pneumoniae* pathogenesis is Outer Membrane Protein-K17 (OMPK17, also known as OMPX). Studies have demonstrated that *K. pneumoniae* OMPK17 can be detected by immune serum from patients with acute *K. pneumoniae* infection[25,26]. In gram-negative bacteria, adhesion is frequently facilitated by fimbriae, filamentous organelles expressed on the bacterial cell surface[27]. Fimbriae play a pivotal role as a virulence factor for *K. pneumoniae*, enabling the pathogen to adhere to and invade host cells, as well as form biofilms on both biotic and abiotic surfaces[28,29]. New-generation vaccines, designed with subunit components or nucleic acids, exhibit lower immunogenicity compared to traditional vaccines containing live-attenuated or inactivated pathogens[30]. The adjuvant augments the strength and longevity of the vaccine, modifying the impact of subsequent adaptive immune responses to vaccine antigens without causing a specific immune response against itself[31–34]. *Escherichia coli (E. coli)* LT has been shown to be a potent adjuvant capable of enhancing the antigenicity of vaccines, as a platform for the display of proteins or antigenic peptides for the development of novel vaccines, and as a vehicle for the delivery of naturally derived proteins to cells[35–37].

We aimed to use immunoinformatics tools to design a multi-epitope protein targeting different mechanisms using OmpA, OMPK17 and fimbriae proteins that play important roles in the survival and function of *K. pneumoniae*, in addition, we designed a multi-epitope protein combined with LT adjuvant to compare its immunogenicity. To develop a robust vaccine against *K. pneumoniae*, we used reverse vaccinology to identify immune epitopes within the OmpA, OMPK17 and fimbriae proteins that are essential for *K. pneumoniae* function. We used computational approaches to evaluate the interaction of the recombinant proteins with TLR2 and TLR4 to assess their biological activity. Finally, we used in silico immune simulations to analyze the impact of the protein vaccine candidates on the induction of immune function.

Results

Genome extraction of OmpA, OMPK17 and fimbriae proteins

The sequences of OmpA (Accession number: W9BFP5), OMPK17 (Accession number: Q48427) and fimbriae (Accession number: W1DI64) proteins were extracted from the Uniprotkb database (https://www.uniprot.org/).

Prediction of B-cell epitopes

B cells, which are potent inducers of humoral immunity, are essential for complete host protection against *Klebsiella* infections. Therefore, for the development of new vaccine candidates against this bacterium, the identification of antigenic B cell epitopes is valuable[38]. Prediction of B-Cell epitopes for all three proteins OmpA, OMPK17 and fimbriae was done using Bepipred 2.0 and Kolaskar and Tongaonkar of IEDB server. Among the epitopes predicted by Kolaskar and Tongaonkar, epitopes with scores higher than the threshold of 1.02 were selected (Supplementary Table S1). The selected epitopes from the BepiPred 2.0 web server also had scores above the threshold of 0.5 (Supplementary Table S1).

Prediction of T-cell epitopes

CD4+ T cells are significantly increased in cells involved in primary *K. pneumoniae* infection, and mice lacking T cells are not protected from reinfection, suggesting that T cells are required for the protective response against this bacterium[39]. For the prediction of T cell epitopes, the IEDB (NetMHCIIpan 4.1 BA) and Rankpep databases

were also used. A set of MHCII alleles DRB1*0101, *0301, *0401, *0701, *0801, *1101, *1301, *1501 was chosen to cover most human genetic backgrounds. The most important epitopes with the highest scores were selected. In the Rankpep database, the binding threshold score was different for each allele. In this prediction, only epitopes categorized in the binding type highlighted in red and having a score higher than the threshold binding were selected (Supplementary Table S2). Epitopes predicted by NetMHCIIpan 4.1 BA with IC50 less than 50 were selected as good binders (Supplementary Table S2).

Selection of epitope-rich regions and vaccine candidates design

To select the domains that make up the vaccine candidate, the regions of these three proteins of OmpA, OMPK17 and fimbriae with the highest epitope abundance and antigenicity are considered the target domain for vaccine design. Under these conditions, five epitope-rich domains from these three proteins were finally selected as vaccine candidates containing many B- and T-cell epitopes (Table 1). By combining the selected domains with EAAAK and GGGGS linkers, 48 protein sequences were designed. The 48 designed sequences were evaluated based on physicochemical, antigenic, secondary structure and tertiary structure characteristics, and finally the most suitable protein was introduced as a vaccine candidate (Fig. 1). We also added *E. coli* LT adjuvant to the selected vaccine candidate structure and by comparing the characteristics of the designed protein with and without adjuvant, we introduced the best vaccine candidate against *K. pneumoniae* (Fig. 1).

Physical and chemical properties of designed structures

ProtParam and Protscale tools (http://expasy.org/tools/protparam.html) were used to obtain the physical and chemical properties of the designed structures. These properties included number of amino acids, molecular weight, PI, number of charged amino acids, amino acid composition, hydrophobicity and hydrophilicity. The physicochemical parameters of the designed protein and the adjuvant protein have been evaluated by means of ProtParam. The results of this investigation are presented in Table 2. Both proteins were found to be stable at different temperatures.

Antigenicity, allergenicity, toxicity and solubility evaluation

The antigenicity of the designed protein was calculated by the VaxiJen v2.0 server. The results indicated that the vaccine candidate had a high potential to stimulate the immune system. The antigenicity of the selected protein 0.6048 was predicted and it was 0.7242 for vaccine candidate + adjuvant. AllerTOP v.2 web server was used to check the allergenicity of the designed structure. The designed protein and protein with adjuvant were not allergenic. Also, the solubility of the vaccine candidates were evaluated using Recombinant Protein Solubility Prediction (https://biotech.ou.edu/), which indicated the both proteins have the potential for solubility when overexpressed in *E. coli*. The designed proteins were submitted to the ToxinPred2 server and the results showed that the vaccine candidates were not-toxic (Supplementary Table S3).

Secondary and tertiary structure prediction and validation

The GOR IV software was used to check the second structure of the designed structures. The amino acids that make up these recombinant proteins are involved in forming random coils, alpha helixes, and beta strands. The results of the prediction of the second structure of the unadjuvanted and the adjuvanted protein are shown in Table 3 and Fig. 2a–b. The I-TASSER and GalaxyWEB web servers for the prediction of unadjuvanted and adjuvanted tertiary structures of protein sequences. All structures were validated and the best structure was selected. GalaxyWeb predicted 3D structures of higher quality than I TASSER. Predicted tertiary structures were evaluated using the MolProbity, ProSA-web, and SAVES v6.1 servers. The MolProbity server was used to assess the structural similarity of new proteins to the best-known structures of similar proteins (http://molprobity.biochem.duke.edu/index.php). In the MolProbity analysis, the protein structure analysis was evaluated based on the Clash and MolProbity scores. The MolProbity score, normalized to a scale similar to X-ray resolution, combines the Clashscore, Rotamer, and Ramachandran scores into a single score. Among structures of similar resolution, the 100th percentile is the best and the 0 percentile is the worst[40]. A good and acceptable MolProbity score is less than 2[41]. The SAVES v6.1 server (https://saves.mbi.ucla.edu/) was also used to check the Ramachandran plot and evaluate the placement of amino acids in the favored, allowed, and disallowed regions. The results of the tertiary structure evaluation of two proteins with and without adjuvants are shown in Table 4. Ramachandran

Antigenic determinant	Position	Antigen
MNFKKSIALLFIALNIASLPTYAGVSKTFKDKCASTTAKLVQSVQLVKLASDTNKD SKGIYITDSTGKTRFIPGGQYYPENYLSNEMRKIAMAAVLSNVRVNICASEAYTPNHVWAIELAAE		LT adjuvant
MLSLGVSYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLK PEGQQALDQLYTQLSNMDPKDGSAVVLGYT	210–290	OmpA 1
AGKISARGMGESTPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAA	320–374	OmpA 2
YNKGQYYGITAGPAYRLNDWASIYGVVGVGYGKFQNNNYPHKSDMSDYGFSYGAGLQFNPIENVALDFSYEQSRIRNVDVG	80–160	OMPK17
DVRLANDNTISTDYRGYAIVPYVTPYRRTDITLDSTTLGEDMELPETTKSVVPTRGAIVRASYDGNIGQRAFVHL KTASGQDVPYGAMVLLVGDSKSQPSIVSDAGMVYMSGLQQTGILNVQWGKSAAQQCNASFTLPTREGKASGISKS KRFAVSHSQGNILMKKIVILITLLMLGGQAA	620–800	fimbriae 1
SCNISTPRSQITLNRIDKGKLMSLSRGATTPSEKTIAMNISCPNDSVGNTLIYWFNPVGGI SASGNGIVDNMLSGSDAANNVGIIFKLSGSPVVFYDTDNYNYKINTSNDLNKTINLTADY	980–1100	fimbriae 2

Table 1. Five epitope-rich domains selected from three proteins OmpA, OMPK17 and fimbriae.

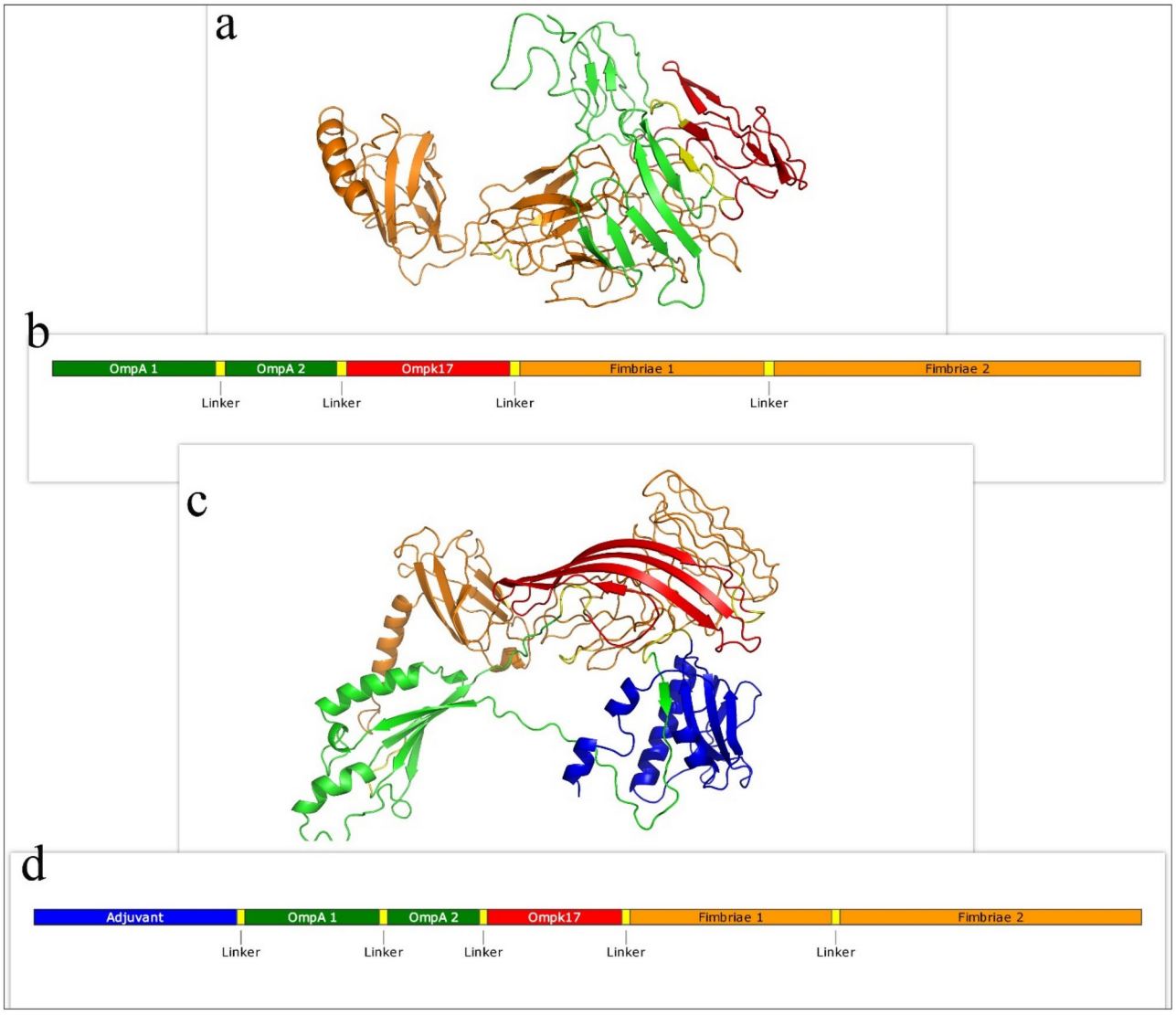

Fig. 1. (**a**) The tertiary structure of the designed protein. (**b**) Schematic representation of the final construct of the multi-epitope protein. (**c**) The tertiary structure of the adjuvanted protein. (**d**) Schematic representation of the final construct of the adjuvanted protein.

Property	Vaccine candidate	Vaccine + adjuvant
Number of amino acids	539	666
Molecular weight (dalton)	57,702.07	70,719.80
Number of negatively charged amino acids	50	55
Number of positively charged amino acids	51	61
Isoelectric point (PI)	7.53	8.64
Instability index	31.77	33.85
Estimated half-life	30 h in mammalian cells, more than 20 h in yeast cells and more than 10 h in *E.coli*	30 h in mammalian cells, more than 20 h in yeast cells and more than 10 h in *E.coli*
Aliphatic index	81.11	80.71

Table 2. Physicochemical characteristics of protein candidate and adjuvanted protein.

Secondary structure	Vaccine candidate	Vaccine + adjuvant
Alpha-helix	133 (24.68%)	125 (18.77%)
Extended-strand	103 (19.11%)	159 (23.87%)
Random coil	303 (56.22%)	382 (57.36%)

Table 3. Evaluation of the second structure of unadjuvanted and adjuvanted proteins using the GOR web server.

plot analysis further validated the reliability of the model, showing that most residues are located in the most favoured and allowed regions for both unadjuvanted (Fig. 2c) and adjuvanted (Fig. 2d) proteins (Table 4). Also, unadjuvanted (Fig. 2e) and adjuvanted (Fig. 2f) proteins had Z-scores within the range of the native protein structure (Table 4). Proteins less than 1000 residues with z-score ≤ 10 are acceptable[42]. These results support the suitability of the models for subsequent molecular docking and dynamic simulation studies, suggesting that the backbone conformation of the generated model is favourable.

Prediction of conformational B-cell epitopes

Analysis of the 3D structure of the vaccine candidates using ElliPro Server revealed conformational epitope regions (Fig. 3a–e). Predicted conformational B cell epitopes ranged from 0.754 to 0.863 for the unadjuvanted protein (Fig. a–c) and from 0.876 to 0.832 for the adjuvanted protein (Fig. d–e) (Table 5).

Protein–protein molecular docking

Cluspro 2.0 was used to analyze protein–protein binding between vaccine candidates designed with and without adjuvants with TLR4 and TLR2. To select the best interaction, weighted score parameters, number of bindings between vaccine candidates and TLR4 and TLR2 receptors, affinity and binding free energy were used. The molecular interactions between the residues in the vaccine-receptor complexes were demonstrated using PDBsum generate web server (Figs. 4 and 5). We considered the lowest energy, the most negative weight score, the lowest Kd for affinity, and the highest number of amino acids involved in binding, especially with active site residues, as essential standards for selecting the strongest complexes. As shown in Figs. 4 and 5, the vaccine candidates formed a strong interaction with the active site of TLR2 and TLR4, and this binding includes the essential amino acids Ile319, Phe322, Phe325, Tyr326, Val348, Phe349, and Pro352 for TLR2 and the amino acids Arg434. Arg380, Lys341, Lys263 and Gln339 for TLR4. The results showed that the binding of adjuvanted vaccine compared to unadjuvanted vaccine has higher affinity and energy with immune system receptors. It also interacts with all amino acids of the active site (Figs. 4 and 5) (Table 6). Residues involved in the interaction of vaccine-receptor complexes were visualized using PyMOL (The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC) and mapped using the PDBsum generate web server.

Evaluate the stability of structures using molecular dynamics simulation

Molecular dynamics simulations up to 100 ns (ns) were performed to verify the stability of the designed protein structures and protein-receptor complexes. The root mean square distance (RMSD) parameter is used when analyzing the results of molecular dynamics simulations of proteins and complexes to obtain the degree of movement of the protein or atoms when the ligand is placed in the active site of the receptor and to evaluate the stability of the structure during the simulation period. The analysis of the results in terms of the RMSD of the adjuvanted and non-adjuvanted proteins and their complexes with the receptor showed that both designed proteins were stabilized during the simulation with an average RMSD of 0.69 and 0.61 nm, respectively (Fig. 6a). The protein-TLR2 complexes with an average RMSD of 1.04 did not reach complete stability and fluctuated. However, the adjuvanted protein-TLR2 complex with an average RMSD of 0.46 nm was completely stable during the simulation. Similarly, the protein-TLR4 and adjuvanted protein-TLR4 complexes were completely stable during the simulation with average RMSD of 0.69 and 0.57, respectively (Fig. 6a). Rg, which evaluates the extent of compressibility changes during MD simulations, is another parameter that has been studied in the evaluation of molecular dynamics simulations. Radius of gyration (Rg) is widely used to calculate the behavior of proteins and is defined as the distribution of protein atoms around their axis. The more stable the compression of the protein during the simulation, the more stable the protein and the complexes are, so this variable allows us to analyze the overall dimensions of the protein. As the graph shows, the compression levels of the adjuvanted and unadjuvanted proteins alone and in interaction with TLR4 and TLR2 were stable during the simulation (Fig. 6b).

An evaluation of the movement and flexibility of the structure during the simulation is provided by the root mean square fluctuations (RMSF) of the amino acids. To investigate the changes in the backbone atoms of the designed adjuvanted and non-adjuvanted proteins, as well as their complexes with TLR4 and TLR2, we performed RMSF analysis. As shown in Fig. 6c, the RMSF values for the designed proteins fluctuated at some points, but when these proteins interacted with the immune system receptors, they reached complete stability. This indicates that the proteins interactions with these receptors are strong and stable.

Evaluate immune response simulations

C-ImmSim Server output provides simulated immune response for producing immunoglobulins and cytokines. Based on the results extracted from laboratory studies, it also analyzes the production of various immune cells, including B cells, T cells. The simulation of the humoral and cellular immune response of the host to the vaccine candidate proteins is shown in Fig. 7. The antigen and immunoglobulin parameters, the B cell population, the

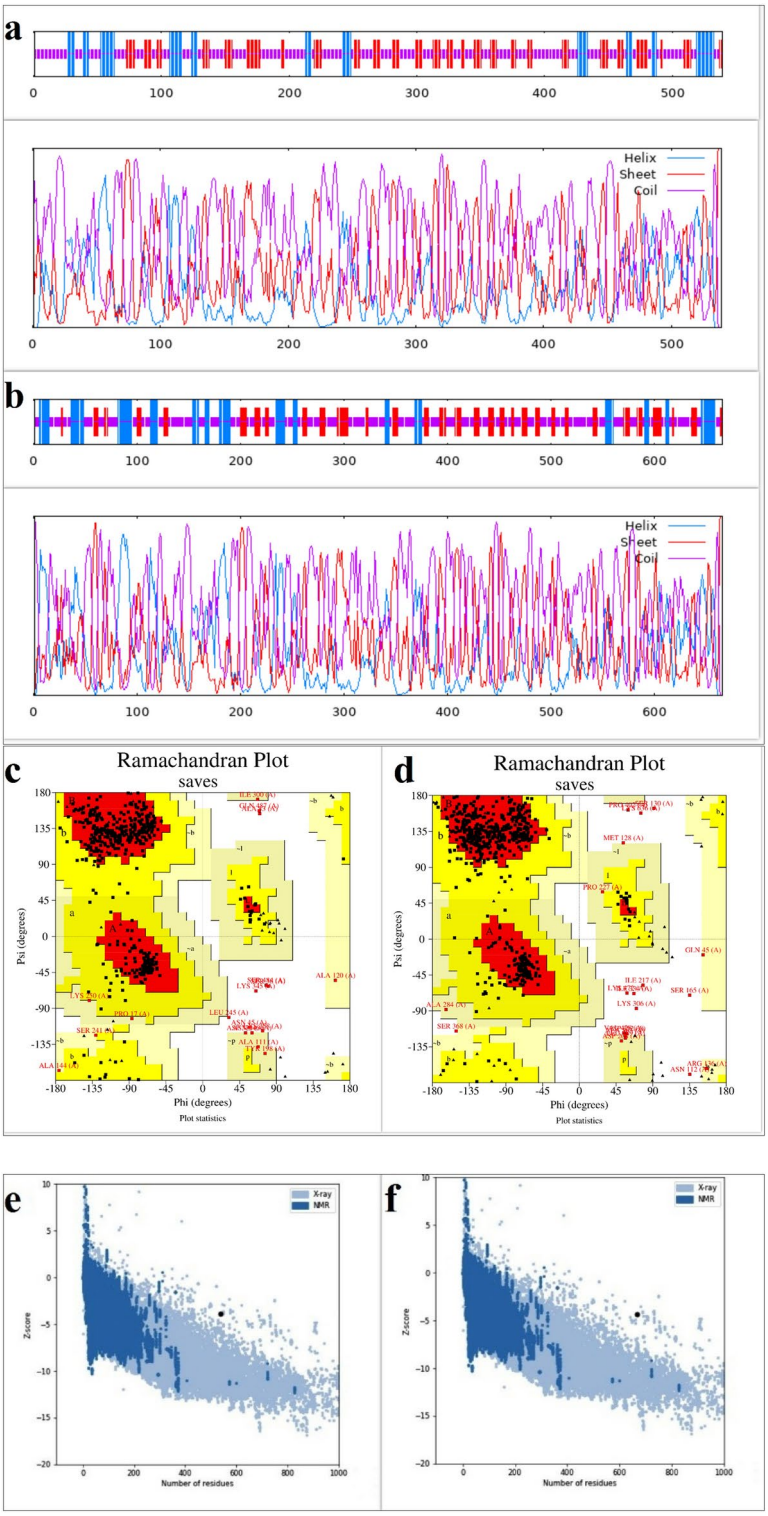

Fig. 2. Prediction and validation of the secondary and tertiary structure of the vaccine candidates (**a**) secondary structure of the unadjuvanted designed protein, (**b**) secondary structure of the adjuvanted protein, (**c**) validation of the tertiary structure of the unadjuvanted protein by Ramachandran plot, (**d**) validation of the tertiary structure of the adjuvanted protein by Ramachandran plot, (**e**) validation of the tertiary structure of the unadjuvanted protein by ProSA-web, (**f**) validation of the tertiary structure of the adjuvanted protein by ProSA-web.

Web server	Score	Vaccine candidate	Vaccine + adjuvant
MolProbity analysis	Clash score	0 (100% best among structures)	7.38 (85% best among structures)
	MolProbity score	1.10 (100% best among structures)	1.90 (81% best among structures)
ProSA-web	z-score	− 3.87	− 4.39
SAVES	Percent of proteins in desired regions	98.3%	97.9%

Table 4. Evaluation of the third structure of unadjuvanted and adjuvanted proteins using MolProbity, ProSA-web, and SAVES servers.

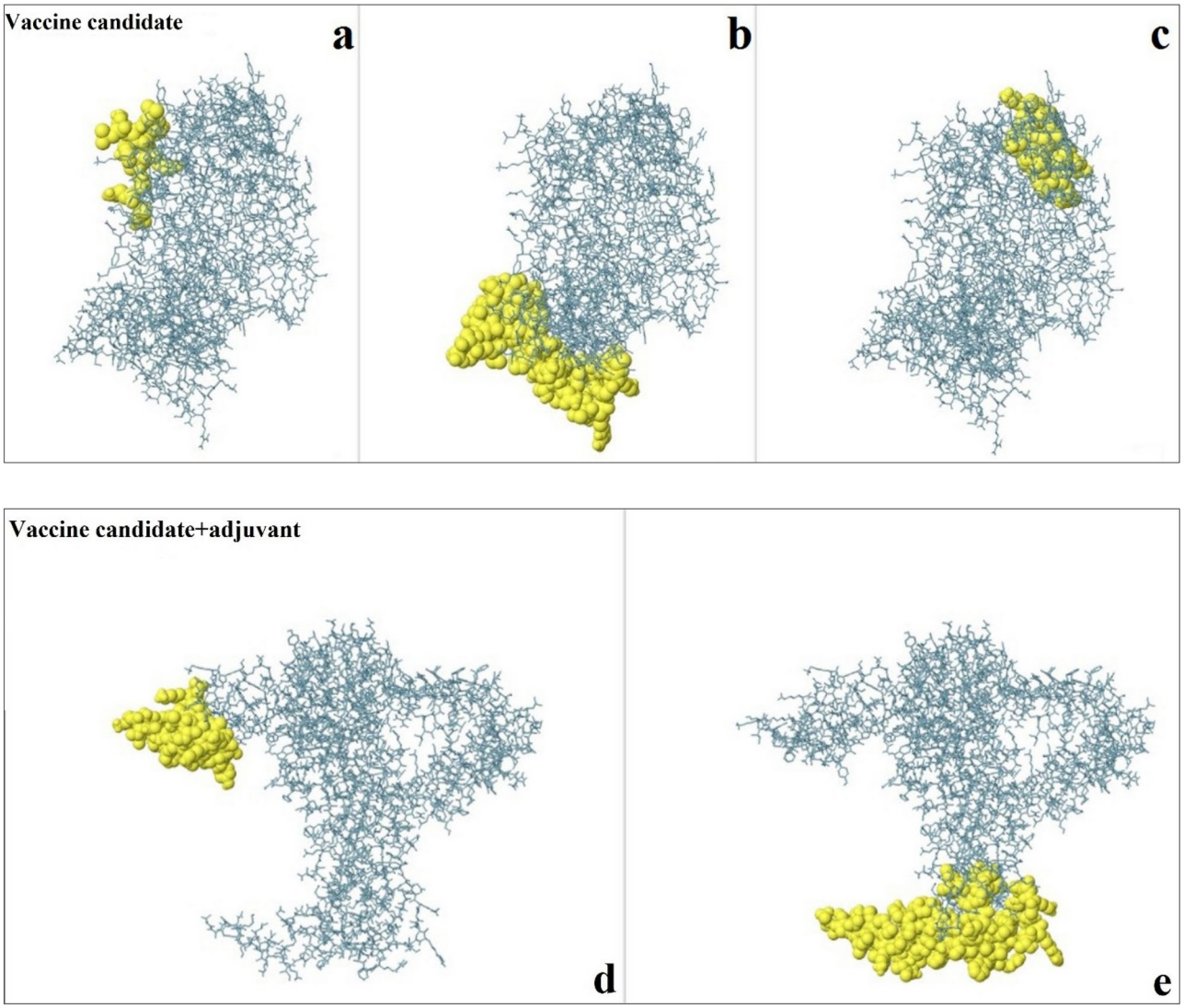

Fig. 3. The most potent vaccine candidate conformational epitopes designed using the Ellipro server. (**a**) 13 residues (AA 1–13) with residue score 0.863; (**b**) 68 residues (AA 472–539) with residue score 0.846; (**c**) 18 residues (AA 123–140) with residue score 0.754; (**d**) 43 residues (AA 1–43) with residue score 0.876; (**e**) 83 residues (AA 584–666) with residue score 0.832.

TH cell population and the cytokine production have been studied in this evaluation. The increase in IgM levels indicates the initial host response, with both unadjuvanted and adjuvanted proteins showing increased titers. In addition, a secondary response to antigenic proteins is demonstrated by increased immunoglobulin levels, B cell populations and TH cell populations (Fig. 7). There was also a significant increase in the levels of cytokines and interleukins after immunization, especially interferon γ. Increases, especially in immunoglobulin levels, B and

Candidate	Start	End	Peptide	Number of residues	Score
Vaccine	1	13	MLSLGVSYRFGQE	13	0.863
	472	539	QQTGILNVQWGKSAAQQCNASFTLPTREGKASGISKS KRFAVSHSQGNILMKKIVILITLLMLGGQAA	68	0.846
	123	140	RRVEIEVKGYKEVVTQPA	18	0.754
Vaccine + adjuvant	1	43	MNFKKSIALLFIALNIASLPTYAGVSKTFKDKCASTTAKLVQS	43	0.876
	584	666	PSIVSDAGMVYMSGLQQTGILNVQWGKSAAQQ CNASFTLPTREGKASGISKSKRFAVSH SQGNILMKKIVILITLLMLGGQAA	83	0.832

Table 5. Prediction of B-cell conformational epitopes of proteins with and without adjuvant by Ellipro.

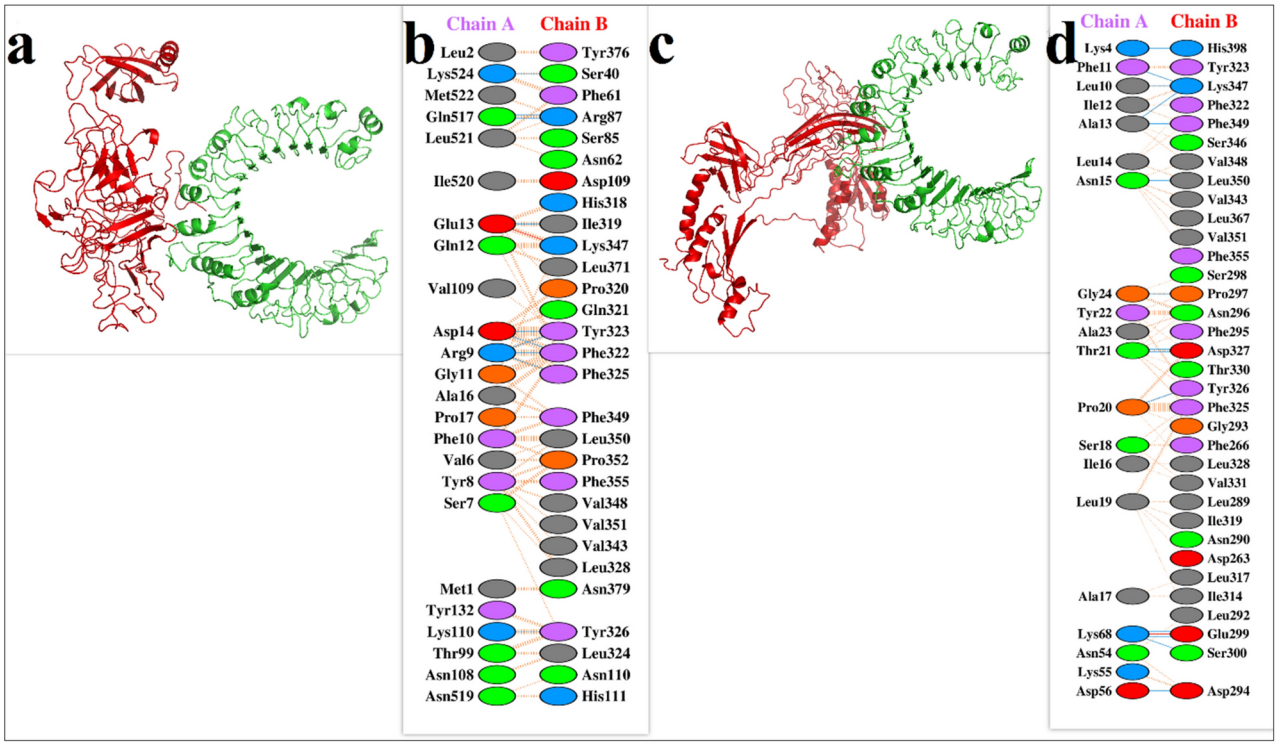

Fig. 4. (**a**) 3D representation of the docking complex of the unadjuvanted vaccine construct (red) with the human TLR2 (green) (**b**) Molecular interactions between chain-A of TLR2 molecule and chain-B of unadjuvanted vaccine construct (**c**) 3D representation of the docking complex of the adjuvanted vaccine construct (red) with the human TLR2 receptor (green) (**d**) Molecular interactions between chain-A of TLR2 molecule and chain-B of adjuvanted vaccine construct.

T cell populations, which are necessary to eliminate Klebsiella infection, showed a higher titer of the adjuvanted protein (Fig. 7 b,d,f,h) compared to the unadjuvanted protein (Fig. 7 a,c,e,g).

Codon optimization and in silico cloning of the designed candidate vaccines

Codon compatibility index and GC content of nucleotide sequence before optimization were 0.56 and 56.95% for unadjuvanted protein and 0.56 and 57.10 for adjuvanted protein, respectively. After codon optimization, the parameters were 1 and 50.64% for unadjuvanted protein and 1 and 50.45% for adjuvanted protein, respectively.

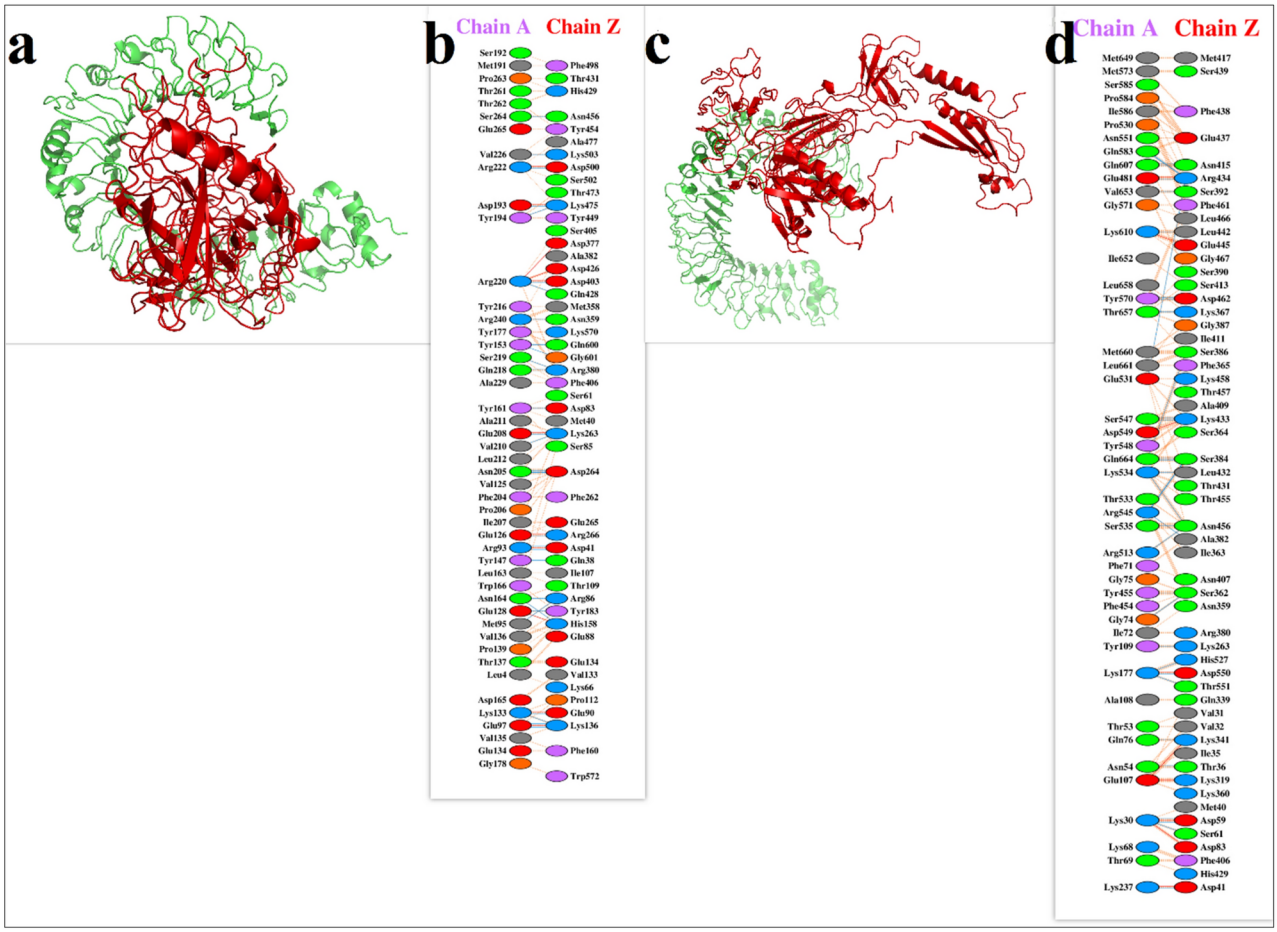

Fig. 5. (**a**) 3D representation of the docking complex of the unadjuvanted vaccine construct (red) with the human TLR4 (green) (**b**) Molecular interactions between chain-Z of TLR4 molecule and chain-A of unadjuvanted vaccine construct (**c**) 3D representation of the docking complex of the adjuvanted vaccine construct (red) with the human TLR4 recepto (**d**) Molecular interactions between chain-Z of TLR4 molecule and chain-A of adjuvanted vaccine construct.

Complex	ΔG (kcal mol^{-1})	Weighted score	No. of hydrogen bonds	PRODIGY (Kd)
Vaccine candidate–TLR4	− 19.9	− 1137.4	30	2.5e−15
Vaccine candidate–TLR2	− 11.5	− 1354.3	10	3.4e−09
Vaccine Candidate + adjuvant–TLR4	− 21.9	− 1059.5	36	8.7e−17
Vaccine Candidate + adjuvant–TLR2	− 15.7	− 1320.0	14	3.1e−12

Table 6. Evaluation of molecular binding results between unadjuvanted and adjuvanted vaccine candidates with TLR4 and TLR2.

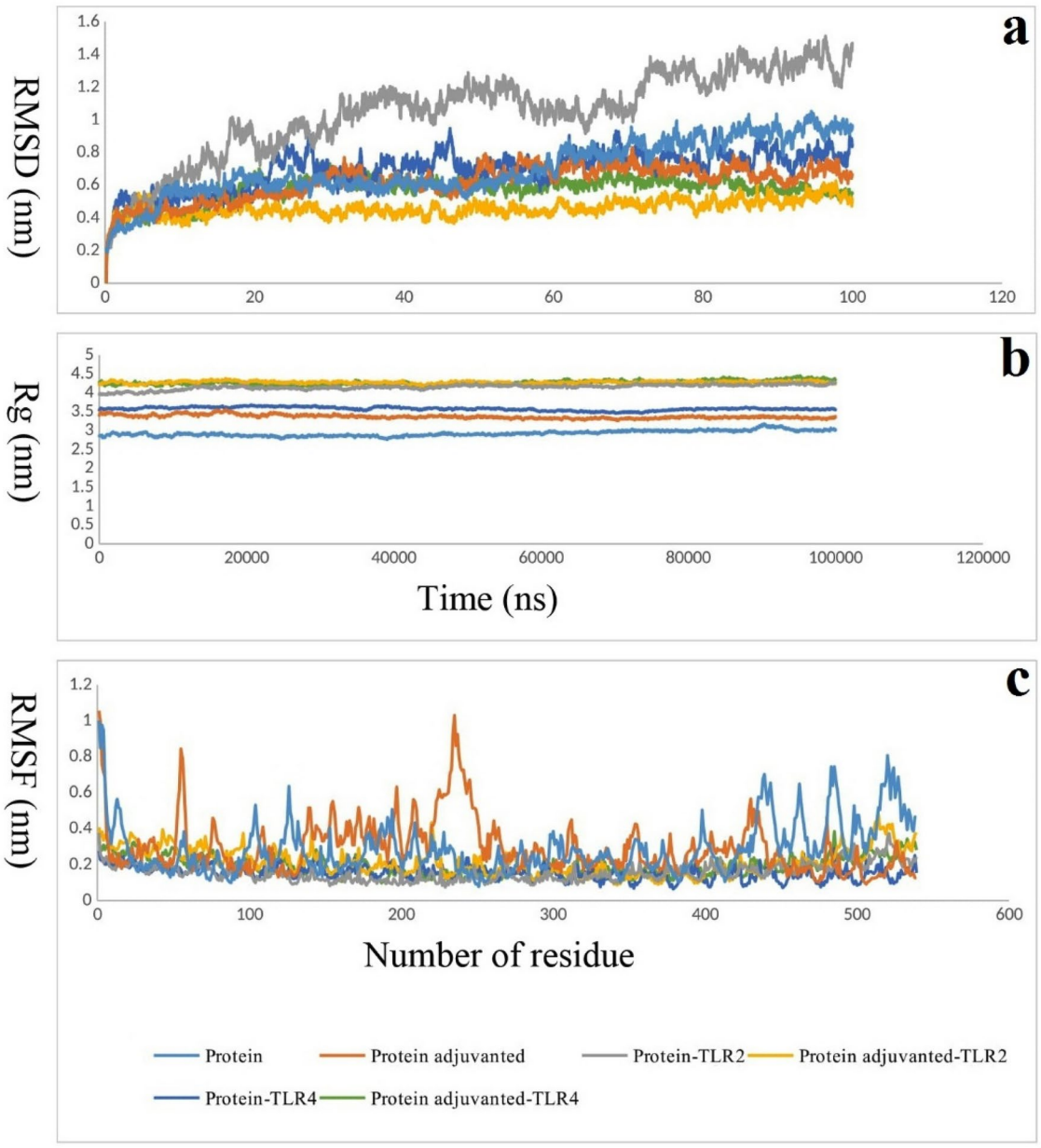

Fig. 6. (**a**) RMSD results of the unadjuvanted and adjuvanted proteins, protein-TLR2, adjuvanted protein-TLR2, protein-TLR4, and adjuvanted protein-TLR4 complexes. (**b**) Rg results of the unadjuvanted and adjuvanted proteins, protein-TLR2, adjuvanted protein-TLR2, protein-TLR4, and adjuvanted protein-TLR4 complexes. (**c**) RMSF results of the unadjuvanted and adjuvanted proteins, protein-TLR2, adjuvanted protein-TLR2, protein-TLR4, and adjuvanted protein-TLR4 complexes.

Cloning of the optimized vaccine candidate sequence in pET-28a $^{(+)}$ using SnapGene Version 3.2.1 showed that the vaccine candidate sequences in pET-28a $^{(+)}$ could be cloned (Fig. 8a,c). Double digestion with NcoI and XhoI enzymes revealed the presence of vaccine candidates of 1623 bp for the unadjuvanted protein and 2005 bp for the adjuvanted protein together with the pET-28a $^{(+)}$ vector of 5231 bp (Fig. 8b,d).

Materials and methods

Retrieval of the OmpA, OMPK17 and fimbriae protein sequences

The protein sequences of OmpA, OMPK17, and fimbriae from *K. pneumoniae* were obtained from the UniProtKB database (https://www.uniprot.org/). UniProtKB offers a user-friendly interface facilitating the retrieval of specific protein sequences and access to protein data[43].

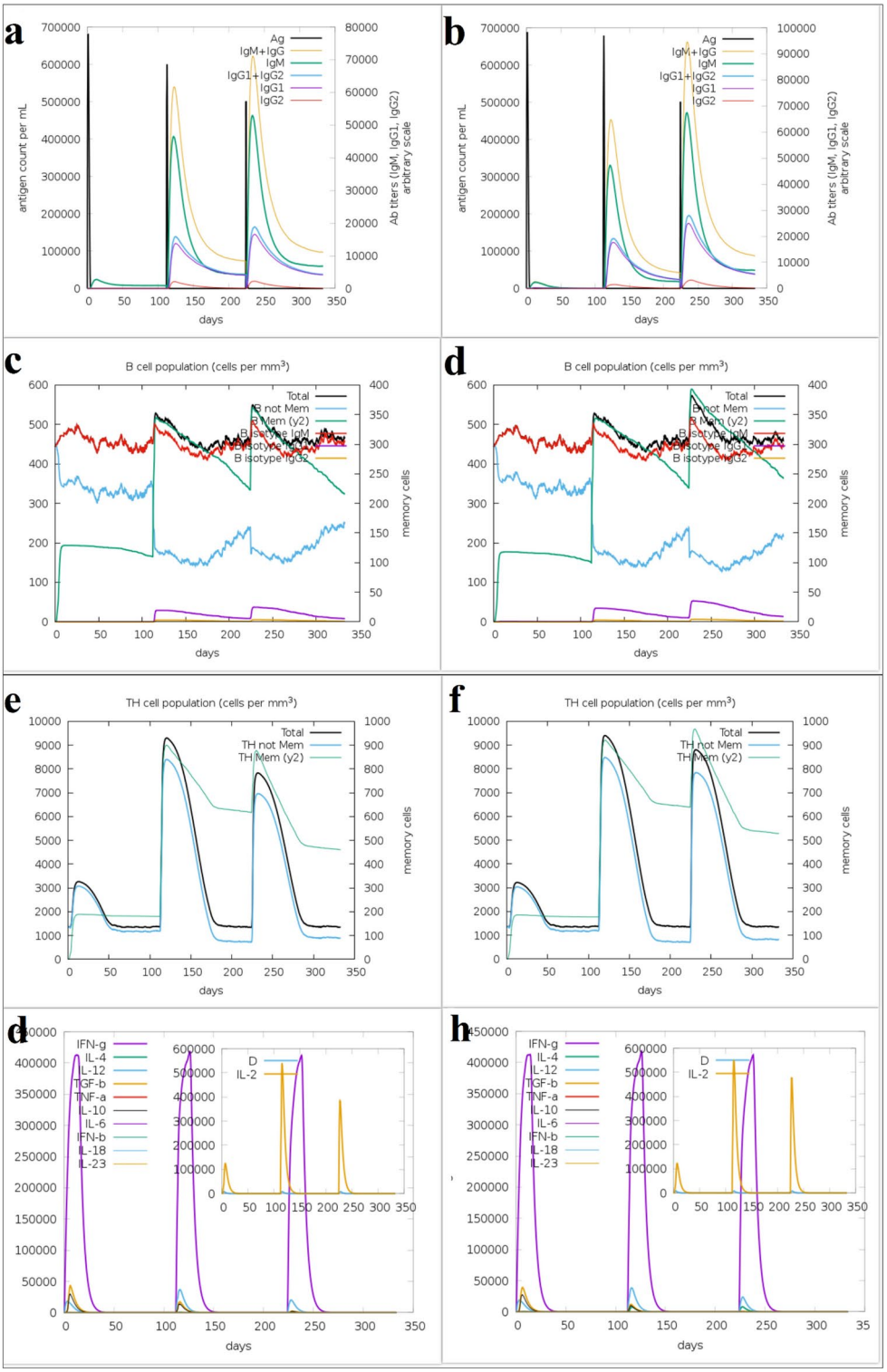

Fig. 7. In silico immunity simulation against protein antigens designed as vaccine candidates using the C-ImmSim web server. Simulations after three injections at steps 1, 336 and 672 are shown. (**a**) Antigen and immunoglobulin simulated by unadjuvanted protein; (**b**) Antigen and immunoglobulin simulated by adjuvanted protein; (**c**) B cell population simulated by unadjuvanted protein; (**d**) B cell population simulated by adjuvanted protein; (**e**) TH cell population simulated by unadjuvanted protein; (**f**) TH cell population simulated by adjuvanted protein; (**g**) cytokine production simulated by unadjuvanted protein; (**h**) cytokine production simulated by unadjuvanted protein.

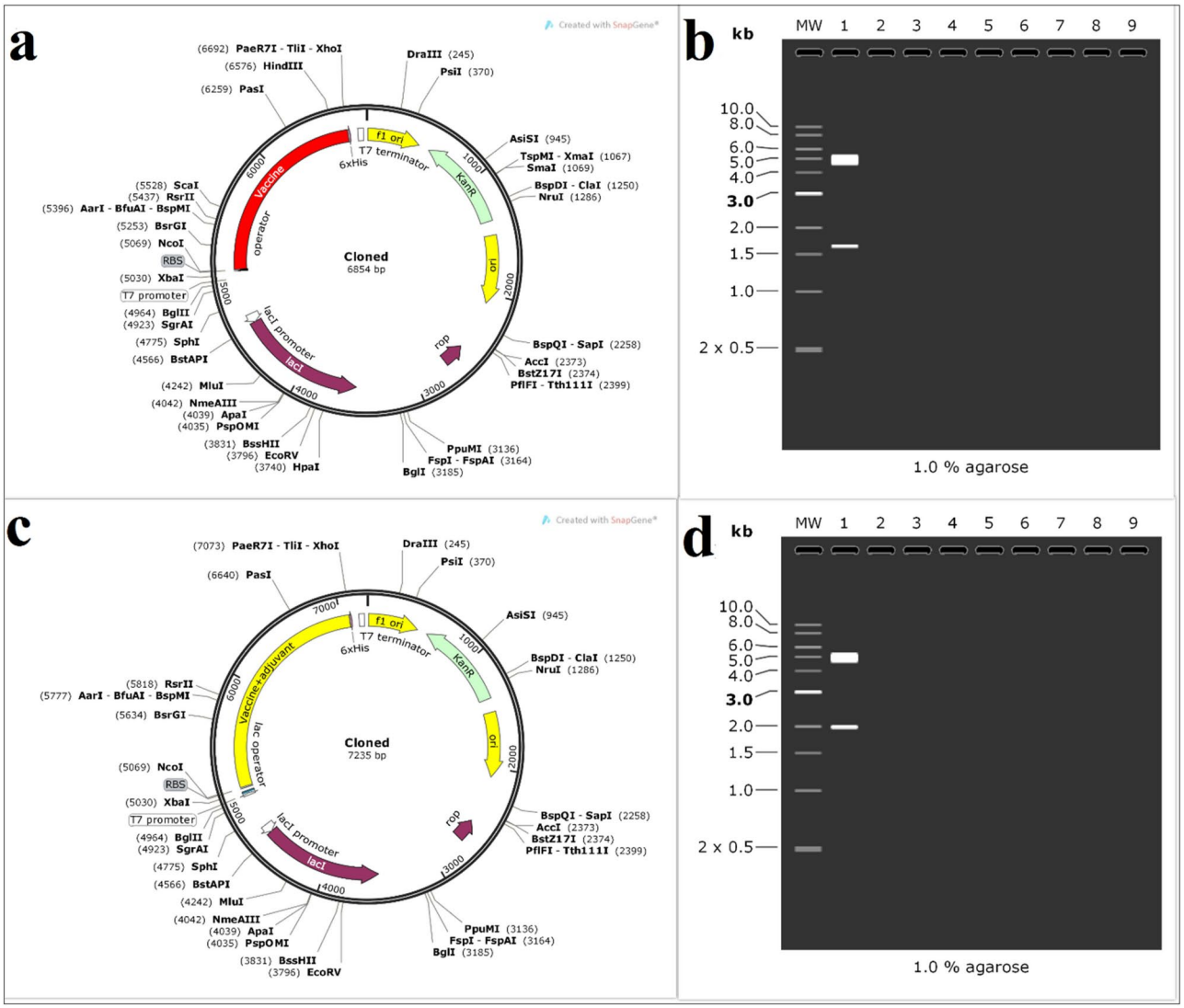

Fig. 8. (**a**) Cloning of the unadjuvanted protein gene into the pET-28a vector (shown in red) (**b**) Informatics evaluation of the cloning of the unadjuvanted protein by double digest. (**c**) Cloning of the adjuvanted protein gene into the pET-28a vector (shown in yellow) (**d**) Informatics evaluation of the cloning of the adjuvanted protein by double digest.

T cell epitope prediction

Prediction of T-cell epitopes is critical for eliciting both humoral and cellular immune responses and is therefore essential for the design of effective prophylactic bacterial vaccines. We used the Rankpep 1D sequence-based screening server to identify T cell epitopes. This server uses position-specific scoring matrices (PSSM) and a binding threshold of 2% to predict immunodominant peptides that interact with MHC molecules[44]. In addition, we used the Immune Epitopes and Analysis Resource (IEDB) server (http://tools.immuneepitope.org/mhcii/) with the NetMHCIIpan 4.1 BA method to predict T cell epitopes. For HLA class II epitope prediction, we considered the DRB1*0101, DRB1*0301, DRB1*0401, DRB1*0701, DRB1*0801, DRB1*1101, DRB1*1301, and DRB1*1501 alleles, which cover more than 95% of the global HLA variability[45]. NetMHCIIpan 4.1 BA prediction evaluates the ability of a peptide to bind an MHC molecule using Artificial Neural Networks (ANNs) and employs custom machine learning strategies to integrate different types of training data[46].

Prediction of linear B cell epitopes

Two databases were used to predict linear B cell epitopes, namely BepiPred 2.0 and Kolaskar & Tongaonkar Antigenicity (http://www.iedb.org/). The BepiPred 2.0 server predicts B cell epitopes from a protein sequence using a random forest algorithm on epitopes and non-epitope amino acids that are determined from crystal structures. Residues are predicted to be part of an epitope if they have a score above a threshold of 0.5[47]. Kolaskar & Tongaonkar is a semi-empirical method for B cell epitope prediction that uses the physicochemical properties of amino acid residues and their frequency of occurrence in empirically known segmental epitopes. Evaluation of this method for a large number of proteins has shown that the accuracy of antigen prediction is about 75%[48].

The IEDB database is a collection of experiments that identify and characterize epitopes and epitope-specific immune receptors, along with various other details such as host organism, immune exposures, and inducible immune responses[49].

Selection of epitope-rich regions and vaccine candidates

In order to select epitope-rich domains, the predicted B-cell and T-cell epitopes were scored, and regions of the proteins that shared B-cell and T-cell epitopes with high scores were selected and used for protein production. Selected epitopes were located within conserved regions of the protein. Multiple epitopes were combined. Finally, the B-cell and T-cell epitope rich domains of the OmpA, OMPK17, and fimb proteins were selected for attachment to various linkers and the candidate vaccine sequence was produced.

Physical and chemical characteristics of the final designed structure

Using the Protparam tool available on the server (http://expasy.org/tools/protparam.html), the physical and chemical properties of the final structure were investigated, including the number of amino acids, molecular weight, PI, number of charged amino acids, amino acid composition, estimated half-life, instability index, aliphatic index, and grand average of hydropathicity. The comparative analysis of these properties is helpful in the understanding of the functional and structural characteristics of the protein, as well as in the exploration of its molecular evolution[50]. Comparative analysis of these properties helps in understanding the functional and structural characteristics of the protein, as well as in exploring its molecular evolution[51].

Investigation of allergenicity, toxicity, solubility and antigenicity of the designed protein

The VaxiJen v2.0 server was employed to analyze the antigenicity of the designed vaccine candidate antigen, classifying antigens based on automatic cross-covariance (ACC), which transforms protein sequences into uniform vectors of key amino acid features[52]. AllerTOP v.2 (http://www.ddg-pharmfac.net/AllerTOP/) was utilized to predict the allergenicity of the designed protein based on key physicochemical properties. Allergens are small antigens that typically elicit an IgE antibody response[53]. Additionally, the Recombinant Protein Solubility Prediction (https://biotech.ou.edu/) was used to predict protein solubility; this server calculates average charge, percentage of turn residues, percentage of cysteine, percentage of proline, hydrophilicity, and total number of residues[54]. The ToxinPred2 server (https://webs.iiitd.edu.in/raghava/toxinpred/) was also used to predict the toxicity of the vaccine candidate[55].

Secondary and tertiary structure prediction

Secondary structure prediction of the candidate protein was conducted using the Garnier-Osguthorpe-Robson IV (GOR IV) server (https://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html), which is based on probability parameters obtained from experimentally solved protein tertiary structures determined by x-ray crystallography[56]. Furthermore, tertiary structure modeling was performed using the I-TASSER (https://zhanggroup.org/I-TASSER/) and GalaxyWEB web servers (https://galaxy.seoklab.org/)[57]. GalaxyRefine server was used to refine the modeled structures[58].

Evaluation of the tertiary structure of designed protein

The predicted 3D structure of the designed protein was assessed using MolProbity (http://molprobity.biochem.duke.edu)[59], ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php)[60] and SAVES v6.1 (https://saves.mbi.ucla.edu)[61] servers were used to evaluate predicted 3D structure. These tools are essential for identifying errors in both experimental and theoretical models of protein structures. Evaluations were based on criteria ensuring proximity to natural protein structures for cellular expression. ProSA-web assessed structure quality using Z-scores and energy plots. An error in the 3D structure is indicated by a Z-score out of range for native proteins[60]. MolProbity, a widely used tool for assessing macromolecular structures, analyzed parameters such as MolProbity score, collision score, and Ramachandran diagram to validate 3D structures[62]. Additionally, the Ramachandran plot from PROCHECK in the SAVES v6.1 web server was utilized to calculate torsion angles of residues, assessing whether the residues in the distant regions are allowed and favorable[61].

Conformational B cell epitope prediction

For the prediction of structural epitopes of the vaccine candidate protein, the ElliPro server (http://tools.iedb.org/ellipro/) was used with a threshold of 0.5. This server predicts B-cell structural epitopes based on the tertiary structure of the designed protein using a modified version of Thornton's method along with MODELLER and Jmol[63].

Investigation of the interaction of recombinant proteins with receptors of the immune system

Crystallographic structures of TLR2 (PDB ID: 53di) and TLR4 (PDB ID: 7mlm) receptors were retrieved from the PDB database (https://www.rcsb.org). After removal of ligands and water molecules, hydrogen atoms and charges were added to both receptors and the unadjuvanted anted proteins structure using the Dock prep tool in UCSF Chimera (Chimera 1.5.3) software[64]. TLR2 and TLR4 are important receptors of the immune system to fight *K. pneumoniae*. Important amino acids in the active site of TLR2 include Leu317, Ile319, Phe322, Leu324, Phe325, Tyr326, Val348, Phe349, and Pro352[65]. Other important amino acids in theactive site of the TLR4 receptor include Arg434, Ser413, Ser386, Arg380, Lys341, Lys263, and Gln339[66]. Designed protein interaction and TLR2 and TLR4 of the immune system using the Cluspro2 server. (https://cluspro.bu.edu/home.php)[67]. The PRODIGY web server was also used to predict the binding energy of the protein complexes. This server focuses on prediction of binding affinity in biological complexes and identification of biological interfaces[68]. Finally,

residues involved in vaccine-receptor interactions were visualized using PyMOL (The PyMOL Molecular Graphics System, Version 1.1, Schrödinger, LLC) and mapped using the PDBsum generate web server (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html)[69].

Molecular dynamics simulation

To evaluate the stability of the strongest complexes, unadjuvanted proten-TLR2, unadjuvanted proten-TLR4, adjuvanted proten-TLR2 and unadjuvanted proten-TLR4, as well as the 3D structure of both proteins, MD simulations were performed up to 100 ns. The simulation was performed using the GROMACS 2022 software and the OPLS-AA force field[70]. The structures were placed in a triaxial box. The distance from all edges was 1 nm. The protein-receptor complexes and unbound protein structures were then placed in a simulation box filled with TIP3P water molecules. The ligand-receptor interaction can change the total charge, and water molecules added to the system change the charge of the system from zero. In the simulation process, if the charge of the system is positive, it is neutralized with chlorine ions, and if the charge of the system is negative, it is neutralized with sodium ions. The energy minimization process for single protein structures and their complexes with simulated immune system receptors was performed in two parts: first the systems were equilibrated using NVT (constant number of particles, volume and temperature) at 300 K and 100 ps, then NPT (number of stationary particles, pressure and temperature) was equilibrated using a Parrinello-Rahman barostat at 300 K, 1.0 bar and 100 ps. The Ewald mesh particle algorithm was used to handle long-range electrostatic charges, and the ineffective distance was considered to be 10 angstroms. Van der Waals interactions were calculated with an effective distance of 1 nm. A linear constraint algorithm was used to limit the length of the covalent bonds. After applying the necessary balances, the complexes and unbound structures were simulated up to 100 ns. Using the simulation data, we performed analyzes RMSF, RMSD, and Rg.

Simulation of the immune system

To evaluate the possibility of inducing the required immune response against the designed proteins unadjuvanted and adjuvanted as antigen, the study of immune simulation and immune response profiling was performed. Immune response induction for the designed proteins was performed using the C-ImmSim simulation server (http://150.146.2.1/C-IMMSIM/index.php). The simulations were performed with the default parameters and the time steps for injection of the antigens were set to 1, 336, and 672. Simulations were performed with 10 µL simulation volume, 1000 simulation steps, LPS-free and adjuvanted vaccine[71].

Codon optimization and cloning simulation

In order to improve recombinant protein production, codon optimization was performed using the JCat tool (https://www.jcat.de/) for the *E. coli* K12 strain. The parameters of CAI and the percentage of GC content were measured. CAI is the most important index for the prediction of gene expression levels. It is a measure of how well the coding sequence describes the use of codons in an organism[72]. The optimized codons of the unadjuvanted and adjuvanted proteins were cloned into pET28a$^{(+)}$ vector and evaluated using the SnapGene Version 3.2.1 tool for double digestion with NcoI and XhoI enzymes[73].

Discussion

Today, several proteins of *K. pneumoniae* are being introduced and investigated as vaccine candidates in many studies[74,75]. Outer membrane proteins (OMPs) are surface proteins that act as membrane channels to allow hydrophilic molecules to pass across the membrane. OMPs of *K. pneumoniae* as subunit vaccine candidates induce host-specific antibodies and have been effective against *Klebsiella* infection in animal models[76,77]. The OmpA protein appears to be one of the best protein candidates for vaccine development. This is due to its immunogenicity and its role in pathogenesis[23]. OmpA is highly conserved among members of the *Enterobacteriaceae* and has important pathogenic roles, including bacterial adhesion, invasion or intracellular survival, as well as evasion of host defenses or stimulation of pro-inflammatory cytokine production[78]. OMPK17 is one of the antigens that can be detected in the immune serum of patients with an acute *K. pneumoniae* infection. A member of the OMP family that plays an important role in the pathogenesis of *K. pneumoniae*, outer membrane protein-K17 (OMPK17, also known as OMPX)[79]. This protein can induce protection against *K. pneumoniae* in mice, mainly through induction of T-helper type 2 immune response[26,80]. Fimbriae, filamentous organelles expressed on the surface of the bacterial cell, often mediate adhesion in Gram-negative bacteria. Two fimbrial adhesins, type 1 fimbriae and type 3 fimbriae, can be produced by most clinical isolates of *K. pneumoniae*[28]. Given the necessity of the aforementioned OmpA, OMPK17 and fimb proteins, we decided in this study to design a vaccine consisting of domains rich in B- and T-cell epitopes. The ultimate goal is to protect these necessary and important proteins in the event of human exposure to *K. pneumoniae*. The bacterium is targeted and destroyed by the immune system. Based on immunoinformatics analysis, 5 epitope-rich domains of these three proteins containing B- and T-cell antigenic epitopes were finally selected and linked with GGGGS and EAAAK bonds. Ideally, epitope vaccines should contain B cell epitopes that induce protective antibody responses, but also essential T cell epitopes that stimulate Th responses. We also added *E. coli* LT adjuvant to the designed construct to select the best candidate with protective efficacy against *K. pneumoniae* by comparing the immunogenicity of two proteins with and without LT adjuvant. Physical and chemical parameters of proteins, secondary and tertiary structure, solubility, allergenicity and antigenicity were analyzed after designing the protein sequences as vaccine candidate. Both protein candidates had the required physicochemical properties, were non-allergenic and antigenic. Also, the third structure quality evaluations of both structures showed that they had an acceptable third structure. Today, bioinformatics studies are used for a variety of purposes, including vaccine design and characterization of new genes. In this context, computational methods for protein–protein interaction prediction and analysis facilitate the evaluation of the interactions of vaccine candidates with immune system receptors[81].

After confirming the physicochemical properties and evaluating the tertiary structure of the designed proteins, we evaluated vaccine candidates with and without adjuvant to interact with the immune system's TLR2 and TLR4. The innate immune response, mediated by TLRs, controls the growth of the bacteria at the site of infection. In *K. pneumoniae* infection, TLR2 and TLR4 are important TLRs. TLR2 transduces the signal for resistance to foreign pathogens. TLR4 can induce host defense against gram-negative bacterial lung infections through recognition of bacterial LPS[82,83]. The result of molecular docking using Cluspro 2. web server showed that both recombinant proteins effectively bind to the active site of TLR2 and TLR4 and can activate innate immune receptors to induce immune response. Analysis of the results showed that the protein designed with LT adjuvant has a stronger and more effective interaction with the active site amino acids of both receptors. To evaluate the stability of the tertiary structure of the designed proteins as well as the stability of their interaction with TLR2 and TLR4, MD simulation was performed up to 100 ns. The analysis of the results in terms of RMSD, Rg and RMSF of the designed proteins showed that the designed proteins with and without adjuvants have a stable tertiary structure during the simulation. The unadjuvanted protein fluctuated during the simulation in the interaction with TLR2, indicating that their interaction is somewhat unstable, but the interaction of the adjuvanted protein with this receptor is quite stable during the simulation, confirming that it is an effective interaction. Both unadjuvanted protein-TLR4 and adjuvanted protein-TLR4 complexes were stable during the simulation. These results show that the adjuvanted protein shows more acceptable results and is more effective in interacting with the receptors of the immune system.

Several studies have been conducted in relation to the immunoinformatic design of *K. pneumoniae* proteins as vaccine candidates[84–86]. Ranjbarian et al. introduced YGKFQNNNYPHKSDMSDYGFSYGAG peptide vaccine candidate, which was B-cell and T-cell overlapping[86]. Our investigation showed that our designed protein also contains this peptide. Furthermore, the simulation results of the immune response against two proteins designed with and without adjuvant as antigens showed that both proteins are suitable candidates as vaccines and stimulate the necessary responses against *K. pneumoniae*, but the responses produced in the adjuvanted protein generate a stronger and more appropriate response, especially the production of immunoglobulins, the B-cell and T-cell responses. These responses are necessary to eliminate *K. pneumoniae*.

In this study, we designed and introduced a multi-epitopic vaccine candidate against *K. pneumoniae* infection using the most accurate and comprehensive immunoinformatics tools. All immunoinformatics analyses indicated that our vaccine candidate is likely to be effective against the *K. pneumoniae*. However, our study has limitations. Experimental testing is the cornerstone of the development and validation of multi-epitopic vaccines. While in silico tools and computational modeling provide valuable insights into epitope selection, vaccine structure design, and immunogenicity potential, these methods rely on predictions that require experimental confirmation. In vitro and in vivo testing bridges the gap between theoretical predictions and practical applications and ensures the safety, efficacy, and immunogenicity of vaccine candidates.

Multi-epitopic vaccines are designed to target specific regions of pathogens, which are computationally predicted to elicit robust immune responses. However, the actual recognition and processing of these epitopes by the immune system needs to be confirmed through in vitro assays and animal models.

Consequently, experimental testing is essential to advance multi-epitope vaccines from conceptual designs to viable clinical solutions. This ensures that theoretical constructs meet the stringent requirements of real-world application, safety, efficacy, and practicality of these innovative vaccines.

Conclusion

In summary, we have analyzed three *K. pneumoniae* proteins, OmpA, OMPK17, and fimb, to select the best immunogenic domains to induce a strong immune response. We designed a candidate protein vaccine against *K. pneumoniae* by selecting 5 B- and T-cell epitope-rich domains. We added the sequence of *E. coli* LT as an adjuvant and evaluated the results, which showed this adjuvant increased the immunogenicity of the candidate vaccine, and its interaction with TLR2 and TLR4 made the immune system stronger and more stable. Finally, the adjuvanted vaccine candidate provides a potentially more appropriate response against MDR *K. pneumoniae.*

Data availability

The sequences of the OmpA (accession number: W9BFP5), OMPK17 (accession number: Q48427), and fimbriae (accession number: W1DI64) proteins were extracted from the Uniprotkb database (https://www.uniprot.org/). Prediction of B cell epitopes for all three proteins was performed using the BepiPred 2.0 and Kolaskar & Tongaonkar Antigenicity (http://www.iedb.org/). The IEDB with the NetMHCIIpan 4.1 BA method (https://www.iedb.org/) and Rankpep (http://imed.med.ucm.es/Tools/rankpep.html) databases were used to predict T cell epitopes. The ProtParam and Protscale (http://expasy.org/tools/protparam.html) tools were used to obtain the physical and chemical properties of the designed structures. The antigenicity of the designed protein was calculated using the VaxiJen v2.0 server (https://www.ddg-pharmfac.net/vaxijen/). The AllerTOP v.2 web server was used to check the allergenicity of the designed structure (https://www.ddg-pharmfac.net/AllerTOP/). The ToxinPred2 server was used to predict the toxicity of the designed protein (https://crdd.osdd.net/raghava/toxinpred/). GOR IV software was used to check the second structure of the designed structures (https://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html). The Recombinant Protein Solubility Prediction (https://biotech.ou.edu/) was used to predict protein solubility. The web servers I-TASSER (https://zhanggroup.org/I-TASSER/) and GalaxyWEB (https://galaxy.seoklab.org/) were used to predict the tertiary structures of the designed structures. The predicted tertiary structures were evaluated using the MolProbity (http://molprobity.biochem.duke.edu/index.php), ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) and SAVES v6.1 (https://saves.mbi.ucla.edu/) servers. The 3D structure of the vaccine candidates was analyzed using the ElliPro server (http://tools.iedb.org/ellipro/). Cluspro 2.0 (https://cluspro.bu.edu/login.php?redir=/queue.php) was

used to analyze protein-protein binding between vaccine candidates designed with TLR2 and TLR4 receptors. Crystallographic structures of TLR2 (PDB ID: 53di) and TLR4 (PDB ID: 7mlm) were obtained from the RCSB protein database (http://www.rcsb.org/). The binding site of all complexes was analyzed using PDBsum generate web server (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html). MD simulations of vaccine candidates-TLRs complexes were performed using the GROMACS 2022 package (https://doi.org/10.5281/zenodo.6103568.). The C-ImmSim server was used to simulate immune responses to vaccine candidates (https://kraken.iac.rm.cnr.it/C-IMMSIM/index.php). Codon optimization was performed using the JCat tool (https://www.jcat.de/). Designed proteins were cloned into the pET28a$^{(+)}$ vector and evaluated using the SnapGene Version 3.2.1 tool. All data generated or analyzed during this study are included in this article.

Received: 20 November 2024; Accepted: 21 February 2025
Published online: 01 March 2025

References

1. Yamasaki, S. et al. Genetic analysis of ESBL-producing *Klebsiella pneumoniae* isolated from UTI patients in Indonesia. *J. Infect. Chemother.* **27**, 55–61 (2021).
2. Salawudeen, A. et al. Epidemiology of multidrug-resistant *Klebsiella pneumoniae* infection in clinical setting in South-Eastern Asia: A systematic review and meta-analysis. *Antimicrob. Resist. Infect. Control* **12**, 142 (2023).
3. Organization, W. H. Antimicrobial resistance: global report on surveillance. (2014).
4. Assoni, L., Girardello, R., Converso, T. R. & Darrieux, M. Current stage in the development of *Klebsiella pneumoniae* vaccines. *Infect. Dis. Ther.* **10**, 2157–2175 (2021).
5. Mueller, H. L. & Lanz, M. Hyposensitization with bacterial vaccine in infectious asthma. A double-blind study and a longitudinal study. *Jama* **208**(1379), 1383 (1969).
6. Opoku-Temeng, C., Kobayashi, S. D. & DeLeo, F. R. *Klebsiella pneumoniae* capsule polysaccharide as a target for therapeutics and vaccines. *Comput. Struct. Biotechnol. J.* **17**, 1360–1366 (2019).
7. Prawiningrum, A., Paramita, R. & Panigoro, S. Immunoinformatics approach for epitope-based vaccine design: Key steps for breast cancer vaccine. *Diagnostics (Basel)* **12**, 2981 (2022).
8. Rawat, S. S., Keshri, A. K., Kaur, R. & Prasad, A. Immunoinformatics approaches for vaccine design: A fast and secure strategy for successful vaccine development. *Vaccines* **11**, 221 (2023).
9. Sarvmeili, J., Baghban Kohnehrouz, B., Gholizadeh, A., Shanehbandi, D. & Ofoghi, H. Immunoinformatics design of a structural proteins driven multi-epitope candidate vaccine against different SARS-CoV-2 variants based on fynomer. *Sci. Rep.* **14**, 10297 (2024).
10. Soria-Guerra, R. E., Nieto-Gomez, R., Govea-Alonso, D. O. & Rosales-Mendoza, S. An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. *J. Biomed. Inform.* **53**, 405–414 (2015).
11. Akter, A. et al. Computational approach for identifying immunogenic epitopes and optimizing peptide vaccine through in-silico cloning against Mycoplasma genitalium. *Heliyon* **10**, e28223 (2024).
12. Dimitrov, I., Bangov, I., Flower, D. R. & Doytchinova, I. AllerTOP v. 2: A server for in silico prediction of allergens. *J. Mol. Model.* **20**, 1–6 (2014).
13. Shetty, S. et al. Immunoinformatics design of a multi-epitope vaccine for Chlamydia trachomatis major outer membrane proteins. *Sci. Rep.* **14**, 1–18 (2024).
14. Kolla, H. B. et al. Immuno-informatics study identifies conserved T cell epitopes in non-structural proteins of Bluetongue virus serotypes: formulation of a computationally optimized next-generation broad-spectrum multi-epitope vaccine. *Front. Immunol.* **15**, 1424307 (2024).
15. Bui, H.-H., Sidney, J., Li, W., Fusseder, N. & Sette, A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinform.* **8**, 1–6 (2007).
16. Hwang, W. et al. Current and prospective computational approaches and challenges for developing COVID-19 vaccines. *Adv. Drug Del. Rev.* **172**, 249–274 (2021).
17. García-Machorro, J., Ramírez-Salinas, G. L., Martinez-Archundia, M. & Correa-Basurto, J. The advantage of using immunoinformatic tools on vaccine design and development for coronavirus. *Vaccines* **10**, 1844 (2022).
18. Shawan, M. et al. Advances in computational and bioinformatics tools and databases for designing and developing a multi-epitope-based peptide vaccine. *Int. J. Pept. Res. Ther.* **29**, 60 (2023).
19. Bengoechea, J. A. & Sa Pessoa, J. *Klebsiella pneumoniae* infection biology: Living to counteract host defences. *FEMS Microbiol. Rev.* **43**, 123–144 (2019).
20. Riwu, K. H. P., Effendi, M. H., Rantam, F. A., Khairullah, A. R. & Widodo, A. A review: virulence factors of Klebsiella pneumonia as emerging infection on the food chain. *Vet. World* **15**, 2172–2179 (2022).
21. Oikonomou, K. G. & Aye, M. *Klebsiella pneumoniae* liver abscess: A case series of six Asian patients. *Am. J. Case Rep.* **18**, 1028–1033 (2017).
22. Llobet, E., March, C., Giménez, P. & Bengoechea, J. A. *Klebsiella pneumoniae* OmpA confers resistance to antimicrobial peptides. *Antimicrob. Agents Chemother.* **53**, 298–302 (2009).
23. March, C. et al. *Klebsiella pneumoniae* outer membrane protein A is required to prevent the activation of airway epithelial cells. *J. Biol. Chem.* **286**, 9956–9967 (2011).
24. Prasadarao, N. V., Blom, A. M., Villoutreix, B. O. & Linsangan, L. C. A novel interaction of outer membrane protein A with C4b binding protein mediates serum resistance of Escherichia coli K1. *J. Immunol.* **169**, 6352–6360 (2002).
25. Kurupati, P., Teh, B. K., Kumarasinghe, G. & Poh, C. L. Identification of vaccine candidate antigens of an ESBL producing *Klebsiella pneumoniae* clinical strain by immunoproteome analysis. *Proteomics* **6**, 836–844 (2006).
26. Ranjbarian, P. et al. Finding epitopes of *Klebsiella pneumoniae* outer membrane protein-K17 (OMPK17) and introducing a 25-mer peptide of it as a vaccine candidate. *Biologia (Bratisl)* **78**(8), 2261–2271 (2023).
27. Hornick, D. B., Allen, B. L., Horn, M. A. & Clegg, S. Adherence to respiratory epithelia by recombinant Escherichia coli expressing *Klebsiella pneumoniae* type 3 fimbrial gene products. *Infect. Immun.* **60**, 1577–1588 (1992).
28. Struve, C., Bojer, M. & Krogfelt, K. A. Characterization of *Klebsiella pneumoniae* type 1 fimbriae by detection of phase variation during colonization and infection and impact on virulence. *Infect. Immun.* **76**, 4055–4065 (2008).
29. Clegg, S. & Murphy, C. N. Epidemiology and virulence of *Klebsiella pneumoniae*. *Microbiol. Spectr.* https://doi.org/10.1128/9781555817404.ch18 (2016).
30. Díaz-Dinamarca, D. A. et al. Protein-based adjuvants for vaccines as immunomodulators of the innate and adaptive immune response: Current knowledge, challenges, and future opportunities. *Pharmaceutics* **14**, 1671 (2022).
31. Wilson-Welder, J. H. et al. Vaccine adjuvants: Current challenges and future approaches. *J. Pharm. Sci.* **98**, 1278–1316 (2009).
32. Awate, S., Babiuk, L. A. & Mutwiri, G. Mechanisms of action of adjuvants. *Front. Immunol.* **4**, 114 (2013).
33. Tavares Da Silva, F., Di Pasquale, A., Yarzabal, J. P. & Garçon, N. Safety assessment of adjuvanted vaccines: Methodological considerations. *Hum. Vaccin. Immunother.* **11**, 1814–1824 (2015).

34. Pulendran, B., Arunachalam, S. P. & O'Hagan, D. T. Emerging concepts in the science of vaccine adjuvants. *Nat. Rev. Drug Discov.* **20**(6), 454–475 (2021).
35. Connell, T. D. Cholera toxin, LT-I, LT-IIa and LT-IIb: The critical role of ganglioside binding in immunomodulation by type I and type II heat-labile enterotoxins. *Expert Rev. Vaccines* **6**, 821–834 (2007).
36. Liang, S. & Hajishengallis, G. Heat-labile enterotoxins as adjuvants or anti-inflammatory agents. *Immunol. Invest.* **39**, 449–467 (2010).
37. Duan, Q., Xia, P., Nandre, R., Zhang, W. & Zhu, G. Review of newly identified functions associated with the heat-labile toxin of enterotoxigenic *Escherichia coli*. *Front. Cell. Infect. Microbiol.* **9**, 292 (2019).
38. Zargaran, F. N. et al. B cell epitopes of four fimbriae antigens of *Klebsiella pneumoniae*: A comprehensive in silico study for vaccine development. *Int. J. Peptide Res. Ther.* **27**, 875–886 (2021).
39. Mackel, J. J., Wasbotten, R., Dahler, A., Smith, C. M. & Rosen, D. A. T cell immunity to a classical strain of *Klebsiella pneumoniae*. *J. Immunol.* **208**, 170.121-170.121 (2022).
40. Williams, C. J. et al. MolProbity: More and better reference data for improved all-atom structure validation. *Prot. Sci.* **27**, 293–315 (2018).
41. Feig, M. Local protein structure refinement via molecular dynamics simulations with locPREFMD. *J. Chem. Inf. Model.* **56**, 1304–1312 (2016).
42. Wiederstein, M. & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407–W410 (2007).
43. Coudert, E. et al. Annotation of biologically relevant ligands in UniProtKB using ChEBI. *Bioinformatics* **39**, btac793 (2023).
44. Reche, P. A. & Reinherz, E. L. Prediction of peptide-MHC binding using profiles. *Methods Mol. Biol.* **409**, 185–200 (2007).
45. Kruiswijk, C. et al. In silico identification and modification of T cell epitopes in pertussis antigens associated with tolerance. *Hum. Vaccin Immunother.* **16**, 277–285 (2020).
46. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, W449–W454 (2020).
47. Jespersen, M. C., Peters, B., Nielsen, M. & Marcatili, P. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* **45**, W24–W29 (2017).
48. Kolaskar, A. S. & Tongaonkar, P. C. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.* **276**, 172–174 (1990).
49. Vita, R. et al. The immune epitope database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, D339-d343 (2019).
50. Gasteiger, E. et al. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook* (ed. Walker, J. M.) 571–607 (Humana Press, 2005).
51. Garg, V. K. et al. MFPPI: Multi FASTA ProtParam interface. *Bioinformation* **12**, 74–77 (2016).
52. Doytchinova, I. A. & Flower, D. R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **8**, 4 (2007).
53. Dimitrov, I., Flower, D. R. & Doytchinova, I. AllerTOP–a server for in silico prediction of allergens. *BMC Bioinform.* **14**(Suppl 6), S4 (2013).
54. Diaz, A. A. et al. Prediction of protein solubility in *Escherichia coli* using logistic regression. *Biotechnol. Bioeng.* **105**, 374–383 (2010).
55. Sharma, N., Naorem, L. D., Jain, S. & Raghava, G. P. ToxinPred2: An improved method for predicting toxicity of proteins. *Brief. Bioinform.* **23**, bbac174 (2022).
56. Kloczkowski, A., Ting, K. L., Jernigan, R. L. & Garnier, J. Combining the GOR V algorithm with evolutionary information for protein secondary structure prediction from amino acid sequence. *Proteins* **49**, 154–166 (2002).
57. Zheng, W. et al. Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Rep Methods* **1**, 100014 (2021).
58. Heo, L., Park, H. & Seok, C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.* **41**, W384–W388 (2013).
59. Williams, C. J. et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
60. Wiederstein, M. & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407-410 (2007).
61. Pontius, J., Richelle, J. & Wodak, S. J. Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J. Mol. Biol.* **264**, 121–136 (1996).
62. Oberholser, K. Proteopedia entry: Ramachandran plots. *Biochem. Mol. Biol. Educ.* **38**, 430 (2010).
63. Ponomarenko, J. et al. ElliPro: A new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform.* **9**, 514 (2008).
64. Pettersen, E. F. et al. UCSF Chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
65. Koymans, K. J. et al. Structural basis for inhibition of TLR2 by staphylococcal superantigen-like protein 3 (SSL3). *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11018–11023 (2015).
66. Su, L. et al. Sulfatides are endogenous ligands for the TLR4-MD-2 complex. *Proc. Natl. Acad. Sci. U. S. A.* https://doi.org/10.1073/pnas.2105316118 (2021).
67. Desta, I. T., Porter, K. A., Xia, B., Kozakov, D. & Vajda, S. Performance and Its limits in rigid body protein–protein docking. *Structure* **28**, 1071-1081.e1073 (2020).
68. Honorato, R. V. et al. Structural biology in the clouds: The WeNMR-EOSC ecosystem. *Front. Mol. Biosci.* **8**, 729513 (2021).
69. Laskowski, R. A. et al. PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci.* **22**, 488–490 (1997).
70. Van Der Spoel, D. et al. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
71. Rapin, N., Lund, O., Bernaschi, M. & Castiglione, F. Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One* **5**, e9862 (2010).
72. Nuryana, I., Laksmi, F. A., Dewi, K. S., Akbar, F. R. & Harmoko, R. Codon optimization of a gene encoding DNA polymerase *from Pyrococcus furiosus* and its expression in *Escherichia coli*. *J. Genet. Eng. Biotechnol.* **21**, 129 (2023).
73. Yamasaki, S. et al. Genetic analysis of ESBL-producing *Klebsiella pneumoniae* isolated from UTI patients in Indonesia. *J. Infect. Chemother.* **27**, 55–61 (2021).
74. Babu, L., Uppalapati, S. R., Sripathy, M. H. & Reddy, P. N. evaluation of recombinant multi-epitope outer membrane protein-based *Klebsiella pneumoniae* subunit vaccine in mouse model. *Front. Microbiol.* https://doi.org/10.3389/fmicb.2017.01805 (2017).
75. Rodrigues, M. X., Yang, Y., de Souza Meira, E. B., do Carmo Silva, J. & Bicalho, R. C. Development and evaluation of a new recombinant protein vaccine (YidR) against *Klebsiella pneumoniae* infection. *Vaccine* **38**, 4640–4648 (2020).
76. Pletz, M. W., Uebele, J., Götz, K., Hagel, S. & Bekeredjian-Ding, I. Vaccines against major ICU pathogens: where do we stand?. *Curr. Opin. Crit. Care* **22**, 470–476 (2016).
77. Zhu, J., Wang, T., Chen, L. & Du, H. Virulence factors in hypervirulent *Klebsiella pneumoniae*. *Front. Microbiol.* **12**, 642484 (2021).
78. Jeannin, P. et al. Outer membrane protein A (OmpA): a new pathogen-associated molecular pattern that interacts with antigen presenting cells—impact on vaccine strategies. *Vaccine* **20**, A23–A27 (2002).

79. Climent, N. et al. Molecular characterization of a 17-kDa outer-membrane protein from *Klebsiella pneumoniae*. *Res. Microbiol.* **148**, 133–143 (1997).
80. Hussein, K. E., Bahey-El-Din, M. & Sheweita, S. A. Immunization with the outer membrane proteins OmpK17 and OmpK36 elicits protection against *Klebsiella pneumoniae* in the murine infection model. *Microb. Pathog.* **119**, 12–18 (2018).
81. Yurina, V. & Adianingsih, O. R. Predicting epitopes for vaccine development using bioinformatics tools. *Ther. Adv. Vaccines Immunother.* **10**, 25151355221100216 (2022).
82. Baral, P., Batra, S., Zemans, R. L., Downey, G. P. & Jeyaseelan, S. Divergent functions of Toll-like receptors during bacterial lung infections. *Am. J. Respir. Crit. Care Med.* **190**, 722–732 (2014).
83. Le, J., Kulatheepan, Y. & Jeyaseelan, S. Role of toll-like receptors and nod-like receptors in acute lung infection. *Front. Immunol.* **14**, 1249098 (2023).
84. Dar, H. A. et al. Immunoinformatics-aided design and evaluation of a potential multi-epitope vaccine against *Klebsiella pneumoniae*. *Vaccines* **7**, 88 (2019).
85. Farhadi, T. et al. Designing of complex multi-epitope peptide vaccine based on omps of *Klebsiella pneumoniae*: an in silico approach. *Int. J. Pept. Res. Ther.* **21**, 325–341 (2015).
86. Ranjbarian, P. et al. Finding epitopes of *Klebsiella pneumoniae* outer membrane protein-K17 (OMPK17) and introducing a 25-mer peptide of it as a vaccine candidate. *Biologia* **78**, 2261–2271 (2023).

Acknowledgements

This work was supported by Hormozgan University of Medical Sciences, Bandar Abbas, Iran (Number: IR.HUMS.REC.1402.014).

Author contributions

K.J. and B.SH.: execution of the programs, preparation of the first draft of the manuscript, analysis, and interpretation of the data; B.SH., and M.B.: execution of the programs and analysis tools; P.S.: study design, supervision of study implementation steps, analysis of data; KH. A.: study design and advice, preparation of the first draft of the manuscript, and final revision.

Funding

Not applicable.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-91602-y.

Correspondence and requests for materials should be addressed to K.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.