

RESEARCH

Open Access



In silico designing of a novel epitope-based candidate vaccine against *Streptococcus pneumoniae* with introduction of a new domain of PepO as adjuvant

Zohreh Bahadori^{1,2,3†}, Mona Shafaghi^{1,2,3†}, Hamid Madanchi^{1,2,4}, Mohammad Mehdi Ranjbar⁵, Ali Akbar Shabani^{1,2*} and Seyed Fazlollah Mousavi^{3*}

Abstract

Background: *Streptococcus pneumoniae* is the leading reason for invasive diseases including pneumonia and meningitis, and also secondary infections following viral respiratory diseases such as flu and COVID-19. Currently, serotype-dependent vaccines, which have several insufficiency and limitations, are the only way to prevent pneumococcal infections. Hence, it is plain to need an alternative effective strategy for prevention of this organism. Protein-based vaccine involving conserved pneumococcal protein antigens with different roles in virulence could provide an eligible alternative to existing vaccines.

Methods: In this study, PspC, PhtD and PsaA antigens from pneumococcus were taken to account to predict B-cell and helper T-cell epitopes, and epitope-rich regions were chosen to build the construct. To enhance the immunogenicity of the epitope-based vaccine, a truncated N-terminal fragment of pneumococcal endopeptidase O (PepO) was used as a potential TLR2/4 agonist which was identified by molecular docking studies. The ultimate construct was consisted of the chosen epitope-rich regions, along with the adjuvant role (truncated N-PepO) and suitable linkers.

Results: The epitope-based vaccine was assessed as regards physicochemical properties, allergenicity, antigenicity, and toxicity. The 3D structure of the engineered construct was modeled, refined, and validated. Molecular docking and simulation of molecular dynamics (MD) indicated the proper and stable interactions between the vaccine and TLR2/4 throughout the simulation periods.

Conclusions: For the first time this work presents a novel vaccine consisting of epitopes of PspC, PhtD, and PsaA antigens which is adjuvanted with a new truncated domain of PepO. The computational outcomes revealed that the suggested vaccine could be deemed an efficient therapeutic vaccine for *S. pneumoniae*; nevertheless, in vitro and in vivo examinations should be performed to prove the potency of the candidate vaccine.

Keywords: *S. pneumoniae*, PspC, PhtD, PsaA, PepO, Protein-based vaccine, Toll-like receptor (TLR) agonist

[†]Zohreh Bahadori and Mona Shafaghi have contributed equally to this work

*Correspondence: shabani@semums.ac.ir; mousavi@pasteur.ac.ir

¹ Department of Medical Biotechnology, Faculty of Medicine, Semnan University of Medical Sciences, Semnan, Iran

³ Department of Bacteriology, Pasteur Institute of Iran, Tehran, Iran
Full list of author information is available at the end of the article

Background

Streptococcus pneumoniae (pneumococcus), an opportunistic Gram+ bacterium, resides normally on the throat or nasopharynx of healthy people and is transmitted through the carriers' respiratory droplets [1]. It can cause life-threatening bacterial infections such as



meningitis, sepsis, and pneumonia if it invades sterile regions of the body [2]. Moreover, pneumococcus is one of the leading causes of severe secondary infections following viral respiratory diseases such as coronavirus [3]. According to reports from the WHO, pneumococcal (Pnc) infections claimed the lives of 1.6 million people in both developed and developing nations in 2005, which happened a lot in the children <2 and elderly ≥ 65 years [4]. Prevention of this considerable mortality rate on one side and progressing resistance of *S. pneumoniae* to existing antimicrobials on the other emphasize the critical need for an efficient vaccine that protects high-risk groups from the majority of clinically invasive serotypes. Presently, there are two types of polysaccharide pneumococcal vaccines on the market, unconjugated (plain) polysaccharide vaccines (PPV) and protein-conjugated polysaccharide vaccines (PCV) [5]. PPV comprising the T-cell independent polysaccharide antigens is ineffective in infants under two-year old who are at the highest risk for severe pneumococcal diseases. PCV protects children, however, it is costly, very difficult to manufacture, requires multiple injections, and needs refrigeration. Moreover, the PCV serotype coverage includes pneumococcal serotypes of developed countries [6, 7]. Both existing vaccines are serotype-dependent and only elicit serotype-specific immunity [7]. Because of the boundaries of available *S. pneumoniae* vaccines, the development of a powerful, broad-spectrum, and cost-effective vaccine that can be effective in preventing infections caused via different pneumococcal serotypes is a significant concern of World Health Organization.

Protein-based vaccine, comprising multiple conserved pneumococcal protein antigens, can give an alternative to serotype-dependent vaccines [8]. Understanding the different roles of virulence and colonization factors that induce pneumococcal infections is fundamental for developing the effective protein-based vaccines. Various pneumococcal proteins have been evaluated as possible vaccine candidates throughout the years, for instance, pneumococcal surface protein A and C (PspA/C) [9, 10], the members of pneumococcal histidine triad (Pht) family A, B, D, and E (PhtA, B, D, and E) [11, 12], and ATP-binding cassette (ABC) transporters such as PsaA, PiuA, and PiaA [13, 14]. Each of the above mentioned virulence proteins could produce various levels of protection against immunological challenges with numerous pneumococcal serotypes in animal model [7].

Pneumococcal surface protein C is one of the most important surface proteins of pneumococcus that is considered as a vaccine candidate [15]. PspC has been shown to have notable roles in adhering of pneumococcus, invading, and evading complement [10, 16]. Hence, it is also called factor H-binding inhibitor of complement

(Hic) [17], C3-binding protein A (PbcA) [18], *S. pneumoniae*-secretory IgA-binding protein (SpsA) [19], and choline-binding protein A (CbpA) [20]. PspC consists of three domains: an exposed N-terminal (N-ter) α -helical domain, a proline-rich domain, and C-terminal (C-ter) repeat region [21]. This antigen shows variability among different serotypes and is classified into 11 groups [22]. The protein PspC as a recombinant vaccine candidate has been indicated to induce antibody-mediated protection against pneumococci [21, 23]. Investigations have revealed that anti-PspC3 (PspC from group 3) antibodies are capable to identify the majority of Pnc clinical isolates [24]. Pht family, which is a group of surface-associated pneumococcal antigens, contains 5 conserved His triad motifs (His-xx-His-x-His) in PhtA, B, and D, as well as 6 motifs in PhtE [12, 25]. A number of roles have been suggested for this family including involvement in the homeostasis of Zinc (Zn^{2+}) which is crucial for the colonization and invasion of pneumococcus [26], and also complement deposition inhibition on the surface of pneumococcus via the recruitment of complement factor H [27]. Among the members of this family, PhtD has been determined to be the most conserved across different strains [28], and so to induce the broadest protection against pneumococcal diseases [29]. Researches have indicated that the C-terminal of PhtD (PhtD-C), compared to the other areas of the protein, is more surface-exposed and could be a more immunogenic part [28, 30]. Another protein candidate is pneumococcal surface adhesion A (PsaA) that has been assessed against Pnc infections with encouraging results in animal models and human clinical trials [31]. The PsaA antigen, a conserved lipoprotein expressed by all pneumococcal strains, is an important virulence factor of pneumococci [32, 33]. PsaA plays vital functions in manganese (Mn) transport [34] and adherence to the cells of the host [35]. In pneumococcus, the acquisition of Mn^{2+} has been found to have significant roles in growth, proliferation, and virulence [36].

Adjuvants are another component of vaccines that could significantly enhance the efficiency of vaccines and have been employed for more than a century [37]. In the past few decades, toll-like receptor (TLR) agonists have been extensively studied as possible adjuvants for vaccines [38]. Most of the TLR agonists currently being investigated are non-protein microbial products such as lipopeptides, oligonucleotides, and lipopolysaccharides [39]. On the other side, a developing number of studies have revealed there are various bacterial proteins that can activate TLR signaling and immune responses. Bacterial proteins that show one or more of the following features can be considered as potential adjuvant candidates: (1) induction of pro-inflammatory cytokines,

(2) up-regulation of the costimulatory molecule expression on antigen presenting cells (APCs), (3) activation of the TLR2/TLR4 signaling, (4) induction of antibody- or cell-mediated immunity [39]. Microbial protein TLR agonists present unique characteristics that are not provided by non-protein TLR agonists, including the ability to modify the structure/function as needed for favorable immunogenicity and minimum toxicity [39]. To ensure the co-delivery of adjuvant and antigen into the target cells, protein TLR agonists could be genetically joined to protein antigens, leading to powerful stimulation of innate and acquired immune response [39]. *S. pneumoniae* endopeptidase O (PepO) is a TLR agonist that is able to induce a strong natural immune response in a TLR2/TLR4-dependent way [40, 41]. This protein which share homology with M13 peptidase family consists of two conserved domains: M13_N domain (residues 3–383) and M13 domain (residues 437–627). It has shown that the TLRs 2 and 4 bind to the Pnc endopeptidase O through the M13_N domain [40, 41].

Recent advances in cost- and time-efficient approaches of immunoinformatics and the availability of a broad array of tools for vaccine designing have notably allowed to develop new stable and accurate epitope vaccines [42–44]. Considering the feasibility and benefits of vaccines generated via immunoinformatics methods, the aim of this research was to employ these approaches to design an epitope-based vaccine against *streptococcus pneumoniae*. Herein, we used several virulence proteins of pneumococcus, i.e. PspC, PhtD, and PsaA, to predict B-cell and helper T-cell epitopes from them. The M13_N domain of PepO was subjected to molecular docking with TLR2/4 to determine the potential regions involved in the PepO-TLR interaction. Furthermore, molecular docking, molecular dynamics simulation, and in silico cloning were done on the final construct. The novelty of this study is the construction of an epitope-based vaccine containing more effective antigenic epitope-rich domains of three pneumococcal proteins, PspC, PhtD, and PsaA, which are capable of production of more diverse and robust responses. Another novelty of our study lies in the use of a new truncated domain of PepO as a TLR agonist adjuvant candidate for promoting the immunogenicity of the vaccine.

Methods

This investigation was accomplished in three steps including the following: (1) immunoinformatics analysis of the proteins PspC, PhtD, and PsaA to predict the epitope-rich regions, (2) analysis of the PepO (as a TLR2/4 agonist) to identify its potential regions involved in the interaction with TLR, in order to be used as an adjuvant candidate, (3) design and construction of final

epitope-based vaccine, and in silico cloning. Figure 1 depicts the process of this research.

Immunoinformatics analysis

Protein sequence retrieval and primary analysis

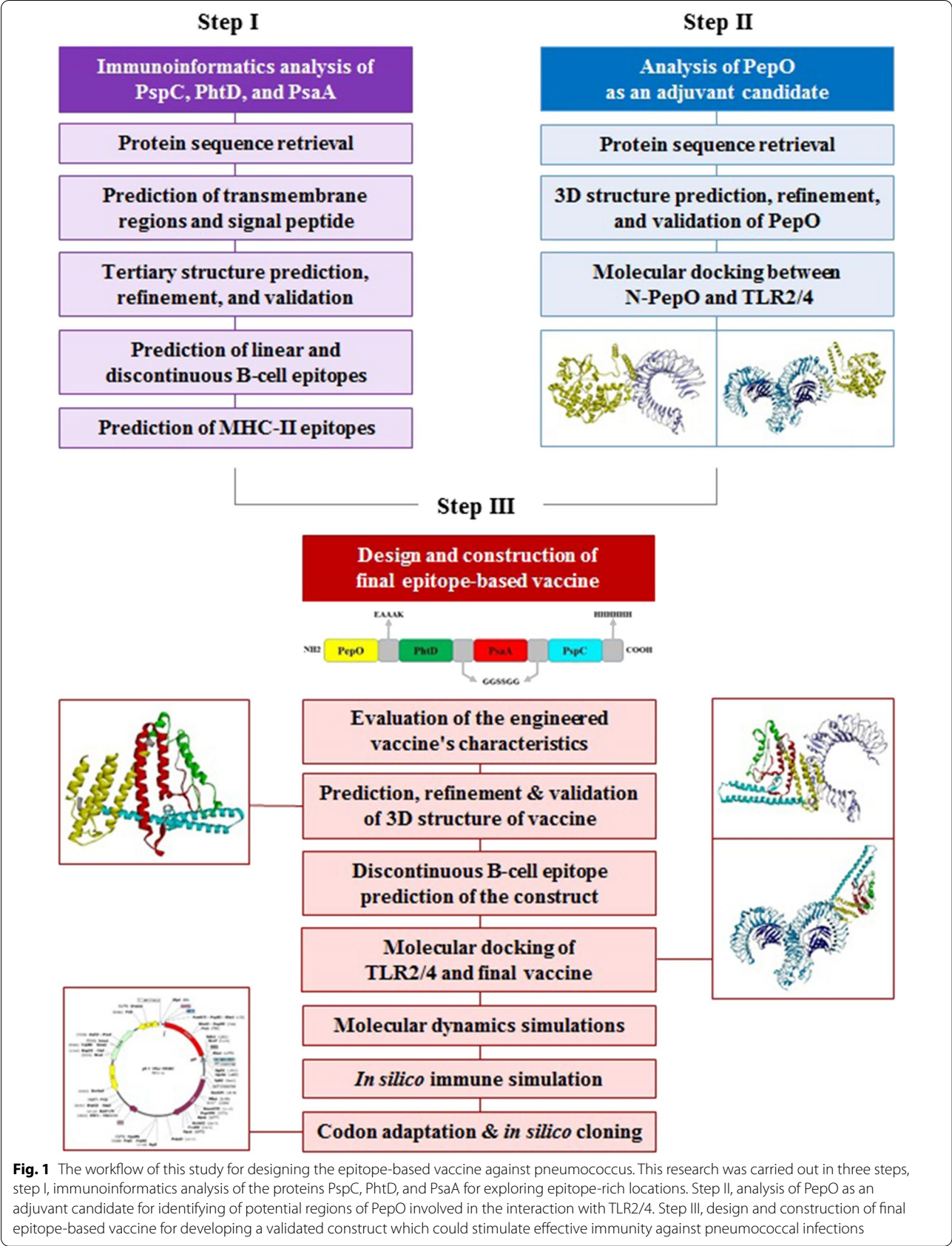
The sequences of PhtD (Accession no. AAK99711), PsaA (Accession no. AAF70663), and PspC from group 3 (PspC3, Accession no. EF424119) were fetched from the NCBI protein database (www.ncbi.nlm.nih.gov/protein). TMHMM Server v. 2.0 was employed to identify the transmembrane helices of proteins (www.cbs.dtu.dk/services/TMHMM-2.0) [45]. The SignalP 5.0 server was applied to detect the signal peptide region (<http://www.cbs.dtu.dk/services/SignalP/>) [46].

Tertiary structure prediction of the selected proteins

Depending on the (un) availability of the structural template in the protein data bank (PDB), the template (free)-based modeling methods are used for the prediction of protein tertiary structure. In this study, the tertiary structures of PhtD and PspC were modeled by I-TASSER, a template-based method of modeling (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) [47]. The I-TASSER suite, which automatically models 3D-structure of a given amino acid sequence, provides the model's ranking scores such as confidence score (C-score), template modeling (TM) score, and Root Mean Square Deviation (RMSD). The score of confidence ranges from -5 to +2, with +2 representing the most appropriate model. Template modeling score is in the range of 0–1, with closer to 1 representing the closeness of structure to global topology [48]. The score of RMSD indicates the similarity degree between a specific protein and the other. The model with a lower RMSD is more acceptable and realistic [49]. I-TASSER usually takes 1–2 days to produce a model depending on the protein size.

Tertiary structure refinement and validation

The most desirable 3D structures modeled by the I-TASSER were refined through the GalaxyRefine server (<http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>) [50]. GalaxyRefine, which can improve the predicted structure's local and global quality, first reconstructs the side chains, and repacks them, and subsequently relaxes the overall structure using MD simulation. The average run time is usually 1 to 2 h. The server shows 5 modeled structures and respective their Mol-Probity, GDT-HA, Clash, Poor rotamers, RMSD along with Rama favored scores. The Z-score of ProSA server (<https://prosa.services.came.sbg.ac.at/prosa.php>) [51] and Ramachandran plot (R-Plot) of PROCHECK (<https://servicesn.mbi.ucla.edu/PROCHECK/>) [52] were computed to assess the correctness of the projected model



and the effectiveness of the refinement process. The Z-score represents how similar the models are to native proteins, while the R-plot depicts the distribution of residues in most favored, allowed, and disallowed zones.

Prediction of B-cell epitopes

A multi-method strategy was used to identify more dependable B-cell epitopes. The sequential B-cell epitopes of the antigens PhtD, PsaA, and PspC were predicted through machine learning-based algorithms such as ABCpred (<http://crdd.osdd.net/raghava/abcpred/>) [53], and LBTope (<http://crdd.osdd.net/raghava/lbtope/>) [54], and also physicochemical-based algorithms e.g. IEDB Emini tool (<http://tools.iedb.org/bcell/result/>) [55], and Ellipro (<http://tools.iedb.org/ellipro/>) [56]. Beside the high accuracy of the epitope prediction servers, one of their important features is their short response time (less than 5 min). ABCpred server is one of the first servers on the basis of the recurrent neural network, which generates epitopes with the length in the range of 12–20-mers. The LBtope server has an accuracy of around 81%. IEDB Emini tool was used to predict the surface accessibility of epitopes. The servers were applied at default settings.

Likewise, conformational B-cell epitopes were predicted based on the tertiary structures of the considered antigens using the servers Ellipro and DiscoTope (<http://tools.iedb.org/discotope/>) [57]. The modeled structures using the I-TASSER server and the crystal structure of PsaA (PDB id: 1PSZ) retrieved from RCSB (<https://www.rcsb.org/>) were used as inputs of these predictor servers.

Prediction of MHC-II epitopes

T cell epitope prediction tool of the IEDB web server (<http://tools.immuneepitope.org/mhcii>) [58] and NetMHCIIpan 4.0 (<http://www.cbs.dtu.dk/services/NetMHCIIpan/>) [59] were used to discover appropriate MHC-II binding epitopes. In IEDB, each projected epitope has a percentile rank that indicates how well the epitope binds to MHC-II. The lower percentile rank is the better rank. We chose the eight HLA-DR1 alleles (DRB1*01:01, DRB1*03:01, DRB1*04:01, DRB1*07:01, DRB1*08:01, DRB1*11:01, DRB1*13:01, and DRB1*15:01) as the most common human alleles for MHC-II [60], and also the 3 mouse alleles (H2-IAb, H2-IAd, and H2-IEd).

Determination of the potential regions involved in PepO-TLR interaction by molecular docking

Sequence retrieval and structure modeling of PepO

The sequence of PepO (Accession no. ABJ55398) was fetched from the NCBI protein database. Since the tertiary structure of Pnc endopeptidase O was not available in PDB, the 3D structure of its M13_N domain was

modeled and refined through the I-TASSER and Galaxy-Refine servers, respectively.

Molecular docking between N-PepO and TLR2/4

The protein PepO is able to enhance macrophage unspecific phagocytosis and bactericidal effect, depending on interactions of its N-terminal with the toll-like receptors 2 and 4. For a more detailed understanding of the interaction of N-PepO with TLR2/4 and prediction of the amino acid residues involved in the interacting interface, molecular docking was performed through the ClusPro v2.0, a totally automated protein–protein docking server (<https://cluspro.org/login.php>) [61]. The average time used for the docking run is about 4 h, depending upon the size of the submitted proteins. This server conducts the following computing steps: rigid-body docking, clustering of the lowest energy conformations, and refinement of chosen complexes through energy minimization. For this docking analysis, 3-D structure of N-PepO was modeled using I-TASSER (as mentioned above) and the crystal structures of human TLR2 (PDB code: 6NIG) and TLR4 (PDB code: 3FXI) were retrieved from the RCSB site. The best docked complexes of the N-PepO-receptor with the lowest energy were chosen. To display and investigate the considered complexes, the Discovery Studio (DS) Visualizer and DIMPLOT tool from LigPlot + v2.2.4 [62] were used.

Design and construction of epitope-based vaccine

On the basis of the predicted findings of the used servers, B cell and T cell epitopes with high scores were considered for generating an epitope-based vaccine. To keep and increase the immunogenicity of the chosen epitopes, a suitable pattern of them should be linked together employing appropriate linkers. The GGSSGG flexible linker was introduced between the proper epitopes of PhtD, PsaA, and PspC to maintain their independent folding and immunological activities [63]. The chosen sequence of the N-PepO as a potential adjuvant candidate was connected to the N-terminal of the above mentioned sequence with the EAAAK helical linker [64, 65] to promote the immunogenicity of the construct. In the end, a hexa-histidine (His6) tag was attached to the C-terminal of the engineered construct to aid the purification.

Evaluation of the engineered vaccine's characteristics

The antigenicity potential of the candidate construct was investigated via the VaxiJen server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) [66] with a threshold of 0.4 and the ANTIGENpro tool from the Scratch Protein Predictor server (<http://scratch.proteomics.ics.uci.edu/>) [67]. To predict allergenicity of the designed construct, AlgPred (<https://www.imtech.res>

in/raghava/algpred/) [68] and AllerTOP (<http://www.ddg-pharmfac.net/AllerTOP/>) [69] servers were utilized. The SOLpro server at (<https://scratch.proteomics.ics.uci.edu/>) [70] was employed for predicting the solubility of the vaccine upon overexpression in *Escherichia coli*. Toxicity of the primary structure of the construct was predicted through the ToxinPred server (<https://webs.iitd.edu.in/raghava/toxinpred/index.html>) [71]. The ProtParam tool from ExPasy web server (<https://web.expasy.org/protparam/>) was applied for determining theoretical isoelectric point value, molecular weight, half-life, aliphatic index, instability index, and other physiochemical properties of the designed construct [72]. Tools such as VaxiJen, AlgPred, AllerTOP, ToxinPred, and ProtParam are very fast and can predict in a few seconds, while other tools such as ANTIGENpro and SOLpro are slower and require about half an hour.

Prediction, refinement, and validation of the construct's tertiary structure

The designed construct's 3D structure was modeled via Robetta server (<http://rosetta.bakerlab.org/>) [73]. The server takes about 4 to 6 h to run a short query of 150 residues. The best three-dimensional model was chosen and refined through GalaxyRefine server. Validating the model is a key step for identifying potential errors in the primary model and comparing the quality of the primary and refined models. The geometry, stereochemistry, and other structural aspects of the models were compared through Ramachandran plot, ProSA Z-score, and ERRAT (<https://servicesn.mbi.ucla.edu/ERRAT/>). R-Plot checked the stereochemical and geometrical restrictions of the construct. The ProSA server computed a Z-score to represent the construct structure's overall quality. The non-bonded atomic interactions were analyzed via ERRAT [74]. A higher ERRAT score represents a more trustworthy model for future investigation.

Prediction of discontinuous B-cell epitopes of the construct

Following the three dimensional modeling of designed construct, prediction of its conformational B-cell epitopes which is a key task in designing of vaccine was performed via Ellipro tool of the IEDB server (<http://tools.iedb.org/ellipro/>) [56].

Analysis of the designed vaccine's molecular interactions with immune receptors

The ClusPro online server was applied to carry out molecular docking between the engineered construct and toll-like receptors (TLR2 and TLR4). The 3D structure of the vaccine modeled by the Robetta server and the crystal structures of human TLR2/4 were used as the inputs for the docking processes. The rest of the steps

were conducted similarly to the above explained docking between N-PepO and TLR2/4, and finally the best complexes were chosen.

Molecular dynamics (MD) simulation

After conducting the docking and determining the finest orientation of the designed vaccine and receptor for interaction with each other, the 2 top-scored complexes were presented to the simulations of MD in 2 separate runs. The MD simulation analysis was performed for 50 ns to compute the docked complex's stability using the GROMACS program [75] with GROMOS96 54A7 force field [76]. The MD run was performed in about two days for each selected complex, using the described conditions in our prior works [77, 78]. Briefly, the target complexes were placed within a proper cubic box, solvated applying the TIP3P water type, and neutralized by adding sufficient Na⁺/Cl⁻ ions. The solvated systems were energy-minimized through the steepest descent method and equilibrated under NVT and NPT. During the MD simulations, the values of RMS Deviation and RMS Fluctuation (RMSF) were calculated through the use of g_rms and g_rmsf, respectively.

Immune system simulation

To further verify the immune response and immunogenic profile of the designed construct, a computational simulation of the immune system was performed utilizing the C-ImmSim server (<https://kraken.iac.rm.cnr.it/C-IMMSIM/>) [79]. The server simulates at the same time three mammalian immune system compartments including bone marrow, thymus, and tertiary lymphatic organs. It utilizes machine learning approaches and a Position Specific Scoring Matrix (PSSM) method to predict immune epitopes and immune interactions, respectively. Three injections of the engineered vaccine were administered at 4-week intervals [80], and all simulation settings were left to default with time steps set at 1, 84, and 168, where each time step is eight hours, and the first time step is injection at the time zero.

Reverse translation and codon optimization of the designed construct

In order to express the engineered vaccine in *Escherichia coli* expression system, the Java Codon Adaptation Tool (JCat) server (<http://www.jcat.de/>) was applied to do reverse translation and codon optimization in a few seconds [81]. The amino acid sequence of designed construct was translated reversely to the DNA sequence by setting the codon usage of *E. coli* k12 host. Additional optimizations were chosen to avoid the restriction enzymes (RE) cleavage sites, rho-independent transcription termination, and prokaryote ribosome binding site.

The output of JCat includes the percent of CG content and codon adaptation index (CAI) that may be utilized to evaluate the levels of protein expression. The sequence's CG content which has favorable effects on translation and transcription should be between 30 and 70 percent. CAI presents details on codon usage biases; the desired score of CAI is >0.8 and ≤ 1.0 . Ultimately, to ensure vaccine expression, the optimized sequence (or gene) was inserted into the pET-28a (+) through SnapGene tool (<https://www.snapgene.com/try-snapgene/>) which provides a fast and easy way for in silico cloning.

Results

Immuno-informatics analyses

Protein sequence retrieval

The protein sequences of PhtD, PsaA, and PspC were retrieved from the NCBI. The C-terminal of PhtD (PhtD-C, 383–853), PsaA, and PspC amino acid chains were applied in the project to predict T helper and B cell epitopes.

Prediction of transmembrane regions and signal peptide

The transmembrane topology investigation showed that the mentioned antigens lacked any transmembrane helices (Additional file 1: Fig. S1). According to the results of the SignalP server, the selected proteins except PsaA did not have a signal peptide (Additional file 1: Fig. S2).

Tertiary structure prediction, refinement, and validation

The 3D structures of PhtD-C and PspC obtained from I-TASSER showed better quality than the models generated by the other servers. Considering the C-score, the model number one of PhtD-C (with C-score: -0.64) and PspC (with C-score: -0.65) were selected as the most suitable 3D-structure among five generated models for each of the proteins. The selected models of PhtD-C and PspC were refined using GalaxyRefine (Additional file 1: Fig. S3). The structure quality of the models was improved following treatment with the server as represented via ProSA Z-score and Ramachandran plot. The results of validation of the protein models before and after refinement are presented in Additional file 1: Table S1. The ProSA server revealed that refined models of PhtD-C and PspC had the Z-score of -9.1 and -0.04, respectively, which are within the range of scores for similar size natural proteins (Additional file 1: Fig. S4 A and B). According to the assessment of the Ramachandran plot of refined models, there were 81.7%, 16.9%, and 1.4% residues of the PhtD-C, 91.5%, 6.6%, and 1.9% residues of the PspC in the most favored, allowed, and disallowed regions, respectively (Additional file 1: Fig. S4 a, b).

Identification of linear and discontinuous B-cell epitopes

As B-cell epitopes have an essential role in humoral responses, the sequential and discontinuous B-cell epitopes of the mentioned proteins were predicted using LBTope, ABCpred, IEDB Emini tool, and Ellipro (for the sequential epitopes), Ellipro and DiscoTope servers (for discontinuous epitopes). The prediction results obtained by the servers are provided in Additional file 1: Tables S2–S7, the first 3 tables are for linear and the last 3 are for discontinuous B-cell epitopes.

Identification of MHCII epitopes

T-helper epitopes from PhtD-C, PsaA, and PspC sequences were predicted using IEDB (percentile rank <10) and NetMHCIIpan (%Rank ≤ 1). The different MHCII epitopes from the protein sequences were derived via the servers according to eight HLA-DRB1 alleles (DRB1*01:01, DRB1*03:01, DRB1*04:01, DRB1*07:01, DRB1*08:01, DRB1*11:01, DRB1*13:01, and DRB1*15:01) and three mouse alleles (H2-IAb, H2-IAd, and H2-IEd). The results of predictions are listed in Additional file 1: Tables S8–S10.

Analysis of the interaction of N-PepO with TLR2/4

Retrieval of sequence and modeling of 3D structure

The protein sequence of PepO was fetched from the NCBI. The tertiary structure of N-PepO obtained from I-TASSER showed better quality than the models generated through the other servers. The 3D model number one of N-PepO (with C-score: 1.42) was chosen as the most proper structure among the generated models. The selected model of N-PepO was refined using GalaxyRefine (Additional file 1: Fig. S5A). The results of validation before and after refinement represented that the structure quality of the model was improved. The ProSA server revealed that the refined model had the Z-score of -2.62, which is within the range of scores for similar size natural proteins (Additional file 1: Fig. S5B). According to the assessment of the Ramachandran plot of the refined model, there were 94.4%, 4.5%, and 1.1% residues in the most favored, allowed, and disallowed zones, respectively (Additional file 1: Fig. S5C).

A more detailed identification of interaction of N-PepO with TLR2/4 by molecular docking

To identify the residual interactions between N-PepO and TLR2/4, the refined modeled structure of N-PepO from the previous step and the crystal structures of TLRs 2 and 4 were docked by the ClusPro 2.0 online server. Following docking, it was seen that the N-PepO possessed good interaction with TLR2/4. The model.000.00 for N-PepO-TLR2 and model.000.02 for N-PepO-TLR4 were chosen with the lowest energy docking mode -859.4 and

-967.1, respectively. The Discovery Studio Visualizer and DIMPLOT from LigPlot⁺v2.2.4 were used for visualizing and analyzing the chosen complexes (Figs. 2 and 3). The N-PepO-TLR2 complex formed 14 hydrogen bonds and 4 salt bridges, Arg³⁵⁴-Glu²⁶⁴, Lys³⁵⁷-Glu²⁶⁴, Asp³⁴-Arg³²¹, Phe⁸-Tyr³²³, Tyr⁹-Tyr³²³, Asn¹³-Tyr³²³, Gln⁵-Leu³²⁸, Arg³-Tyr³³², Gln⁵-Tyr³³², Asp⁶-Tyr³³², Glu¹⁵-Lys³⁴⁷, and Arg³-Leu³⁵⁰ (Table 1). Comprehensive insights into intermolecular interactions of the N-PepO-TLR2 showed that all the hydrogen bonds except Asn¹³-Tyr³²³ formed at distances below 3 Å (Table 1). Likewise, the N-PepO-TLR4 complex made 10 hydrogen bonds and 2 salt bridges, Thr¹⁹-Arg⁶⁷, Glu²¹-Arg⁶⁷, Ala¹¹-Arg⁸⁷, Ile¹²-Arg⁸⁷, Glu¹⁵-Arg⁸⁷, Asn¹³-Gln⁹¹, Tyr⁴-Ser¹⁸⁴, Tyr⁴-Lys¹⁸⁶, Gln⁵-Lys¹⁸⁶, and Gln⁵-Glu²⁶⁶. All the hydrogen bonds, excluding Ile¹²-Arg⁸⁷, formed at distances less than 3 Å (Table 1). Considering the obtained results, a truncated fragment of N-PepO from amino acids 1 to 112, which was potentially involved in the interactions between N-PepO and TLRs, was selected as a potential TLR2/4 agonist candidate for further stage of the research.

Designing of epitope-based vaccine

Based on the results of the immuno-informatics investigation, epitopes of B-cell and MHC-II with the best scores were considered. The final construct of the vaccine consisting of four domains: the truncated N-PepO (1–112), PhtD-C (196–256), PsaA (100–187), and PspC (276–363) was designed. The chosen domains of PhtD, PsaA, and PspC were connected together via GGSSGG linkers. The selected fragment of N-PepO as a potential candidate adjuvant was joined to the N-terminal of the considered domains with the EAAAK linker. Following attachment of the His6 (HHHHHH) tag to the C-terminal of the construct, the epitope-based vaccine (ODAC) was finally designed to be 372 residues in length. The ODAC sequence is as follows:

MTRYQDDFYDAINGEWQQTAEIPADKSQTG-GFVDLDQEIEDLMLATTDKWLAGEEVPEDAILN-FVKYHRLVRDFDKREADGITPVLPLLKEFQELET-FADFTAKLAEFELAEAAKAAQAYAKEKGLTPPST-DHQDSGNTEAKGAEAIYNRVKAAKKVPLDRMPY-NLQYTVEVKNGSGGSSGSDGVDVIYLEGQNEK-GKEDPHAWLNLENGIIFAKNIAKQLSAKDPNNKE-FYEKNLKEYTDKLDKLDKESKDKFNKIPAEEKKLIVT-SEGGSSGGRRNYPSTNTYFSLELEISESDVEVKKAEFEL-

VKEEAKEPRNEEKVKQAKAKVESKKAVATRLENIKT-DRKKAEEEEAKRKAAEEDKVKEKGHHHHHHH.

The details of the chosen epitopes in the designed construct and schematic diagram of the vaccine are presented in Table 2 and Fig. 4.

Evaluation of the sequence properties of the designed vaccine

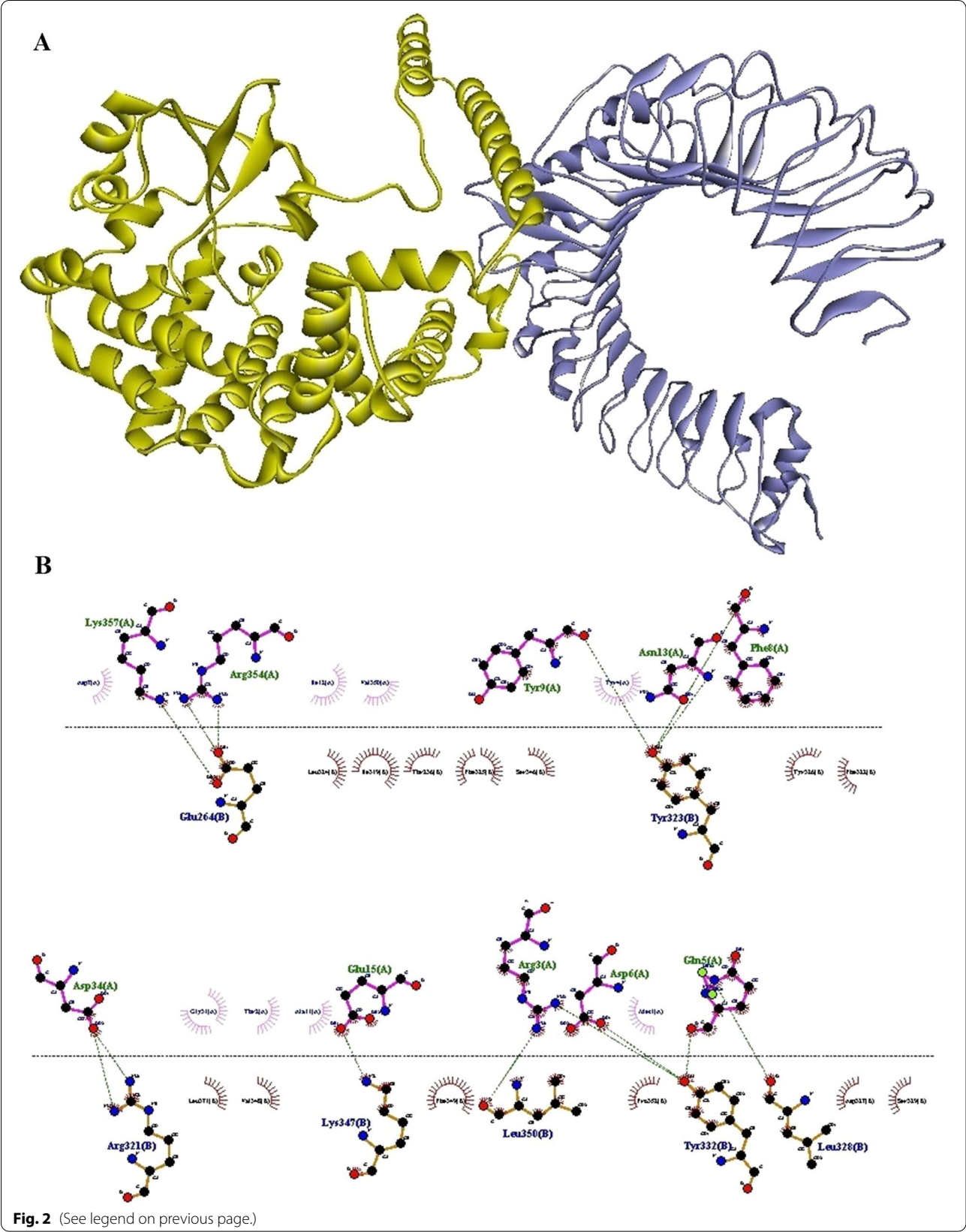
The sequence antigenicity probability was estimated to be 0.748 and 0.926 using the VaxiJen and the ANTIGEN-pro, respectively. The engineered sequence was predicted to be non-allergenic by Alg Pred and Allertop online servers. According to SOLpro results, the solubility of the candidate vaccine upon overexpression in *Escherichia coli* was 0.893. ToxinPred was employed for predicting the toxicity of the designed sequence and it was deemed as non-toxic. Furthermore, the ProtParam results indicated that theoretical isoelectric point value (pI) and molecular weight (Mw) of the construct were 5.29 and 41.939 kDa, respectively. Half-life was computed to be 30 h in mammalian reticulocytes, >20 h in yeast, and more than 10 h in *E. coli*. Aliphatic index (AI) and grand average of hydropathicity (GRAVY) values were defined as 68.01 and -0.954, respectively. The instability index (II) of the designed sequence was estimated to be 36.83, indicating that it is stable. The total number of positive (+R) and negative (-R) residues were computed 60 and 70, respectively. The details of the sequence properties of the designed vaccine are presented in Table 3.

Prediction, refinement, and validation of tertiary structure of the vaccine

Five 3-dimensional structures of the designed vaccine sequence were generated by the Robetta server, and the best model was chosen. The R-plot revealed that 92.5, 6.9, and 0.6 percent of amino acids were located in the favored, allowed, and disallowed zones (Fig. 5A). Moreover, the ProSA Z-score of -6.78 was inside the area of scores for natural proteins, confirming the overall quality and reliability of the designed model (Fig. 5B). The ERRAT score of the chosen model was 94.92 percent (Fig. 5C). Following the refinement of the structure using Galaxy Refine, the model properties were improved. The primary and the refined structures are compared in Fig. 5 and Table 4. The R-plot of the refined model showed that 94.00, 5.1, and 0.9 percent of residues are located in favored, allowed, and disallowed zones, respectively

(See figure on next page.)

Fig. 2 Docking output of the N-PepO and TLR2 receptor. **A** The 3D structure shows the interaction between TLR2 and PepO as the receptor and ligand represented in harbor blue and yellow, respectively. **B** The formation of hydrogen bonds between the Glu²⁶⁴, Arg³²¹, Tyr³²³, Leu³²⁸, Tyr³³², Lys³⁴⁷, and Leu³⁵⁰ of TLR2 and the Arg³, Gln⁵, Asp⁶, Phe⁸, Tyr⁹, Asn¹³, Glu¹⁵, Asp³⁴, Arg³⁵⁴, and Lys³⁵⁷ residues of PepO is shown by Dimplot 2D-interaction plot. TLR2 residues, PepO residues, hydrogen bonds, and non-bonded residues are presented in blue, green, green dashed lines, and red/pink eyelashes, respectively



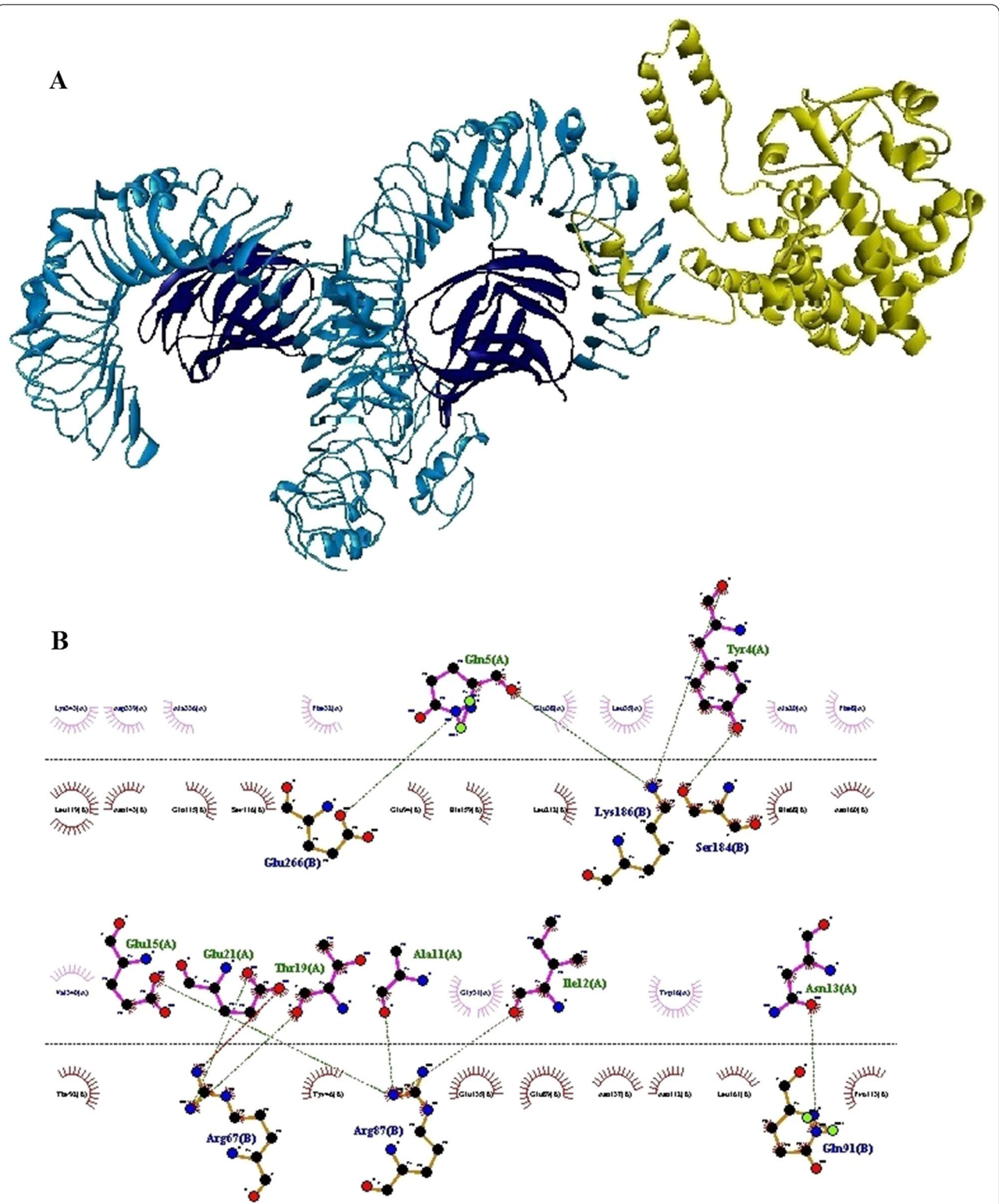


Fig. 3 The result of docking of the N-PepO and TLR4 receptor. **A** The model demonstrates the interaction between PepO and receptor depicted in yellow and blue, respectively. **B** Dimplot 2D-interaction plot depicts the creation of hydrogen bonds between PepO residues (Tyr⁴, Gln⁵, Ala¹¹, Ile¹², Asn¹³, Glu¹⁵, Thr¹⁹, and Glu²¹) and TLR4 residues (Arg⁶⁷, Arg⁸⁷, Gln⁹¹, Ser¹⁸⁴, and Val²⁶⁶). PepO residues, TLR4 residues, hydrogen bonds, and non-bonded residues are shown in green, blue, green dashed lines, and red/pink eyelashes, respectively

Table 1 Interactions of the N-PepO with TLR2 and TLR4. Investigation of interactions was performed through DIMPLOT program from LigPlot⁺v2.2.4

| N-PepO-TLR2 | | | | | N-PepO-TLR4 | | | | |
|-----------------|-----|---------------|-----|-----------------|-----------------|----|---------------|-----|-----------------|
| N-PepO residues | | TLR2 residues | | H-bond distance | N-PepO residues | | TLR4 residues | | H-bond distance |
| Name | No | Name | No | | Name | No | Name | No | |
| Arg | 354 | Glu | 264 | 2.74 | Thr | 19 | Arg | 67 | 2.79 |
| Arg | 354 | Glu | 264 | 2.71 | Glu | 21 | Arg | 67 | 2.84 |
| Lys | 357 | Glu | 264 | 2.75 | Ala | 11 | Arg | 87 | 2.77 |
| Asp | 34 | Arg | 321 | 2.73 | Ile | 12 | Arg | 87 | 3.06 |
| Asp | 34 | Arg | 321 | 2.69 | Glu | 15 | Arg | 87 | 2.78 |
| Phe | 8 | Tyr | 323 | 2.79 | Asn | 13 | Gln | 91 | 2.93 |
| Tyr | 9 | Tyr | 323 | 2.68 | Tyr | 4 | Ser | 184 | 2.97 |
| Asn | 13 | Tyr | 323 | 3.09 | Tyr | 4 | Lys | 186 | 2.59 |
| Gln | 5 | Leu | 328 | 2.95 | Gln | 5 | Lys | 186 | 2.59 |
| Arg | 3 | Tyr | 332 | 2.63 | Gln | 5 | Glu | 266 | 2.84 |
| Gln | 5 | Tyr | 332 | 2.75 | | | | | |
| Asp | 6 | Tyr | 332 | 2.60 | | | | | |
| Glu | 15 | Lys | 347 | 2.51 | | | | | |
| Arg | 3 | Leu | 350 | 2.66 | | | | | |

(Fig. 5a). Furthermore, the Z-score (Fig. 5b), and ERRAT score (Fig. 5c) were improved to -6.96 and 96.11 in the refined model, respectively (Table 4). The 3-dimensional structure of the refined model was visualized via Discovery Studio Viewer (Fig. 6).

Discontinuous B-cell epitope prediction of the engineered construct

Given the significance of conformational epitopes in immunization, discontinuous B cell epitopes of the 3D structure of engineered construct were predicted based on the interaction of protein and antibody using the Ellipro server. Conformational peptides with a score of >0.5 were chosen, and the scores showed the surface atoms of the construct that are responsible for antibody binding. Table 5 summarizes the number of amino acids, the amino acid compositions, the score values, and the sequence location.

Molecular docking of TLR2/4 and ODAC vaccine

As mentioned above, the first 112 residues of the N-PepO may act as a potential binding site for TLR2/4. The construct of the vaccine was engineered in an effective way to interact with the TLR2 and TLR4 receptors. The process of protein–protein docking between TLR2/4 receptor and vaccine was performed by Cluspro 2.0 online server after eliminating the ligand and water molecules from the crystallographic structure of receptor. Several complexes were generated by the docking, and it was observed that the engineered vaccine could

interact well with the TLR2/TLR4. The model.000.00 and model.000.01 were chosen for ODAC-TLR2 and ODAC-TLR4 complexes with the lowest energy docking mode -922.1 and -798.3, respectively. For displaying and evaluating the considered complexes, the DS Visualizer and DIMPLOT tools from LigPlot + v2.2.4 were used (Figs. 7 and 8). According to a detailed review of the results, the designed vaccine and TLR2/4 interacted properly via the truncated region of N-PepO. The engineered construct-TLR2 formed 15 hydrogen bonds and 3 salt bridges, Met¹-Pro³²⁰, Trp⁵⁰-Arg³²¹, Thr²-Arg³²¹, Trp⁵⁰-Tyr³²³, Asp⁶-Lys³⁴⁷, Glu⁴⁰-Glu³⁷⁴, Tyr⁹-Gln³⁹⁶, Asp¹⁰-Gln³⁹⁶, Asp³⁴-Lys⁴⁰⁴, Gln³⁷-Lys⁴⁰⁴, Asp³⁶-Ser⁴²⁴, Asn¹³-Arg⁴⁴⁷, and Glu⁵⁸-Arg³²¹. Analysis of ODAC-TLR2 intermolecular interaction revealed that only Glu40-Glu374 hydrogen bond formed at distance above 3 Å (Table 5). Also, the designed vaccine-TLR4 made 12 hydrogen bonds and 4 salt bridges, Thr²-Asp⁹⁵, Arg³-Leu¹¹⁹, Asp⁵⁹-Asn¹⁴³, Arg⁷⁰-Glu¹⁶⁹, Trp¹⁶-Asn¹⁷³, His³⁶⁸-His¹⁹⁹, Arg⁷⁰-Gln²⁰⁰, His³⁶⁸-Gln²⁰⁰, His³⁶⁹-Glu²²⁵, His³⁷²-Arg²²⁷, Asp¹⁰-Lys¹⁵⁰, and Glu¹⁵-Lys¹⁵⁰. Investigation of ODAC-TLR4 intermolecular interaction exhibited that all the hydrogen bonds formed at distances between 2.59 and 2.99 Å (Table 6). Both chosen complexes were subjected as the initial structures in MD simulation processes.

MD Simulations of selected complexes

The chosen docked structures of the designed vaccine-TLR2/4 were subjected to simulations of MD to calculate the level of their structural stability. On the basis of the

Table 2 Details of the chosen domains of PhtD, PsaA and PspC for designing the final construct

| Protein | PhtD-C | PsaA | PspC |
|-------------------------------|---|--|--|
| Final selected domain | (Residues 196–256) AAQAYAKEKGLTPP-STDHQDSGNTAKGAEAIYNRVKAACKVPL-DRMPYNLQYTVKNGS | (Residues 100–187) SDGVDVIYLEGQNEK-GKEDPHAWLNLENGIFAKNIAKQLSAK-DPNNKEFYEKNLKEYTDKLDKDKESKDK-FNKIPAEEKLIVTSEG | (Residues 276–363) RRNYPSN-TYFSLELEISESDVEVKKAEEFELV-KEEAKPRNEEKVKQAKAKVESKKA-VATRLENIKTDRKKAEEEEAKRKAEEED-KVKEK |
| Final B-cell epitopes | 224–238 225–239 226–240 234–248 237–251 238–252 239–253 240–255 241–256 | 103–117 104–118 105–119 106–120 107–121 108–122 109–123 110–124 113–127 133–147 141–155 148–162 | 285–299 288–302 294–308 295–309 324–339 327–341 |
| Final MHC-II Epitopes (Mouse) | 189–203 190–204 217–231 218–232 223–237 224–238 225–239 226–240 227–241 | 174–188 175–199 176–190 | 288–302 289–303 290–304 298–312 299–313 300–314 301–315 316–330 333–347 323–337 322–336 321–335 324–338 320–334 |
| Final MHC-II epitopes (Human) | 222–236 223–237 224–238 225–239 227–241 228–242 240–254 241–255 | 103–117 104–118 126–140 167–181 168–182 169–183 | 303–317 334–348 335–349 336–350 337–351 333–347 |

GROMOS96 54A7 force field, we conducted a separate 50 ns MD simulation for each docked complex.

Investigation of stability

The RMS Deviation was applied for evaluating the vaccine-TLR2/-TLR4 complex stability. The backbone

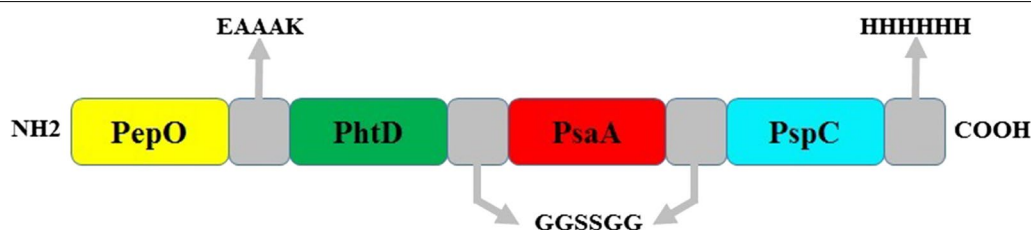


Fig. 4 Schematic diagram of the candidate construct. The epitope-rich domains of PhtD, PsaA, and PspC are shown in green, red, and turquoise blue, respectively, which are linked together with GGSSGG. The truncated fragment of PepO shown in yellow as a potential adjuvant candidate was joined to the N-terminal of the considered domains through the EAAAK linker. A hexa-histidine (His6) tag was fused at the C-terminal of the candidate vaccine to facilitate purification

Table 3 Analysis of sequence characteristics of the ODAC construct. Physicochemical parameters, antigenicity, solubility and toxicity of the sequence were computed through ExPASy's ProtParam, ANTIGENpro/Vaxijen, SOLpro and Toxin Pred, respectively. Also, the allergenicity of the desired sequence was evaluated using AllerTOP and AlgPred

| ProtParam tool | | Vaxi Jen | | | | | | | | | | ToxinPred | Aller TOP | Alg Pred | Scratch protein predictor | | |
|-----------------|-------|----------|-----|------|--------------|--------|--------|-------|-------|--------|-------|-----------|--------------|--------------|---------------------------|------|----------------|
| Sequence Length | M.wt | + R | - R | pI | Half-life in | | | II | AI | GRAVY | | | | | | | |
| | | | | | Mammalian | | Yeast | | | | | | | | | | <i>E. coli</i> |
| | | | | | | | | | | | | | | | | | |
| 372 | 41.93 | 60 | 75 | 5.29 | 30 h | > 20 h | > 10 h | 36.83 | 68.01 | -0.954 | 0.748 | Non toxin | Non allergen | Non allergen | 0.89 | 0.92 | |

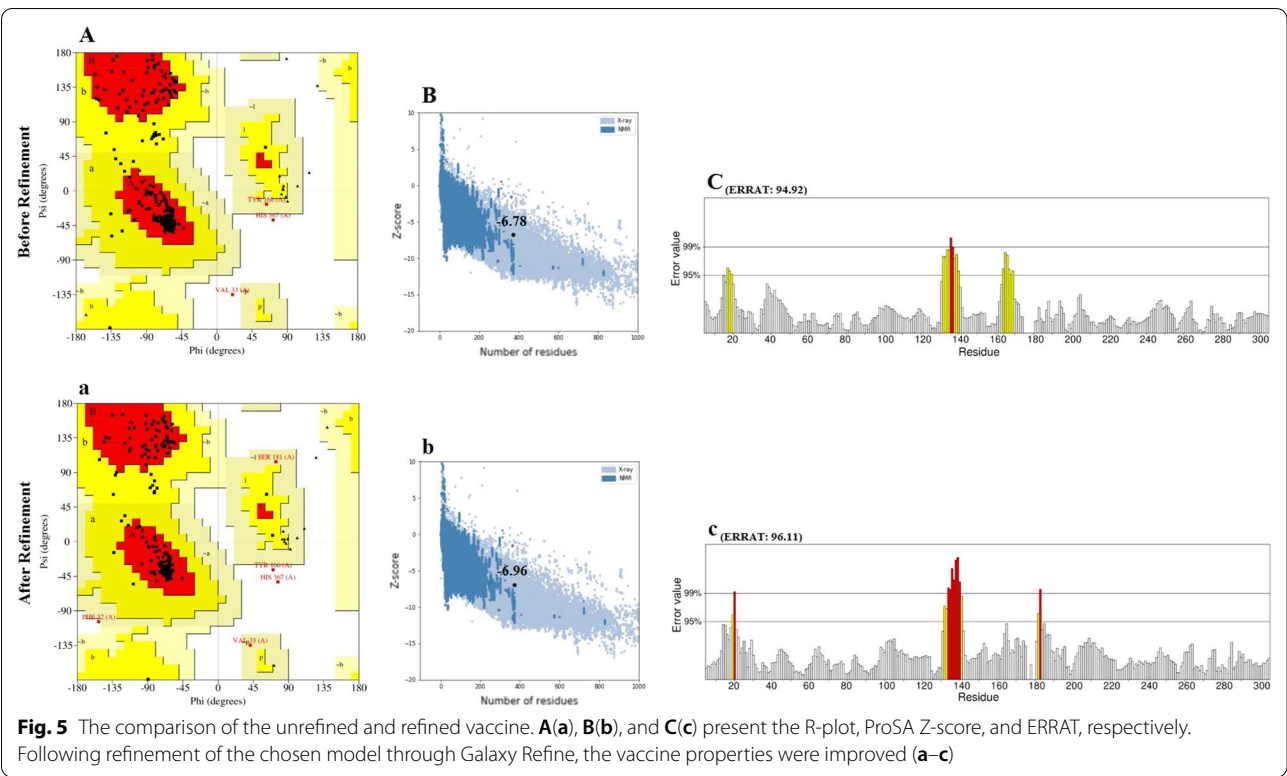


Table 4 Evaluation of the modeled vaccine before and after refinement. Analysis of the models was performed by the PROCHECK Ramachandran plot, ERRAT and ProSA-web Z-score

| Step | Ramachandran plot | | | ERRAT | Z-score |
|-------------------|--------------------------|---------------------|------------------------|-------|---------|
| | Most favored regions (%) | Allowed regions (%) | Disallowed regions (%) | | |
| Before refinement | 92.5 | 6.9 | 0.6 | 94.92 | − 6.78 |
| After refinement | 94.00 | 5.1 | 0.9 | 96.11 | − 6.96 |

deviations of chosen complexes of the initial conformations were plotted as a function of time. The ODAC-TLR2 and ODAC-TLR4 showed RMSD values in the range of 0.6–0.8 nm after 20 ns and 0.57–0.72 nm after 15 ns, respectively, as they are displayed in Fig. 9.

Analysis of flexibility

The RMS fluctuation is another way to anticipate the dynamic stability (DS) of the docked complex that evaluates fluctuations of the residues. The RMSF values showed that PepO residues in ODAC had almost the same degree of flexibility as TLR2 amino acids, while the rest of ODAC residues had a greater degree of flexibility than TLR2 residues (Fig. 10A). Likewise, as shown in

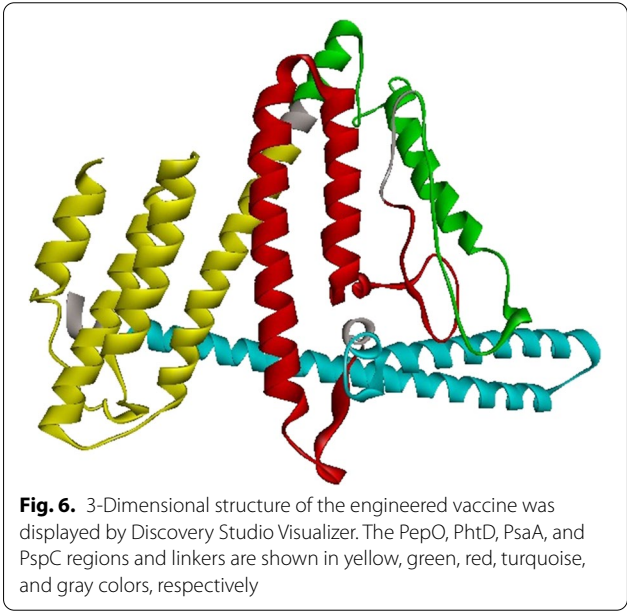


Fig. 10B, the RMSF value of residues in ODAC and TLR4 revealed no notable flexibility.

The simulation of the immune system

The outcomes of the immune simulations provided by the C-ImmSim online server were consistent with strong

Table 5 Prediction of discontinuous B-cell epitopes of ODAC vaccine. The conformational epitopes were predicted using ElliPro

| Epitopes | Residues | Number of residues | Score |
|----------|--|--------------------|-------|
| 1 | S296, E299, V300, K301, K302, A303, E304, F305, E306, L307, V308, K309, E310, E311, A312, K313, E314, P315, R316, N317, E318, E319, K320, V321, K322, Q323, A324, K325, A326, K327, V328, E329, S330, K331, K332, A333, V334, A335, T336, R337, N340, I341, D344, K347 | 44 | 0.819 |
| 2 | M1, T2, R3, Y4, Q5, D6, D7, Y9, D10, A11, I12, N13, G14, E15, W16, Q17, Q18, T19, A20, E21, I22, P23, A24, D25, K26, S27, Q28, T29, G30, G31, F32, V33, D34, P57, E58, H69, R70, V72, R73, D74, F75, D76, K77, R78, E79, A80, D81, G82, I83, T84, P85, V86, A357, E358, E359, K361, V362, K363, E364, K365, G366, H367, H368, H369, H370, H371, H372 | 67 | 0.724 |
| 3 | L51, A52, G53, E54, A114, A115, A116, K117, A118, A119, Q120, A121, Y122, A123, K124, E125, K126, G127, L128, T129, P130, P131, S132, T133, D134, H135, Q136, D137, S138, G139, N140, T141, E142, A143, K144, G145, E147, V174, K175, N176, G177, S178, G179, G180, S181, S182, S225, A226, K227, D228, P229, N230, N231, K232, E233, F234, E236, K237 | 58 | 0.715 |
| 4 | V159, P160, D162, R163 | 4 | 0.595 |
| 5 | D255, N258, K259 | 3 | 0.591 |
| 6 | I260, P261, A262, E263, K264, K265, P282, S283 | 8 | 0.579 |

immune reactions as shown in Fig. 11. It was found that the developed vaccine construct could able to elicit the primary response and powerful secondary and tertiary responses. The primary response is illustrated by the raised levels of IgM antibodies after a 5–7-day delay of antigen exposure. The secondary and tertiary responses are characterized by the increase in the population of B-cells and expression of immunoglobulins (IgM, IgG1 + IgG2, and IgG + IgM), resulting in a subsequent decline in the concentration of antigen (Fig. 11 A, B). The suggested vaccine, in addition to inducing high proliferation of B cells, also leads to memory B-cell generation. Likewise, the outcomes demonstrated that the construct could increase the population of T helper cells with memory development (Fig. 11 C). As depicted in Fig. 11 there is also an increase in the IFN- γ concentration following each injection. These results clearly display that the designed vaccine could efficiently elicit immune responses and provide the basis for immunity against pneumococcal infection.

Reverse translation, codon adaptation, and in silico cloning

JCat tool was employed for reverse translation and codon adaptation of the devised vaccine candidate in order to maximal expression of it in *E. coli* K-12. The vaccine sequence after the reverse translation was 1116 nucleotides long. According to the codon optimization findings,

the CAI value and CG content were 0.99 and 46.41%, respectively, which are accepted since they are within the permitted limit. These outcomes confirm a proper expression of the engineered vaccine in *E. coli* K-12. Ultimately, the ODAC vaccine sequence was successfully cloned into the pET-28a(+) between *Nde*I and *Xho*I enzymes by SnapGene software free-trial (Fig. 12).

Discussion

Streptococcus pneumoniae is a powerful pathogen owing to its ability to adhere to cells, invasion to host tissues, and escape the immune system of the host [82], and can cause fatal diseases such as sepsis, pneumonia, and meningitis with a significant death rate worldwide [83]. The currently existing pneumococcal vaccines have several limitations including serotype dependence [7, 84] and so they are not capable of protecting subjects against all Pnc serotypes. A hopeful alternative can be the utilization of proteins and peptide fragments, some of which have been investigated as vaccine candidates in pre-clinical trials, either alone or in combination with each other. Pneumococcal proteins commonly expressed across all serotypes have been regarded as a potential candidate for developing serotype-independent Pnc vaccines [7]. In many biological subjects, in silico techniques could be highly useful to decrease the cost and time of experimental researches [37, 85, 86]. Epitope identification by

(See figure on next page.)

Fig. 7 The outcome of docking of the vaccine construct and TLR2 receptor. **A** The three-dimensional structure illustrates the interaction of the construct and TLR2 that are shown in colored parts and harbor blue, respectively. TLR2 and vaccine interacted with each other via the truncated N-PepO (yellow part). **B** The Dimplot 2D-interaction plot demonstrates the formation of hydrogen bonds between the construct residues (Met¹, Thr², Asp⁶, Tyr⁹, Asp¹⁰, Asn¹³, Asp³⁴, Asp³⁶, Gln³⁷, Glu⁴⁰, Trp⁵⁰, and Glu⁵⁸) and TLR2 residues (Pro³²⁰, Arg³²¹, Tyr³²³, Lys³⁴⁷, Glu³⁷⁴, Gln³⁹⁶, Lys⁴⁰⁴, Ser⁴²⁴, and Arg⁴⁴⁷). PepO residues, TLR2 residues, hydrogen bonds, and non-bonded residues are indicated in green, blue, green dashed lines, and red/pink eyelashes, respectively

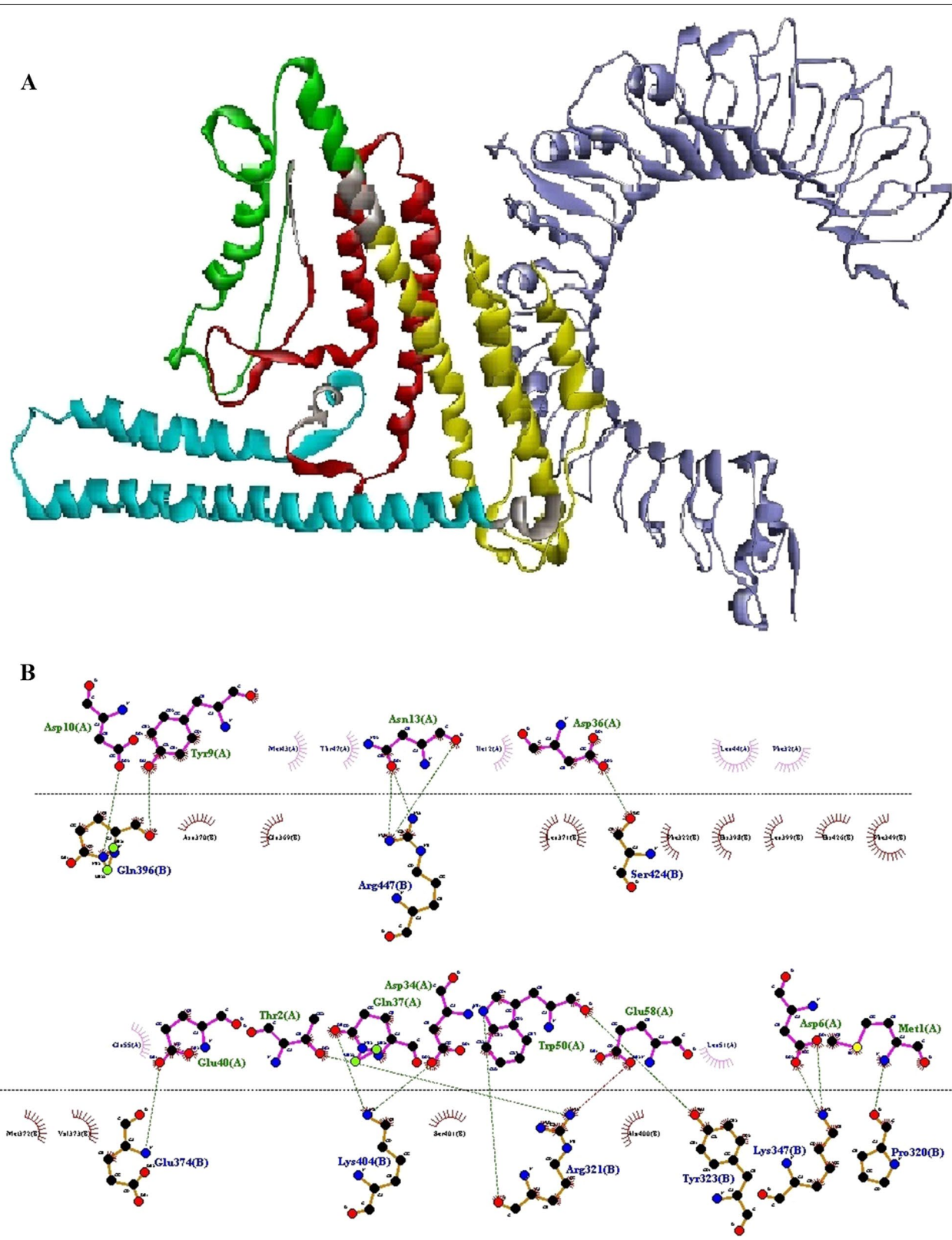


Fig. 7 (See legend on previous page.)

immunoinformatics tools could be quite helpful in the different applications within the field of epitope mapping, such as designing peptide-based vaccines, identifying immunological processes, predicting epitopes used in the disease diagnosis, and so on [87–90]. Peptide-based vaccines are one of the alternative and appealing approaches to develop vaccines, which only include peptide fragments with the highest antigenicity and the most ability to induce immune responses [91, 92]. In reality, the immune responses are only formed against immunogenic epitopes, and other parts of the protein virulence factor causing unwanted reactions can be eliminated [93]. The combining of several highly-conserved peptides can be employed to induce both native and acquired immune responses, and prevent various stages of Pnc infection, as well as present broader serotype coverage [7].

PspC protein which is almost found in all pneumococcal strains is one of the major and surface-exposed virulence factors [21]. A study showed that the survival time for mice vaccinated with a combination of PspC and a genetic toxoid derivative of pneumolysin (PdB) was considerably longer than that for mice treated only with PdB, but was not remarkably different from that for mice immunized only with PspC [94]. In another study in 2014, the PspC peptides which participate in adherence and invasion were genetically fused to L460D pneumolysoid protein. The fusion construct was more immunogenic than L460D pneumolysoid alone, and anti-fusion antibodies were active against pneumococcus. Immunization of mice with the fusion construct protected them from the Pnc carriage, pneumonia, meningitis, bacteremia, and sepsis [95]. PhtD and PhtE antigens, two members of the conserved Pht family, have been found to be surface-exposed and expressed almost in all pneumococcal strains [96]. It has been exhibited that both antigens induce efficient protection against colonization, pneumonia and sepsis in animal models [97]. Godfroid et al. indicated that mice vaccinated with this family were protected more effective against pneumonia compared to those immunized with other Pnc vaccine candidates such as PsaA and PspA [96]. In the research of Verhoeven et al., the efficacy of a multivalent vaccine containing PhtD, choline-binding protein A (PcpA), and detoxified pneumolysin (PlyD1) was evaluated and it was shown that the vaccine protected the infant mouse model against serotypes 3 and 6A pneumococcus [98]. Plumtre

et al., demonstrated the truncated derivatives of PhtD as vaccine antigens in animal models were more effective than the whole PhtD protein [12]. Pneumococcal surface adhesion A is another conserved and immunogenic antigen that is expressed by all Pnc strains [99]. Lin et al. reported that mice vaccinated with the 23-valent capsular polysaccharide (CPS)–rPsaA conjugate showed rapid pneumococcal clearance from blood and could provide the best protection against Pnc challenge [100]. Lu et al., prepared a fusion conjugate consisting of a fusion protein of pneumolysin nontoxic derivative (PdT) with PsaA coupled to cell wall polysaccharide (CWPS). The fusion conjugate protected mice against colonization, while mice vaccinated with the mixture of the three antigens were not protected [101]. In the study of Lu et al., in 2015, the fusion protein PspA-PsaA was assessed in an animal model against fatal challenges with pneumococci. The results showed that levels of antibodies against both proteins increased in mice vaccinated by the fusion protein. Also following fatal challenge, immunized mice revealed decreased levels of pneumococci in the lungs and blood compared with the control group and were protected against pneumococcal challenge [32]. Another study showed that PsaA antigen prevented pneumococcal carriage more than PspA [102].

The principal drawback of epitope vaccines is that they could be degraded quickly by proteinase in the body, making them hard to be recognized by immune cell receptors [103, 104]. One of the approaches recommended for enhancing the immune responses elicited by epitope vaccines is to utilize adjuvants in the vaccine constructs [104]. In recent decades, TLR agonists have been broadly examined as candidate adjuvants for vaccine constructs [38]. There are various TLR agonists with different sources which a group of them are bacterial proteins capable of inducing TLR signaling and native immune responses [39]. Pneumococcal endopeptidase O is one of the bacterial TLR agonists that can be considered as a potential candidate adjuvant [39]. Shu and colleagues identified the Pnc endopeptidase O to be a TLR2/TLR4 bi-ligand. To distinguish the binding site of TLRs in PepO, they prepared two recombinant truncated forms of it, PepO 1–430 and PepO 431–630, according to two conserved domains of PepO named M13_N domain and M13 domain, respectively. Their result revealed that TLR2 and 4 bind to the M13_N domain of this protein

(See figure on next page.)

Fig. 8 Molecular docking of the vaccine with TLR4 receptor. **A** The complex shows the interaction between the engineered construct and TLR4 that are represented in colored parts and blue, respectively. TLR4 and ODAC interacted with each other by the truncated N-PepO (yellow part). **B** The Dimplot 2D-interaction plot shows the creation of hydrogen bonds between the designed construct residues (Thr², Arg³, Asp¹⁰, Trp¹⁶, Asp⁵⁹, Arg⁷⁰, His³⁶⁸, His³⁶⁹, and His³⁷²) and TLR4 residues (Asp⁹⁵, Leu¹¹⁹, Asn¹⁴³, Lys¹⁵⁰, Glu¹⁶⁹, Asn¹⁷³, His¹⁹⁹, Gln²⁰⁰, Glu²²⁵, Arg²²⁷). TLR4 residues, construct residues, hydrogen bonds, and non-bonded residues are depicted in blue, green, green dashed lines, and red/pink eyelashes, respectively

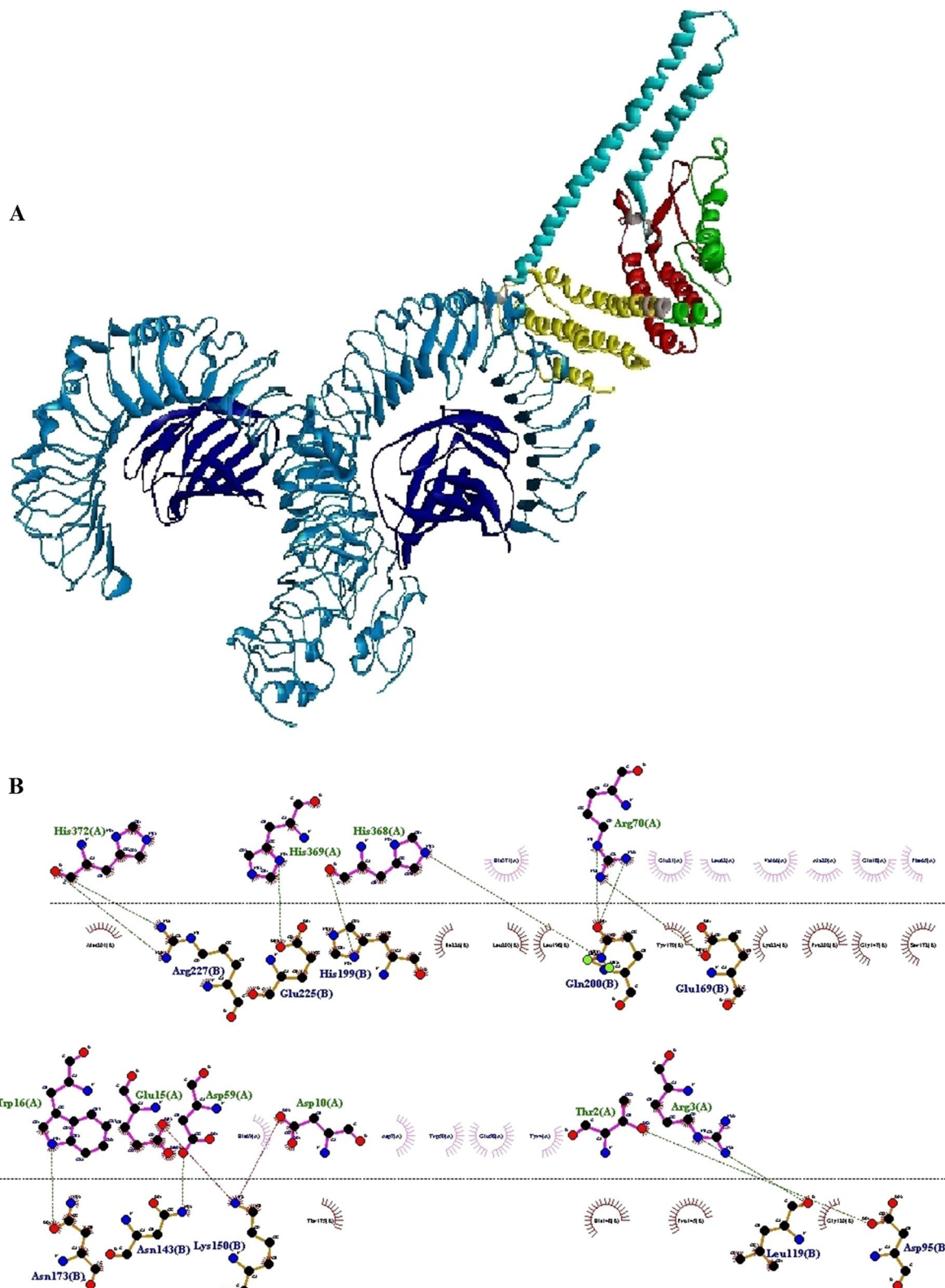


Fig. 8 (See legend on previous page.)

Table 6 Interactions of the engineered vaccine with TLR2 and TLR4. Analysis of interactions was performed by DIMPLOT program from LigPlot + v2.2.4

| ODAC -TLR2 | | | | | ODAC -TLR4 | | | | |
|---------------|----|---------------|-----|-----------------|---------------|-----|---------------|-----|-----------------|
| ODAC residues | | TLR2 residues | | H-bond distance | ODAC residues | | TLR4 residues | | H-bond distance |
| Name | No | Name | No | | Name | No | Name | No | |
| Met | 1 | Pro | 320 | 2.88 | Thr | 2 | Asp | 95 | 2.90 |
| Trp | 50 | Arg | 321 | 2.97 | Arg | 3 | Leu | 119 | 2.83 |
| Thr | 2 | Arg | 321 | 2.61 | Asp | 59 | Asn | 143 | 2.91 |
| Trp | 50 | Tyr | 323 | 2.81 | Arg | 70 | Glu | 169 | 2.78 |
| Asp | 6 | Lys | 347 | 2.64 | Trp | 16 | Asn | 173 | 2.86 |
| Asp | 6 | Lys | 347 | 2.54 | Arg | 70 | Gln | 200 | 2.77 |
| Glu | 40 | Glu | 374 | 3.08 | Arg | 70 | Gln | 200 | 2.62 |
| Tyr | 9 | Gln | 396 | 2.74 | His | 368 | Gln | 200 | 2.99 |
| Asp | 10 | Gln | 396 | 2.75 | His | 369 | Glu | 225 | 2.78 |
| Asp | 34 | Lys | 404 | 2.75 | His | 372 | Arg | 227 | 2.84 |
| Gln | 37 | Lys | 404 | 2.77 | Asp | 10 | Lys | 150 | 2.64 |
| Asp | 36 | Ser | 424 | 2.82 | Glu | 15 | Lys | 150 | 2.71 |
| Asn | 13 | Arg | 447 | 2.69 | | | | | |
| Asn | 13 | Arg | 447 | 2.71 | | | | | |
| Glu | 58 | Arg | 321 | 2.69 | | | | | |

[41]. In this regard, in the present study, to identify the main residues potentially involved in the interaction between the N-PepO and TLRs, we employed molecular docking using ClusPro 2.0. Following docking, the best complexes were chosen with the lowest energy docking mode -859.4 (for N-PepO-TLR2) and -967.1 (for N-PepO-TLR4). The DIMPLOT from LigPlot + v2.2.4 was used to analyze the chosen complexes (see Figs. 2 and 3). The N-PepO residues Asp³⁴, Arg³, Gln⁵, Asp⁶, Phe⁸, Tyr⁹, Asn¹³, Glu¹⁵, Arg³⁵⁴, and Lys³⁵⁷, and likewise, the N-PepO amino acids Tyr⁴, Gln⁵, Ala¹¹, Ile¹², Asn¹³, Glu¹⁵, Thr¹⁹, and Glu²¹ were predicted to play roles in the interactions with TLR2 and TLR4, respectively (see Table 1). Ultimately, to maintain the domain structure, a truncated fragment from N-PepO (residues 1–112) was selected as a potential TLR2/4 agonist candidate in this research.

The strategies applied in this investigation were the use of derivatives of three important virulence proteins PspC, PhtD, and PsaA into the vaccine construct which could help at preventing infection by pneumococci, and the identification of epitope-rich domains that could provoke both robust B- and T-cell immune responses. An efficient epitope vaccine should be engineered to include B and T cell epitopes capable of providing effective reactions to a specific infection [105]. Since the recognition of putative B and T cell epitopes using the wet-lab approach requires experimental screening of a large number of active epitopes against inactive epitopes, the computational

approach can act as a cost-effective, rapid, reliable, and accurate alternative [106, 107]. In this regard, the use of a consensus prediction strategy is more reliable and robust providing better results than any of the individual prediction methods [108]. Hence in this study, the B- and helper T-cell epitopes were predicted through various reliable databases and servers including ABCpred, LBtope, IEDB, Ellipro, DiscoTope, and NetMHCIIpan servers to achieve more reliable results. Finally, three epitope-rich regions containing high-ranked and shared epitopes, i.e. residues 196–256, 100–187, and 276–363 from PhtD-C, PsaA, and PspC, respectively, were chosen (Table 2). According to IEBD database, the selected region of PspC was consistent with two significant epitopes EDRRNYPST and EAKEPRNEEKVKQAK experimentally reported as B cell epitopes by Vadesilho et al. The outcomes of recent experimental study could support our predictions. The physical–chemical features and 3-D structure of the epitope-based vaccines correlate with the nature of the epitopes, linkers, and adjuvants, as well as their place and order in the vaccine construct [109]. Herein, the chosen regions of PhtD, PsaA, and PspC were joined together in an appropriate pattern using suitable linkers. In a similar manner to Tian et al., the GGSSGG flexible linkers were applied to keep the independent folding of domains and their immunological activities [63]. In addition, the EAAAK helical linker was employed for connecting the truncated N-PepO, as an adjuvant candidate, to the N-terminal of the above-quoted sequence to boost the

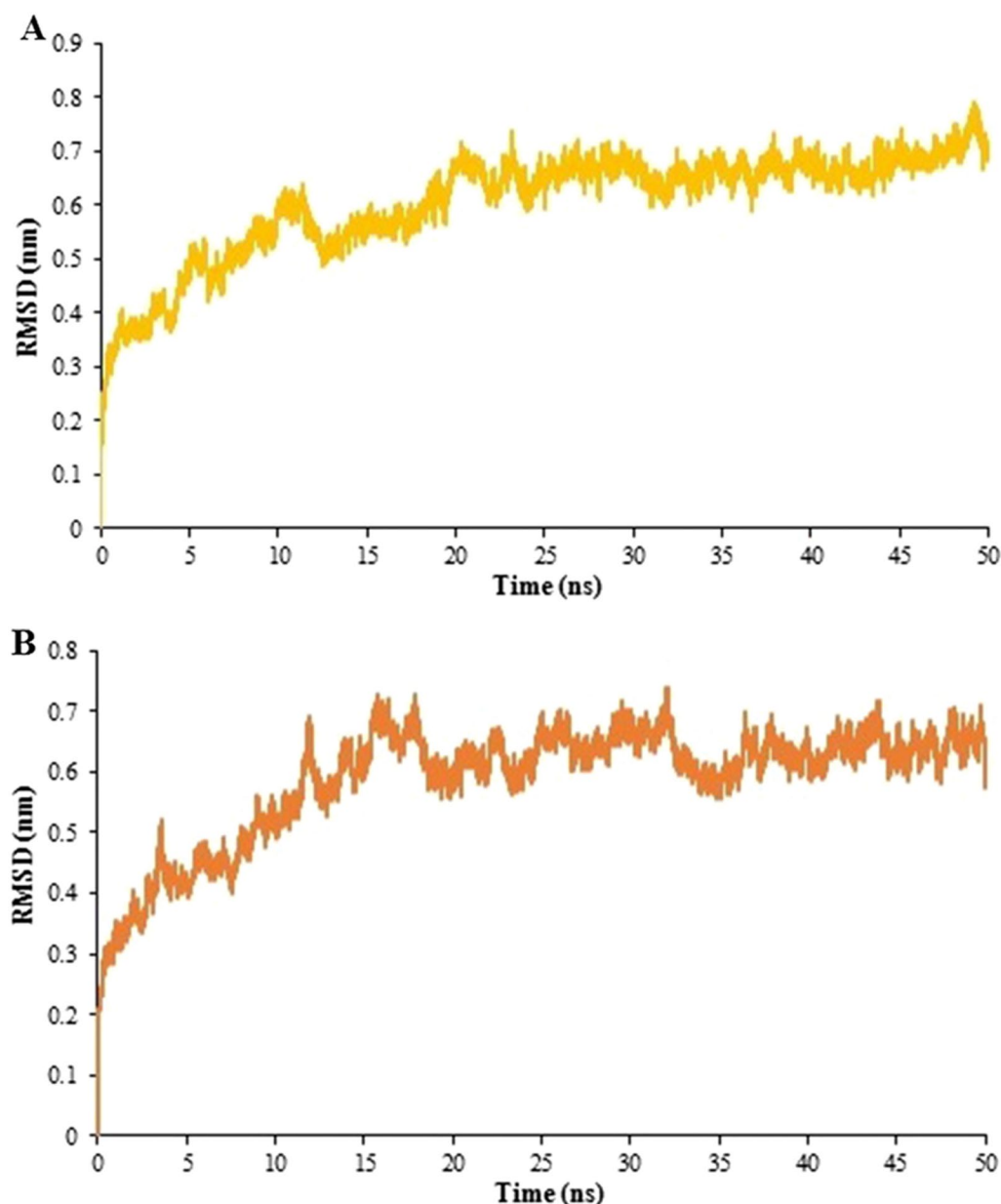


Fig. 9 RMSD plots for 50 ns simulation runs for vaccine-receptor complexes. The changes in the RMSD values were in the range of 0.6–0.8 nm after 20 ns for (A) ODAC-TLR2 and 0.57–0.72 nm after 15 ns for (B) ODAC-TLR4. The results showed that the binding of the vaccine to both receptor is stable

stability and immunogenicity of the construct [64, 65]. None of the selected three domains (196–256 of PhtD-C, 100–187 of PsaA, and 276–363 of PspC) has been used before and also the truncated fragment of PepO (residues 1 to 112) was not introduced earlier as a TLR agonist adjuvant candidate, showing the novelty of our study. The final engineered vaccine construct (ODAC) was assessed for toxicity, antigenicity, and allergenicity. It was antigenic, non-toxic, and non-allergic, which indicated its

effectiveness in evoking sturdy immune responses without no harmful reactions. According to SOLpro results, the solubility of vaccine upon overexpression in *E. coli* host, which is one of the fundamental requirements of most functional and biochemical investigations, was 0.893 [110, 111]. The theoretical pI of the ODAC vaccine was 5.29 which indicates the ODAC is basic in nature. The molecular weight of the construct was 41.93 kDa, which is ideal because the purification of protein constructs

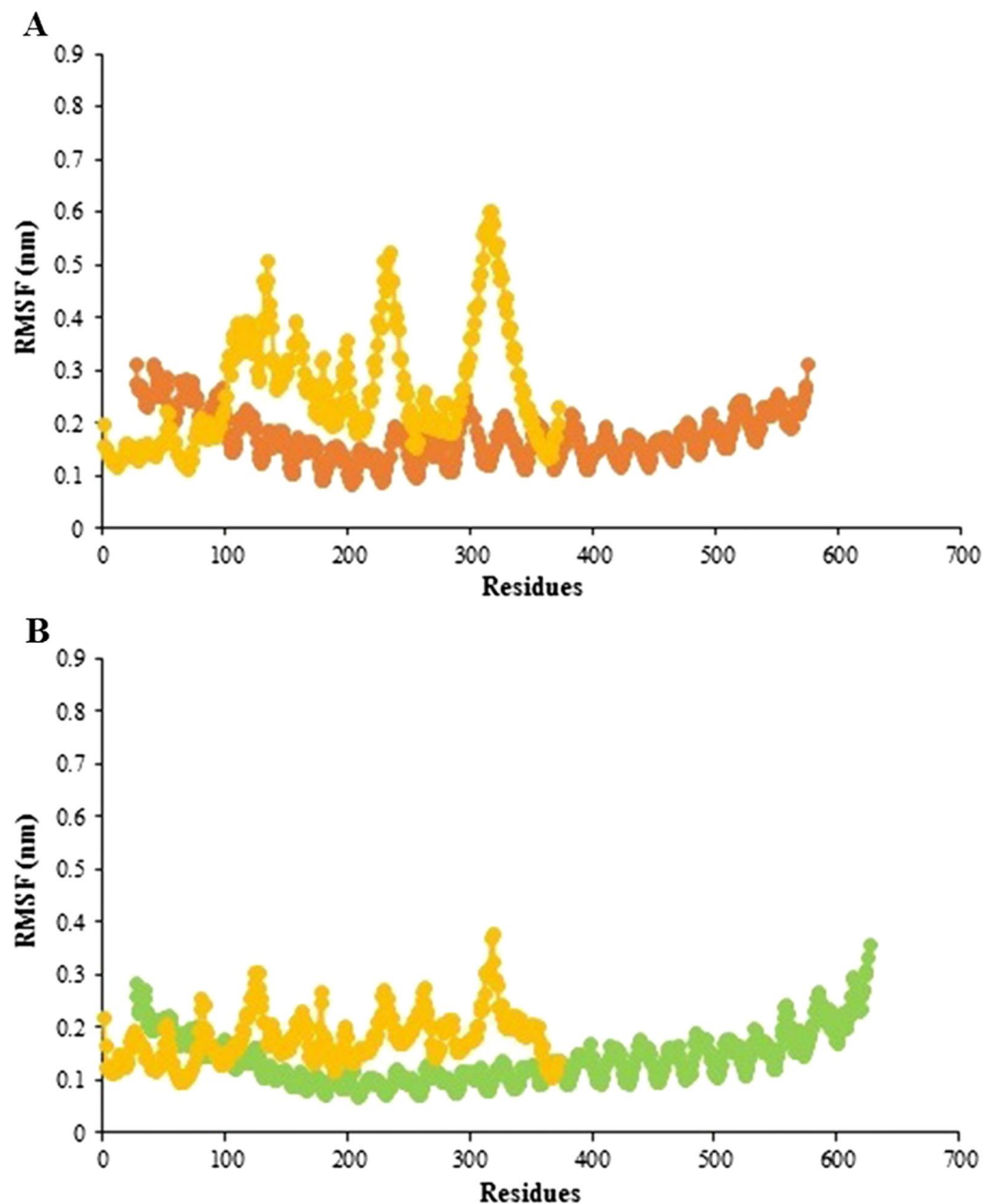
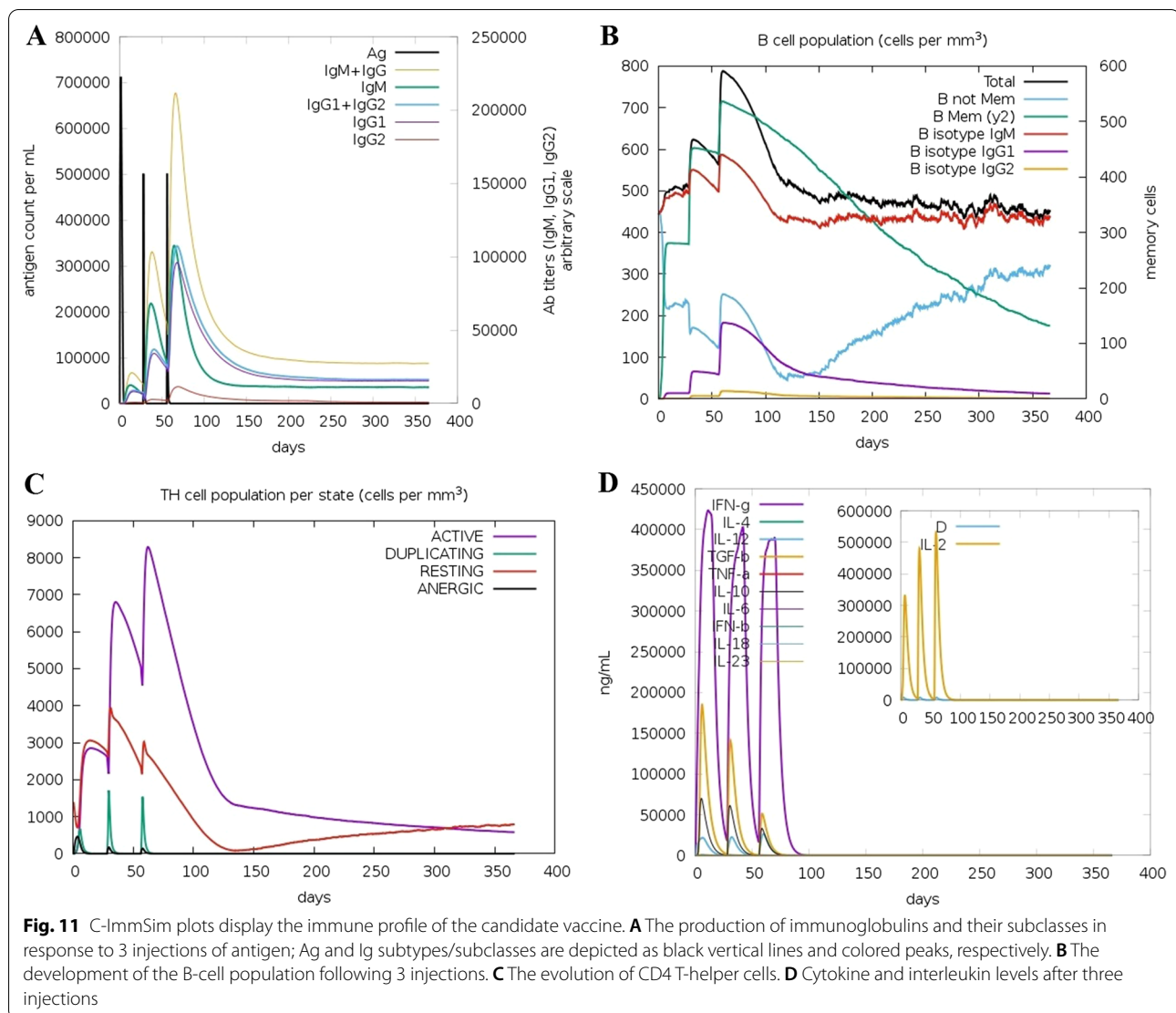


Fig. 10 RMS fluctuation values as a function of residue number for the vaccine-receptor complexes. **A** RMSF plots for the vaccine and TLR2 are depicted in yellow and orange, respectively. The plots reveal that the vaccine residues except for 112 residues of PepO have a greater degree of flexibility than that of TLR2. **B** RMSF plots for the vaccine and TLR4 are represented in yellow and green, respectively. The ODAC and TLR4 residues show no significant flexibility

with Mws fewer than 110 kDa is faster and easier [104, 112]. The half-life ($T_{1/2}$) of the engineered vaccine was computed to be 30 h, > 20 h, and more than 10 h, in mammalian reticulocytes, yeast, and *E. coli*, respectively. The half-life is the time which takes for 50% of the protein amount inside a cell to eliminate following its synthesis in the cell. ProtParam uses the "N-end rule" which links

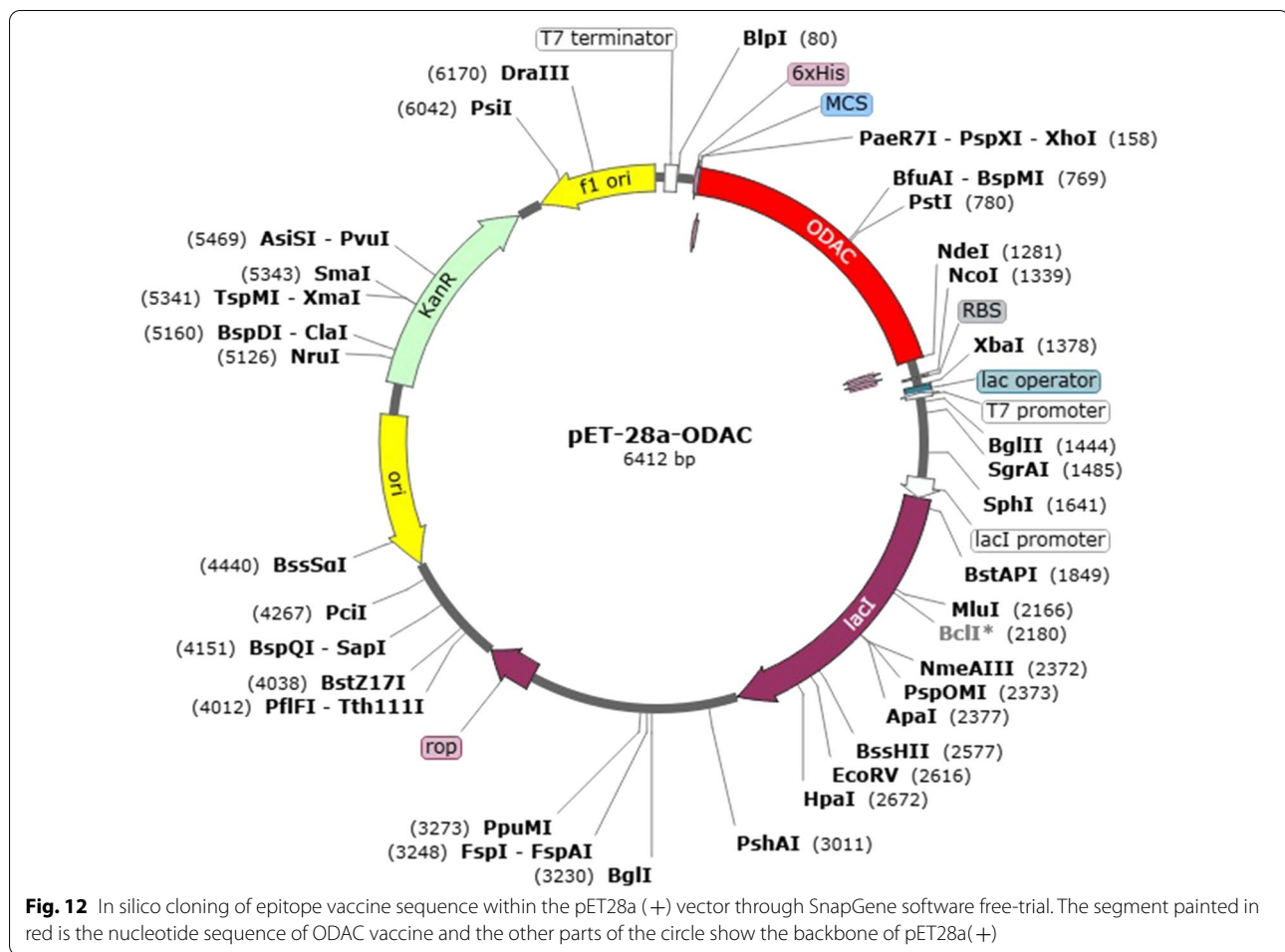
the protein's half-life to the identity of its amino-terminal residues that applies to both prokaryotic and eukaryotic organisms [113]. The aliphatic index of the designed vaccine was computed to be 68.01, implying that it could be a thermostable vaccine [114]. The GRAVY index of ODAC vaccine was -0.954, and the negative value of this parameter demonstrates that the construct is hydrophilic



and could well interact with molecules of water [115]. The instability index is assessed to estimate whether a vaccine is (un)stable, if this index of a protein is below (above) 40 then it's called stable (unstable) [72, 115]. The instability index of the ODAC construct was computed to be 36.83, and as the value is < 40 , the engineered vaccine is deemed stable (see Table 3).

The initial tertiary structure of the engineered vaccine was modeled via the Robetta server. Following this step, the refining process was utilized to promote the quality of the 3D structure, bringing it closer to the native structure. The qualities of the unrefined and refined models were compared through the R-plot, ProSA Z-score, and ERRAT score. The results showed that all scores related to the structure of the refined model improved well (Fig. 5). The refined 3D structure was subjected to discontinuous B-cell epitopes identification and molecular

docking process. A proper vaccine ought to include B-cell epitopes, besides to T-cell epitopes, with the purpose of ensuring the induction of humoral immunity [116]. The conformational B-cell epitopes were predicted on the basis of the interaction of vaccine-antibody through the Ellipro server, and the epitopes with a score of > 0.5 were picked (see Table 4). The refined model had a plurality of discontinuous B-cell epitopes, demonstrating that the engineered vaccine is quite potent in evoking humoral immunity. The docking processes between the ODAC vaccine and TLRs 2 and 4 were performed by Cluspro 2.0. The best complex models were considered with the lowest energy docking mode -922.1 and -798.3 for ODAC-TLR2 and -TLR4, respectively. The DIMPLLOT was utilized for analyzing the chosen models (Figs. 7 and 8). The ODAC construct residues Met¹, Thr², Asp⁶, Tyr⁹, Asp¹⁰, Asn¹³, Asp³⁴, Asp³⁶, Gln³⁷, Glu⁴⁰,



Trp⁵⁰, and Glu⁵⁸, and likewise, the ODAC vaccine amino acids Thr², Arg³, Asp¹⁰, Glu¹⁵, Trp¹⁶, Asp⁵⁹, Arg⁷⁰, His³⁶⁸, His³⁶⁸, His³⁶⁹, and His³⁷² were predicted to perform roles in the interactions with TLR2 and TLR4, respectively (see Table 5). The outcomes of the molecular docking showed that the truncated N-PepO of designed vaccine interacted favorably with TLRs, indicating it could act as a possible adjuvant for TLR2/TLR4 activation. With the purpose of assessing the molecular stability of ODAC-TLR2/4 docked complexes, simulations of MD were conducted through GROMACS. The structural fluctuations of the ODAC-TLR2 and ODAC-TLR4 reached the range of 0.6–0.8 nm and 0.57–0.72 nm after 20 and 15 ns, respectively (Fig. 9). The outcomes revealed that the binding of the designed vaccine to the TLR receptors was stable during simulation times. The truncated region of N-PepO as the TLR agonist was capable to keep well these connections. RMS Fluctuation is the other tool for computing the dynamic stability of complexes. The findings of RMSF showed that the residues of ODAC, except the truncated N-PepO residues, were more flexible than

those of TLR2, while the engineered vaccine and TLR4 amino acids showed no notable flexibility (Fig. 10). Conforming to the RMSF analyses, the truncated N-PepO had the lowest fluctuations in the areas having the most interactions with the TLR2/4 receptors. The results from the C-ImmSim server were consistent with actual immune responses previously reported in pneumococcal infections [117]. The Induction of B and T cell memory responses is deemed as one of the criteria for being an effective candidate vaccine [118]. It was found that populations of memory B and T cells were increased by each injection of ODAC and these levels were maintained after the third injection (Fig. 11), indicating our construct as a suitable one. Another observation of interest was that the level of IFN- γ was elevated following the first injection and remained at the peaks after repeated the injections. This shows a high concentration of T helper cells and thus the efficient production of immunoglobulin, which supports humoral responses [119]. To enhance the expression of the designed vaccine in *E. coli* K-12, codon adaptation was conducted through JCat tool. The

ODAC construct possessed the CAI value of 0.99 and GC content of 46.41%. Since CAI values > 0.8 and GC content > 30 and < 70% are considered to be desirable for expression [104, 115], our outcomes are satisfactory.

In numerous recent studies, the reliable computational procedures have been used to select efficient epitopes and design novel vaccines against various pathogens such as *Mycobacterium tuberculosis* [118], *Helicobacter Pylori* [42], *Onchocerca volvulus* [111], *Plasmodium* [120], and Crimean-Congo hemorrhagic fever virus [119]. In this respect, the present study recommends a final peptide construct as the best epitope vaccine based on the strategies used in the research to be an attractive intervention against the *Streptococcus pneumoniae* pathogen. Furthermore, antigenic epitope-rich portions from each of the proteins involved in the construct could also be employed in future studies to develop the other new sub-unit vaccines consisting of other virulence proteins. In addition, the new truncated domain of PepO could play as a new adjuvant candidate in other studies to increase the potency of vaccines against pneumococci or other pathogens.

Conclusions

Current pneumococcal vaccines, although, have been useful in decreasing the rate of invasive Pnc diseases, they have some issues. The main goal of this research was designing a novel epitope-based vaccine that could cover the limitations of existing Pnc vaccines. In this regard, the immunodominant epitope-rich regions from highly protective Pnc antigens, namely PhtD, PsaA, PspC, were predicted by immunoinformatics tools and joined together using proper linkers. Then, the truncated fragment of N-PepO (as the potential adjuvant) was linked to the engineered construct in order to enhance the vaccine's immunogenicity. The final designed construct was determined to be antigenic and non-allergenic, and its physical-chemical characteristics were satisfactory. Molecular docking was conducted for evaluating the binding affinity of the vaccine to TLR2/4 receptors. Simulations of MD were applied to confirm the stability of the considered docked complexes. Finally, the expression efficiency of the epitope-based vaccine was assessed. Although the outcomes of this research were quite remarkable, the potency of the engineered vaccine ought to be confirmed in laboratory and animal models.

Abbreviations

Pnc: Pneumococcal; PPV: Pneumococcal polysaccharide vaccines; PCV: Protein-conjugated polysaccharide vaccines; PspA: Pneumococcal surface protein A; PspC: Pneumococcal surface protein C; Pht: Pneumococcal histidine triad; ABC: ATP-binding cassette; Hic: H-binding inhibitor of complement; PbcA: C3-binding protein A; SpsA: *S. pneumoniae*-Secretory IgA-binding protein;

CbpA: Choline-binding protein A; N-ter: N-terminal; C-ter: C-terminal; Zn²⁺: Zinc; Mn: Manganese; TLR: Toll-like receptor; APCs: Antigen presenting cells; PepO: *S. pneumoniae* Endopeptidase O; PDB: Protein data bank; C-score: Confidence score; TM: Template modeling; RMSD: Root mean square deviation; DS: Discovery studio; His6: Hexa-histidine; JCat: Java codon adaptation tool; CAI: Codon adaptation index; pI: Isoelectric point; Mw: Molecular weight; GRAVY: Grand average of hydropathicity; + R: Total number of positive residues; - R: Total number of negative residues; PcpA: Choline-binding protein A; PlyD1: Detoxified pneumolysin; CWPS: Cell wall polysaccharide; PhtA: Pneumococcal histidine triad A; PhtB: Pneumococcal histidine triad B; PhtD: Pneumococcal histidine triad D; PhtE: Pneumococcal histidine triad E; PspC3: PspC from group 3; PhtD-C: C-terminal of PhtD; PsaA: Pneumococcal surface adhesion A; RMSF: Root Mean Square Fluctuation; RE: Restriction Enzymes; ODAC: The final designed vaccine in this study; AI: Aliphatic index; PdB: Toxoid derivative of pneumolysin; CPS-rPsaA: 23-valent capsular polysaccharide-rPsaA conjugate; PdT: Pneumolysin nontoxic derivative.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-022-03590-6>.

Additional file 1: Fig. S1. TMHMM result for PhtD-C (A), PsaA (B), and PspC (C). The TMHMM plots (A, B, and C) show no existence of transmembrane helices. **Fig. S2.** SignalP 4.0 web-server output for the sequences of PsaA and PspC. The C-, S-, and Y-scores represent the predicted cleavage site value, the projected signal peptide value, and a combination of C- and S-scores, respectively. (A) The anticipated cleavage site of PsaA is located between positions 19 and 20. (B) In PspC, there is not any cleavage site. **Fig. S3.** Three-dimensional structures of proteins PhtD-C (A) and PspC (B). The structures were predicted by I-TASSER. **Fig. S4.** Validation of the refined tertiary structures. A and B show the results of ProSA web for PhtD-C and PspC, respectively. The a and b display Ramachandran plots of PhtD-C and PspC, respectively. **Fig. S5.** Three-dimensional structure of TLR agonist N-PepO. The structure of N-PepO was predicted and refined by I-TASSER and GalaxyRefine, respectively. **Table S1.** The results of validation before and after refinement. Ramachandran plot statistics from PROCHECK and Z-score from ProSA web server. **Table S2.** Prediction of linear B-cell epitopes from PhtD-C by LBTope, ABCpred, IEDB Emini tool, and Ellipro. **Table S3.** Prediction of linear B-cell epitopes from PsaA by LBTope, ABCpred, Emini, and Ellipro. **Table S4.** Prediction of linear B-cell epitopes from PspC by LBTope, ABCpred, Emini, and Ellipro. **Table S5.** Prediction of conformational epitopes from PhtD-C via Ellipro and Discotope. **Table S6.** Prediction of conformational epitopes from PsaA via Ellipro and Discotope. **Table S7.** Prediction of conformational epitopes from PspC via Ellipro and Discotope. **Table S8.** Prediction of MHCII epitopes from PhtD-C by IEDB and NetMHCIIpan. **Table S9.** Prediction of MHCII epitopes from PsaA by IEDB and NetMHCIIpan. **Table S10.** Prediction of MHCII epitopes from PspC by IEDB and NetMHCIIpan.

Acknowledgements

We are grateful to Dr. Mohammad Reza Pourshafie, Professor of Microbiology, Pasteur Institute of Iran, for his valuable help in completing this work. This work was supported by a grant number of 1790 from the Semnan University of Medical Sciences, Semnan, Iran.

Author contributions

ZB: Conceptualizing, designing and study implementation, data collection and analysis, and writing the manuscript. MSH: Cooperation in data collection, writing and editing of the manuscript. HM and MMR: Consulting in immunoinformatics. AASH and SFM: Supervision and monitoring of implementation of the study. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Not applicable.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

Not applicable.

Author details

¹Department of Medical Biotechnology, Faculty of Medicine, Semnan University of Medical Sciences, Semnan, Iran. ²Research Center of Biotechnology, Semnan University of Medical Sciences, Semnan, Iran. ³Department of Bacteriology, Pasteur Institute of Iran, Tehran, Iran. ⁴Drug Design and Bioinformatics Unit, Department of Medical Biotechnology, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran. ⁵Agricultural Research, Education, and Extension Organization (AREEO), Razi Vaccine and Serum Research Institute, Karaj, Iran.

Received: 14 February 2022 Accepted: 14 August 2022

Published online: 04 September 2022

References

- Weiser JN, Ferreira DM, Paton JC. *Streptococcus pneumoniae*: transmission, colonization and invasion. *Nat Rev Microbiol*. 2018;16(6):355. <https://doi.org/10.1038/s41579-018-0001>.
- Liu X, Kimmey JM, Matarazzo L, de Bakker V, Van Maele L, Sirard J-C, et al. Exploration of bacterial bottlenecks and *Streptococcus pneumoniae* pathogenesis by CRISPRi-Seq. *Cell Host Microbe*. 2021;29(1):107–20. <https://doi.org/10.1016/j.chom.2020.10.001>.
- Brueggemann AB, van Rensburg MJJ, Shaw D, McCarthy ND, Jolley KA, Maiden MC, et al. Changes in the incidence of invasive disease due to *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Neisseria meningitidis* during the COVID-19 pandemic in 26 countries and territories in the Invasive Respiratory Infection Surveillance Initiative: a prospective analysis of surveillance data. *Lancet Digital Health*. 2021;3(6):e360–70. [https://doi.org/10.1016/S2589-7500\(21\)00077-7](https://doi.org/10.1016/S2589-7500(21)00077-7).
- Organization WH. Pneumococcal conjugate vaccine for childhood immunization—WHO position paper. The weekly epidemiological record. 2007;82(12):93–104. <https://apps.who.int/iris/handle/10665/240895>.
- Malley RJJ. Antibody and cell-mediated immunity to *Streptococcus pneumoniae*: implications for vaccine development. *J Molecular Med*. 2010;88:135–42. <https://doi.org/10.1007/s00109-009-0579-4>.
- Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, Bennett NM, et al. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *J Infect Dis*. 2010;201(1):32–41. <https://doi.org/10.1086/648593>.
- Lagousi T, Basdeki P, Routsias J, Spoulou V. Novel protein-based pneumococcal vaccines: assessing the use of distinct protein fragments instead of full-length proteins as vaccine antigens. *Vaccines*. 2019;7(1):9. <https://doi.org/10.3390/vaccines7010009>.
- Alderson MRJV. Status of research and development of pediatric vaccines for *Streptococcus pneumoniae*. *Vaccine*. 2016;34(26):2959–61. <https://doi.org/10.1016/j.vaccine.2016.03.107>.
- Berry AM, Paton JC. Immunity Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. *Infect Immun*. 2000;68(1):133–40. <https://doi.org/10.1128/IAI.68.1.133-140.2000>.
- Kerr AR, Paterson GK, McCluskey J, Iannelli F, Oggioni MR, Pozzi G, et al. The contribution of PspC to pneumococcal virulence varies between strains and is accomplished by both complement evasion and complement-independent mechanisms. *Infect Immun*. 2006;74(9):5319–24. <https://doi.org/10.1128/IAI.00543-06>.
- Guerra AJ, Dann III CE, Giedroc DPJ. Crystal structure of the zinc-dependent MarR family transcriptional regulator AdcR in the Zn (II)-bound state. *J Am Chem Soc*. 2011;133(49):19614–7. <https://doi.org/10.1021/ja2080532>.
- Plumptre CD, Ogunniyi AD, Paton JC. Vaccination against *Streptococcus pneumoniae* using truncated derivatives of polyhistidine triad protein D. *PLoS ONE*. 2013;8(10):e78916. <https://doi.org/10.1371/journal.pone.0078916>.
- Cundell D, Pearce B, Sandros J, Naughton A, Masure HJ. Immunity peptide permeases from *Streptococcus pneumoniae* affect adherence to eucaryotic cells. *Infect Immun*. 1995;63(7):2493–8. <https://doi.org/10.1128/iai.63.7.2493-2498.1995>.
- Jomaa M, Yuste J, Paton JC, Jones C, Dougan G, Brown JSJ, et al. Antibodies to the iron uptake ABC transporter lipoproteins PiaA and PiuA promote opsonophagocytosis of *Streptococcus pneumoniae*. *Infect Immun*. 2005;73(10):6852–9. <https://doi.org/10.1128/IAI.73.10.6852-6859.2005>.
- Daniels CC, Coan P, King J, Hale J, Benton KA, Briles DE, et al. The proline-rich region of pneumococcal surface proteins A and C contains surface-accessible epitopes common to all pneumococci and elicits antibody-mediated protection against sepsis. *Infect Immun*. 2010;78(5):2163–72. <https://doi.org/10.1128/IAI.01199-09>.
- Masomian M, Ahmad Z, Ti Gew L, Poh CLJV. Development of next generation *Streptococcus pneumoniae* vaccines conferring broad protection. *Vaccines*. 2020;8(1):132. <https://doi.org/10.3390/vaccines8010132>.
- Janulczyk R, Iannelli F, Sjöholm AG, Pozzi G, Björck LJ. Hic, a novel surface protein of *Streptococcus pneumoniae* that interferes with complement function. *J Biol Chem*. 2000;275(47):37257–63. <https://doi.org/10.1074/jbc.M004572200>.
- Cheng Q, Finkel D, Hostetter MKJB. Novel purification scheme and functions for a C3-binding protein from *Streptococcus pneumoniae*. *Biochemistry*. 2000;39(18):5450–7. <https://doi.org/10.1021/bi992157d>.
- Hammerschmidt S, Talay SR, Brandtzaeg P, Chhatwal GSJM. SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component. *Mol Microbiol*. 1997;25(6):1113–24. <https://doi.org/10.1046/j.1365-2958.1997.5391899.x>.
- Rosenow C, Ryan P, Weiser JN, Johnson S, Fontan P, Ortqvist A, et al. Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Mol Microbiol*. 1997;25(5):819–29. <https://doi.org/10.1111/j.1365-2958.1997.mmi494.x>.
- Brooks-Walter A, Briles DE, Hollingshead SKJ. Immunity the pspC gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia. *Infection Immunity*. 1999;67(12):6533–42. <https://doi.org/10.1128/IAI.67.12.6533-6542.1999>.
- Iannelli F, Oggioni MR, Pozzi GJG. Allelic variation in the highly polymorphic locus pspC of *Streptococcus pneumoniae*. *Gene*. 2002;284(1–2):63–71. [https://doi.org/10.1016/S0378-1119\(01\)00896-4](https://doi.org/10.1016/S0378-1119(01)00896-4).
- Balachandran P, Brooks-Walter A, Virolainen-Julkunen A, Hollingshead SK, Briles DEJ. Immunity role of pneumococcal surface protein C in nasopharyngeal carriage and pneumonia and its ability to elicit protection against carriage of *Streptococcus pneumoniae*. *Infection Immunity*. 2002;70(5):2526–34. <https://doi.org/10.1128/IAI.70.5.2526-2534.2002>.
- Moreno AT, Oliveira MLS, Ho PL, Vadesilho CF, Palma GM, Ferreira JM Jr, et al. Cross-reactivity of antipneumococcal surface protein C (PspC) antibodies with different strains and evaluation of inhibition of human complement factor H and secretory IgA binding via PspC. *Clin Vaccine Immunol*. 2012;19(4):499–507. <https://doi.org/10.1128/CI.05706-11>.
- Malekan M, Siadat SD, Aghasadeghi M, Shahrokhi N, Afrough P, Behrouzi A, et al. Evaluation of protective immunity responses against pneumococcal PhtD and its C-terminal in combination with outer-membrane vesicles as adjuvants. *J Med Microbiol*. 2020;69(3):465–77. <https://doi.org/10.1099/jmm.0.001103>.
- Bersch B, Bougault C, Roux L, Favier A, Vernet T, Durmort C. New insights into histidine triad proteins: solution structure of a *Streptococcus*

- pneumoniae* PhtD domain and zinc transfer to AdcAll. *PLoS ONE*. 2013;8(11):e81168. <https://doi.org/10.1371/journal.pone.0081168>.
27. Ogunniyi AD, Grabowicz M, Mahdi LK, Cook J, Gordon DL, Sadlon TA, et al. Pneumococcal histidine triad proteins are regulated by the Zn²⁺-dependent repressor AdcR and inhibit complement deposition through the recruitment of complement factor H. *FASEB J*. 2009;23(3):731–8. <https://doi.org/10.1096/fj.08-119537>.
 28. Rioux S, Neyt C, Di Paolo E, Turpin L, Charland N, Labbe S, et al. Transcriptional regulation, occurrence and putative role of the Pht family of *Streptococcus pneumoniae*. *Microbiology*. 2011;157(2):336–48. <https://doi.org/10.1099/mic.0.042184-0>.
 29. Kallio A, Sepponen K, Hermand P, Denoël P, Godfroid F, Melin M. Role of Pht proteins in attachment of *Streptococcus pneumoniae* to respiratory epithelial cells. *Infect Immun*. 2014;82(4):1683–91. <https://doi.org/10.1128/IAI.00699-13>.
 30. Beghetto E, Gargano N, Ricci S, Garufi G, Peppoloni S, Montagnani F, et al. Discovery of novel *Streptococcus pneumoniae* antigens by screening a whole-genome λ -display library. *FEMS Microbiol Lett*. 2006;262(1):14–21. <https://doi.org/10.1111/j.1574-6968.2006.00360.x>.
 31. Yu J, Li B, Chen X, Lu J, Wang D, Gu T, et al. Comparison of immunogenicity and protection of two pneumococcal protein vaccines based on PsaA and PspA. *Infect Immun*. 2018;86(6):e00916–e917. <https://doi.org/10.1128/IAI.00916-17>.
 32. Lu J, Sun T, Wang D, Dong Y, Xu M, Hou H, et al. Protective immune responses elicited by fusion protein containing PsaA and PspA fragments. *Immunol Invest*. 2015;44(5):482–96. <https://doi.org/10.3109/08820139.2015.1037956>.
 33. Sampson JS, Furlow Z, Whitney AM, Williams D, Facklam R, Carlone GMJ, et al. Limited diversity of *Streptococcus pneumoniae* psaA among pneumococcal vaccine serotypes. *Infect Immun*. 1997;65(5):1967–71. <https://doi.org/10.1128/iai.65.5.1967-1971.1997>.
 34. Rajam G, Anderton JM, Carlone GM, Sampson JS, Ades EW. Pneumococcal surface adhesin a (PsaA): a review. *Crit Rev Microbiol*. 2008;34(3–4):131–42. <https://doi.org/10.1080/10408410802275352>.
 35. Wang S, Li Y, Shi H, Scarpellini G, Torres-Escobar A, Roland KL, et al. Immune responses to recombinant pneumococcal PsaA antigen delivered by a live attenuated *Salmonella* vaccine. *Infect Immun*. 2010;78(7):3258–71. <https://doi.org/10.1128/IAI.00176-10>.
 36. McDevitt CA, Ogunniyi AD, Valkov E, Lawrence MC, Kobe B, McEwan AG, et al. A molecular mechanism for bacterial susceptibility to zinc. *PLoS Pathog*. 2011;7(11):e1002357. <https://doi.org/10.1371/journal.ppat.1002357>.
 37. Dorosti H, Eslami M, Negahdaripour M, Ghoshon MB, Gholami A, Heidari R, et al. Vaccinomics approach for developing multi-epitope peptide pneumococcal vaccine. *J Biomol Struct Dyn*. 2019;37(13):3524–35. <https://doi.org/10.1080/07391102.2018.1519460>.
 38. Tartey S, Takeuchi OJ. Pathogen recognition and toll-like receptor targeted therapeutics in innate immune cells. *Internat Rev Immunol*. 2017;36(2):57–73. <https://doi.org/10.1080/08830185.2016.1261318>.
 39. Kumar S, Sunagar R, Gosselin EJ. Bacterial protein toll-like-receptor agonists: a novel perspective on vaccine adjuvants. *Front Immunol*. 2019;10:1144. <https://doi.org/10.3389/fimmu.2019.01144>.
 40. Zhang H, Kang L, Yao H, He Y, Wang X, Xu W, et al. *Streptococcus pneumoniae* endopeptidase O (PepO) elicits a strong innate immune response in mice via TLR2 and TLR4 signaling pathways. *Front Cell Infect Microbiol*. 2016;6:23. <https://doi.org/10.3389/fcimb.2016.00023>.
 41. Shu Z, Yuan J, Wang H, Zhang J, Li S, Zhang H, et al. *Streptococcus pneumoniae* PepO promotes host anti-infection defense via autophagy in a Toll-like receptor 2/4 dependent manner. *Virulence*. 2020;11(1):270–82. <https://doi.org/10.1080/21505594.2020.1739411>.
 42. Khan M, Khan S, Ali A, Akbar H, Sayaf AM, Khan A, et al. Immunoinformatics approaches to explore *helicobacter pylori* proteome (Virulence Factors) to design B and T cell multi-epitope subunit vaccine. *Sci Rep*. 2019;9(1):1–13. <https://doi.org/10.1038/s41598-019-49354-z>.
 43. Tahirul QM, Bari A, Adeel MM, Maryam A, Ashfaq UA, Du X, et al. Peptide vaccine against chikungunya virus: immuno-informatics combined with molecular docking approach. *Transl Med*. 2018;16(1):1–14. <https://doi.org/10.1186/s12967-018-1672-7>.
 44. Fereshteh S, Abdoli S, Shahcheraghi F, Ajdary S, Nazari M, Badmasti FJMp. New putative vaccine candidates against *Acinetobacter baumannii* using the reverse vaccinology method. *Micro Pathogen*. 2020;143:104114. <https://doi.org/10.1016/j.micpath.2020.104114>.
 45. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELJ. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Molecular Biol*. 2001;305(3):567–80. <https://doi.org/10.1006/jmbi.2000.4315>.
 46. Armenteros JJA, Salvatore M, Emanuelsson O, Winther O, Von Heijne G, Elofsson A, et al. Detecting sequence signals in targeting peptides using deep learning. *Life Sci Allian*. 2019;2(5):1–14. <https://doi.org/10.26508/lsa.201900429>.
 47. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang YJNm. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 2015;12(1):7–8. <https://doi.org/10.1038/nmeth.3213>.
 48. Zhang YJB. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 2008;9(1):1–8. <https://doi.org/10.1186/1471-2105-9-40>.
 49. Kabsch WJACSAC. Diffraction, theoretical, crystallography g a discussion of the solution for the best rotation to relate two sets of vectors. *Acta Cryst Section A*. 1978;34(5):827–8. <https://doi.org/10.1107/S0567739478001680>.
 50. Shin W-H, Lee GR, Heo L, Lee H, Seok CJB. Prediction of protein structure and interaction by GALAXY protein modeling programs. *Bio Design*. 2014;2(1):1–11.
 51. Wiederstein M, Sippl MJNar. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35(2):W407–10. <https://doi.org/10.1093/nar/gkm290>.
 52. Laskowski RA, MacArthur MW, Moss DS, Thornton MJJoac. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystall*. 1993;26(2):283–91. <https://doi.org/10.1107/S0021889892009944>.
 53. Saha S, Raghava GPSJPS. Function prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins: Struct Funct Bioinform*. 2006;65(1):40–8. <https://doi.org/10.1002/prot.21078>.
 54. Singh H, Ansari HR, Raghava GP. Improved method for linear B-cell epitope prediction using antigen's primary sequence. *PLoS ONE*. 2013;8(5):e62216. <https://doi.org/10.1371/journal.pone.0062216>.
 55. Emini EA, Hughes JV, Perlow D, Boger J. Induction of hepatitis a virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol*. 1985;55(3):836–9. <https://doi.org/10.1128/jvi.55.3.836-839.1985>.
 56. Ponomarenko J, Bui H-H, Li W, Füsseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*. 2008;9(1):1–8. <https://doi.org/10.1186/1471-2105-9-514>.
 57. Kringelum JV, Lundegaard C, Lund O, Nielsen MJPCb. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol*. 2012;8(12):e1002829. <https://doi.org/10.1371/journal.pcbi.1002829>.
 58. Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters BJPCb. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Computational Biol*. 2008;4(4):e1000048. <https://doi.org/10.1371/journal.pcbi.1000048>.
 59. Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology*. 2018;154(3):394–406. <https://doi.org/10.1111/imm.12889>.
 60. van de Garde MDB, van Westen E, Poelen MCM, Rots NY, van Els CACM. Prediction and validation of immunogenic domains of pneumococcal proteins recognized by human CD4⁺ T cells. *Infect Immun*. 2019;87(6):e00098–19. <https://doi.org/10.1128/IAI.00098-19>.
 61. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The cluspro web server for protein–protein docking. *Nat Protoc*. 2017;12(2):255–78. <https://doi.org/10.1038/nprot.2016.169>.
 62. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng Des Sel*. 1995;8(2):127–34. <https://doi.org/10.1093/protein/8.2.127>.
 63. Tian H, Hou X, Liu X. Real-time SPR characterization of the interactions between multi-epitope proteins and antibodies against classical swine fever virus. *Biochem Biophys Res Commun*. 2013;431(2):315–20. <https://doi.org/10.1016/j.bbrc.2012.12.104>.
 64. Ahmad S, Shahid F, Tahirul QM, Abbasi SW, Sajjad W, Ismail S, et al. Immuno-informatics analysis of pakistan-based HCV subtype-3a for

- chimeric polypeptide vaccine design. *Vaccines*. 2021;9(3):293. <https://doi.org/10.3390/vaccines9030293>.
65. Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng Des Sel*. 2001;14(8):529–32. <https://doi.org/10.1093/protein/14.8.529>.
 66. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinf*. 2007;8(1):1–7. <https://doi.org/10.1186/1471-2105-8-4>.
 67. Magnan CN, Zeller M, Kayala MA, Vigil A, Randall A, Felgner PL, et al. High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics*. 2010;26(23):2936–43. <https://doi.org/10.1093/bioinformatics/btq551>.
 68. Saha S, Raghava GPSJNar. AllgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res*. 2006;34(2):W202–9. <https://doi.org/10.1093/nar/gkl343>.
 69. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v 2—a server for in silico prediction of allergens. *J Molecular Model*. 2014;20(6):1–6. <https://doi.org/10.1007/s00894-014-2278-5>.
 70. Magnan CN, Randall A, Baldi PJB. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics*. 2009;25(17):2200–7. <https://doi.org/10.1093/bioinformatics/btp386>.
 71. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R. In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE*. 2013;8:e73957. <https://doi.org/10.1371/journal.pone.0073957>.
 72. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch AJTpph. Protein identification and analysis tools on the ExPASy server. *Proteomics Proto Handbook*. 2005. <https://doi.org/10.1385/1-59259-890-0/571>.
 73. Kim DE, Chivian D, Baker DJNar. Protein structure prediction and analysis using the robetta server. *Nucl Acids Res*. 2004;32(2):W526–31. <https://doi.org/10.1093/nar/gkh468>.
 74. Colovos C, Yeates TOJPs. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*. 1993;2(9):1511–9. <https://doi.org/10.1002/pro.5560020916>.
 75. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 2008;4(3):435–47. <https://doi.org/10.1021/ct700301q>.
 76. Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J*. 2011;40(7):843–56. <https://doi.org/10.1007/s00249-011-0700-9>.
 77. Bahadori Z, Shabani AA, Minucheir Z. Rational design of hyper-glycosylated human follicle-stimulating hormone analogs (a bioinformatics approach). *J Biomol Struct Dyn*. 2021;40(16):1–12. <https://doi.org/10.1080/07391102.2021.1924268>.
 78. Nabizadeh Z, Minucheir Z, Shabani AA. Rational design of hyper-glycosylated human chorionic gonadotropin analogs (a bioinformatics approach). *Lett Drug Des Discovery*. 2020;17(8):1001–14. <https://doi.org/10.2174/1570180817666200225101938>.
 79. Rapin N, Lund O, Bernaschi M, Castiglione FJPo. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLoS ONE*. 2010;5(4):e9862. <https://doi.org/10.1371/journal.pone.0009862>.
 80. Castiglione F, Mantile F, De Berardinis P, Prisco AJC, medicine mmi. How the interval between prime and boost injection affects the immune response in a computational model of the immune system. 2012;2012.
 81. Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucl Acids Res*. 2005;33(2):W526–31. <https://doi.org/10.1093/nar.gki376>.
 82. Loughran AJ, Orihuela CJ, Tuomanen EI. *Streptococcus pneumoniae*: Invasion and inflammation. *Microbiol Spect*. 2019;7(2):1–31. <https://doi.org/10.1128/microbiolspec.GPP3-0004-2018>.
 83. Petukhova E, Vorobyev D, Sidorov A, Semenova I, Volokh Y, Leonova A, et al. Immunization with recombinant pneumolysin induces the production of antibodies and protects mice in a model of systemic infection caused by *Streptococcus pneumoniae*. *Bull Exp Biol Med*. 2020;168(4):485–7. <https://doi.org/10.1007/s10517-020-04736-6>.
 84. Yu J, Chen X, Li B, Gu T, Meng X, Kong W, et al. A pneumococcal vaccine combination with two proteins containing PspA families 1 and 2 can potentially protect against a wide range of *Streptococcus pneumoniae* strains. *Immunol Res*. 2018;66(4):528–36. <https://doi.org/10.1007/s12026-018-9016-5>.
 85. Shafaghi M, Shabani AA, Minucheir Z. Rational design of hyper-glycosylated human luteinizing hormone analogs (a bioinformatics approach). *Comput Biol Chem*. 2019;79:16–23. <https://doi.org/10.1016/j.compbiolchem.2019.01.002>.
 86. Gautam A, Chaudhary K, Kumar R, Sharma A, Kapoor P, Tyagi A, et al. In silico approaches for designing highly effective cell penetrating peptides. *Transl Med*. 2013;11(1):1–12. <https://doi.org/10.1186/1479-5876-11-74>.
 87. Tomar N, De RK. Immunoinformatics: an integrated scenario. *Immunology*. 2010;131(2):153–68. <https://doi.org/10.1111/j.1365-2567.2010/03330.x>.
 88. Idrees S, Ashfaq UA, Khaliq SJJotm. HCV Envelope protein 2 sequence comparison of Pakistani isolate and In-silico prediction of conserved epitopes for vaccine development. *Transl Med*. 2013;11(1):1–9. <https://doi.org/10.1186/1479-5876-11-105>.
 89. Huang P, Yu S, Wu C, Liang LJotm. Highly conserved antigenic epitope regions of hemagglutinin and neuraminidase genes between 2009 H1N1 and seasonal H1N1 influenza: vaccine considerations. *Transl Med*. 2013;11(1):1–8. <https://doi.org/10.1186/1479-5876-11-47>.
 90. Tischer S, Geyerregger R, Kwoczek J, Heim A, Figueiredo C, Blasczyk R, et al. Discovery of immunodominant T-cell epitopes reveals penton protein as a second immunodominant target in human adenovirus infection. *Transl Med*. 2016;14(1):1–16. <https://doi.org/10.1186/s12967-016-1042-2>.
 91. Tahir ul QM, Saleem S, Ashfaq UA, Bari A, Anwar F, Alqahtani SJJotm. Epitope-based peptide vaccine design and target site depiction against middle east respiratory syndrome coronavirus: an immunoinformatics study. *Transl Med*. 2019;17(1):1–14. <https://doi.org/10.1186/s12967-019-2116-8>.
 92. Banerjee A, Santra D, Maiti SJJotm. Energetics and IC50 based epitope screening in SARS CoV-2 (COVID 19) spike protein by immunoinformatic analysis implicating for a suitable vaccine development. *Transl Med*. 2020;18(1):1–14. <https://doi.org/10.1186/s12967-020/02435-4>.
 93. Raoufi E, Hemmati M, Eftekhari S, Khaksaran K, Mahmodi Z, Farajollahi MM, et al. Epitope prediction by novel immunoinformatics approach: a state-of-the-art review. *Int J Pept Res Ther*. 2020;26(2):1155–63. <https://doi.org/10.1007/s10989-019-09918-z>.
 94. Ogunniyi AD, Woodrow MC, Poolman JT, Paton JC. Protection against *Streptococcus pneumoniae* elicited by immunization with pneumolysin and CbpA. *Infect Immun*. 2001;69(10):5997–6003. <https://doi.org/10.1128/IAI.69.10.5997-6003.2001>.
 95. Mann B, Thornton J, Heath R, Wade KR, Tweten RK, Gao G, et al. Broadly protective protein-based pneumococcal vaccine composed of pneumolysin toxin–CbpA peptide recombinant fusion protein. *J Infect Dis*. 2014;209(7):1116–25. <https://doi.org/10.1093/infdis/jit502>.
 96. Godfroid F, Hermand P, Verlant V, Denoël P, Poolman JT. Preclinical evaluation of the Pht proteins as potential cross-protective pneumococcal vaccine antigens. *Infect Immun*. 2011;79(1):238–45. <https://doi.org/10.1128/IAI.00378-10>.
 97. Khan MN, Pichichero ME. Vaccine candidates PhtD and PhtE of *Streptococcus pneumoniae* are adhesins that elicit functional antibodies in humans. *Vaccine*. 2012;30(18):2900–7. <https://doi.org/10.1016/j.vaccine.2012.02.023>.
 98. Verhoeven D, Xu Q, Pichichero MEJV. Vaccination with a *Streptococcus pneumoniae* trivalent recombinant PcpA, PhtD and PlyD1 protein vaccine candidate protects against lethal pneumonia in an infant murine model. *Vaccine*. 2014;32(26):3205–10. <https://doi.org/10.1016/j.vaccine.2014.04.004>.
 99. De B, Sampson J, Ades E, Huebner R, Jue D, Johnson S, et al. Purification and characterization of *Streptococcus pneumoniae* palmitoylated pneumococcal surface adhesin expressed in *Escherichia coli*. *Vaccine*. 2000;18(17):1811–21. [https://doi.org/10.1016/S0264-410X\(99\)00481-8](https://doi.org/10.1016/S0264-410X(99)00481-8).
 100. Lin H, Lin Z, Meng C, Huang J, Guo Y. Preparation and immunogenicity of capsular polysaccharide–surface adhesin (PsaA) conjugate of *Streptococcus pneumoniae*. *Immunobiology*. 2010;215(7):545–50. <https://doi.org/10.1016/j.jimbio.2009.08.008>.

101. Lu Y-J, Forte S, Thompson CM, Anderson PW, Malley R. Protection against Pneumococcal colonization and fatal pneumonia by a trivalent conjugate of a fusion protein with the cell wall polysaccharide. *Infect Immun*. 2009;77(5):2076–83. <https://doi.org/10.1128/IAI.01554-08>.
102. Briles DE, Ades E, Paton JC, Sampson JS, Carlone GM, Huebner RC, et al. Intranasal immunization of mice with a mixture of the pneumococcal proteins PsaA and PspA is highly protective against nasopharyngeal carriage of *Streptococcus pneumoniae*. *Infect Immun*. 2000;68(2):796–800. <https://doi.org/10.1128/IAI.68.2.796-800.2000>.
103. Rueckert C, Guzmán CA. Vaccines: from empirical development to rational design. *PLoS Pathog*. 2012;8(11):e1003001. <https://doi.org/10.1371/journal.ppat.1003001>.
104. Sanami S, Azadegan-Dehkordi F, Rafeian-Kopaei M, Salehi M, Ghasemi-Dehnoo M, Mahooti M, et al. Design of a multi-epitope vaccine against cervical cancer using immunoinformatics approaches. *Scientific Rep*. 2021;11(1):1–15. <https://doi.org/10.1038/s41598-021-91997-4>.
105. Sayed SB, Nain Z, Khan MSA, Abdulla F, Tasmin R, Adhikari, UKJljojr, et al. Exploring lassa virus proteome to design a multi-epitope vaccine through immunoinformatics and immune simulation analyses. *Int J Pept Res Therap*. 2020;26(4):2089–107. <https://doi.org/10.1007/s10989-019-10003-8>.
106. Bukhari SNH, Jain A, Haq E, Mehbodniya A, Webber JJP. Machine learning techniques for the prediction of B-Cell and T-Cell epitopes as potential vaccine targets with a specific focus on SARS-CoV-2 pathogen: a review. *Pathogens*. 2022;11(2):146. <https://doi.org/10.3390/pathogens11020146>.
107. Khan MT, Islam MJ, Parihar A, Islam R, Jerin TJ, Dhote R, et al. Immunoinformatics and molecular modeling approach to design universal multi-epitope vaccine for SARS-CoV-2. *Inf Med*. 2021;24:100578. <https://doi.org/10.1016/j.jimu.2021.100578>.
108. Rai J, Lok KI, Mok CY, Mann H, Noor M, Patel P, et al. Immunoinformatic evaluation of multiple epitope ensembles as vaccine candidates: *E coli* 536. *Bioinformatics*. 2012;8(6):272. <https://doi.org/10.6026/97320630008272>.
109. Negahdaripour M, Nezafat N, Eslami M, Ghoshoon MB, Shoolian E, Najafipour S, et al. Structural vaccinology considerations for in silico designing of a multi-epitope vaccine. *Infect Genet Evol*. 2018;58:96–109. <https://doi.org/10.1016/j.meegid.2017.12.008>.
110. Khatoun N, Pandey RK, Prajapati VKJSr. Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci Rep*. 2017;7(1):1–12. <https://doi.org/10.1038/s41598-017-08842-w>.
111. Shey RA, Ghogomu SM, Esh KK, Nebangwa ND, Shintouo CM, Nongley NF, et al. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci Rep*. 2019;9(1):1–18. <https://doi.org/10.1038/s41598-019-40833-x>.
112. Barh D, Barve N, Gupta K, Chandra S, Jain N, Tiwari S, et al. Exoproteome and secretome derived broad spectrum novel drug and vaccine candidates in *Vibrio cholerae* targeted by piper betel derived compounds. *PLoS ONE*. 2013;8(1):e52773. <https://doi.org/10.1371/journal.pone.0052773>.
113. Varshavsky AJGtc. The N-end rule pathway of protein degradation. *Genes Cells*. 1997;2(1):13–28. <https://doi.org/10.1046/j.1365-2443.1997.1020301.x>.
114. Ikai AJTJoB. Thermostability and aliphatic index of globular proteins. *Biochemistry*. 1980;88(6):1895–8. <https://doi.org/10.1093/oxfordjournals.jbchem.a133168>.
115. Ali M, Pandey RK, Khatoun N, Narula A, Mishra A, Prajapati VKJSr. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Sci Rep*. 2017;7(1):1–13. <https://doi.org/10.1038/s41598-017-09199-w>.
116. Zhang X, Issagholian A, Berg EA, Fishman JB, Nesburn AB, BenMohamed LJJov. Th-cytotoxic T-lymphocyte chimeric epitopes extended by N ϵ -palmitoyl lysines induce herpes simplex virus type 1-specific effector CD8+Tc1 responses and protect against ocular infection. *J Virol*. 2005;79(24):15289–301. <https://doi.org/10.1128/JVI.79.24.15289-15301.2005>.
117. Brooks LR, Mias GI. *Streptococcus pneumoniae*'s virulence and host immunity: aging, diagnostics, and prevention. *Front Immunol*. 2018;9:1366. <https://doi.org/10.3389/fimmu.2018.01366>.
118. Bibi S, Ullah I, Zhu B, Adnan M, Liaqat R, Kong W-B, et al. In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. *Sci Rep*. 2021;11(1):1–16. <https://doi.org/10.1038/s41598-020-80899-6>.
119. Omoniyi AA, Adebisi SS, Musa SA, Nzalak JO, Bauchi ZM, Bako KW, et al. In silico design and analyses of a multi-epitope vaccine against Crimean-Congo hemorrhagic fever virus through reverse vaccinology and immunoinformatics approaches. *Sci Rep*. 2022;12(1):1–17. <https://doi.org/10.1038/s41598-022-12651-1>.
120. Pandey RK, Bhatt TK, Prajapati VKJSr. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. *Sci Rep*. 2018;8(1):1–11. <https://doi.org/10.1038/s41598-018-19456-1>.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

