


RESEARCH

Open Access



# In silico design of an epitope-based vaccine against PspC in *Streptococcus pneumoniae* using reverse vaccinology

Md. Nahian<sup>1</sup>, Muhammad Shahab<sup>2</sup>, Lincon Mazumder<sup>1,3</sup>, Jonas Ivan Nobre Oliveira<sup>4</sup>, Tanjina Akhtar Banu<sup>5</sup>, Murshed Hasan Sarkar<sup>5</sup>, Barna Goswami<sup>5</sup>, Ashashan Habib<sup>5</sup>, Shamima Begum<sup>1</sup> and Shahina Akter<sup>5\*</sup> 

## Abstract

**Background** *Streptococcus pneumoniae* is a major pathogen that poses a significant hazard to global health, causing a variety of infections including pneumonia, meningitis, and sepsis. The emergence of antibiotic-resistant strains has increased the difficulty of conventional antibiotic treatment, highlighting the need for alternative therapies such as multi-epitope vaccines. In this study, immunoinformatics algorithms were used to identify potential vaccine candidates based on the extracellular immunogenic protein Pneumococcal surface protein C (PspC).

**Method** The protein sequence of PspC was retrieved from NCBI for the development of the multi-epitope vaccine (MEV), and potential B cell and T cell epitopes were identified. Linkers including EAAAK, AAY, and CPGPG were used to connect the epitopes. Through molecular docking, molecular dynamics, and immunological simulation, the affinity between MEV and Toll-like receptors was determined. After cloning the MEV construct into the PET28a (+) vector, SnapGene was used to achieve expression in *Escherichia coli*.

**Result** The constructed MEV was discovered to be stable, non-allergenic, and antigenic. Microscopic interactions between ligand and receptor are confirmed by molecular docking and molecular dynamics simulation. The use of an in-silico cloning approach guarantees the optimal expression and translation efficiency of the vaccine within an expression vector.

**Conclusion** Our study demonstrates the potential of in silico approaches for designing effective multi-epitope vaccines against *S. pneumoniae*. The designated vaccine exhibits the required physicochemical, structural, and immunological characteristics of a successful vaccine against SPN. However, laboratory validation is required to confirm the safety and immunogenicity of the proposed vaccine design.

**Keywords** Epitope, Vaccine, B cell, T cell, Immunoinformatics, Reverse vaccinology

\*Correspondence:

Shahina Akter  
shupty2010@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Background

Pneumonia is a respiratory infection that specifically affects one or both lungs. The lungs are comprised of small sacs called alveoli, which are essential for the breathing process. However, in cases of persistent pneumonia, these alveoli can become filled with pus and fluid, causing painful breathing and reduced oxygen intake. As a result, it can lead to the development of acute respiratory syndrome [1]. *Streptococcus pneumoniae*, also known as pneumococcus, a bacterium that is gram-positive and non-motile, is the main cause of community-acquired pneumonia. Pneumococcal pneumonia predominantly affects infants below the age of 2, older adults, and people with weakened immune systems [2]. Pneumococcal invasive diseases comprise various illnesses such as invasive pneumococcal disease (IPD), bacteremia, sepsis, meningitis, empyema, sinusitis, and middle ear otitis media [2, 3]. As a consequence, almost 14 million cases of pneumococcal disease have been reported around the world, with an estimated 1.6 million fatalities owing to this infection annually [4]. The incidence of pneumonia among children in undeveloped nations was around 0.29 cases per person per year, whereas it was just 0.05 cases per person per year in developed nations [4, 5]. Hence, it is the top cause of mortality among children worldwide.

Research conducted in the USA in 2008 showed that pneumococci had evolved resistance to many kinds of antibiotics. These include quinolones, penicillin, macrolides, and cephalosporins [6]. Evidence from China indicates that *S. pneumoniae* is very resistant to many antibiotics, with rates of resistance of 95.8% to clindamycin, 95.2% to erythromycin, 93.6% to tetracycline, and 66.7% to trimethoprim/sulfamethoxazole [7]. Furthermore, the existing pneumococcal vaccines do not provide full protection against all strains of *Streptococcus pneumoniae* since there are more than 100 different serotypes of this bacterium. At present, there are two categories of pneumococcal vaccines: plain polysaccharide vaccines (PPV) and protein-conjugated polysaccharide vaccines (PCV) [8]. Due to its reliance on T cell-independent polysaccharide antigens, PPV does not effectively protect babies younger than two, who are at the highest risk for serious pneumococcal infections. On the other hand, PCV provides protection to children, but it has drawbacks such as a high cost, being difficult to manufacture, requiring multiple injections, and needing to be refrigerated. Additionally, PCV only covers certain pneumococcal serotypes found in developed countries [9, 10]. Both vaccines are limited to specific serotypes and can only induce immunity against those specific serotypes [9]. To address this limitation and protect against a wider range of *S. pneumoniae* serotypes, there is an urgent need for

a vaccination approach that is not limited to specific serotypes.

An effective approach for vaccination against *Streptococcus pneumoniae* is to use an epitope-based strategy that targets specific epitopes using a group of conserved pneumococcal protein antigens. This method can serve as a feasible alternative to vaccines that are designed solely for specific serotypes [11]. By focusing on conserved protein antigens, epitope-based vaccines have the potential to provide broad protection against various strains or serotypes of pneumococcus [12, 13]. Over the years, researchers have extensively studied different pneumococcal proteins with the aim of creating a vaccine that can protect against pneumococcal diseases. These include pneumococcal surface proteins A and C (PspA/C) [14, 15], pneumococcal histidine triad (Pht) family members A, B, D, and E (PhtA, B, D, and E) [16, 17], and ATP binding cassette (ABC) transporters PsaA, PiuA, and PiaA [18, 19]. Among these antigens, PspC is one of the most attractive candidates because it is highly prevalent and highly conserved among *Streptococcus pneumoniae* strains [20]. PspC is a surface-exposed protein that can bind to human complement factor H and secretory immunoglobulin A, thereby evading the innate and adaptive immune systems. PspC also acts as an adhesin that mediates the attachment of the pneumococcus to epithelial cells and facilitates colonization and invasion [21]. These properties make PspC a promising target for the development of a vaccine that can provide broader protection against *S. pneumoniae*.

Recent advancements in immunoinformatics have revolutionized the development of epitope vaccines, offering cost-effective and time-efficient approaches along with an expanded range of vaccine design tools [22–24]. In recent years, scientists have extensively relied on computational approaches for selecting efficient epitopes and developing innovative vaccinations against a wide variety of diseases, including SARS-CoV-2 [25], *Burkholderia pseudomallei* [26], *Mycobacterium tuberculosis* [27], and *Staphylococcus aureus* [28]. Given the practicality and advantages of immunoinformatics-based vaccine design, the primary objective of this study is to progress the creation of a multi-epitope vaccine to safeguard against *Streptococcus pneumoniae* by specifically targeting the PspC protein. To create the vaccine, multiple B and T cell epitopes will be included, which were produced by utilizing various in silico methods. Vaccine structure were modeled and docked with TLR4 to obtain an effective immune response. In addition, molecular dynamics simulations were employed to validate the bound complex and investigate its properties. Lastly, the process of in silico cloning and codon optimization

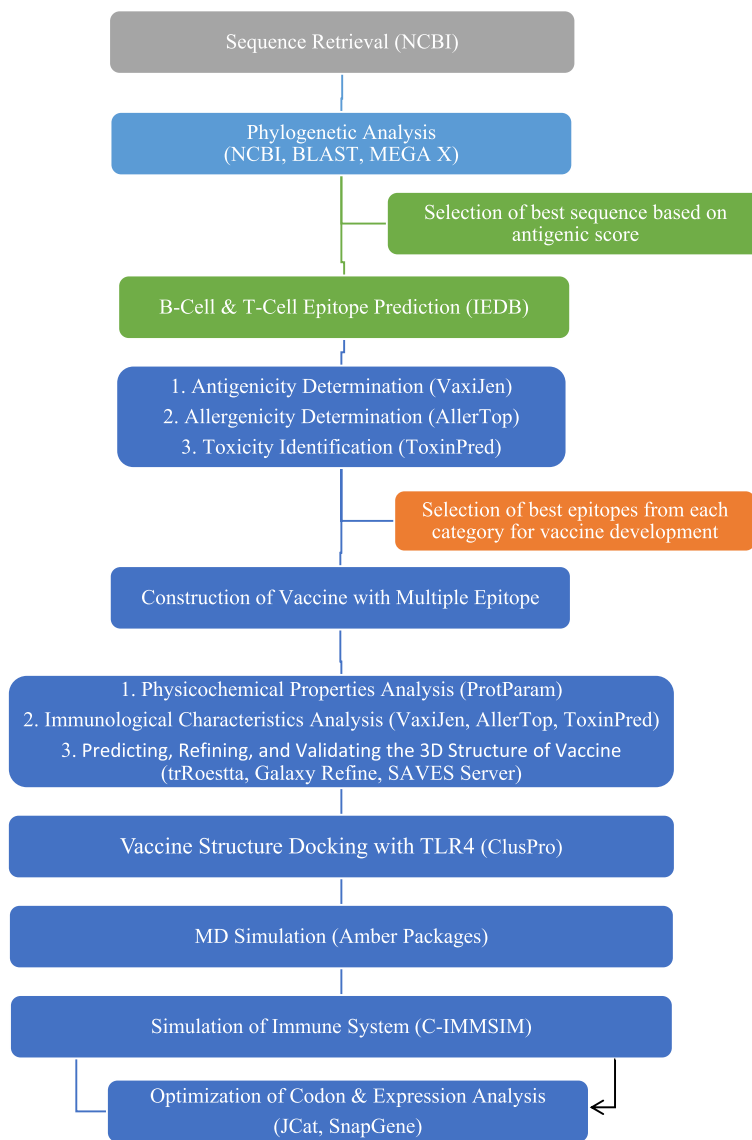
was conducted to evaluate the expression of chimeric proteins in a suitable host. The main strength of this study is the inclusion of more potent antigenic epitope-rich proteins in the vaccine, which can elicit diverse and robust immune responses. The significance of this approach has already been demonstrated in combating diseases like malaria [29], multiple sclerosis [30], and tumors [31].

**Methods**

Figure 1 illustrates the complete workflow of this scientific study. The figure also displays the diverse set of tools utilized in this approach.

**Retrieval and initial analysis of protein sequences**

The research began by acquiring the protein sequence of PspC from NCBI (<https://www.ncbi.nlm.nih.gov/protein>). Afterward, this sequence was submitted to protein–protein blast (Blastp) [32]. MUSCLE v3.6 program [33] was used to perform the multiple sequence alignment. Mega X was utilized to determine the evolutionary relationships between the sequences [34]. Subsequently, the antigenic characteristics of each protein sequence have been analyzed through the online tool VaxiJen version 2.0 [35], accessible at <http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>. Subsequently, an allergic prediction scan was conducted on the protein sequences utilizing the AllerTOP v. 2.0 [36]



**Fig. 1** The procedure adopted in this research to develop the pneumococcal epitope-based vaccination

web tool, accessible at <https://www.ddg-pharmfac.net/AllerTOP/>.

#### **Analysis of physical and chemical properties**

ExPasy ProtParam was used to investigate the protein's physical and chemical characteristics [37], accessible at <https://www.expasy.org/resources/protparam/>. The purpose of this web tool is to provide a comprehensive analysis of the protein sequence by calculating various physical and chemical parameters. These parameters provide insights into the protein's structure, function, stability, and interactions. For instance, the molecular weight reflects the protein's size and complexity, the half-life estimates the protein's longevity and degradation rate, and the GRAVY measures the protein's hydrophobicity or hydrophilicity. These parameters help us determine the protein's suitability for further analysis and application [38].

#### **B cell epitopes prediction**

The identification of B-cell epitopes (BCEs), which are responsible for the formation of antibodies that confer humoral immunity, is an essential initial step in the construction of an epitope-based vaccine. Therefore, the IEDB Linear Epitope Prediction Tool v2 (<http://tool.iedb.org/bcell/>) [39] was used to make predictions about B cell epitopes. This tool uses a combination of amino acid scales and hidden Markov models to score the epitope potential of each residue in a protein sequence. The tool also incorporates solvent-accessible surface area calculations, as well as contact distances into its prediction of B cell epitope potential [40]. To ensure accuracy, the final selection of B cell epitopes involved a meticulous screening process using VaxiJen v2.0, AllerTOP v2.0, and ToxinPred servers.

#### **Prediction of the MHC-specific epitopes**

Using the IEDB website (<http://tool.iedb.org/main/>), T cell epitopes that trigger an effective immune response were analyzed by utilizing the commonly occurring fragment with the help of the MHC-1 (<http://tools.iedb.org/mhci/>) and MHC-2 tools (<http://tools.iedb.org/mhcii/>). To identify MHC-I restricted epitopes, the ANN 4.0 algorithm [41] was utilized, and for MHC-II restricted epitopes, NN-align 2 [42] was used. Then the epitopes were ranked according to their VaxiJen score, allergenicity, and toxicity.

#### **Determination of population coverage**

Between populations of various ancestries, there are major differences in the expression and distribution of HLA alleles [43]. In this work, we estimated the proportion of the population that would be covered by MHC-I

and MHC-II epitopes using the IEDB Analysis Resource, which can be found at <http://tools.iedb.org/population/>, to make an estimation of the percentage of the population that would be covered by our selected epitopes.

#### **Finalizing the construct**

The most effective B and T cell epitopes have been combined sequentially using an appropriate linker to create a multi-subunit vaccine. In the creation of vaccines, linkers such as AAY, EAAAK, and CPGPG were used. Since the peptides used to make vaccines are not highly immunogenic, adjuvants must be used to elicit an immune response [44]. To facilitate purification experiments, the construct was modified by adding a hexa-histidine tail, also known as a poly-histidine tag.

#### **Physicochemical and immunological property analysis**

The vaccine's physicochemical characteristics were determined using ExPasy of the ProtParam (<https://web.expasy.org/protparam/>) [37]. This included calculating its isoelectric pH, instability, aliphatic index, molecular weight, in vitro and in vivo half-life, as well as its GRAVY index. To evaluate the immunological properties of the vaccine, including antigenicity, allergenicity, and toxigenicity, comprehensive assessments were conducted using VaxiJen (<https://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) [35], AllerTOP ([https://webs.iitd.edu.in/raghava/toxinpred/multi\\_submit.php](https://webs.iitd.edu.in/raghava/toxinpred/multi_submit.php)) [36], and ToxinPred ([https://webs.iitd.edu.in/raghava/toxinpred/multi\\_submit.php](https://webs.iitd.edu.in/raghava/toxinpred/multi_submit.php)) [45].

#### **Predicting, refining and validating vaccine 3D structure**

To estimate the 3D structure of our build sequence, we utilized the popular and publicly accessible trRosetta online tool (<http://yanglab.nankai.edu.cn/trRosetta/>) [46], a widely used and freely available tool for quick and accurate protein structure prediction. It constructs protein structure by using constrained Rosetta and direct energy minimizations, with the help of a deep neural network to determine the locations and orientations of the residues. GalaxyRefine [47] was utilized to refine the design of the vaccine and its receptor, which is accessible at <http://galaxy.seoklab.org/>. Five high-performing models were evaluated, and the one with the biggest increase in RAMA score was chosen. The RAMPAGE server, which is accessible at <https://www.ccp4.ac.uk/html/rampage.html> [48] was utilized to validate the vaccination model and make predictions about the Ramachandran plot. The Ramachandran plot showcases the representation of individual amino acids in relation to their F and Y values. The top-right quadrant of the plot is indicative of a greater probability of amino acids forming left-handed alpha helices, while the lower-left quadrant

suggests right-handed alpha helices. Amino acids are represented by several  $\beta$ -sheet types in the upper-left quadrant, such as twisted, parallel, and anti-parallel strands [48]. The darker sections indicate the amino acids that are highly favored, while the lighter regions signify that the amino acids are acceptable. On the other hand, the white regions suggest that the amino acids are either low in quality or strictly prohibited.

#### Vaccine structure docking with TLR4

The vaccination has to engage the immune cell receptor in order to stimulate the immune system effectively. Therefore, we conducted docking research to evaluate the interaction between the constructed MEV and TLR4. The crystal structure of TLR4, identified by its PDB ID of 4G8A, was obtained through the protein data library at RCSB. The vaccine structure and TLR4 receptor were docked using Cluspro2.0, a publicly accessible web server, which can be accessed at <https://cluspro.org/tutdock.php>. [49]. It is a web-based program that performs protein–protein docking, to predict the three-dimensional structure of protein complexes based on interaction energies and centers. It is a reliable and widely used tool that can compete with the best human predictor groups in the CAPRI (Critical Assessment of Predicted Interactions) experiment, which is a community-wide blind test of protein–protein docking methods [50].

#### MD simulation

Molecular dynamics studies serve a critical part in any in silico investigation by evaluating the stability of the protein–protein complex [51]. We utilized MD simulations using the Amber 22 setup to investigate the TLR4-vaccine complex's stability. This software allows the exploration of the structural aspects, dynamic behaviors, and molecular interactions inherent to biomolecules across various environmental contexts and under diverse conditions. It facilitates accelerated simulations through the utilization of parallel central processing unit (CPU) or graphics processing unit (GPU) hardware, while also offering access to advanced force fields and computational methods to enhance the level of analysis and accuracy of simulations [52].

#### Immune simulation

The C-ImmSim server (<https://kraken.iac.rm.cnr.it/C-IMMSIM/>) was utilized to simulate the immune system computationally in order to validate the immunological response and immunogenic profile of the constructed design [53]. This online server predicts the immune response in a manner similar to the innate response of the body.

#### Expression analysis

Using the Java Codon Adaptation Tool (JCat) server (<http://www.jcat.de/>) [54], the designed vaccine was reverse-translated and codon-optimized, which facilitated its expression in an *E. coli* expression system. JCat includes three additional factors: the locations of restriction enzyme cleavage, the sites where the bacterial ribosome binds, and the sites where transcription is terminated independently of Rho. JCat has provided indications of protein expression levels in the form of percentages of CG content and the codon adaptation index (CAI). The CAI must be greater than 0.8 and less than or equal to 1.0, and the CG percentage must be between 30 and 70%. Using SnapGene (<https://www.snapgene.com/>), the generated sequence was then inserted into pET-28a (+) to ensure the expression of the vaccine [55].

## Results

#### Retrieval of sequences, phylogenetic categorization and sequence prioritization

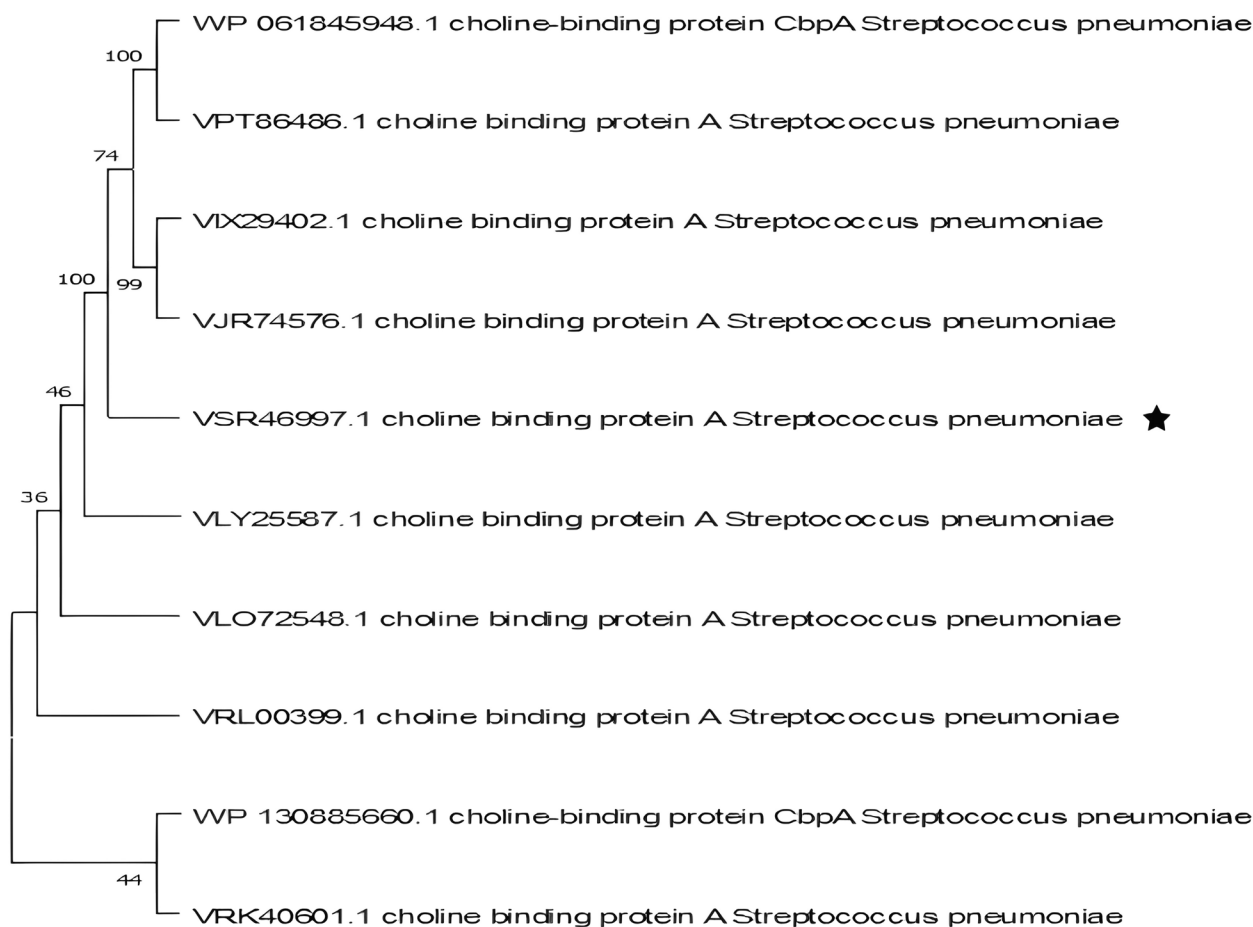
The PspC reference sequence was obtained from the NCBI database (accession number VSR46997.1). Using Blastp, the top 10 sequences were extracted, and then multiple sequence alignment was performed using the MUSCLE v3.6 software. As shown in Fig. 2, a phylogenetic tree was constructed in the software MEGA X in order to visualize the evolutionary relationships between the sequences. Amongst the compared sequences, the protein sequence obtained from NCBI (accession number VSR46997.1) exhibited the highest antigenic potency or immunogenicity, with a notable VaxiJen score of 0.8245. In addition, the Aller-TOP server concurred that no allergens were present in this sequence. The combined use of VaxiJen and AllerTOP confirms this protein's viability as a vaccine target.

#### Physiochemical characterization

With the use of the ExPasy ProtParam program, we were able to determine that this sequence consists of 251 amino acids and has a molecular weight of 28,465.62 [37]. The pI value was 8.52. The computed value of the instability index is 38.17, which indicates the protein is stable. Aliphatic index of 55.82 determined that our protein is a stable one, along with temperature assortment [56]. The  $C_{3465}H_{5437}N_{963}O_{1116}S_{13}$  formula identified the number of sulfur (S), oxygen (O), nitrogen (N), carbon (C), and hydrogen (H). The value of GRAVY was  $-1.083$ .

#### Anticipating B cell epitopes

The prevention of microbial infections is largely dependent on the presence of B-cell epitopes. These epitopes possess modified traits that guide B cells in identifying



**Fig. 2** Phylogenetic tree illustrating the relationship among the top 10 sequences, including our reference protein sequence (marked with a ★ symbol). These sequences were identified using the BlastP search algorithm in a non-redundant database. The Poisson correction method was utilized to compute the evolutionary distances, and these distances are conveyed through the average number of alterations in amino acids per location

and triggering diverse immune responses that facilitate the recognition of particular microbial infections. Using the IEBD analysis tool, 23 linear B cell epitopes were computed with a cutoff score of 0.400. Allergenic and poisonous epitopes were eliminated, and only antigenic, non-allergic, and non-toxic epitopes were chosen. On average, 12 out of 23 B cell epitopes were considered effective. As shown in Table 1, antigenicity ranged from a maximum of 1.1522 to a minimum of 0.1247, with an average of 0.66. The threshold value for antigenic determination of this protein was 0.4, so any number greater than that may be regarded as an antigenic determinant. Twelve antigenic epitopes were selected in the end.

**Anticipating T cell epitopes**

Numerous data entries with IC50 values ranging from 2.86 to 48,867.22 were generated by the calculations of MHC-I and MHC-II-restricted epitopes. These entries were further analyzed, and potential epitopes were

identified by filtering them based on an IC50 value of 250 or less [58]. A low IC50 value indicates that the vaccine epitope can be active at sub-lethal doses, leading to reduced systemic toxicity after administration. This means that the vaccine’s epitopes can produce a strong immune response with a smaller amount of the vaccine [59].

**MHC-I-restricting epitopes**

Table 2 lists the eight epitopes selected as target epitopes from a pool of 450 anticipated MHC-I epitopes. The IC50 values (below 250) of these epitopes were used as selection criteria, along with their high antigenic potential and lack of allergenicity and toxicity.

**MHC-II-restricting epitopes**

Table 3 displays the 15 MHC-II epitopes with the lowest IC50 values among the 586 total. While selecting these epitopes, we only kept the ones that were antigenic,

**Table 1** Using the Kolaskar and Tongaonkar antigenicity approach [57], B-cell epitopes were anticipated with the help of the IEDB B-cell epitope prediction program

| Start | End | Peptide                              | Length | Antigenicity | Allergenicity | Toxicity  |
|-------|-----|--------------------------------------|--------|--------------|---------------|-----------|
| 510   | 515 | WYYLNA                               | 6      | 1.1522       | Non-allergen  | Non-toxin |
| 282   | 297 | LSPSLKLGKKVAEAE                      | 16     | 1.0333       | Non-allergen  | Non-toxin |
| 650   | 656 | WYYLEAS                              | 7      | 1.0143       | Non-allergen  | Non-toxin |
| 69    | 75  | NKKLQLD                              | 7      | 0.7865       | Non-allergen  | Non-toxin |
| 109   | 125 | AELPSKIKAKLDAAFDQ                    | 17     | 0.743        | Non-allergen  | Non-toxin |
| 404   | 414 | EQQPAPAPQP                           | 11     | 0.6585       | Non-allergen  | Non-toxin |
| 7     | 38  | ERKVHYSIRKFSVGVAS<br>VVVASLVMGSVWHAT | 32     | 0.6142       | Non-allergen  | Non-toxin |
| 569   | 575 | SWYYLNA                              | 7      | 0.6079       | Non-allergen  | Non-toxin |
| 549   | 554 | SWYYLN                               | 6      | 0.5103       | Non-allergen  | Non-toxin |
| 529   | 535 | SWYYLNS                              | 7      | 0.3833       | Non-allergen  | Non-toxin |
| 663   | 685 | QWFKVSDKWYYVNLGA<br>LAVNTT           | 23     | 0.3554       | Non-allergen  | Non-toxin |
| 97    | 105 | LRELVLED                             | 9      | 0.1247       | Non-allergen  | Non-toxin |

**Table 2** Epitopes that are restricted by MHC-I and have been predicted using IEDB

| No. | length | Peptide    | All alleles   | Antigenicity | Allergenicity | Toxicity  |
|-----|--------|------------|---|--------------|---------------|-----------|
| 1   | 9      | KVSDKWYYV  | HLA-A*02:06,HLA-A*02:01,HLA-A*02:03,HLA-A*30:01,HLA-A*31:01,HLA-A*68:02,HLA-A*32:01 | 1.2339       | Non-allergen  | Non-Toxin |
| 2   | 9      | RNYPTNTYK  | HLA-A*30:01,HLA-A*31:01,HLA-A*03:01,HLA-A*11:01                                     | 0.6313       | Non-allergen  | Non-toxin |
| 3   | 9      | TLIIKLSAI  | HLA-A*02:03,HLA-B*08:01,HLA-A*02:06   | 0.4851       | Non-allergen  | Non-toxin |
| 4   | 10     | SSNMAKTEYR | HLA-A*31:01,HLA-A*68:01,HLA-A*11:01   | 1.3595       | Non-allergen  | Non-toxin |
| 5   | 9      | SNMAKTEYR  | HLA-A*31:01,HLA-A*68:01,HLA-A*33:01   | 1.1598       | Non-allergen  | Non-toxin |
| 6   | 10     | NANGAMATGW | HLA-B*58:01,HLA-B*57:01,HLA-B*53:01   | 0.7305       | Non-allergen  | Non-toxin |
| 7   | 10     | QVATSSNMAK | HLA-A*11:01,HLA-A*68:01,HLA-A*03:01   | 0.7747       | Non-allergen  | Non-toxin |
| 8   | 9      | LIIKLSAIK  | HLA-A*68:01,HLA-A*11:01,HLA-A*03:01   | 0.641        | Non-allergen  | Non-toxin |
| 9   | 10     | TLIIKLSAIK | HLA-A*68:01,HLA-A*11:01,HLA-A*03:01   | 0.485        | Non-allergen  | Non-toxin |

non-allergenic, and non-toxic and discarded the ones that did not meet these criteria.

**Population coverage**

Using the IEDB tool, the coverage of MHC-I and MHC-II alleles with different epitopes was assessed. The analysis revealed that the chosen MHC-I epitopes had a population coverage of 76.34%, while the chosen MHC-II epitopes covered 69.26%. After a comprehensive evaluation of every epitope, the total coverage was an impressive 92.73%, as illustrated in Fig. 3. These results suggest that the chosen epitopes for the vaccine have the potential to be effective worldwide, with only minor variations observed in different ethnic groups.

**Construction of the vaccine**

We opted to use a combination of 12 B cell epitopes, 9 MHC-I epitopes, and 15 MHC-II epitopes to build the MEV. To increase the vaccine’s effectiveness, we also

employed linkers (EAAAK, CPGPG, and AAY) to join the adjuvant to the B cell epitopes, the B cell epitopes to the MHC-I, and the MHC-II epitopes to each other. Moreover, a 6×His tag was included in the vaccine’s sequence to aid in protein purification and characterization. The vaccine’s structure is depicted in the following Table 4.

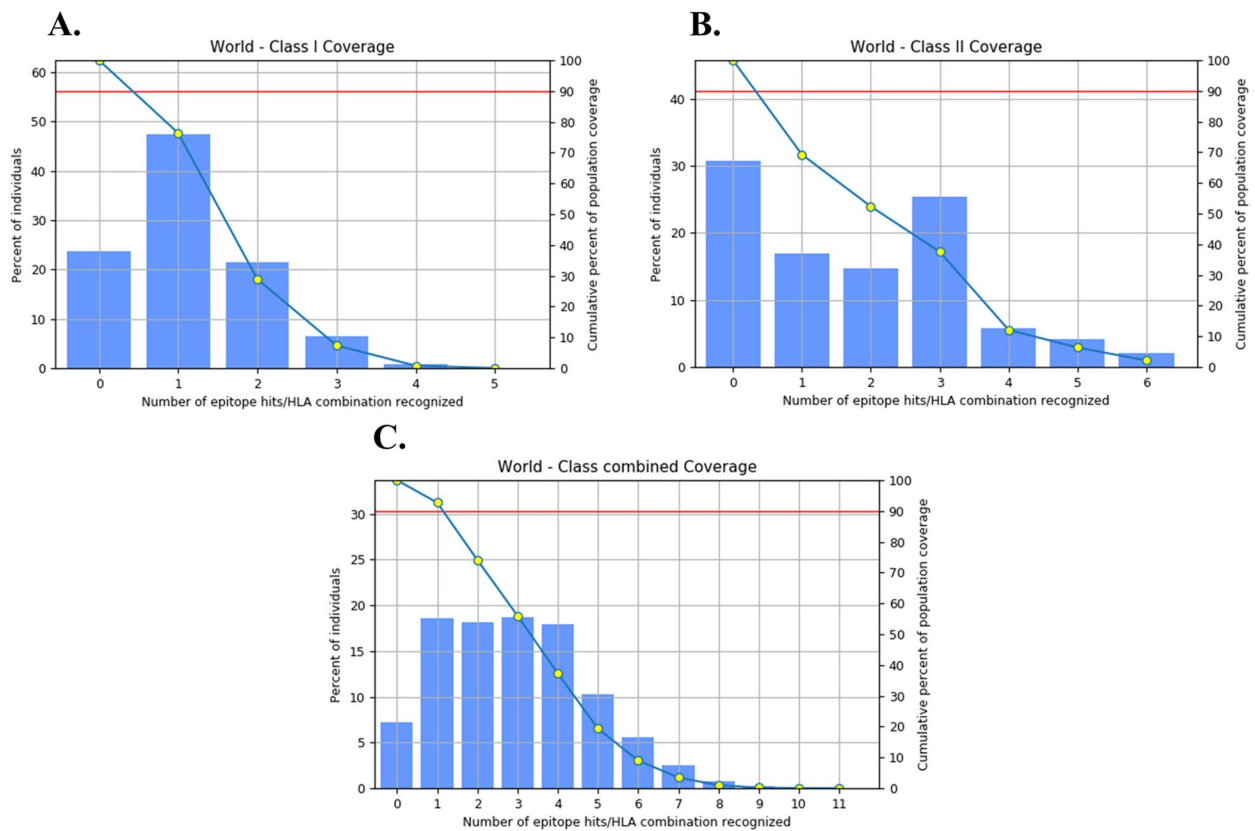
> vaccine protein.

E A A A K W Y Y L N A L P S P S L K L G K K V A E A E W -  
Y Y L E A S N K K L Q L D A E L P S K I K A K L D A A F D Q E Q P Q -  
P A P A P Q P E R K V H Y S I R K F S V G V A S V V V A S L V M G S -  
V V H A T S W Y Y L N A S W Y Y L N S W Y Y L N S Q W F K V S -  
D K W Y Y V N L G A L A V N T T L R E L N V L E D C P G P G -  
K V S D K W Y Y V R N Y P T N T Y K T L I I K L S A I S S N M A K -  
T E Y R S N M A K T E Y R N A N G A M A T G W Q V A T S S N -  
M A K L I I K L S A I K T L I I K L S A I K A A Y S W Y Y L N A N G A -  
M A T G W S W Y Y L N A S G A M A T G W K W Y Y V N L G A -  
L A V N T G S W Y Y L N A N G A M A T G G S W Y Y L N S N G A -  
M A T G S W Y Y L N S N G A M A T G W W Y Y L N A S G A M A T G

**Table 3** IEDB-predicted MHC-II-restricted epitope

| No. | Peptide         | All alleles   | Antigenicity | Allergenicity | Toxicity  |
|-----|-----------------|---|--------------|---------------|-----------|
| 1   | SWYYLNANGAMATGW | HLA-DRB1*01:01,HLA-DRB3*02:02,HLA-DRB1*13:02,HLA-DRB1*09:01,HLA-DRB1*04:01,HLA-DRB1*11:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB5*01:01,HLA-DRB3*01:01,HLA-DRB1*04:05,HLA-DRB1*07:01   | 0.6068       | Non-allergen  | Non-toxin |
| 2   | SWYYLNASGAMATGW | HLA-DRB1*01:01,HLA-DRB3*02:02,HLA-DRB1*09:01,HLA-DRB1*11:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*07:01,HLA-DRB1*13:02,HLA-DRB3*01:01,HLA-DRB1*04:01,HLA-DRB5*01:01,HLA-DRB1*04:05   | 0.502        | Non-allergen  | Non-toxin |
| 3   | KWYYVNLGALAVNT  | HLA-DRB1*01:01,HLA-DRB1*09:01,HLA-DRB3*02:02,HLA-DRB1*13:02,HLA-DPA1*03:01/DPB1*04:02,HLA-DRB5*01:01,HLA-DRB1*12:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*04:05,HLA-DRB1*07:01,HLA-DRB1*04:01,HLA-DRB1*11:01,HLA-DPA1*01:03/DPB1*02:01,HLA-DQA1*01:02/DQB1*06:02 | 0.433        | Non-allergen  | Non-toxin |
| 4   | GSWYYLNANGAMATG | HLA-DRB1*01:01,HLA-DRB3*02:02,HLA-DRB1*13:02,HLA-DRB1*09:01,HLA-DRB1*04:01,HLA-DRB1*11:01,HLA-DRB3*01:01,HLA-DRB5*01:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*04:05,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*07:01,HLA-DRB1*15:01                                      | 0.6041       | Non-allergen  | Non-toxin |
| 5   | GSWYYLNSNGAMATG | HLA-DRB3*02:02,HLA-DRB1*01:01,HLA-DRB1*04:01,HLA-DRB1*09:01,HLA-DRB1*13:02,HLA-DRB1*11:01,HLA-DRB3*01:01,HLA-DRB1*04:05,HLA-DRB5*01:01,HLA-DRB1*07:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*15:01                                      | 0.5133       | Non-allergen  | Non-toxin |
| 6   | SWYYLNSNGAMATGW | HLA-DRB3*02:02,HLA-DRB1*01:01,HLA-DRB1*13:02,HLA-DRB1*09:01,HLA-DRB1*04:01,HLA-DRB1*11:01,HLA-DRB3*01:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB5*01:01,HLA-DRB1*07:01,HLA-DRB1*04:05   | 0.516        | Non-allergen  | Non-toxin |
| 7   | WYYLNASGAMATGWA | HLA-DRB1*01:01,HLA-DRB3*02:02,HLA-DRB1*09:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*11:01,HLA-DRB1*07:01,HLA-DRB1*13:02,HLA-DRB3*01:01,HLA-DRB1*04:01,HLA-DRB5*01:01,HLA-DRB1*04:05   | 0.668        | Non-allergen  | Non-toxin |
| 8   | WYYLNANGAMATGWL | HLA-DRB1*01:01,HLA-DRB3*02:02,HLA-DRB1*13:02,HLA-DRB1*09:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*04:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB1*11:01,HLA-DRB5*01:01,HLA-DRB3*01:01,HLA-DRB1*07:01  | 0.5875       | Non-allergen  | Non-toxin |
| 9   | WYYLNSNGAMATGWL | HLA-DRB3*02:02,HLA-DRB1*01:01,HLA-DRB1*13:02,HLA-DRB1*09:01,HLA-DRB1*04:01,HLA-DRB1*11:01,HLA-DQA1*05:01/DQB1*03:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DRB3*01:01,HLA-DRB5*01:01,HLA-DRB1*07:01  | 0.4967       | Non-allergen  | Non-toxin |
| 10  | TWYYLEASGAMKASQ | HLA-DRB1*01:01,HLA-DRB5*01:01,HLA-DRB1*09:01,HLA-DRB3*01:01,HLA-DRB1*07:01,HLA-DRB1*04:01,HLA-DQA1*01:02/DQB1*06:02,HLA-DQA1*05:01/DQB1*03:01,HLA-DRB1*13:02,HLA-DRB1*11:01,HLA-DRB1*04:05  | 0.734        | Non-allergen  | Non-toxin |
| 11  | VNGSWYYLNASGAMA | HLA-DRB1*09:01,HLA-DRB1*09:01,HLA-DRB3*02:02,HLA-DRB3*01:01,HLA-DRB1*07:01,HLA-DRB1*11:01,HLA-DRB1*13:02,HLA-DRB1*04:01,HLA-DRB5*01:01,HLA-DRB1*04:05,HLA-DQA1*05:01/DQB1*03:01   | 0.5358       | Non-allergen  | Non-toxin |
| 12  | TLIKLSAIKTEYLR  | HLA-DRB1*01:01,HLA-DRB4*01:01,HLA-DRB1*12:01,HLA-DRB1*09:01,HLA-DRB1*15:01,HLA-DRB5*01:01,HLA-DRB1*04:05,HLA-DRB1*13:02,HLA-DRB1*07:01,HLA-DRB1*08:02,HLA-DPA1*03:01/DPB1*04:02,HLA-DPA1*02:01/DPB1*05:01,HLA-DRB1*11:01,HLA-DRB1*04:01                       | 0.5855       | Non-allergen  | Non-toxin |
| 13  | LIKLSAIKTEYLRE  | HLA-DRB1*01:01,HLA-DRB4*01:01,HLA-DRB1*12:01,HLA-DRB1*09:01,HLA-DRB1*04:05,HLA-DRB1*15:01,HLA-DRB5*01:01,HLA-DRB1*13:02,HLA-DRB1*07:01,HLA-DPA1*03:01/DPB1*04:02,HLA-DPA1*02:01/DPB1*05:01,HLA-DRB1*08:02,HLA-DPA1*02:01/DPB1*01:01,HLA-DRB1*04:01            | 0.5094       | Non-allergen  | Non-toxin |
| 14  | IKLSAIKTEYLRELN | HLA-DRB1*01:01,HLA-DRB4*01:01,HLA-DPA1*03:01/DPB1*04:02,HLA-DRB1*04:05,HLA-DRB1*15:01,HLA-DRB1*07:01,HLA-DRB1*09:01,HLA-DPA1*02:01/DPB1*01:01,HLA-DPA1*01:03/DPB1*04:01,HLA-DPA1*02:01/DPB1*05:01,HLA-DPA1*01:03/DPB1*02:01                                   | 0.7839       | Non-allergen  | Non-toxin |
| 15  | KLSAIKTEYLRELVN | HLA-DRB1*01:01,HLA-DPA1*03:01/DPB1*04:02,HLA-DRB4*01:01,HLA-DRB1*15:01,HLA-DRB1*07:01,HLA-DRB1*09:01,HLA-DPA1*01:03/DPB1*04:01,HLA-DPA1*02:01/DPB1*01:01,HLA-DPA1*01:03/DPB1*02:01,HLA-DPA1*02:01/DPB1*05:01  | 0.7314       | Non-allergen  | Non-toxin |





**Fig. 3** The coverage of the target alleles with the potential epitopes was evaluated for the population. The figure depicts the extent of coverage for MHC-I epitopes (A), MHC-II epitopes (B), and both types of epitopes together (C)

WAWYYLNANGAMATGWLWYYLNSNGAMATGWLWYYLEASGAMKASQVNGSWYYLNASGAMATLIKLSAIKTEYLRLIILKLSAIKTEYLREIKLSAIKTEYLRELNKLSAIKTEYLRELVNHHHHH.

**Analysis of physicochemical and immunological properties**

The vaccine construct consists of 438 building blocks, and the computed molecular weight was 49,134.28. The predicted isoelectric point (PI) of the protein is 9.51, indicating a positive charge, as isoelectric points above 7.0 are positively charged. The protein was classified as stable with an instability index (II) of 25.61, as determined by ProtParam. The aliphatic index is 82.62, meaning that it can withstand a wide range of temperatures. The calculated grand average of hydropathicity (GRAVY) using the chemical formula  $C_{2482}H_{3708}N_{626}O_{684}S_{16}$  was -0.170. According to the Vexijen v2.0 server, the vaccine sequence we created is likely an antigen with an antigenic score of 0.6872. Afterward, Allertop v2.0 verifies that the structure is not an allergen, while Toxinpred confirms that it is not toxic.

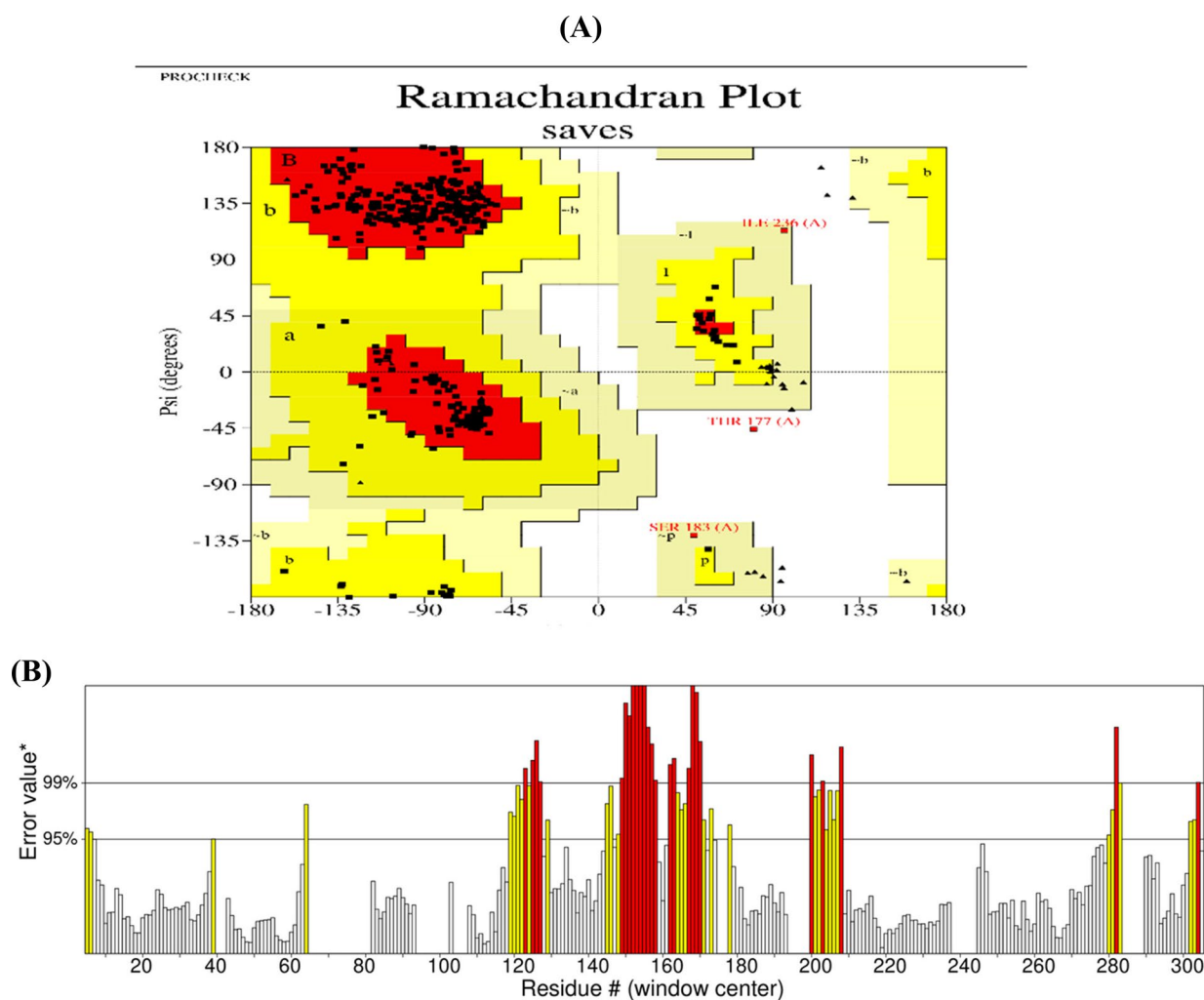
**Predicting, improving and verifying the 3D structure of vaccine**

The trRosetta server created five different 3-dimensional structures of the intended vaccination sequence, and the best one was selected. The protein structure was stabilized and given a higher quality score on the SAVES server after being refined by the Galaxy server. Analyzing the Ramachandran plot for the revised structure (Fig. 4A) showed that 91.4% of the 3D residues were in the optimal area. Similar patterns can be seen in the ERRAT program, where the total quality factor was raised to 84.53 (Fig. 4B). PyMol 2 was used to generate the 3D representation of the final vaccine structure shown in Fig. 5.

**Molecular docking of the TLR4 receptor-vaccine construct**

Ten different models were analyzed using Cluspro docking analysis predictions. After a visual comparison of all ten docking models using Pymol, the model with the least amount of energy consumption and the highest number of contributors to the cluster’s formation was selected (depicted in Fig. 6). This model produced a satisfactory



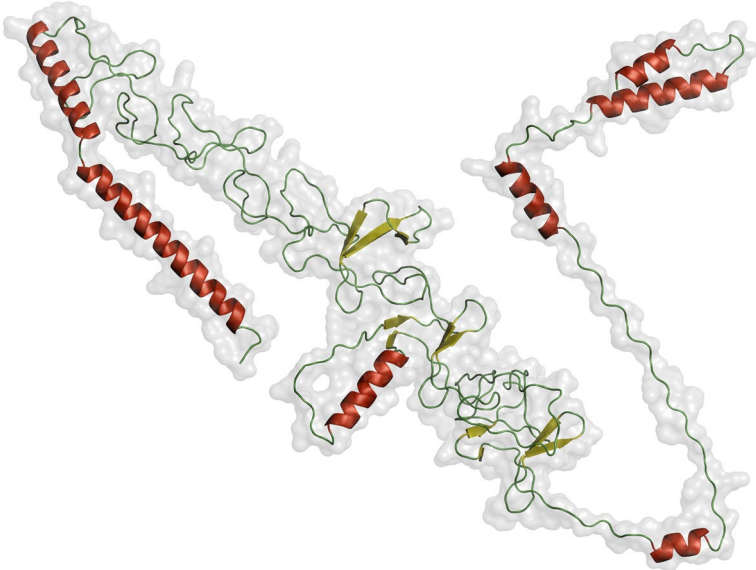


**Fig. 4** Tertiary structural verification of the vaccine construct. The Ramachandran plot of the modified model (A) indicates that 91.4% of the 3D residues are situated within the optimal region, while (B) represents an assessment of the improved model's ERRAT quality factor (84.53)

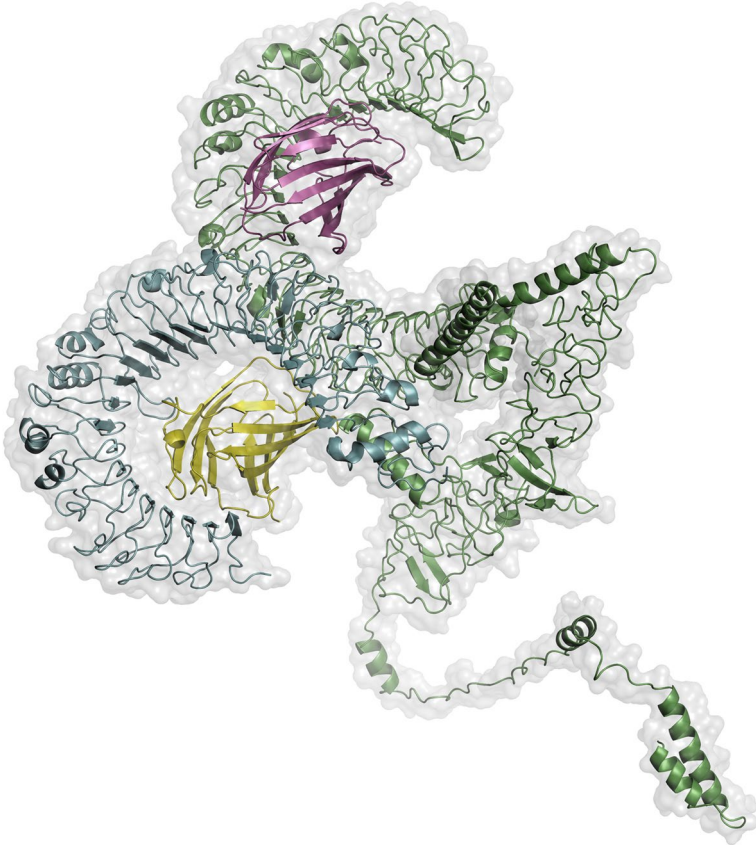
among children and the elderly, and is categorized as a high-burden disease [61]. Although currently available pneumococcal vaccines, such as PPV and PCV, can prevent various types of pneumococcal disease, they have been known to fail in some cases due to serotype replacement [62]. The production complications and high costs associated with PCVs have rendered them unaffordable, particularly in developing nations [63]. Furthermore, there has been an increase in antibiotic resistance among the serotype replacement strains. However, immunoinformatic approaches can be used to address these issues. Using *in silico* methods can reduce the amount of time and money required for experiments. Immunoinformatics-assisted epitope identification has multiple applications in epitope mapping, including advancements in peptide-based vaccine research, characterization of immunological processes, and prediction of epitopes

used in the diagnosis of disease [64]. Epitope-based vaccines are an attractive and prospective new method for developing vaccines, as they employ only fragments of peptides that are known to be highly immunogenic and capable of evoking immune responses [65]. Therefore, our approach aimed to identify a potential epitope-based pneumococcal vaccine candidate that would ideally be conserved across most serotypes, have broad population coverage, and induce T cell-dependent immune responses.

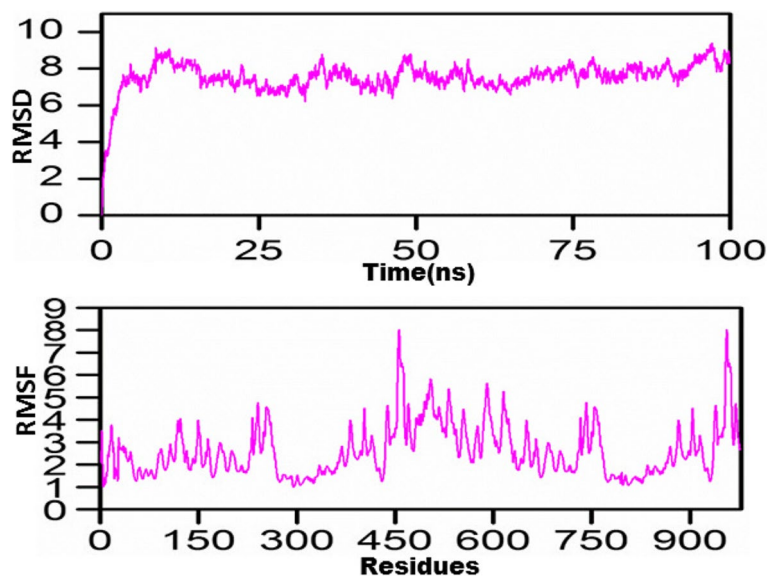
Vaccines based on epitopes must contain B and T cell epitopes that induce potent immune responses against a specific infection [66]. Traditional wet-lab approaches for identifying potential B and T cell epitopes involve experimental screening of numerous active and inactive epitopes, which can be time-consuming and costly. As an alternative, computational methods offer a cost-effective,



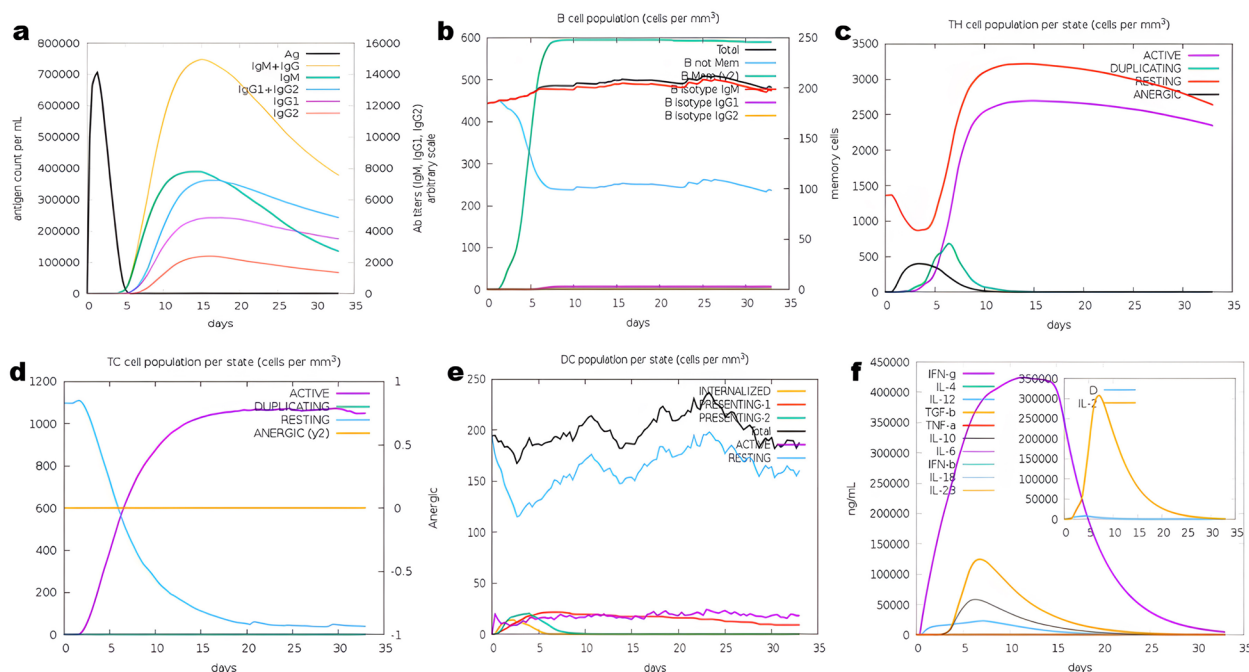
**Fig. 5** Rendering of vaccine's 3D structure in PyMol 2



**Fig. 6** PyMol visualization of cluspro-derived docked complex with 32 clusters and a minimum energy score of  $-1133.3$



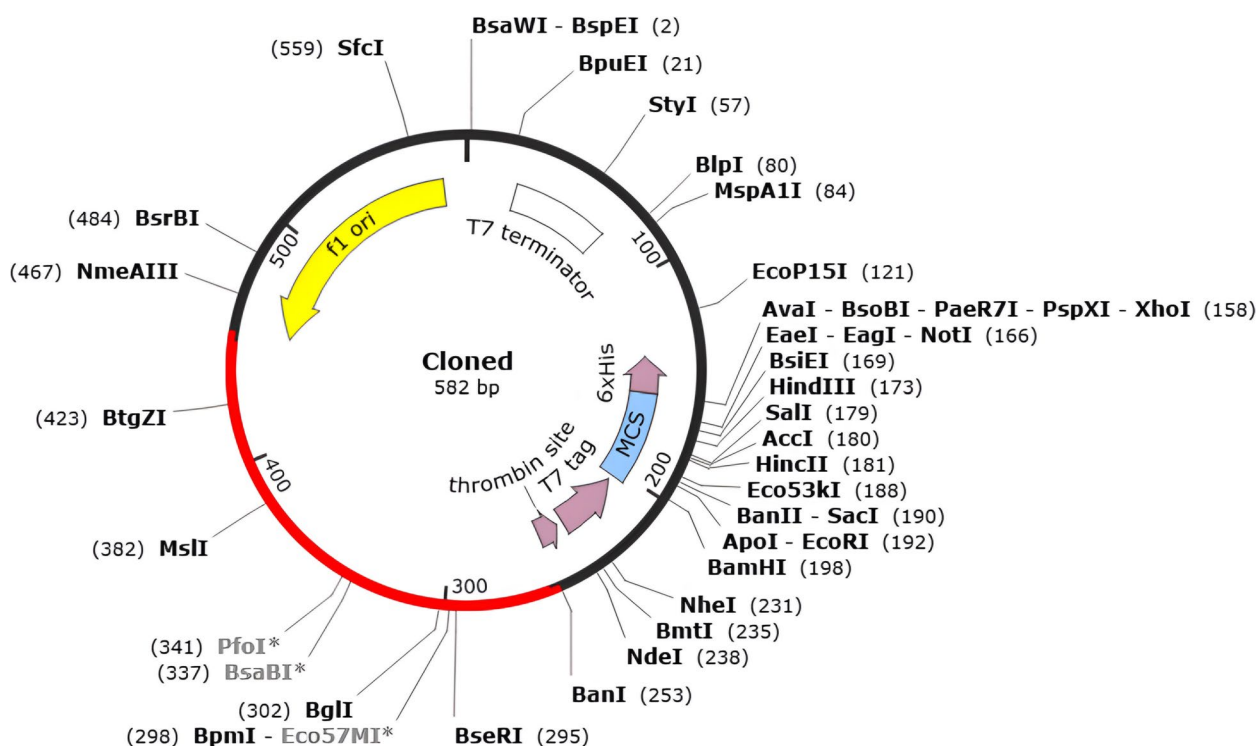
**Fig. 7** MD simulation findings of the vaccine construct for 100 ns. RMSD (A) and RMSF (B) were elucidated



**Fig. 8** Vaccine immune simulation through the C-ImmSim server

rapid, reliable, and accurate approach [67, 68]. In this regard, the use of a consensus prediction strategy is proven to be more reliable and robust compared to individual prediction methods [69]. In order to construct a MEV, first B and T-cell epitopes were anticipated using reliable databases. Then the vaccine formulation included linkers such as EAAAK, AAY, and CPGPG, which facilitated better and longer-lasting protection. A significant

challenge with epitope vaccines is their vulnerability to degradation by proteases in the body [70]. To address this issue, the vaccine sequences were inserted into the 50S ribosomal chromosome as an adjuvant. According to computational analysis, the manufactured vaccine has been shown to be non-allergenic and highly antigenic (0.68). The proposed vaccine has a high probability of being thermostable, as estimated by the aliphatic index



**Fig. 9** The final MEV was in silico cloned using pET28a (+). The vector is depicted by a black circle, while the insertion location for the vaccination is shown by the red area

formula [71]. The proposed vaccine exhibits hydrophilicity and preferred interactions with water molecules, as shown by its GRAVY index of  $-0.162$  [72]. The designed vaccine is stable, as shown by its instability index of 25.62, which is lower than 40. The utilization of a 3D structural model provides valuable insights into protein dynamics, ligand interactions, function, and spatial organization. Substantial refinement of the vaccine's formulation has notably improved its desired properties. The majority of the residues were found to be within the permissible range when the vaccine design was examined using a Ramachandran plot. This indicates that the design of the vaccine is of high quality. The subsequent critical step in validating a vaccine involves molecular docking. In order to generate a powerful immune response in the host, a low binding energy score between the receptor and ligand is required. The docking studies conducted in this investigation revealed a significantly reduced binding energy of  $-1133.3$  kcal/mol, indicating a strong interaction between the manufactured vaccine and the TLR4 receptor molecule. We also conducted MD simulations, and the analysis showed our vaccine maintains its structural stability. The immune simulation graph demonstrates a notable increase in IgM production following the administration of our designed vaccine, indicating the occurrence of a primary immune response. Additionally,

the enhanced expression of immunoglobulins in B cells corresponded to a reduction in antigen concentration. JCat was utilized to modify the codons of the vaccine to improve its expression in *E. coli*. The vaccine's construction has a CAI value of 0.95 and a GC content of 64.42%. Our findings are excellent since CAI values over 0.8 and GC contents between 30 and 70% are thought to be good for expression. The purpose of the *Escherichia coli* in silico cloning was to lay the groundwork for later wet laboratory studies by other researchers attempting to build an effective vaccine. However, more validation via in vitro and in vivo experiments using animal models is required to confirm the effectiveness of the developed MEV.

### Conclusion

In this current study, we have developed a vaccine candidate against *Streptococcus pneumoniae* using immunoinformatic techniques. We used various online tools and databases to select suitable proteins, predict B and T cell epitopes to construct vaccine candidates and evaluate their physicochemical properties, antigenicity, allergenicity, and stability. In order to evaluate the vaccine's interaction and binding affinity, we also ran molecular docking and molecular dynamics simulations with human toll-like receptor 4 (TLR-4). Our results suggested that our vaccine candidate has high immunogenicity, broad

population coverage, and a low risk of adverse effects. It also showed a strong and stable interaction with human TLR-4, indicating the potential to elicit both adaptive and innate immunity. However, these results are based on computational predictions and need to be validated experimentally. Therefore, we recommend further testing of our vaccine candidate in appropriate tissue culture and animal models to confirm its efficacy and safety before proceeding to clinical trials.

#### Abbreviations

|      |   |
|------|---|
| MHC  | Major histocompatibility complex              |
| TLR  | Toll-like receptor                            |
| NCBI | National Center for Biotechnology Information |
| IEDB | Immune Epitope Database                       |
| CAI  | Codon Adaptation Index                        |
| GC   | Guanine-cytosine                              |

#### Acknowledgements

The authors acknowledge the encouragement and support of the Government of the People's Republic of Bangladesh. We would like to express special thanks of gratitude to Architect Yeafesh Osman, honorable minister of Science and Technology. We are very much thankful to the secretary, Ministry of Science and Technology, and Chairman Bangladesh Council of Scientific and Industrial Research.

#### Authors' contributions

MN designed the study and experimental work. MN, MS, and SA conceptualized the study, provided overall guidance, and wrote the manuscript. LM and JINO participated in the drafting of the manuscript. TAB, MHS, BG, AH, SB, and SA read and participated in supervising and reviewing the draft, thoroughly checking and revising the manuscript for necessary changes in format. All authors read and approved the final version of the manuscript. SA submits the final manuscript.

#### Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

#### Availability of data and materials

The datasets supporting the conclusions of this article are included within the article.

#### Declarations

##### Ethics approval and consent to participate

Not applicable. No impact on ethical standards in this study, and there is no human or animal involvement.

##### Consent for publication

The submitting research article "In Silico Design of an Epitope-Based Vaccine against PspC in *Streptococcus pneumoniae* Using Reverse Vaccinology" for publication in your journal of repute, is a unique article and nobody did it earlier. The authors also declared that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

##### Competing interests

All authors declare that they have no competing interests.

##### Author details

<sup>1</sup>Department of Microbiology, Jagannath University, Dhaka 1100, Bangladesh. <sup>2</sup>State key laboratories of chemical Resources Engineering, Beijing University of Chemical Technology, Beijing 100029, China. <sup>3</sup>Department of Biology, Indiana State University, Terre Haute, United States. <sup>4</sup>Departamento de Biofísica e Farmacologia, Universidade Federal do Rio Grande do Norte, 59072-970, Natal,

RN, Brazil. <sup>5</sup>Bangladesh Council of Scientific and Industrial Research (BCSIR), Dhaka-1205, Bangladesh.

Received: 6 July 2023 Accepted: 14 November 2023

Published online: 12 December 2023

#### References

- K. Enneth M C and I. Ntosh, Community acquired pneumonia in children, 2002. [Online]. Available: [www.nejm.org](http://www.nejm.org)
- D. Bogaert, R. de Groot, and P. W. M. Hermans, Dynamics of nasopharyngeal colonisation, 2004. [Online]. Available: <http://infection.thelancet.com>
- Kadioglu A, Weiser JN, Paton JC, Andrew PW (2008) The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* 6(4):288–301. <https://doi.org/10.1038/nrmicro1871>
- Johnson HL et al (2010) Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* 375:1969–1987. <https://doi.org/10.1016/S0140>
- Rudan I, Boschi-Pinto C, Biloglav Z, Mulholland K, Campbell H (2008) Epidemiology and etiology of childhood pneumonia. *Bull World Health Organ* 86(5):408–416. <https://doi.org/10.2471/BLT.07.048769>
- S. G. Jenkins, S. D. Brown, and D. J. Farrell, Trends in antibacterial resistance among *Streptococcus pneumoniae* isolated in the USA: Update from PROTEKT US years 1–4, *Ann Clin Microbiol Antimicrob*, vol. 7, Jan. 2008, <https://doi.org/10.1186/1476-0711-7-1>
- C. Y. Wang et al., Antibiotic resistance profiles and multidrug resistance patterns of *Streptococcus pneumoniae* in pediatrics: a multicenter retrospective study in mainland China, *Medicine (United States)*, vol. 98, no. 24, Jun. 2019, <https://doi.org/10.1097/MD.00000000000015942>
- Malley R (2010) Antibody and cell-mediated immunity to *Streptococcus pneumoniae*: Implications for vaccine development. *J Mol Med* 88(2):135–142. <https://doi.org/10.1007/s00109-009-0579-4>
- T. Lagousi, P. Basdeki, J. Routsias, and V. Spoulou, Novel protein-based pneumococcal vaccines: Assessing the use of distinct protein fragments instead of full-length proteins as vaccine antigens, *Vaccines*, vol. 7, no. 1. MDPI AG, 2019. <https://doi.org/10.3390/vaccines7010009>
- Pilishvili T et al (2010) Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *J Infect Dis* 201(1):32–41. <https://doi.org/10.1086/648593>
- Oyarzún P, Kobe B (2016) Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. *Hum Vaccin Immunother* 12(3):763–767. <https://doi.org/10.1080/21645515.2015.1094595>
- J. Aceil and F. Y. Avci, Pneumococcal surface proteins as virulence factors, immunogens, and conserved vaccine targets, *Frontiers in Cellular and Infection Microbiology*, vol. 12, Frontiers Media S.A., May 12, 2022. <https://doi.org/10.3389/fcimb.2022.832254>
- C. C. Daniels, P. D. Rogers, and C. M. Shelton, A review of pneumococcal vaccines: current polysaccharide vaccine recommendations and future protein antigens, 2016. [Online]. Available: [www.jppt.org](http://www.jppt.org)
- A. M. Berry and J. C. Paton, Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence Proteins, 2000. [Online]. Available: <http://iai.asm.org/>
- Kerr AR et al (2006) The contribution of PspC to pneumococcal virulence varies between strains and is accomplished by both complement evasion and complement-independent mechanisms. *Infect Immun* 74(9):5319–5324. <https://doi.org/10.1128/IAI.00543-06>
- Guerra AJ, Dann CE, Giedroc DP (2011) Crystal structure of the zinc-dependent MarR family transcriptional regulator AdcR in the Zn(II)-bound state. *J Am Chem Soc* 133(49):19614–19617. <https://doi.org/10.1021/ja2080532>
- C. D. Plumptre, A. D. Ogunniyi, and J. C. Paton, Vaccination against *Streptococcus pneumoniae* using truncated derivatives of polyhistidine triad protein D, *PLoS One*, vol. 8, no. 10, Oct. 2013, <https://doi.org/10.1371/journal.pone.0078916>
- D. R. Cundell, B. J. Pearce, J. Sandros, A. M. Naughton, and H. R. Masure, Peptide permeases from *Streptococcus pneumoniae* affect adherence to

- eucaryotic cells, 1995. [Online]. Available: <https://journals.asm.org/journal/iai>
19. Jomaa M, Yuste J, Paton JC, Jones C, Dougan G, Brown JS (2005) Antibodies to the iron uptake ABC transporter lipoproteins PiaA and PiuA promote opsonophagocytosis of *Streptococcus pneumoniae*. *Infect Immun* 73(10):6852–6859. <https://doi.org/10.1128/IAI.73.10.6852-6859.2005>
  20. Giefing C et al (2008) Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. *J Exp Med* 205(1):117–131. <https://doi.org/10.1084/jem.20071168>
  21. Daniels CC et al (2010) 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* 78(5):2163–2172. <https://doi.org/10.1128/IAI.01199-09>
  22. A. Parihar, S. Malviya, and R. Khan, Immunoinformatics and reverse vaccinomic approaches for effective design, in *Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV2 Infection: Revolutionary Strategies to Combat Pandemics*, Elsevier, 2022, pp. 357–378. <https://doi.org/10.1016/B978-0-323-91172-6.00004-2>
  23. M. Shahab et al., Computational design of medicinal compounds to inhibit RBD-hACE2 interaction in the Omicron variant: unveiling a vulnerable target site, *Inform Med Unlocked*, vol. 40, Jan. 2023, <https://doi.org/10.1016/j.jimu.2023.101281>
  24. U. Farooq et al., Arbutin stabilized silver nanoparticles: synthesis, characterization, and its catalytic activity against different organic dyes, *Catalysts*, vol. 12, no. 12, Dec. 2022, <https://doi.org/10.3390/catal12121602>
  25. S. Akter et al., Immunoinformatics approach to epitope-based vaccine design against the SARS-CoV-2 in Bangladeshi patients, *J Genet Eng Biotechnol* vol. 20, no. 1, Dec. 2022, <https://doi.org/10.1186/s43141-022-00410-8>
  26. M. Shahab, C. Hayat, R. Sikandar, G. Zheng, and S. Akter, In silico designing of a multi-epitope vaccine against *Burkholderia pseudomallei*: reverse vaccinology and immunoinformatics, *J Genet Eng Biotechnol* vol. 20, no. 1, Dec. 2022, <https://doi.org/10.1186/s43141-022-00379-4>
  27. S. Bibi et al., In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology, *Sci Rep*, vol. 11, no. 1, Dec. 2021, <https://doi.org/10.1038/s41598-020-80899-6>
  28. F. Etmiani, A. Etmiani, S. O. Hasson, H. K. Judi, S. Akter, and M. Saki, In silico study of inhibition effects of phytochemicals from four medicinal plants against the *Staphylococcus aureus*  $\beta$ -lactamase, *Inform Med Unlocked*, vol. 37, Jan. 2023, <https://doi.org/10.1016/j.jimu.2023.101186>
  29. J. Jos' et al., A synthetic malaria vaccine elicits a potent CD8 + and CD4 + T lymphocyte immune response in humans. Implications for vaccination strategies, 2001
  30. Bourdette DN et al (2005) A highly immunogenic trivalent T cell receptor peptide vaccine for multiple sclerosis. *Mult Scler* 11(5):552–561. <https://doi.org/10.1191/1352458505ms1225oa>
  31. L. Knutson Keith, Schiffman Kathy, and Mary L. Disis, Immunization with a HER-2/neu helper peptide vaccine generates HER-2/neu CD8 T-cell immunity in cancer patients, *J Clin Invest*, 2001
  32. S. F. Altschup, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, Basic local alignment search tool, 1990
  33. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797. <https://doi.org/10.1093/nar/gkh340>
  34. Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* 38(7):3022–3027. <https://doi.org/10.1093/molbev/msab120>
  35. I. A. Doytchinova and D. R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, *BMC Bioinformatics*, vol. 8, Jan. 2007, <https://doi.org/10.1186/1471-2105-8-4>
  36. I. Dimitrov, D. R. Flower, and I. Doytchinova, AllerTOP - a server for in silico prediction of allergens, *BMC Bioinformatics*, vol. 14, no. SUPPL6, Apr. 2013, <https://doi.org/10.1186/1471-2105-14-S6-S4>
  37. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31(13):3784–3788. <https://doi.org/10.1093/nar/gkg563>
  38. E. Gasteiger et al., Protein Identification and Analysis Tools on the ExPASy Server, in *The Proteomics Protocols Handbook*, Humana Press, 2005, pp. 571–607. <https://doi.org/10.1385/1-59259-890-0.571>
  39. Erik J, Larsen P, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. *Immunome Res* 24:2. <https://doi.org/10.1186/1745-7580-2-2>
  40. W. Fleri et al., The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design, *Front Immunol*, vol. 8, no. MAR. Frontiers Research Foundation, Mar. 14, 2017. <https://doi.org/10.3389/fimmu.2017.00278>
  41. Nielsen M et al (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 12(5):1007–1017. <https://doi.org/10.1110/ps.0239403>
  42. Jensen KK et al (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154(3):394–406. <https://doi.org/10.1111/imm.12889>
  43. Adhikari UK, Rahman MM (2017) Overlapping CD8 + and CD4 + T-cell epitopes identification for the progression of epitope-based peptide vaccine from nucleocapsid and glycoprotein of emerging Rift Valley fever virus using immunoinformatics approach. *Infect Genet Evol* 56:75–91. <https://doi.org/10.1016/j.meegid.2017.10.022>
  44. W. Li, M. D. Joshi, S. Singhanian, K. H. Ramsey, and A. K. Murthy, Peptide vaccine: progress and challenges, *Vaccines*, vol. 2, no. 3. MDPI AG, pp. 515–536, Jul. 02, 2014. <https://doi.org/10.3390/vaccines2030515>
  45. S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, and G. P. S. Raghava, In silico approach for predicting toxicity of peptides and proteins, *PLoS One*, vol. 8, no. 9, Sep. 2013, <https://doi.org/10.1371/journal.pone.0073957>
  46. Z. Du et al., The trRosetta server for fast and accurate protein structure prediction, *Nature Protocols*, vol. 16, no. 12. Nature Research, pp. 5634–5651, Dec. 01, 2021. <https://doi.org/10.1038/s41596-021-00628-9>
  47. J. Ko, H. Park, L. Heo, and C. Seok, GalaxyWEB server for protein structure prediction and refinement, *Nucleic Acids Res*, vol. 40, no. W1, Jul. 2012, <https://doi.org/10.1093/nar/gks493>
  48. S. C. Lovell et al., Structure validation by C geometry: , and C Deviation, 2003. [Online]. Available: <http://www.crysl>
  49. Desta IT, Porter KA, Xia B, Kozakov D, Vajda S (2020) Performance and its limits in rigid body protein-protein docking. *Structure* 28(9):1071–1081. <https://doi.org/10.1016/j.str.2020.06.006>
  50. Kozakov D et al (2017) The ClusPro web server for protein-protein docking. *Nat Protoc* 12(2):255–278. <https://doi.org/10.1038/nprot.2016.169>
  51. O. M. H. Salo-Ahen et al., Molecular dynamics simulations in drug discovery and pharmaceutical development, *Processes*, vol. 9, no. 1. MDPI AG, pp. 1–63, 2021. <https://doi.org/10.3390/pr9010071>
  52. Case DA et al (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26(16):1668–1688. <https://doi.org/10.1002/jcc.20290>
  53. F. Castiglione, D. Deb, A. P. Srivastava, P. Liò, and A. Liso, From infection to immunity: understanding the response to SARS-CoV2 through in-silico modeling, *Front Immunol*, vol. 12, Sep. 2021, <https://doi.org/10.3389/fimmu.2021.646972>
  54. Grote A et al (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 33(SUPPL):2. <https://doi.org/10.1093/nar/gki376>
  55. SnapGene | Software for everyday molecular biology. <https://www.snapgene.com/> (accessed Apr. 05, 2023)
  56. J. Kyte and R. F. Doolittle, A simple method for displaying the hydropathic character of a protein, 1982
  57. A. S. Kolaskar and P. C. Tongaonkar, A semi-empirical method for prediction of antigenic determinants on protein antigens, 1990
  58. C. Berrouet, N. Dorilas, K. A. Rejniak, and N. Tuncer, Comparison of drug inhibitory effects (IC 50) in monolayer and spheroid cultures, *Bull Math Biol*, vol. 82, no. 6, Jun. 2020, <https://doi.org/10.1007/s11538-020-00746-7>
  59. A. Banerjee, D. Santra, and S. Maiti, Energetics and IC50 based epitope screening in SARS CoV-2 (COVID 19) spike protein by immunoinformatic analysis implicating for a suitable vaccine development, *J Transl Med*, vol. 18, no. 1, Jul. 2020, <https://doi.org/10.1186/s12967-020-02435-4>
  60. A. J. Loughran, C. J. Orihuela, and E. I. Tuomanen, *Streptococcus pneumoniae*: Invasion and Inflammation, 2019, <https://doi.org/10.1128/microbiolspec>



61. Weinberger M, Weinberger DM, Malley R, Lipsitch M (2011) Serotype replacement in disease after pneumococcal vaccination. *Lancet* 378:1962–1973. <https://doi.org/10.1016/S0140>
62. O'Brien KL et al (2009) Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. *Lancet* 374(9693):893–902. [https://doi.org/10.1016/S0140-6736\(09\)61204-6](https://doi.org/10.1016/S0140-6736(09)61204-6)
63. Herbert JA et al (2018) Production and efficacy of a low-cost recombinant pneumococcal protein polysaccharide conjugate vaccine. *Vaccine* 36(26):3809–3819. <https://doi.org/10.1016/j.vaccine.2018.05.036>
64. Mazumder L et al (2023) An immunoinformatics approach to epitope-based vaccine design against PspA in Streptococcus pneumoniae. *J Genet Eng Biotechnol* 21(1):57. <https://doi.org/10.1186/s43141-023-00506-9>
65. M. Tahir Ul Qamar, S. Saleem, U. A. Ashfaq, A. Bari, F. Anwar, and S. Alqahtani, Epitope-based peptide vaccine design and target site depiction on Middle East Respiratory Syndrome Coronavirus: An immunoinformatics study, *J Transl Med*, vol. 17, no. 1, Nov. 2019, <https://doi.org/10.1186/s12967-019-2116-8>
66. J. Rai et al., Hypothesis Immunoinformatic evaluation of multiple epitope ensembles as vaccine candidates: E coli 536, *print* Bioinformation, vol. 8, no. 6, p. 272, 2012, [Online]. Available: [www.bioinformation.net](http://www.bioinformation.net)
67. S. N. H. Bukhari, A. Jain, E. Haq, A. Mehbodniya, and J. Webber, 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*, vol. 11, no. 2. MDPI, Feb. 01, 2022. <https://doi.org/10.3390/pathogens11020146>
68. Z. Bahadori, M. Shafaghi, H. Madanchi, M. M. Ranjbar, A. A. Shabani, and S. F. Mousavi, In silico designing of a novel epitope-based candidate vaccine against Streptococcus pneumoniae with introduction of a new domain of PepO as adjuvant, *J Transl Med*, vol. 20, no. 1, Dec. 2022, <https://doi.org/10.1186/s12967-022-03590-6>
69. S. Bin Sayed, Z. Nain, M. S. A. Khan, F. Abdulla, R. Tasmin, and U. K. Adhikari, Exploring lassa virus proteome to design a multi-epitope vaccine through immunoinformatics and immune simulation analyses, *Int J Pept Res Ther*, vol. 26, no. 4, pp. 2089–2107, Dec. 2020, <https://doi.org/10.1007/s10989-019-10003-8>
70. M. T. Khan et al., Immunoinformatics and molecular modeling approach to design universal multi-epitope vaccine for SARS-CoV-2, *Inform Med Unlocked*, vol. 24, Jan. 2021, <https://doi.org/10.1016/j.jimu.2021.100578>
71. Ikai A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem* 88(6):1895–1898. <https://doi.org/10.1093/oxfordjournals.jbchem.a133168>
72. M. Ali, R. K. Pandey, N. Khatoon, A. Narula, A. Mishra, and V. K. Prajapati, Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection, *Sci Rep*, vol. 7, no. 1, Dec. 2017, <https://doi.org/10.1038/s41598-017-09199-w>

## Publisher's Note

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

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---