Computers in Biology and Medicine 133 (2021) 104390

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Exploring SARS-COV-2 structural proteins to design a multi-epitope vaccine using immunoinformatics approach: An *in silico* study

Samira Sanami [a,1], Morteza Alizadeh [b,1], Masoud Nosrati [c], Korosh Ashrafi Dehkordi [d,**], Fatemeh Azadegan-Dehkordi [e], Shahram Tahmasebian [a], Hamed Nosrati [f], Mohammad-Hassan Arjmand [g], Maryam Ghasemi-Dehnoo [h], Ali Rafiei [e], Nader Bagheri [e,*]

[a] Department of Medical Biotechnology, School of Advanced Technologies, Shahrekord University of Medical Sciences, Shahrekord, Iran
[b] Department of Tissue Engineering, School of Medicine, Shahroud University of Medical Sciences, Shahroud, Iran
[c] Department of Computer Science, Iowa State University, Ames, IA, USA
[d] Department of Molecular Medicine, School of Advanced Technologies, Shahrekord University of Medical Sciences, Shahrekord, Iran
[e] Cellular and Molecular Research Center, Basic Health Sciences Institute, Shahrekord University of Medical Sciences, Shahrekord, Iran
[f] Department of Tissue Engineering and Applied Cell Sciences, School of Advanced Technologies, Shahrekord University of Medical Sciences, Shahrekord, Iran
[g] Cancer Research Center, Shahrekord University of Medical Sciences, Shahrekord, Iran
[h] Medical Plants Research Center, Basic Health Sciences Institute, Shahrekord University of Medical Sciences, Shahrekord, Iran

ARTICLE INFO

Keywords:
SARS-COV-2
Vaccine
Reverse vaccinology
Epitope
CTxB

ABSTRACT

In December 2019, a new virus called SARS-CoV-2 was reported in China and quickly spread to other parts of the world. The development of SARS-COV-2 vaccines has recently received much attention from numerous researchers. The present study aims to design an effective multi-epitope vaccine against SARS-COV-2 using the reverse vaccinology method. In this regard, structural proteins from SARS-COV-2, including the spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins, were selected as target antigens for epitope prediction. A total of five helper T lymphocytes (HTL) and five cytotoxic T lymphocytes (CTL) epitopes were selected after screening the predicted epitopes for antigenicity, allergenicity, and toxicity. Subsequently, the selected HTL and CTL epitopes were fused via flexible linkers. Next, the cholera toxin B-subunit (CTxB) as an adjuvant was linked to the N-terminal of the chimeric structure. The proposed vaccine was analyzed for the properties of physicochemical, antigenicity, and allergenicity. The 3D model of the vaccine construct was predicted and docked with the Toll-like receptor 4 (TLR4). The molecular dynamics (MD) simulation was performed to evaluate the stable interactions between the vaccine construct and TLR4. The immune simulation was also conducted to explore the immune responses induced by the vaccine. Finally, *in silico* cloning of the vaccine construct into the pET-28 (+) vector was conducted. The results obtained from all bioinformatics analysis stages were satisfactory; however, *in vitro* and *in vivo* tests are essential to validate these results.

1. Introduction

In December 2019, a new virus called SARS-CoV-2 was reported in China and quickly spread to other parts of the world. The disease caused by this virus was officially named COVID-19 by the World Health Organization (WHO) [1]. The WHO has reported 119030459 cases and 2 640349 deaths due to COVID 19, as of March 14, 2021 [2]. SARS-COV-2 transmission pathways include direct and indirect transmission. Direct transmission occurs through respiratory droplets and saliva when coughing, sneezing, and close contact. Indirect transmission is also possible due to contaminated surfaces where the virus may remain and act as a secondary source [3,4]. The most common symptoms of COVID-19 are fever, cough, muscle soreness, and fatigue [5,6]. Patients with moderate to severe disease suffer from shortness of breath [5]. Moreover, bloody sputum was reported in a small percentage of patients [5]. Headaches and gastrointestinal symptoms such as nausea, vomiting,

** Corresponding author.
* Corresponding author.
E-mail addresses: kdashrafi@gmail.com (K.A. Dehkordi), n.bagheri1985@gmail.com (N. Bagheri).
[1] Samira Sanami and Morteza Alizadeh contributed equally to this work as first authors.

https://doi.org/10.1016/j.compbiomed.2021.104390
Received 13 January 2021; Received in revised form 7 April 2021; Accepted 7 April 2021
Available online 20 April 2021
0010-4825/© 2021 Elsevier Ltd. All rights reserved.

and diarrhea are not common but may occur [7].

The Coronaviridae family consists of four genera, including alpha, beta, delta, and gamma [8]. The human coronavirus HCoV-NL63 and HCoV-229E belong to the alphacoronavirus genus, while HCoV-HKU-1, HCoV-MERS, HCoV-SARS, and HCoV-OC-43 belong to the betacoronavirus genus [9]. SARS-COV-2 was also classified as a betacoronavirus after the identification of its gene sequence [10]. Alpha and betacoronaviruses mainly infect mammals and cause disease in humans and animals [11]. On the other hand, gamma and delta coronaviruses mainly infect wild birds [12].

SARS-CoV-2 is an enveloped RNA virus, the genome of that is a positive single-stranded ribonucleic acid about 29.9 kb in length [13]. The virus genome consists of 14 open reading frames (ORFs) that encode 27 proteins [14]. About two-thirds of the virus genome includes ORF1a and ORF1b, which encode 16 nonstructural proteins (NSPs) [15], while the remaining third of the genome contains eight accessory proteins, including orf14, 9b, 8b, 7b, 7a, p6, 3b, and 3a, and four structural proteins, including N, M, E, and S.

The entry of SARS-COV-2 into the host cells requires the identification of specific proteins of the cell surface that act as receptors for the S protein. The virus uses angiotensin-converting enzyme 2 (ACE2) as a cellular receptor to enter the host cell [16]. The E protein is an integral protein in the membrane that plays a crucial role in various stages such as virulence, assembly, and morphogenesis of SARS-COV-2 [17]. One of the significant functional components of SARS-COV-2 is the M protein, which plays an essential role in maintaining the size and shape of the virus [18]. The N protein helps to replicate and package viral RNA in the virus [19].

Despite advances in treating infectious diseases, pathogenic microorganisms still appear to be the greatest threat to public health. Although conventional vaccines have played a key role in eradicating some pathogens, their production methods require the cultivation of microorganisms [20], a time-consuming process that can only detect antigens that are produced in large quantities, while protein abundance does not mean that they are immunogenic [21], on the other hand, these methods will not work for uncultivable microorganisms [22]. Recent advancements in genome sequencing and the advent of computer biotechnology have led to new methods, including reverse vaccinology [23], which uses genetic data and computer algorithms to develop vaccines [24]. This method has several advantages over conventional

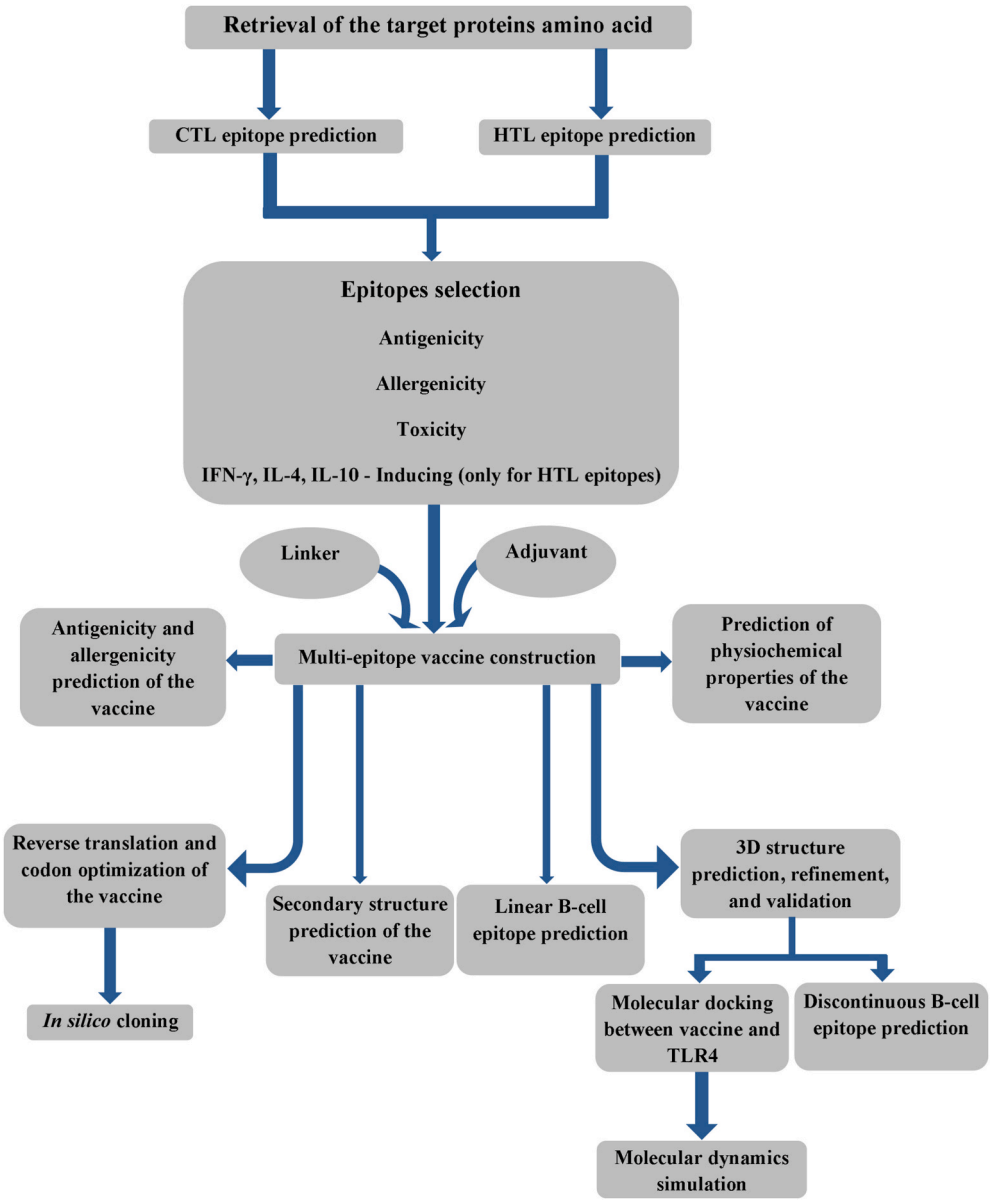

Fig. 1. Flow chart of methods used for *in silico design* of the multi-epitope vaccine.

vaccinology, such as reducing the cost and time of vaccine development, detecting antigens found in small amounts, and allowing for the study of uncultivable or risky pathogens [22].

The development of SARS-COV-2 vaccines has recently received much attention from numerous researchers. This study aims to design an effective multi-epitope vaccine against SARS-COV-2 using the reverse vaccinology method. In this regard, CTL and HTL epitopes of S, M, N, and E proteins were predicted and linked together by appropriate linkers. CTxB as an adjuvant was linked to the N-terminal of the chimeric structure using one EAAAK linker. Subsequently, the vaccine candidate was evaluated for the physicochemical properties, antigenicity, and allergenicity. Furthermore, several analyzes were also performed, including the prediction of secondary and 3D models of the vaccine construct, molecular docking, MD simulation, immune simulation, and *in silico* cloning of the vaccine construct (Fig. 1).

2. Methodology

2.1. Retrieval of the amino acid sequence of the target proteins

Amino acid sequences of the target proteins and adjuvant were retrieved from the NCBI database in FASTA format; their accession number are presented in Table 1.

2.2. T-cell epitopes prediction and selection

NetCTL 1.2 server (http://www.cbs.dtu.dk/services/NetCTL/) was used to predict the CTL epitopes for target proteins (S, M, N, and E) [25]. The sensitivity and specificity of the epitope predicted by the server are 54–89% and 94–99%, respectively. The server predicted epitopes based on three strategies, including TAP transport efficiency, MHC-I binding peptides, and proteasomal C-terminal cleavage. In this study, 12 MHC class I supertypes were selected to predict CTL epitopes, and the threshold score for weight on C terminal cleavage, weight on TAP transport efficiency, and epitope identification were set at 0.15, 0.05, and 0.75, respectively. Using NetMHCII 2.3 server (http://www.cbs.dtu.dk/services/NetMHCII/), HTL epitopes were predicted [26]. The server used ANN to predict epitopes. In this study, HLA-DP, HLA-DR, and HLA-DQ alleles were selected to predict HTL epitopes, and the threshold was set at 2% and 10% for strong binder and weak binder, respectively.

Since there were a large number of predicted epitopes and the vaccine constructs is limited in size, the best epitope should be selected from among them. In this regard, epitopes that had a binding affinity with more MHC alleles and thus cover a major population were selected [27, 28]. A higher antigenicity score for an epitope indicates that it can play a significant role in initiating the immune response [29]. To ensure the safety of epitopes, their toxicity and allergenicity should be checked [30, 31]. Therefore, the selected epitopes from the previous step were screened for antigenicity, toxicity, and allergenicity using VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html), ToxinPred (https://webs.iiitd.edu.in/raghava/toxinpred/design.php), and AllerTOP v. 2.0 (https://www.ddg-pharmfac.net/AllerTOP/), respectively. VaxiJen v2.0 classifies antigens solely based on protein physicochemical characteristics without using sequence alignment. The accuracy of the antigenicity prediction by the VaxiJen v2.0 server is between 70% and 89% [32–34]. In this analysis, the virus was selected as the target organism, and the antigenicity threshold was set at 0.4. AllerTOP v2.0 server uses machine learning approaches, including auto and cross variance transformation, amino acid E-descriptors, and the k nearest neighbors to predict the allergenicity of proteins and peptides [35]. ToxinPred server predicts the toxicity of peptides along with their physicochemical properties such as molecular weight, hydropathicity, pI, charge, hydrophilicity, hydrophobicity, and amphipathicity [36].

Not all HTL epitopes are capable of inducing the production of cytokines, and the cytokines produced by each may be different. IFN-γ is an important cytokine in innate and adaptive immune systems and activates immune responses against bacteria, parasites, and particularly viruses [37,38]. Induction of IL-4 and IL-10 by HTL epitopes in the vaccine construct also helps activate cytotoxic T cells, macrophages, and B cells [39]. Therefore, the HTL epitopes were evaluated for induction of IFN-γ, IL-4, and IL-10. IFNepitope server (https://webs.iiitd.edu.in/raghava/ifnepitope/design.php) was used to predict IFN-γ-inducing HTL epitopes [40]. Furthermore, the SVM-based approach and IFN-γ versus other cytokine were selected to predict IFN-γ inducing epitopes. The ability of HTL epitopes to induce IL-4 was assessed using the IL4pred server (https://webs.iiitd.edu.in/raghava/il4pred/design.php) [41]. Subsequently, the IL-10Pred server (https://webs.iiitd.edu.in/raghava/il10pred/predict3.php) was used to evaluate HTL epitopes for IL-10 induction [42].

Table 1
The accession number of the selected proteins for vaccine design.

Protein	Accession number
Spike	YP_009724390.1
Envelope	YP_009724392.1
Membrane	YP_009724393.1
Nucleocapsid	YP_009724397.2
CTxB	ACT78953.1

2.3. Vaccine design

The selected CTL and HTL epitopes were used as the basic components of the vaccine. The HTL epitopes were fused via GPGPG linkers, whereas KK linkers were used for linking CTL epitopes. The linkers improve the presentation of epitopes [43]. Glycine-rich linkers, such as GPGPG, help improve solubility and allow adjacent domains to act freely [44]. Furthermore, CTxB was attached by an EAAAK linker as an adjuvant to the N-terminal of the chimeric sequences to enhance the immunogenicity of the vaccine candidate. CTxB is the nontoxic segment (124 amino acids) of cholera toxin that can improve the cellular and humoral immune responses [45].

2.4. Evaluation of the physicochemical parameters, allergenicity, and antigenicity of the proposed vaccine

ProtParam server (https://web.expasy.org/protparam/) was used to predict the physicochemical parameters of the vaccine, such as the number of amino acids, amino acid composition, molecular weight (MW), theoretical pI, half-life, instability index, aliphatic index, and grand average of hydropathicity (GRAVY) [46]. The AllerTOP v. 2.0 server was used to evaluate the allergenicity of the vaccine. VaxiJen v2.0 and ANTIGENpro (http://scratch.proteomics.ics.uci.edu/) were used to predict the antigenicity of the vaccine construct. ANTIGENpro predicts protein antigenicity independent of pathogens and based on the protein sequence. The prediction of this server is performed in two stages, based on five machine learning algorithms and multiple representations of the initial sequence [47].

2.5. Prediction of secondary structure

The Prabi server (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html) was used to predict the percentage of the secondary structure elements of the multi-epitope vaccine. This server uses the GOR IV method to predict a secondary structure with a mean accuracy of 64.4% [48].

2.6. 3D model prediction, refinement, and validation of multi-epitope vaccine

I-TASSER server (https://zhanglab.ccmb.med.umich.edu/

I-TASSER/) was used to predict the 3D structure of the vaccine construct. I-TASSER provides a platform for the prediction of high-quality tertiary structure models from the amino acid sequences. This server generates the 3D structures from the amino acid sequence by reassembling excised fragments from threading templates and reports a C-score to evaluate the accuracy of the predicted models [49–51]. 3D refine server (http://sysbio.rnet.missouri.edu/3Drefine/) was used to improve the quality of the 3D structure of the selected model [52–54]. Model validation is an essential step in comparing the quality of the unrefined model with the refined model [55]. PROCHECK (https://saves.mbi.ucla.edu/), ERRAT (https://saves.mbi.ucla.edu/), and ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) were used to validate the 3D model of the vaccine construct. PROCHECK generates the Ramachandran plot, showing the percentage of amino acids presented in the favoured, allowed, and disallowed regions [56,57]. ERRAT conducts a statistical analysis of non-bonded interactions between various atoms in protein structures [58]. ProSA-web calculates the overall quality score for the protein structure. If the computed score is outside the range characteristic of native proteins, it indicates potential errors in the protein structure [59,60].

2.7. B-cell epitopes prediction

The BCPred server (http://ailab-projects1.ist.psu.edu:8080/bcpred/predict.html) was used to predict linear B-cell epitopes. The server predicts linear B-cell epitopes using a subsequence kernel-based SVM classifier with an accuracy of 74.57% [61,62]. In the present study, the vaccine sequence was used as input without changing other default parameter values. The Ellipro server (http://tools.iedb.org/ellipro/) was used to predict discontinuous B-cell epitopes utilizing residue clustering algorithms and Tornton's approach [63]. In this analysis, the refined 3D model of the vaccine was used as input, and the minimum score and maximum distance were set at 0.5 and 6 Å, respectively.

2.8. Molecular docking

Molecular docking is a computational method used to predict the complex structure and interaction of ligands and receptors at the atomic level. Accordingly, ClusPro 2.0 server (https://cluspro.org/login.php) was used to perform molecular docking between the vaccine construct and TLR4 (PDB ID: 4G8A) [64–66]. The refined models of vaccine construct and TLR4 were submitted to the server as ligand and receptor, respectively. LigPlot software was used to map bonds formed between residues of vaccine construct and TLR4 in the docked complex [67].

2.9. MD simulation

The GROMACS 2019.6 software was used to perform the MD simulation. This software predicts the behavior of ligands and receptors over a period of time using Newton's laws of atomic and molecular motion [68–70]. The ff99SB force field was used for the preparation of the input structure. Next, sodium and chlorine ions were used to neutralize the surface charge of the structure. Moreover, the protein was inserted into a layer of TIP3P water molecules with a thickness of 10 Å using gmx solvate software. The energy minimization of the structures was carried out using the steepest descent method to eliminate van der Waals interactions and hydrogen bonds forming between water and complex molecules. Subsequently, the system temperature was gradually increased from 0 to 300 K with a duration of 200 ps in constant volume, and the system was then equilibrated at constant pressure. The MD simulation was performed at a temperature of 300 K and a period of 40 ns Finally, the root mean square deviation (RMSD) and the root mean square fluctuation (RMSF) of the ligand and receptor were calculated.

2.10. Immune simulation

C-ImmSim server (http://150.146.2.1/C-IMMSIM/index.php?page=1) was used to characterize the immune response profile of the designed multi-epitope vaccine. This server is an agent-based simulator that predicts cellular and humoral immune responses to specific antigens using the position-specific scoring matrix (PSSM) and machine learning approaches [71]. In this analysis, random seed, simulation volume, and simulation step parameters were set at 12345, 10, 270, respectively. Two injections with time steps of 1 and 63 were considered for the proposed vaccine. The rest of the parameters were kept as the default values [72].

2.11. Codon optimization and in silico cloning of the vaccine construct

Java Codon Adaptation Tool (JCat) was used to reverse translation and codon optimization of the multi-epitope vaccine [73]. *E. coli* (strain K12) was selected as the host organism to express the vaccine construct. Moreover, cDNA sequence, codon adaptation index (CAI), and GC content were included in the server output. The CAI and GC content are essential parameters to evaluate the protein expression levels. Finally, the sequence of restriction sites for the *Xho*I and *Bam*HI restriction enzyme was added at 5′ and 3′ ends of the nucleotide sequence of the vaccine, respectively, and inserted into the pET-28 (+) vector by SnapGene tool.

3. Results

3.1. T-cell epitopes prediction and selection

A total of 443 CTL epitopes for E, M, N, and S proteins were predicted by NetCTL 1.2 server. Among them, epitopes with the ability to bind to at least four different human MHC class subtypes were assessed for allergenicity, antigen, and toxicity. Finally, one CTL epitope was selected for each of E, M, and N proteins and two CTL epitopes were selected for S protein (Supplementary File. 1). HTL epitopes were predicted for E, M, N, and S proteins using the NetMHCII 2.3 server. A total of 75 high score HTL epitopes (15-mer) with high binding potential for at least six human MHC II alleles were screened for antigenicity, allergenicity, toxicity, and induction of IFN-γ, IL-4, and IL-10. Induction of IL-4 and IL-10 by HTL epitopes in the vaccine construct also helps activate cytotoxic T cells, macrophages, and B cells [39]. One HTL epitope was selected for each of M, N, and S proteins, and two HTL epitopes were selected for E protein based on the mentioned parameters (Supplementary File. 2).

3.2. Vaccine design

A total of five CTL epitopes and five HTL epitopes were used to construct multi-epitope vaccines (Table 2). HTL and CTL epitopes were linked together by GPGPG and KK linkers, respectively. A sequence of 150 amino acids was obtained after epitope fusion. The adjuvant sequence, with a length of 124 amino acids (MIKLKFGVFFTVLLSSAYAHGTPQNITDLCAEYHNTQIHTLNDKIFSYTESLAGKRE-

Table 2
The selected epitopes for the final vaccine construct.

protein	CTL epitope	HTL epitope
E	FLAFVVFLL (20–28)	KPSFYVYSRVKNLNS (53–67) SFYVYSRVKNLNSSR (55–69)
M	ATSRTLSYY (171–179)	WITGGIAIAMACLVG (75–89)
N	LSPRWYFYY (104–112)	QIGYYRRATRRIRGG (83–97)
S	FVFLVLLPL (2-10) WTAGAAAYY (258–266)	NFTISVTTEILPVSM (717–731)

MAIITFKNGATFQVEVPGSQHIDSQKKAIERMKDTLRIAYLTEAK-VEKLCVWNNKTPHAIAAISMAN), was attached to the N-terminal of the vaccine sequence by one EAAAK linker (Fig. 2).

3.3. Evaluation of the physicochemical parameters, allergenicity, and antigenicity of the proposed vaccine

The physicochemical characteristics of the multi-epitope vaccine were computed by the ProtParam server. The final composition of the vaccine included 279 amino acids. Cysteine and alanine were the lowest (1.1%) and the highest (9.3%) residues in the vaccine construct, respectively (Fig. 3). The molecular weight of the vaccine was 30.99 kDa, and its theoretical pI was estimated to be 10.01. The half-life of the vaccine was calculated to be 30 h in mammalian reticulocytes, more than 20 h in yeast, and more than 10 h in *E. coli*. The predicted values for the instability index, aliphatic index, and GRAVY were 39.92, 82.9, and −0.088, respectively. The non-allergic nature of the vaccine construct was defined by AllerTOP v. 2.0 server. The antigenicity of the vaccine was predicted using VaxiJen v2.0 and ANTIGENpro. The antigenicity of the vaccine construct predicted by the VaxiJen v2.0 server was 0.5638 and 0.6015 in the virus and bacterial models, respectively. The ANTIGENpro server also estimated the probability of vaccine antigenicity at 0.8456. These results confirmed the antigenic nature of the proposed vaccine.

3.4. Prediction of secondary structure

The percentage of the secondary structure elements of the vaccine was estimated using the Prabi server. The predicted structure included 34.77% alpha-helix, 24.37% extended strand, and 40.86% random coil (Fig. 4).

3.5. 3D model prediction, refinement, and validation of multi-epitope vaccine

Five 3D structure models of the vaccine were generated by the I-TASSER server using the threading templates (PDB Hit: 1ltrA, 3chbD, 1g55A, 4l6t, 6elcA, 1ltrF, and 5elbA). The C-scores of the models 1 to 5 were −4.18, −3.73, −4.35, −3.36, and −4.67, respectively. The C-score is usually in the range of [-5 to 2], and a high value C-score indicates a high confidence model [49]. Consequently, model 4 with a C-score of −3.36 was selected and refined using the 3D refine server (Fig. 5), which generated five refined models and calculated the 3D refine score, GDT-TS, GDT-HA, RMSD, MolProbity, and RWPlus parameters for each model (Table 3). The lower values of the 3D refine score, RWplus, and MolProbity, and the higher values of the GDT-TS, GDT-HA, and RMSD, indicate higher quality models. Refined model 5 was selected based on the parameters mentioned above (Fig. 5). PROCHECK, ERRAT, and ProSA-web were used to compare the overall quality of the protein structure of the vaccine before and after the refining process. The Ramachandran plot of the initial model showed that 74.4%, 20.2%, 2.9%, and 2.5% of the residues were present in the favoured, additional allowed, generously allowed, and disallowed regions, respectively (Fig. 6a). After refinement, 81.5%, 13.9%, 2.9%, and 1.7% of residues were present in the favoured, additional allowed, generously allowed, and disallowed regions, respectively (Fig. 6b). The ERRAT predicted the overall quality factor to be 79.705 for the unrefined model and 81.413 for the refined model (Fig. 6c and d). The Z-score of the initial model was determined to be −7.6 (Fig. 6e), which reached −7.91 after the refinement (Fig. 6f).

3.6. B-cell epitopes prediction

The BCPred server predicted five linear B-cell epitopes (20-mer) with scores ranging from 0.735 to 1 (Table 4). The place of the linear B-cell epitopes in the 3D model of the vaccine construct is shown in Fig. 7. Ten discontinuous B-cell epitopes were predicted by the ElliPro server (Fig. 8). The size of the predicted epitopes ranged from 6 to 38 amino acids. The lowest and highest scores for the epitopes were 0.621 and 0.814, respectively (Table 5).

3.7. Molecular docking

The molecular docking was performed between the vaccine construct and TLR4 using the ClusPro server, which generated 26 clusters. Cluster 2 was used to perform molecular dynamics simulation analysis due to its negative energy score (−952.8 kcal/mol) compared to other clusters, indicating the highest binding affinity. In this cluster, the two chains, B and D from TLR4, interacted with the vaccine construct (Fig. 9). A total of 9 hydrogen bonds were formed between the residues of the chain B from TLR4 and the vaccine residues, while the number of hydrogen bonds formed between the residues of the chain D from TLR4 and the residues of the vaccine was 11. (Fig. 10a and b). The amino acids that formed these hydrogen bonds and their distances are presented in Tables 6 and 7.

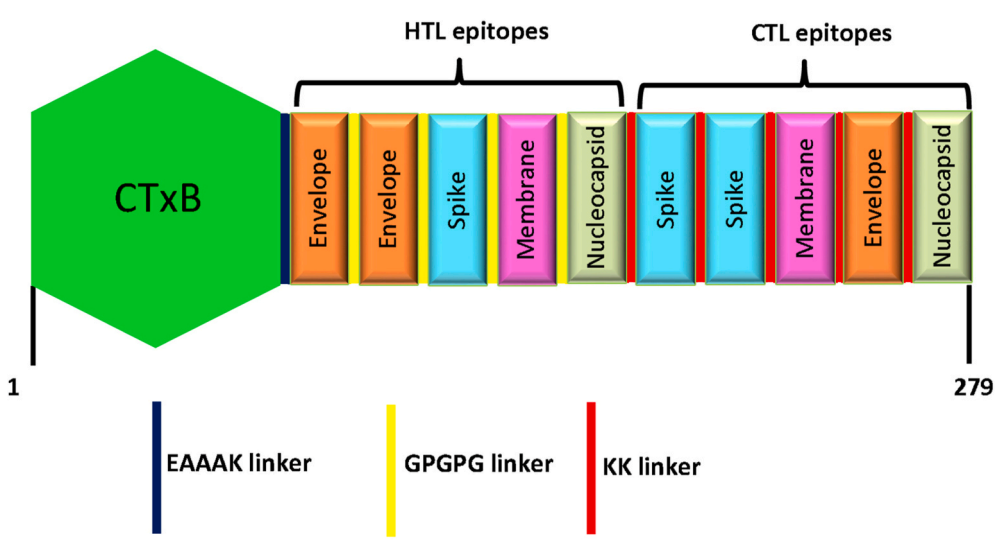

Fig. 2. Schematic illustration of the multi-epitope vaccine construct.

Fig. 3. Amino acid composition in the vaccine construct. The vaccine sequence contains 279 amino acids, and the highest and lowest residues in the vaccine construct are alanine and cysteine, respectively.

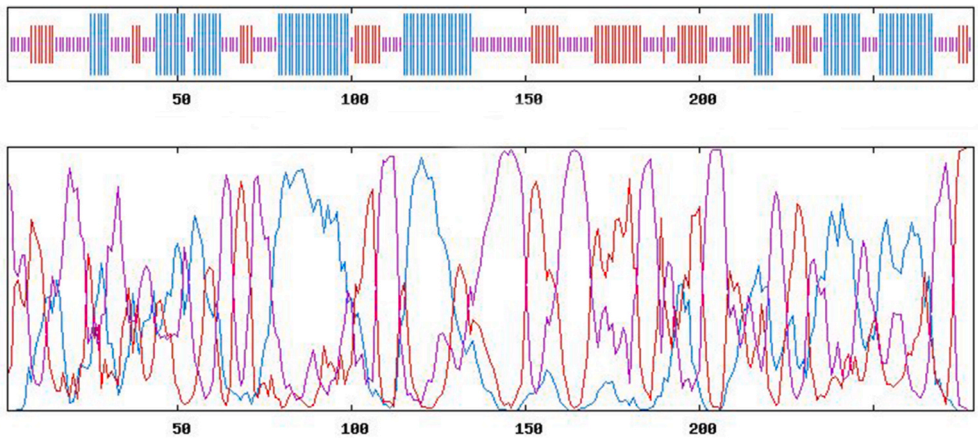

Fig. 4. Schematic illustration of the secondary structure elements of the vaccine construct. The predicted secondary structure consists of alpha-helix (34.77%) in blue, extended strand (24.37%) in red, and random coil (40.86%) in purple.

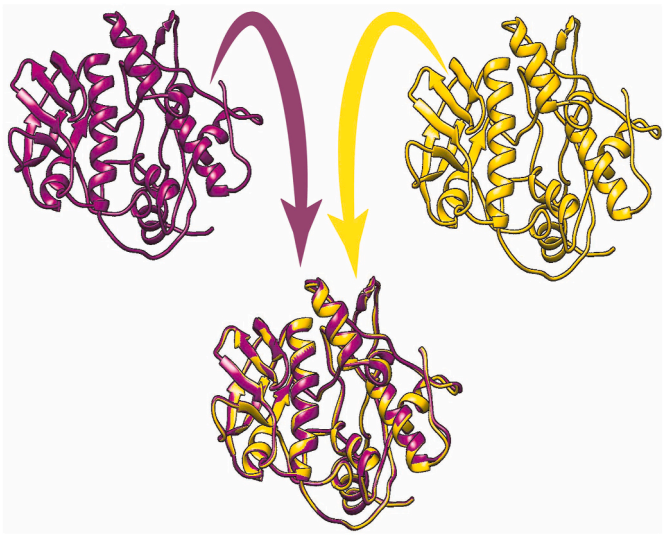

Fig. 5. The initial and refined models are shown in dark magenta and gold colors, respectively. To compare the initial and refined models, the models were superimposed.

3.8. MD simulation

MD simulation was performed using GROMACS 2019.6 software to evaluate stable interactions between the vaccine candidate and TLR4 in the docked complex. The RMSD of the structures generated during the MD simulation in the time dimension is an appropriate and common standard to verify the structural stability of the protein. The RMSD value of TLR4 had an increasing trend at the beginning of the simulation period and reached 0.6 nm after 17000 ps, and then decreased to 0.3 nm in 24,000 ps After this time, the RMSD value increased to approximately 0.4 nm and remained constant until the end of the simulation time. The RMSD plot of the vaccine showed a continuous increase up to 33,000 ps and reached about 0.8 nm, and remained at the same value until the end of the simulation (Fig. 11a). The RMSF of Cα atoms was determined to evaluate the motion and structural flexibility of the ligand and receptor in the docked complex. The RMSF plot of the vaccine showed very sharp fluctuations, in which the highest peaks with an RMSF value of 0.8 nm in the plot indicated a highly flexible region (amino acids 205–215) in the vaccine construct. The fluctuations in the RMSF plot of TLR4 were very mild. The 20 amino acids showed high flexibility at the end of chains A and B (Fig. 11b).

3.9. Immune simulation

The immune response elicitation of the multi-epitope vaccine was evaluated using the C-ImmSim server. The level of antigen was significantly decreased after the first and second injections of the proposed vaccine. The primary response was characterized by an increase in the levels of IgM + IgG, IgM, and IgG1, and in the secondary response the levels of these antibodies increased relative to their concentrations in the primary response, and an increase in IgG1 + IgG2 and IgG2 was also observed (Fig. 12a). The B cell memory and B cell isotype IgM and IgG1 populations were increased after each injection, and a very small population of B cell isotype IgG2 was also observed after the second injection (Fig. 12b). The TH cell population reached 5000 cells per mm^3 from the first injection to day 10 and remained at the same level until the second injection, after that increased again, reaching 9000 on day 30, and then slowly decreased until day 90. Moreover, the level of TH memory cells increased after each injection, and the increase was higher

Table 3
Results of the 3D refine server. Model 5 was selected as the best refined model based on the 3Drefine score, GDT-TS, GDT-HA, RMSD, MolProbity, and RWPlus parameters.

Model	$3D^{refine}$ Score	GDT-TS	GDT-HA	RMSD (Å)	MolProbity	RWPlus
5	17245.5	0.9991	0.9713	0.353	3.256	−53105.552557
4	17446.4	1.0000	0.9803	0.331	3.230	−53051.739246
3	17792.9	1.0000	0.9875	0.301	3.180	−52958.210160
2	18422.5	1.0000	0.9964	0.257	3.213	−52815.241513
1	20587.0	1.0000	0.9982	0.193	3.266	−52650.359028

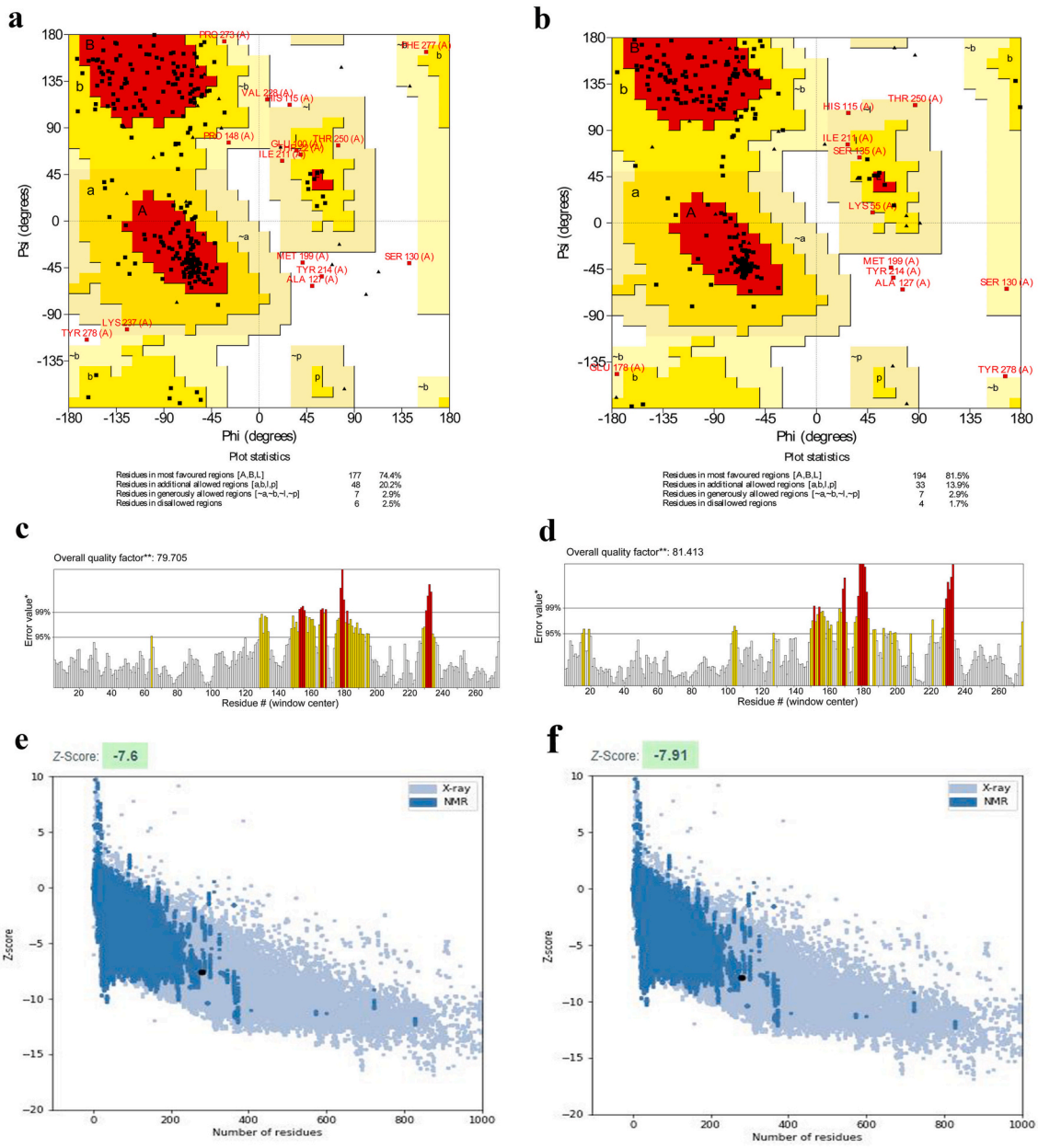

Fig. 6. Validation of the tertiary structure of the vaccine using the Ramachandran plot, ERRAT, and ProSA-web before and after the refining process. (a) Ramachandran plot of the initial model shows that 74.4% of residues to be in the favoured region, (b) while in the refined model, 81.5% of residues to be in the favoured region. (c) The ERRAT results indicate that the overall quality factor for the initial model is 79.705, (d) while this value for the refined model is 81.413. (e) The z-score obtained from the ProSA-web is −7.6 in the initial model and (f) and the z-score of the refined model is −7.91.

after the second injection (Fig. 12c). After the first injection, the TC cell population increased, reaching 1150 cells per mm^3 on the 12th day, which decreased slowly until the 90th day (Fig. 12d). A significant increase was observed in IFN-γ and IL-2 levels following each vaccine injection (Fig. 12e). The active macrophage population also increased after each vaccination (Fig. 12f).

Table 4
Linear B-cell epitope predicted by the BCPred server.

Position	Linear B-cell epitope	Score
203	VGGPGPGQIGYYRRATRRIR	1
139	NLNSSRGPGPGKPSFYVYSR	1
175	VTTEILPVSMGPGPGWITGG	0.999
62	TFKNGATFQVEVPGSQHIDS	0.947
99	TEAKVEKLCVWNNKTPHAIA	0.735

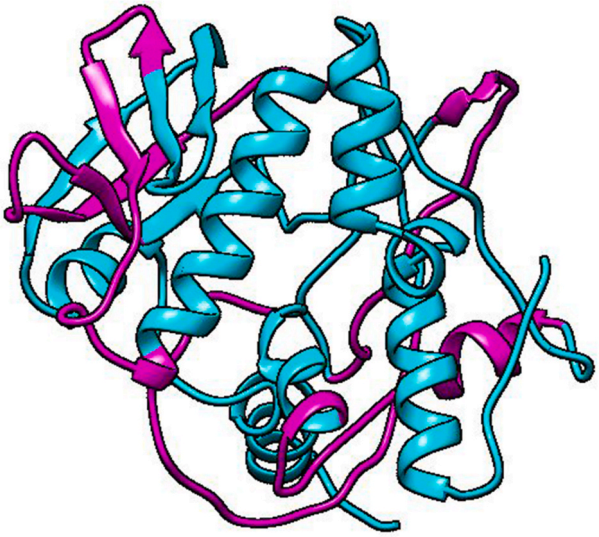

Fig. 7. Linear B-cell epitopes (magenta) are shown in the 3D model of the vaccine construct (blue).

3.10. Codon optimization and in silico cloning of the vaccine construct

Back translation and codon optimization of the vaccine construct was conducted using JCat. There were 837 nucleotides in the optimized codon sequence. The CAI of the optimized nucleotide sequence was 0.95, and the GC content was 49.46%. Finally, the nucleotide sequence of the vaccine was cloned into the pET-28 (+) vector using the SnapGene tool (Fig. 13)..

4. Discussion

The advent of the twentieth century has been accompanied by the emergence of new viruses, the latest case of emerging viruses was the SARS-COV-2 epidemic in Wuhan, China [74]. The threat of any new virus requires us to quickly develop a vaccine. Indeed, vaccines are one of the most effective ways to prevent viral infections because, in most cases, other treatment options are limited or non-existent, and infections lead to clinical deterioration, which decreases the therapeutic effect [75].

In recent months, many efforts have been made by researchers to develop an effective vaccine against SARS-COV-2. As of November 18, 2020, approximately 164 SARS-CoV-2 vaccine candidates were in pre-clinical phases, and 44 vaccine candidates were in clinical trials (phase I–III) [76]. The US Food and Drug Administration (FDA) granted the emergency use authorization for Pfizer/BioNTech and Moderna

Table 5
Discontinuous B-cell epitope predicted by the ElliPro server.

Position	Discontinuous B-cell epitope	Score
267	LLKKLSPRWYFYY	0.814
110	NNKTPH	0.812
1	MIKLKFGVFFTV	0.787
219	RRIRGGKKFV	0.733
72	EVPGSQHID	0.715
138	KNLNSSRGPGPG	0.693
235	LKKWTAGAAAYYK	0.652
31	AEYHNTQIHTLNDKIFSYTESLAGKREMAIITFKNGAT	0.648
155	VYSRVKNLNSGPGPGNFTISVTTEILPVSMGPGPG	0.633
124	NEAAAKS	0.621

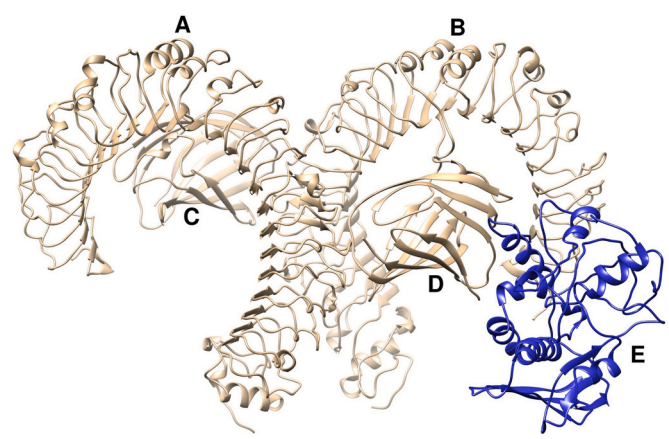

Fig. 9. Docked complex of vaccine construct with TLR4. The vaccine construct (ligand) is shown in blue, whereas tan color represents TLR4 (receptor) in the docked complex.

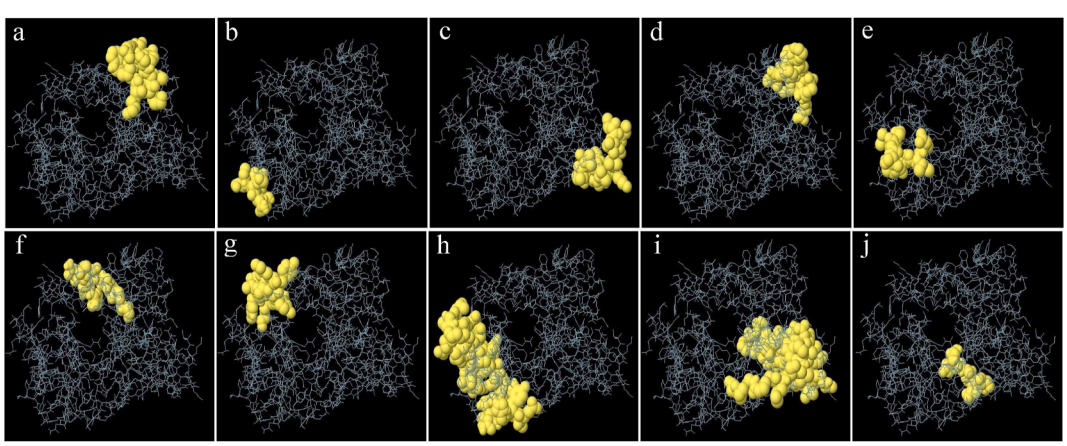

Fig. 8. 3D model representation of discontinuous B-cell epitopes (a–j). The vaccine construct is shown in gray sticks, and the B-cell epitopes are depicted by yellow surface.

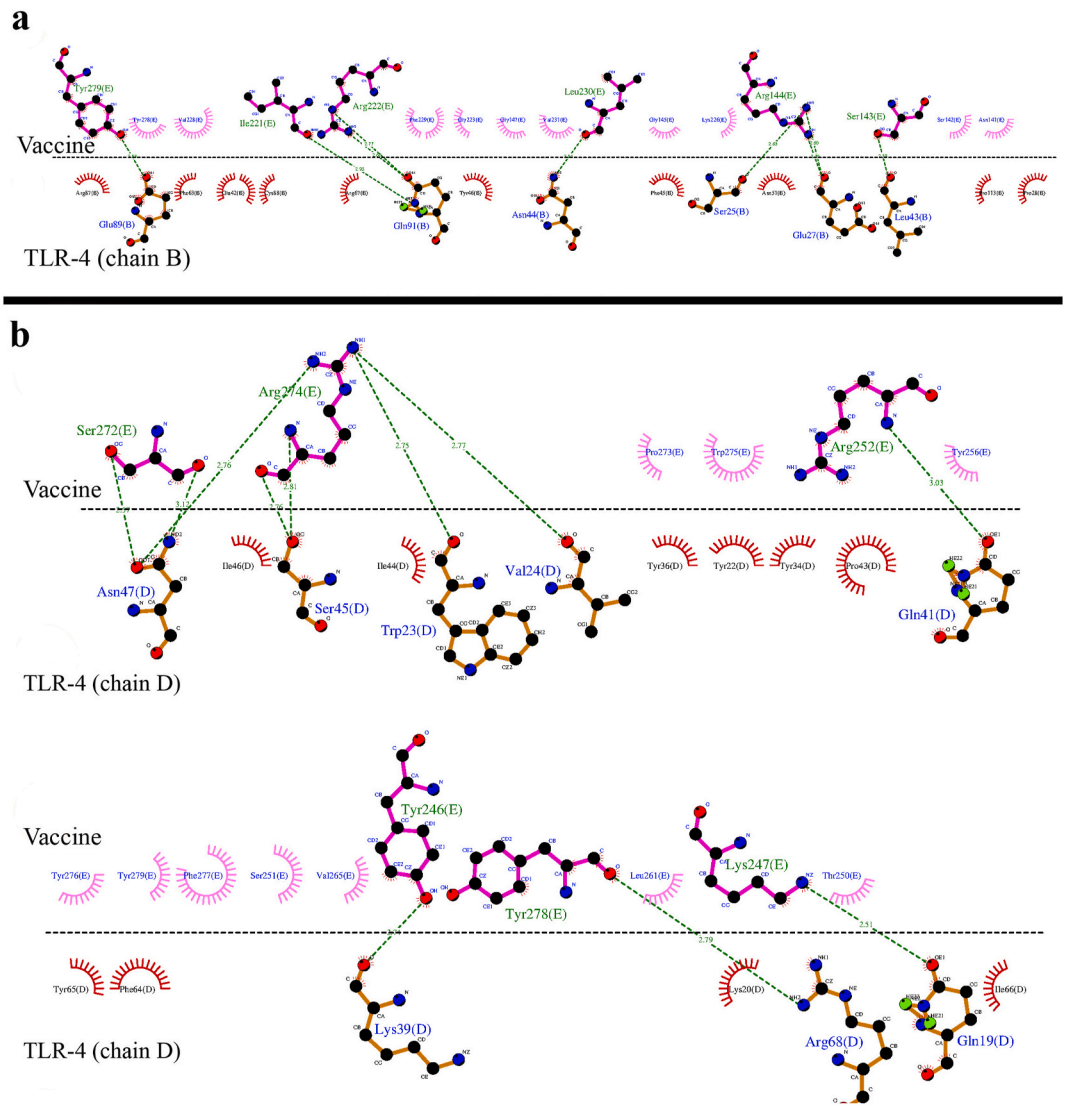

Fig. 10. Map of interactions between vaccine construct and TLR4 generated by the LigPlot software. (a) Interactions between the vaccine construct and chain B from TLR4. (b) Interactions between the vaccine construct and chain D from TLR4. The hydrogen bonds are shown by the green dotted lines.

Table 6
List of amino acids of the vaccine and TLR4 (chain B) that formed hydrogen bonds.

TLR4 (chain B)-Vaccine	Number of hydrogen bonds	Distance (Å)
Glu89-Tyr279	1	2.88
Asn44-Leu230	1	2.84
Leu43-Ser143	1	2.88
Gln91- Ile221	1	2.93
Gln91- Arg222	2	2.65
		2.77
Glu27-Arg144	2	2.60
		2.70
Ser25-Arg144	1	2.65

Table 7
List of amino acids of the vaccine and TLR4 (chain D) that formed hydrogen bonds.

TLR4 (chain D)-Vaccine	Number of hydrogen bonds	Distance (Å)
Lys39-Tyr246	1	2.74
Arg68-Tyr278	1	2.79
Gln41-Arg252	1	3.03
Gln19-Lys247	1	2.51
Asn47-Ser272	2	2.37
		3.12
Asn47-Arg274	1	2.76
Val24-Arg274	1	2.77
Ser45-Arg274	2	2.76
		2.81
Trp23-Arg274	1	2.75

COVID-19 vaccines on December 11 and 18, 2020, respectively [77]. Simultaneously with these clinical trials, another group of researchers used the immunoinformatics approaches to design epitope-based vaccines against SARS-COV-2. In bioinformatics studies, various proteins of SARS-COV-2 have been used as target antigens to predict epitope. In studies conducted by Kar et al. [78], Dar et al. [79], and Kumar et al. [80], S protein alone was used as a target for epitope prediction. However, in other pieces of research conducted by Kalita et al. [81] and Khairkhah et al. [82], S, M, and N protein were identified as target antigens. Furthermore, S, E, and N proteins were reported as target antigens by Kumar et al. [83]. In addition to structural proteins, Zaheer et al. [84], Chauhan et al. [85], Enayatkhani et al. [86], and Jain et al. [87] used some accessory proteins as antigen targets in their studies.

In the present study, as in the study of Singh et al. [88], four structural proteins, including S, M, N, and E, were used to predict epitope.

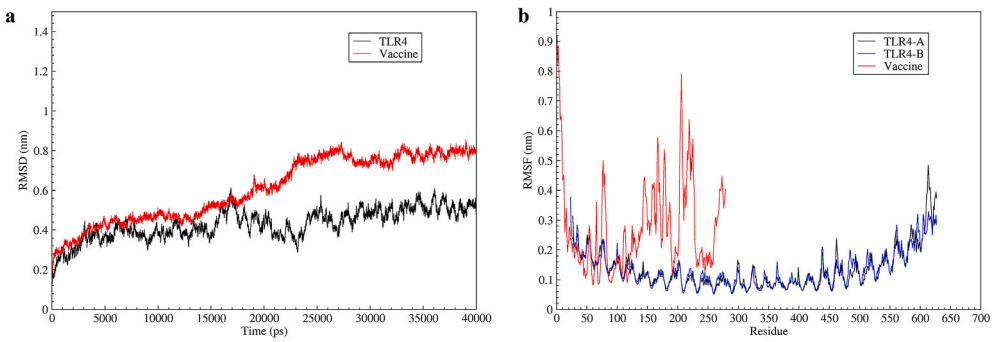

Fig. 11. MD simulation of the vaccine - TLR4 complex. (a) The RMSD plot of TLR4 shows an increasing trend from the beginning of the simulation and remains at about 0.4 nm from 30,000 ps to the end of the simulation. The RMSD plot of the vaccine increases from the beginning of the simulation until 33,000 ps, after that time it remains at approximately 0.8 nm to the end of the simulation. (b) In the RMSF plot, the peaks represent regions with a high degree of flexibility.

After screening the predicted epitopes for allergenicity, antigen, and toxicity, five HTL epitopes and five CTL epitopes were selected for use in the vaccine construct (Table 2). According to a report by Nguyen et al. [89], the epitopes selected in the present study were found in conserved regions of structural proteins and were potentially appropriate for vaccine development. The selected HTL and CTL epitopes were arranged in different orders and positions in the vaccine construct using flexible linkers such as GPGPG, AYY, and KK. After evaluating the physicochemical properties, particularly the stability of the obtained constructs, it was concluded that GPGPG and KK linkers are appropriate linkers for linking HTL and CTL epitopes, respectively. CTxB was added to the N-terminal of the vaccine structure as an adjuvant by the EAAAK linker to increase the immunogenicity of the vaccine. This adjuvant was selected based on previous research, which has shown that it can be used as a viral adjuvant to boost both systemic and mucosal immunity to the respiratory syncytial virus by nasal vaccination [90]. In this work, for the first time, the CTxB sequence and epitopes of structural proteins of M, N, E, and S were organized by appropriate linkers within a vaccine construct.

The molecular weight of the vaccine was 30.99 kDa, due to the easier purification process of proteins with a molecular weight of less than 110 kDa in the laboratory, the predicted value for this parameter is satisfactory [91]. The theoretical pI of the vaccine was estimated to be 10.01, indicating that the vaccine was basic in nature. The half-life of the vaccine was estimated to be 30 h in mammalian, allowing the vaccine with a long half-life to be more exposed to the immune system [92]. The instability index of the vaccine candidate was estimated at 39.92, and since this value is less than 40, the vaccine is classified as a stable protein [92]. The aliphatic index was computed 82.9, indicating that the vaccine is thermostable [93]. The hydrophilicity of the vaccine construct increases its delivery and processing, formulation, and dissolution in an aqueous environment [94]. The negative and positive values of GRAVY indicate that the vaccine construct is hydrophilic and hydrophobic, respectively [95]. The GRAVY score was −0.088, indicating the hydrophilic nature of the vaccine due to its negative value; therefore, it can interact better with water molecules [96]. However, in the study of Qamar et al. [97], GRAVY was calculated to be 0.395, and the use of micelles to improve the interaction of the vaccine in the polar environment of the body seems to be necessary due to the hydrophobic nature of the vaccine.

According to these results, the physicochemical properties of the vaccine candidate are in accordance with the standard criteria required for vaccine formulation. Furthermore, the suggested vaccine was found to be antigenic and non-allergenic, indicating that the multi-epitope vaccine without an allergic reaction would have a strong immune response. After 3D modeling, the structure of the vaccine was refined to achieve a high-quality structure. The evaluation of the Ramachandran plot of the initial and refined model of the vaccine showed that 74.4% of the residues in the initial model were present in the favoured region, which increased to 81.5% after refining, confirming the improvement in the refined model. Approximately 35.8% and 22% of vaccine construct residues were found to be present in the linear and discontinuous B-cell epitopes, respectively, indicating an antiviral antibody response.

Hu et al. showed that TLR4 expression is upregulated within 24 h after SARS-CoV infection, indicating its role in the generation of immune responses [98]. Other studies have also demonstrated the influential role of TLR4 in inducing immune responses against coronavirus [99–101]. Consequently, TLR4 was selected as the immune receptor for molecular docking in the present study, such as studies by Yang et al. [102] and Safavi et al. [103]. The molecular docking analysis showed that our vaccine has the ability to bind TLR4 and induce a strong immune response. In this regard, MD simulation of the docked complex was carried out for 40 ns to evaluate the stability of the vaccine structure. The RMSD plot of the vaccine-TLR4 complex showed the stability of the complex (Fig. 11a). The RMSF plot showed that the residues present in the HTL epitopes fluctuate more than the residues in CTL epitopes (Fig. 11b), which is due to the more interaction of CTL epitope residues with TLR4. The immune simulation results indicated that the final structure of the vaccine was able to activate both cellular and humoral immune responses against SARS-COV-2. Moreover, the secondary response generated by the vaccine was significantly higher than the primary response. Codon optimization was performed to assure the maximum expression of the designed vaccine in *E. coli* (strain K12). The CAI and GC content of the optimized nucleotide sequence were 0.95 and 49.46%, respectively. CAI greater than 0.8 is optimal for expression in the target organism [104], and the GC content between 30 and 70% is considered ideal [96]. The results of the computational evaluation of the vaccine candidate were satisfactory; however, these results must be validated in both *in vitro* and *in vivo* experiments.

5. Conclusion

The SARS-COV-2 outbreak, first reported in China, was spread rapidly to other parts of the world and has recently caused heavy casualties. The vaccine appears to be the most effective way to combat this virus. The reverse vaccinology method can be used to design a safe and high potency vaccine in a short time and at a low cost. In the present study, bioinformatics software and online servers were used to design a multi-epitope vaccine against SARS-COV-2. S, E, N, and M proteins were selected from SARS-COV-2 as the target proteins for epitope prediction. After screening epitopes for antigenicity, allergenicity, and toxicity, a total of five CTL and five HTL epitopes and CTxB (as adjuvant) were organized in a multi-epitope vaccine structure. The evaluation of the vaccine using the ProtParam server showed that the proposed vaccine is of good quality regarding its physicochemical properties. The results of the molecular docking showed the high affinity of the vaccine construct with TLR4. The stability of the vaccine construct was confirmed by MD simulation. Furthermore, the efficacy of the vaccine to elicit an immune

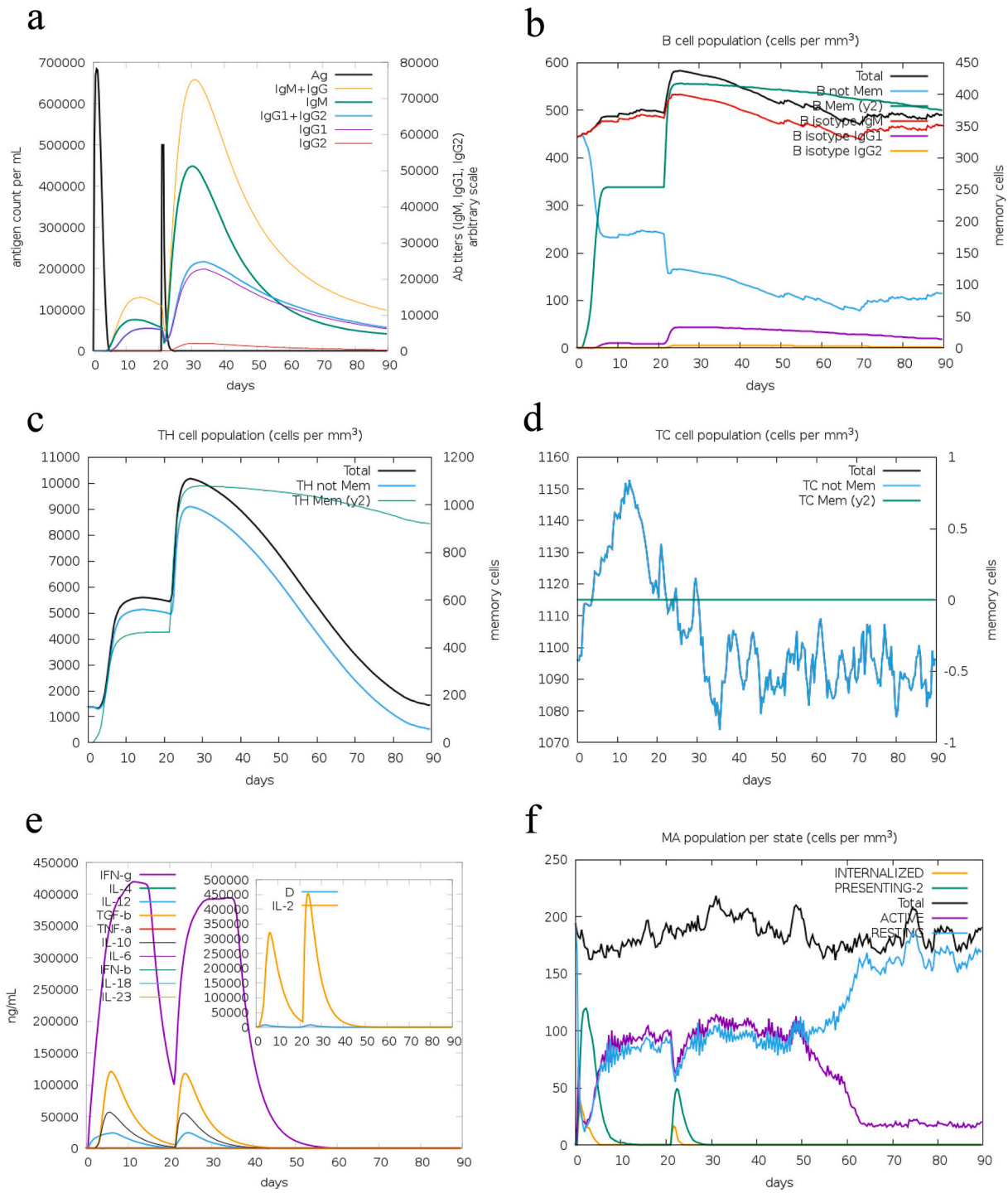

Fig. 12. Simulation of the immune response of the proposed vaccine. (a) Immunoglobulins production after the vaccine injection. (b) B cell population. (c) TH cell population. (d) TC cell population. (e) Concentration of cytokines and interleukins. (f) Macrophages population per state.

response was confirmed by immune simulation. However, these results must be validated in both *in vitro* and *in vivo* tests.

Funding

This study was financially supported by the Research Deputy of Shahrekord University of Medical Sciences with grant number 5659.

Ethical approval

The ethical committee of Shahrekord University of Medical Sciences approved this study with the number: IR.SKUMS.REC.1399.278.

Declaration of competing interest

The authors declare that they have no conflict of interest.

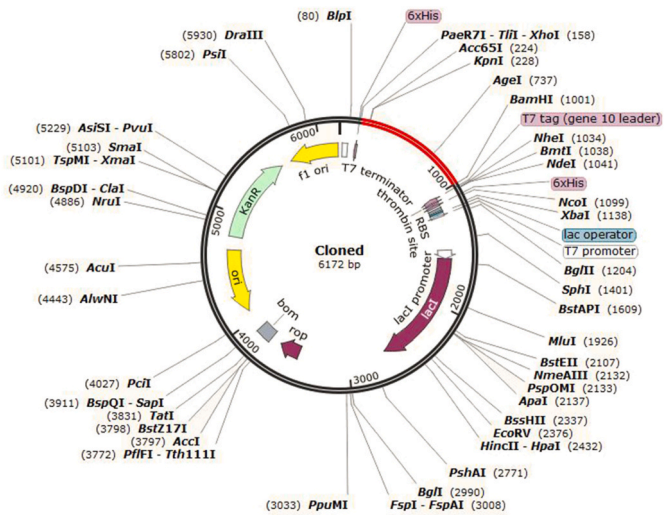

Fig. 13. *In silico* cloning of the vaccine construct. The pET-28 (+) backbone is shown in black and the vaccine construct is shown in red, surrounded between *Xho*I (158) and *Bam*HI (1001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104390.

References

[1] H. Zhu, L. Wei, P. Niu, The novel coronavirus outbreak in Wuhan, China, Global health research and policy 5 (2020) 1–3.

[2] WHO, Coronavirus disease (COVID-19) pandemic. https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=CjwKCAiAhbeCBhBcEiwAkv2cYxWLgs5QXus55bWR3hu2VVK6OSg4p8o4CdsckMqZk4oj3hiDlqiH3RoCheQQAvD_BwE. (Accessed 14 March 2021) accessed.

[3] C.-w. Lu, X.-f. Liu, Z.-f. Jia, 2019-nCoV transmission through the ocular surface must not be ignored, Lancet 395 (2020) e39.

[4] K.K.-W. To, O.T.-Y. Tsang, C.C.-Y. Yip, K.-H. Chan, T.-C. Wu, J.M.-C. Chan, W.-S. Leung, T.S.-H. Chik, C.Y.-C. Choi, D.H. Kandamby, Consistent Detection of 2019 Novel Coronavirus in Saliva, Clinical Infectious Diseases, 2020.

[5] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (2020) 497–506.

[6] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, Lancet 395 (2020) 507–513.

[7] K.S. Cheung, I.F. Hung, P.P. Chan, K. Lung, E. Tso, R. Liu, Y. Ng, M.Y. Chu, T. W. Chung, A.R. Tam, Gastrointestinal manifestations of SARS-CoV-2 infection and virus load in fecal samples from the Hong Kong cohort and systematic review and meta-analysis, Gastroenterology 159 (1) (2020) 81–95.

[8] F.K. Ayittey, C. Dzuvor, M.K. Ayittey, N.B. Chiwero, A. Habib, Updates on Wuhan 2019 novel coronavirus epidemic, J. Med. Virol. 92 (4) (2020) 403–407.

[9] Q. Tang, Y. Song, M. Shi, Y. Cheng, W. Zhang, X.-Q. Xia, Inferring the hosts of coronavirus using dual statistical models based on nucleotide composition, Sci. Rep. 5 (2015) 17155.

[10] D. Kim, J.-Y. Lee, J.-S. Yang, J.W. Kim, V.N. Kim, H. Chang, The architecture of SARS-CoV-2 transcriptome, Cell 181 (4) (2020) 914–921.

[11] C.A. Narh, Genomic cues from beta-coronaviruses and mammalian hosts sheds light on probable origins and infectivity of SARS-CoV-2 causing COVID-19, Front. Genet. 11 (2020) 902.

[12] M. Wille, E.C. Holmes, Wild birds as reservoirs for diverse and abundant gamma- and deltacoronaviruses, FEMS Microbiol. Rev. 44 (2020) 631–644.

[13] R.d.C.A. Tavares, G. Mahadeshwar, H. Wan, N.C. Huston, A.M. Pyle, The global and local distribution of RNA structure throughout the SARS-CoV-2 genome, J. Virol. (2020).

[14] A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China, Cell Host Microbe 27 (3) (2020) 325–328.

[15] Z. Miao, A. Tidu, G. Eriani, F. Martin, Secondary structure of the SARS-CoV-2 5'-UTR, RNA Biol. (2020) 1–10.

[16] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor, Cell 181 (2) (2020) 271–280.

[17] A. Toto, S. Ma, F. Malagrinò, L. Visconti, L. Pagano, K. Stromgaard, S. Gianni, Comparing the binding properties of peptides mimicking the Envelope protein of SARS-CoV and SARS-CoV-2 to the PDZ domain of the tight junction-associated PALS1 protein, Protein Sci. 29 (2020) 2038–2042.

[18] R. Mahtarin, S. Islam, M.J. Islam, M.O. Ullah, M.A. Ali, M.A. Halim, Structure and dynamics of membrane protein in SARS-CoV-2, J. Biomol. Struct. Dyn. (2020) 1–14.

[19] S.M. Cascarina, E.D. Ross, A proposed role for the SARS-CoV-2 nucleocapsid protein in the formation and regulation of biomolecular condensates, Faseb. J. 34 (2020) 9832–9842.

[20] G. Abbas, I. Zafar, S. Ahmad, S.S. Azam, Immunoinformatics design of a novel multi-epitope peptide vaccine to combat multi-drug resistant infections caused by Vibrio vulnificus, Eur. J. Pharmaceut. Sci. 142 (2020) 105160.

[21] R. Arnon, T. Ben-Yedidia, Old and new vaccine approaches, Int. Immunopharm. 3 (2003) 1195–1204.

[22] R. Rappuoli, Reverse vaccinology, Curr. Opin. Microbiol. 3 (2000) 445–450.

[23] A.-R.M. Faisal, S.H. Imtiaz, T. Zerin, T. Rahman, H.U. Shekhar, Computer aided epitope design as a peptide vaccine component against Lassa virus, Bioinformation 13 (2017) 417.

[24] D.R. Flower, I.K. Macdonald, K. Ramakrishnan, M.N. Davies, I.A. Doytchinova, Computer aided selection of candidate vaccine antigens, Immunome Res. 6 (2010) 1–16.

[25] M.V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction, BMC Bioinf. 8 (2007) 424.

[26] K.K. Jensen, M. Andreatta, P. Marcatili, S. Buus, J.A. Greenbaum, Z. Yan, A. Sette, B. Peters, M. Nielsen, Improved methods for predicting peptide binding affinity to MHC class II molecules, Immunology 154 (2018) 394–406.

[27] C. Foged, J. Hansen, E.M. Agger, License to kill: formulation requirements for optimal priming of CD8+ CTL responses with particulate vaccine delivery systems, Eur. J. Pharmaceut. Sci. 45 (2012) 482–491.

[28] L.-L. An, J.L. Whitton, A multivalent minigene vaccine, containing B-cell, cytotoxic T-lymphocyte, and Th epitopes from several microbes, induces appropriate responses in vivo and confers protection against more than one pathogen, J. Virol. 71 (1997) 2292–2302.

[29] M.T. Ul Qamar, S. Saleem, U.A. Ashfaq, A. Bari, F. Anwar, S. Alqahtani, Epitope-based peptide vaccine design and target site depiction against Middle East Respiratory Syndrome Coronavirus: an immune-informatics study, J. Transl. Med. 17 (2019) 1–14.

[30] R. Forster, Study designs for the nonclinical safety testing of new vaccine products, J. Pharmacol. Toxicol. Methods 66 (2012) 1–7.

[31] G.N. Konstantinou, T-cell epitope prediction, Food Allergens, Springer2017, pp. 211-222.

[32] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinf. 8 (2007) 4.

[33] I.A. Doytchinova, D.R. Flower, Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties, Vaccine 25 (2007) 856–866.

[34] I.A. Doytchinova, D.R. Flower, Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines, Open Vaccine J. 1 (2008) 4.

[35] S. Wold, J. Jonsson, M. Sjöström, M. Sandberg, S. Rännar, DNA and peptide sequences and chemical processes multivariately modelled by principal component analysis and partial least-squares projections to latent structures, Anal. Chim. Acta 277 (1993) 239–253.

[36] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P. Raghava, O.S.D.D. Consortium in silico approach for predicting toxicity of peptides and proteins, PloS One 8 (2013), e73957.

[37] T. Decker, S. Stockinger, M. Karaghiosoff, M. Müller, P. Kovarik, IFNs and STATs in innate immunity to microorganisms, J. Clin. Invest. 109 (2002) 1271–1277.

[38] D. Griffin, Cytokines and Chemokines, 2008.

[39] R.V. Luckheeram, R. Zhou, A.D. Verma, B. Xia, CD4+ T cells: differentiation and functions, Clin. Dev. Immunol. (2012) 2012.

[40] S.K. Dhanda, P. Vir, G.P. Raghava, Designing of interferon-gamma inducing MHC class-II binders, Biol. Direct 8 (2013) 30.

[41] S.K. Dhanda, S. Gupta, P. Vir, G. Raghava, Prediction of IL4 inducing peptides, Clin. Dev. Immunol. (2013) 2013.

[42] G. Nagpal, S.S. Usmani, S.K. Dhanda, H. Kaur, S. Singh, M. Sharma, G.P. Raghava, Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential, Sci. Rep. 7 (2017) 42851.

[43] M. Saadi, A. Karkhah, H.R. Nouri, Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches, Infect. Genet. Evol. 51 (2017) 227–234.

[44] M. Kavoosi, A.L. Creagh, D.G. Kilburn, C.A. Haynes, Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli, Biotechnol. Bioeng. 98 (2007) 599–610.

[45] H. Madanchi, H. Honari, M. Sadraeian, M. Hesaraki, Fusion of CtxB with StxB, cloning and expression of in esherichia coli: a challenge for improvement of immune response against StxB, Iran. J. Pharm. Sci. 7 (2011) 185–190.

[46] E. Gasteiger, C. Hoogland, A. Gattiker, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, the Proteomics Protocols Handbook, Springer2005, pp. 571-607.

[47] C.N. Magnan, M. Zeller, M.A. Kayala, A. Vigil, A. Randall, P.L. Felgner, P. Baldi, High-throughput prediction of protein antigenicity using protein microarray data, Bioinformatics 26 (2010) 2936–2943.

[48] J. Garnier, GOR secondary structure prediction method version IV, Meth. Enzym., RF Doolittle Ed. 266 (1998) 540–553.

[49] A. Roy, A. Kucukural, Y. Zhang, I. Tasser, A unified platform for automated protein structure and function prediction, Nat. Protoc. 5 (2010) 725–738.
[50] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction, Nat. Methods 12 (2015) 7–8.
[51] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinf. 9 (2008) 40.
[52] D. Bhattacharya, J. Cheng, i3Drefine software for protein 3D structure refinement and its assessment in CASP10, PloS One 8 (2013), e69648.
[53] D. Bhattacharya, J. Nowotny, R. Cao, J. Cheng, 3Drefine: an interactive web server for efficient protein structure refinement, Nucleic Acids Res. 44 (2016) W406–W409.
[54] D. Bhattacharya, J. Cheng, 3Drefine, Consistent protein structure refinement by optimizing hydrogen bonding network and atomic-level energy minimization, Proteins: Structure, Function, and Bioinformatics 81 (2013) 119–131.
[55] N. Khatoon, R.K. Pandey, V.K. Prajapati, Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach, Sci. Rep. 7 (2017) 1–12.
[56] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[57] R.A. Laskowski, J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, J. Biomol. NMR 8 (1996) 477–486.
[58] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519.
[59] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) W407–W410.
[60] M.J. Sippl, Recognition of errors in three-dimensional structures of proteins, Proteins: Structure, Function, and Bioinformatics 17 (1993) 355–362.
[61] J. Chen, H. Liu, J. Yang, K.-C. Chou, Prediction of linear B-cell epitopes using amino acid pair antigenicity scale, Amino Acids 33 (2007) 423–428.
[62] Y. EL-Manzalawy, D. Dobbs, V. Honavar, Predicting linear B-cell epitopes using string kernels, J. Mol. Recogn.: An Interdisciplinary Journal 21 (2008) 243–255.
[63] J. Ponomarenko, H.-H. Bui, W. Li, N. Fusseder, P.E. Bourne, A. Sette, B. Peters, ElliPro: a new structure-based tool for the prediction of antibody epitopes, BMC Bioinf. 9 (2008) 514.
[64] D. Kozakov, D.R. Hall, B. Xia, K.A. Porter, D. Padhorny, C. Yueh, D. Beglov, S. Vajda, The ClusPro web server for protein–protein docking, Nat. Protoc. 12 (2017) 255.
[65] S. Vajda, C. Yueh, D. Beglov, T. Bohnuud, S.E. Mottarella, B. Xia, D.R. Hall, D. Kozakov, New additions to the C lus P ro server motivated by CAPRI, Proteins: Structure, Function, and Bioinformatics 85 (2017) 435–444.
[66] D. Kozakov, D. Beglov, T. Bohnuud, S.E. Mottarella, B. Xia, D.R. Hall, S. Vajda, How good is automated protein docking? Proteins: Structure, Function, and Bioinformatics 81 (2013) 2159–2166.
[67] A.C. Wallace, R.A. Laskowski, J.M. Thornton, LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, Protein Eng. Des. Sel. 8 (1995) 127–134.
[68] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, Software 1 (2015) 19–25.
[69] S. Páll, M.J. Abraham, C. Kutzner, B. Hess, E. Lindahl, Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS, International Conference on Exascale Applications and Software, Springer, 2014, pp. 3–27.
[70] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J. C. Smith, P.M. Kasson, D. van der Spoel, Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (2013) 845–854.
[71] N. Rapin, O. Lund, M. Bernaschi, F. Castiglione, Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system, PloS One 5 (2010), e9862.
[72] M. Faiza, T. Abdullah, J.F. Calderon-Tantalean, M.R. Upadhyay, A. H. Abdelmoneim, F. Akram, B.S. Thakur, I. Abdulaziz, C.J. Ononamadu, D. A. Ghoraba, in: Silico Multi-Epitope Vaccine against Covid19 Showing Effective Interaction with HLA-B* 15: 03, bioRxiv, 2020.
[73] A. Grote, K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D.C. Hempel, D. Jahn, JCat: a novel tool to adapt codon usage of a target gene to its potential expression host, Nucleic Acids Res. 33 (2005) W526–W531.
[74] S.F. Ahmed, A.A. Quadeer, M.R. McKay, Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies, Viruses 12 (2020) 254.
[75] B. Afrough, S. Dowall, R. Hewson, Emerging viruses and current strategies for vaccine intervention, Clin. Exp. Immunol. 196 (2019) 157–166.
[76] E. Tumban, Lead SARS-CoV-2 candidate vaccines: expectations from phase III trials and recommendations post-vaccine approval, Viruses 13 (2021) 54.
[77] W. Zhang, K. Zheng, D. Geng, European Review for Medical and Pharmacological SciencesLetter to the Editor, European Review for Medical and Pharmacological Sciences vol. 24, 2020.
[78] T. Kar, U. Narsaria, S. Basak, D. Deb, F. Castiglione, D.M. Mueller, A.P. Srivastava, A candidate multi-epitope vaccine against SARS-CoV-2, Sci. Rep. 10 (2020) 1–24.
[79] H.A. Dar, Y. Waheed, M.H. Najmi, S. Ismail, H.F. Hetta, A. Ali, K. Muhammad, Multiepitope subunit vaccine design against COVID-19 based on the spike protein of SARS-CoV-2: an in silico analysis, Journal of immunology research (2020) 2020.
[80] N. Kumar, D. Sood, R. Chandra, Design and optimization of a subunit vaccine targeting COVID-19 molecular shreds using an immunoinformatics framework, RSC Adv. 10 (2020) 35856–35872.
[81] P. Kalita, A.K. Padhi, K.Y. Zhang, T. Tripathi, Design of a peptide-based subunit vaccine against novel coronavirus SARS-CoV-2, Microb. Pathog. 145 (2020) 104236.
[82] N. Khairkhah, M.R. Aghasadeghi, A. Namvar, A. Bolhassani, Design of novel multiepitope constructs-based peptide vaccine against the structural S, N and M proteins of human COVID-19 using immunoinformatics analysis, PloS One 15 (2020), e0240577.
[83] A. Kumar, P. Kumar, K.U. Saumya, S.K. Kapuganti, T. Bhardwaj, R. Giri, Exploring the SARS-CoV-2 structural proteins for multi-epitope vaccine development: an in-silico approach, Expet Rev. Vaccine 19 (2020) 887–898.
[84] T. Zaheer, M. Waseem, W. Waqar, H.A. Dar, M. Shehroz, K. Naz, Z. Ishaq, T. Ahmad, N. Ullah, S.M. Bakhtiar, Anti-COVID-19 multi-epitope vaccine designs employing global viral genome sequences, PeerJ 8 (2020), e9541.
[85] V. Chauhan, T. Rungta, M. Rawat, K. Goyal, Y. Gupta, M.P. Singh, Excavating SARS-coronavirus 2 genome for epitope-based subunit vaccine synthesis using immunoinformatics approach, J. Cell. Physiol. 236 (2021) 1131–1147.
[86] M. Enayatkhani, M. Hasaniazad, S. Faezi, H. Gouklani, P. Davoodian, N. Ahmadi, M.A. Einakian, A. Karmostaji, K. Ahmadi, Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study, J. Biomol. Struct. Dyn. (2020) 1–16.
[87] N. Jain, U. Shankar, P. Majee, A. Kumar, Scrutinizing the SARS-CoV-2 protein information for designing an effective vaccine encompassing both the T-cell and B-cell epitopes, Infect. Genet. Evol. 87 (2021) 104648.
[88] A. Singh, M. Thakur, L.K. Sharma, K. Chandra, Designing a multi-epitope peptide based vaccine against SARS-CoV-2, Sci. Rep. 10 (2020) 1–12.
[89] T.T. Nguyen, P.N. Pathirana, T. Nguyen, Q.V.H. Nguyen, A. Bhatti, D.C. Nguyen, D.T. Nguyen, N.D. Nguyen, D. Creighton, M. Abdelrazek, Genomic mutations and changes in protein secondary structure and solvent accessibility of SARS-CoV-2 (COVID-19 virus), Sci. Rep. 11 (2021) 1–16.
[90] H.J. Kim, J.-K. Kim, S.B. Seo, H.J. Lee, H.-J. Kim, Intranasal vaccination with peptides and cholera toxin subunit B as adjuvant to enhance mucosal and systemic immunity to respiratory syncytial virus, Arch Pharm. Res. (Seoul) 30 (2007) 366–371.
[91] D. Barh, N. Barve, K. Gupta, S. Chandra, N. Jain, S. Tiwari, N. Leon-Sicairos, A. Canizalez-Roman, A.R. dos Santos, S.S. Hassan, Exoproteome and secretome derived broad spectrum novel drug and vaccine candidates in Vibrio cholerae targeted by Piper betel derived compounds, PloS One 8 (2013), e52773.
[92] E. Gasteiger, C. Hoogland, A. Gattiker, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, the Proteomics Protocols Handbook, 2005, pp. 571–607.
[93] A. Ikai, Thermostability and aliphatic index of globular proteins, J. Biochem. 88 (1980) 1895–1898.
[94] C. Saroja, P. Lakshmi, S. Bhaskaran, Recent trends in vaccine delivery systems: a review, International journal of pharmaceutical investigation 1 (2011) 64.
[95] P. Kalita, D.L. Lyngdoh, A.K. Padhi, H. Shukla, T. Tripathi, Development of multi-epitope driven subunit vaccine against Fasciola gigantica using immunoinformatics approach, Int. J. Biol. Macromol. 138 (2019) 224–233.
[96] M. Ali, R.K. Pandey, N. Khatoon, A. Narula, A. Mishra, 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. 7 (2017) 1–13.
[97] M.T. ul Qamar, F. Shahid, S. Aslam, U.A. Ashfaq, S. Aslam, I. Fatima, M. M. Fareed, A. Zohaib, L.-L. Chen, Reverse vaccinology assisted designing of multiepitope-based subunit vaccine against SARS-CoV-2, Infectious diseases of poverty 9 (2020) 1–14.
[98] W. Hu, Y.-T. Yen, S. Singh, C.-L. Kao, B.A. Wu-Hsieh, SARS-CoV regulates immune function-related gene expression in human monocytic cells, Viral Immunol. 25 (2012) 277–288.
[99] L.E. Gralinski, V.D. Menachery, A.P. Morgan, A.L. Totura, A. Beall, J. Kocher, J. Plante, D.C. Harrison-Shostak, A. Schäfer, F. Pardo-Manuel de Villena, Allelic variation in the toll-like receptor adaptor protein ticam2 contributes to SARS-coronavirus pathogenesis in mice, G3: Genes, Genomes, Genetics 7 (2017) 1653–1663.
[100] A. Mubarak, W. Alturaiki, M.G. Hemida, Middle East respiratory syndrome coronavirus (MERS-CoV): infection, immunological response, and vaccine development, Journal of immunology research (2019) 2019.
[101] M. Shah, M.A. Anwar, J.-H. Kim, S. Choi, Advances in antiviral therapies targeting toll-like receptors, Expet Opin. Invest. Drugs 25 (2016) 437–453.
[102] Z. Yang, P. Bogdan, S. Nazarian, An in silico deep learning approach to multi-epitope vaccine design: a SARS-CoV-2 case study, Sci. Rep. 11 (2021) 1–21.
[103] A. Safavi, A. Kefayat, E. Mahdevar, A. Abiri, F. Ghahremani, Exploring the out of sight antigens of SARS-CoV-2 to design a candidate multi-epitope vaccine by utilizing immunoinformatics approaches, Vaccine 38 (2020) 7612–7628.
[104] S. Morla, A. Makhija, S. Kumar, Synonymous codon usage pattern in glycoprotein gene of rabies virus, Gene 584 (2016) 1–6.