Heliyon 10 (2024) e24186

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Multi-epitopes vaccine design for surface glycoprotein against SARS-CoV-2 using immunoinformatic approach

Sarmad Frogh Arshad [a], Rehana Rehana [b], Muhammad Asif Saleem [c], Muhammad Usman [a], Hasan Junaid Arshad [d], Rizwana Rizwana [e], Shakeela Shakeela [f], Asma Shah Rukh [g], Imran Ahmad Khan [h], M. Ali Hayssam [i], Muhammad Anwar [j,*]

[a] Department of Biochemistry and Biotechnology, Muhammad Nawaz Shareef University of Agriculture, Multan, 66000, Pakistan
[b] Institute of Plant Breeding & Biotechnology (IPBB), Muhammad Nawaz Shareef University of Agriculture, Multan, 66000, Pakistan
[c] Department of Plant Breeding and Genetics, Bahauddin Zakaria University, Multan, 60800, Pakistan
[d] Centre of Agricultural Biochemistry and Biotechnology, University of Agriculture, Faisalabad, 38000, Pakistan
[e] Department of Biochemistry, Bahauddin Zakaria University, Multan, 60800, Pakistan
[f] Collage of Nursing, D.H.Q. Hospital, Chakwal, Punjab, Pakistan
[g] Department of Pharmacy, College of Pharmacy Punjab University, Lahore, 54590, Pakistan
[h] Department of Pharmacy, MNS University of Agriculture, Multan, 54590, Pakistan
[i] Department of Botany and Microbiology, College of Science, King Saud University, Riyadh, 1145, Saudi Arabia
[j] School of Tropical Agriculture and Forestry, Hainan University, Haikou, PR China

ARTICLE INFO

Keywords:
Antigenicity
Epitopes
Immune
Prediction
Vector
Construct
B cell
T cell

ABSTRACT

Background: The recent COVID vaccinations have successfully reduced death and severity but did not stop the transmission of viruses by the emerging SARS-CoV-2 strain. There is a need for better and long-lasting dynamic vaccines for numerous prevailing strains and the evolving SARS-CoV-2 virus, necessitating the development of broad-spectrum strains being used to stop infection by reducing the spread rate and re-infection. The spike (S) glycoprotein is one of the proteins expressed commonly in the early phases of SARS-CoV-2 infection. It has been identified as the most immunogenic protein of SARS-CoV-2.

Methods: In this study, advanced bioinformatics techniques have been exploited to design the novel multi-epitope vaccine using conserved S protein portions from widespread strains of SARS-CoV-2 to predict B cell and T cell epitopes. These epitopes were selected based on toxicity, antigenicity score and immunogenicity. Epitope combinations were used to construct the maximum potent multi-epitope construct with potential immunogenic features. EAAAK, AAY, and GPGPG were used as linkers to construct epitopes.

Results: The developed vaccine has shown positive results. After the chimeric vaccine construct was cloned into the PET28a (+) vector for expression screening in *Escherichia coli*, the potential expression of the construct was identified.

Conclusion: The construct vaccine performed well in computer-based immune response simulation and covered a variety of allelic populations. These computational results are more helpful for further analysis of our contract vaccine, which can finally help control and prevent SARS-CoV-2 infections worldwide.

* Corresponding author.
E-mail address: 184370@hainanu.edu.cn (M. Anwar).

https://doi.org/10.1016/j.heliyon.2024.e24186
Received 24 October 2023; Received in revised form 3 January 2024; Accepted 4 January 2024
Available online 10 January 2024

1. Introduction

Based on their genetic characteristics, the four genera that comprise the coronavirus family are Gamma, Beta, Alpha and Delta coronaviruses [1]. A positive-stranded RNA coronavirus primarily affects individuals and wildlife's central nervous and respiratory systems [2]. Coronavirus can overcome spectral obstacles and develop into highly lethal viruses, as demonstrated by the SARS-CoV pandemic 2002 and the MERS-CoV epidemic in 2012 [3]. In addition to cats, dogs, rats, bats, rabbits, ferrets, mink, camels, horses, cattle, swine and other wildlife species, coronavirus can infect human beings [4]. The Viral pneumonia epidemic in Wuhan City on January 7, 2020, is known as a novel coronavirus (2019-nCov) [5]. The viral genome for the 2020 influenza virus, designated Wuhan-Hu-1, was published on NCBI Genbank on January 10. The genomic sequencing of this novel virus reveals that it is closely related to other members of the viral species that cause severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome coronavirus (MERS-CoV). The virus was deemed an international health emergency by WHO on December 31, 2020, with 213 fatalities and about 10,000 cases worldwide [6]. CoV-19 is composed of three viral proteins: an envelope protein (E), a membrane protein (M) and a spike Protein (S) [7]. SARS-CoV2 particles are characterized by spike protein of 180 kDa that protrudes from their surfaces to facilitate host-cell entry and adhesion [8]. The natural cover S glycoprotein is integrated into a virus-like particle when expressed to increase infection. The crown-shaped spikes protein form homotrimers on the viral surface, which helps link the viral body to the host cell's ACE2 receptor. Recent studies claim that the priming of SARS-CoV-2 occurs by the cellular protease TMPRSS2, a protein on the host surface. As a result, many methods, such as RT-PCR and CT scans, can be used to diagnose the harmful pathogen SARS-CoV-2 [9]. Because of its high sensitivity, CT scan results are considered more accurate and reliable than RT-PCR tests. Compared to RT-PCR (83.3 %), the CT-scan's sensitivity was 97.2 %. Therefore, it is essential to separate patients with positive PCR results but promising CT-scan results [10]. Although the novel coronavirus disease (COVID-19) currently has no effective treatment option. However, according to certain published medications, patients notice the effects [11]. Remdesivir has been discovered beneficial in treating coronavirus patients because of its positive results. Remdesivir is an adenosine correspondent that integrates into developing viral RNA chains to trigger premature termination [12].

Due to its acceptable safety and efficacy, the most recent version of the guidelines advises using another medication called hydroxychloroquine. Ivermectin, a parasite medication, is recognized by the FDA as an inhibitor of COVID-19 after numerous in-vitro experiments revealed a nearly 5000-fold decrease in viral replication [13]. While the pharmaceutical sector is doing much research, no treatment has yet been developed to recover this illness. In order to attack the virus at the molecular level, a vaccine must be developed. However, there have been a few advances in vaccine development. New topics of conversation in this period include coronavirus vaccinations based on homology and nanotechnology [14]. A potential vaccination will play a part in starting the herd immune response. Due to recent progress in the sequencing of numerous viral genomes and the development of sequence databases, the quick technique based on *in-silico* informatics received much popularity [15]. The objective was to develop a multi-epitope vaccine against the novel coronavirus (SARS-CoV-2); we selected S protein due to height antigenic and conserved characteristics. As a critical element of the pathological envelope, the S protein plays in viral morphogenesis and interacts through further cell bodies during assembly. It thus plays a crucial part in pathological immune evasion. This new vaccine indicates potential outcomes by meeting all-natural requirements in a synthetic immunological milieu. The vaccine will be manufactured on an engineering scale in a proper expression host. As a result, the peptide is optimized for *Escherichia coli*, the practical and cost-effective expression system for generating vaccines. It is cloned in the PET28a (+) vector for further expression study. This vaccination may be used as a treatment in the future after being tested in vivo and in vitro.

2. Materials and Methods

2.1. Sequence retrieves and structure analysis

The whole genome sequence was retrieved from the NCBI (https://www.ncbi.nlm.nih.gov/) website by using accession number ID: (UGL69821.1) with locus tag GU280_gp02 was identified to be the most antigenic viral protein. Using the ExPASy ProtPram tool, the chosen protein's physiochemical characteristics have been studied [16]. The Vaxijen 2.0 was used to determine the antigenic protein [17]. For further assessment, the protein with the highest level of antigenicity was chosen. The AllerTOP v2.0 web server was used to assess allergenicity [18]. Using the PSI-blast approach for studying peptide structure, the PsiPred tool has been used to predict the secondary structure of a chosen S protein. However, the S protein's tertiary structure was identified using Swiss modelling, and a natural structure with the greatest C-score was chosen [19].

2.2. Predication of B cell epitopes

The production of humoral immune system and antibody production. The IEDB Linear Epitope prediction tool v2.0 (https://tools.iedb.org/main/bcell) used for the B cell epitopes [20].

2.3. Prediction of linear B cell epitopes

The IEBB (http://tools.iedb.org/main/tcell) server were used to predict the CTL and HTL activating epitopes that bind to MHC-I and MHC-II. The most common alleles in the human population were used to identify the MHC-I epitopes using the SMM technique

[21].Based on the SMM-align prediction technique, the most common alleles for MHC-II epitopes had been selected [22]. Finally, the epitopes that indicated a great affinity (IC50 < 500 nM) were closed for further study [23].

2.4. Profiling of specific T and B cell epitopes using unique features

The Vaxigen v2.0 server (https://www.ddg-pharmfac.net/vaxijen/html) was used to determine particular T and B cell epitope antigenicity. It is based on the physiological chemical properties of proteins and is an alignment-independent server [24]. The FASTA sequence was given, and the default values for the parameters were used. The threshold value has been set to 0.4, Vaxijen 2.0's default value. AllerTOP (https://tools.ddg-pharmfac.net/AllerTOP/) was used to evaluate the allergenicity of T and B cell epitopes [25]. ToxinPred server (http://crdd.osdd.net/raghava/toxinpred/) was also utilized for the toxicity study [17].

2.5. Population coverage analysis

Shortlisted epitope population coverage was calculated by default parameters in IEDB's population coverage analysis tool [26]. Allocations of the world were selected, and an individual total of epitopes was further.

2.6. Analysis of conservation

The sequence of the specified S protein was identified to be conserved. The IEDB web-based system determined the S protein's epitope conservation [27].

2.7. Conservancy analysis

The identified S protein was found conserved. IEDB web-based server was used to discover the S protein's epitope conservancy.

2.8. Vaccine construct

The adjuvant sequence was retrieved from the NCBI to produce the multiepitope vaccination construct, and all accessible epitopes were used. The adjuvant 50S ribosomal protein L7/L12 (UniProt ID: P9WHE3) was used to link the vaccine's N terminal. Three primary linker types were utilized—EAAAK, GPGPG, and AAY. To carry out the formation and binding of histidine protein, a 6× His tag was added at the C-terminal [15].

2.9. Assemblage of multi-epitopes vaccine candidate sequences

A potential vaccination sequence was generated by manually arranging adjuvant sequence, EAAAK, AAY, GPGPG linkers, B cell epitopes, and MHC-I and MHC-II binding epitopes [28]. To minimize the size of the vaccine design, the overlapping epitopes from the CTL, HLT, and B-cell epitopes were combined. The lists of the sequence of epitopes that were joined to make the final vaccine construct and area in protein were shown in (Table 3) [29].

2.10. Antigenicity and allergenicity evaluation of vaccine protein

The Vaxigen 2.0 tool was used to estimate the antigenicity of the vaccine construct [30], while the AllerTOP web server was used to identify the allergenicity of the multi-epitope vaccine construct [31]. Moreover, ToxinPred (http://crdd.osdd.net/raghava/toxinpred/) was used for toxicity analysis [32].

2.11. Physiochemical and solubility analysis of construct vaccine

ExPASy Protparam (https://web.expasy.org/protparam-doc.html) was used to analyze the vaccine's physiochemical features [33]. Analysis of solubility was used to identify the purity of a chemical quantitatively. SOLpro (http://scratch.proteomics.ics.uci.edu/)was used to analyze the solubility of the multi-epitope vaccine construct [33].

2.12. Analysis of secondary and tertiary structure

For secondary and tertiary structural extrapolation of multi-epitope vaccine construct, PsiPredwere utilized [34]. This tool provides information on the relevant protein's significant coils, plates, and helix. The protein's tertiary structure was evaluated using Scratch Protein editor's well-referenced and reliable tool 3D PRO [35].

2.13. Refinement of tertiary structure

GalaxyRefine server (https://galaxy.seoklab.Org) was used to improve the quality of the 3D structure and refine the vaccine model). This server used molecular dynamics simulation to reconstruct and repack the side chains [36].

2.14. Validation of 3D structure

The Ramachandran Plot Assessment (RAMPAGE) server was used to evaluate the quality of 3D structure [37]. PROCHCK further validated the model.

2.15. Docking analysis

The molecular docking analysis among binding of TLR3 receptor (PDB ID: 2A0Z) and vaccine construct online docking server Cluspro2.0 (https://cluspro.bu.edu/) was used [38].

2.16. Molecular dynamics simulations

iModS online tool (http://imods.chaconlab.org) was used for molecular dynamics simulation of construct vaccine. It is based on an analysis of the torsional angle of the complex. It was used to analyze the RMSD values, the eigenvalue of interacting residues, the deformation of structure, and co-variance among individual residues. It discovers the stability of complexes based on detailed coordinate investigation [39].

2.17. Codon optimization of construct vaccine peptides for expression analysis

The EMBOSS 6.1 (https://www.ebi.ac.uk/emboss/) tool was used to reverse translate the construct sequence of a protein to express the chimeric protein in an expression vector [40]. Java codon Adaption tool (JCate) (http://www.jcat.de/) server was used for codon optimization [41]. *Escherichia coli*, specifically the (E. coli-k12 strain) was used to express the chimeric peptide. Codon adaption index (CAI) values (>0.8) and GC contents (30–70) are used to calculate the levels of protein expression, and the optimized vaccine sequence was cloned in *E. coli* plasmid vector pET-28a (+) (within HindIII and BamHI restriction sites) to confirm the expression of the vaccine.

3. Results

3.1. Sequence retrieves and structure analysis

To develop a possible vaccine, the surface glycoprotein protein of SARS-CoV-2 was obtained from the NCBI database. The S surface protein on the surface of SARS-CoV-2 plays a vital role in receptor identification, fusion during viral infection, and cell attachment, similar to the other viruses. We selected the S protein with NCBI ID (UGL69821.1) with locus tag GU280_gp02 as an antigenic viral protein. The viral protein's antigenicity was calculated using the Vaxijen 2.0 web server. The threshold value was set at $\geq$ 0.4 for specificity. The S protein showed an antigenicity of 0.556 in the full-length protein antigenicity analysis, indicating it was likely an

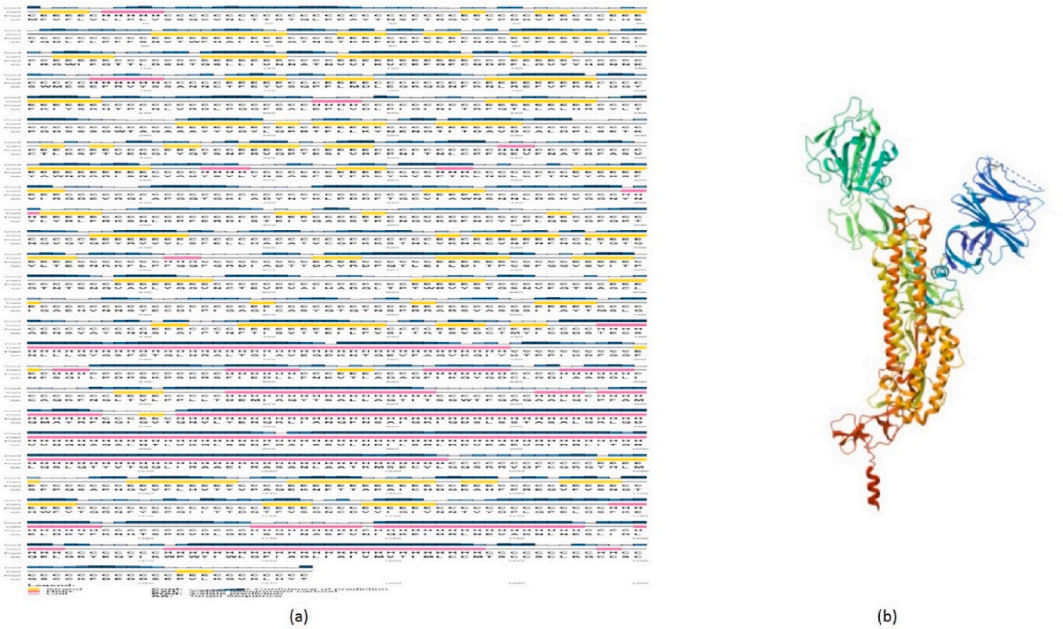


Fig. 1. (a) Secondary structure prediction of SARS-CoV-2 S protein by PSIPERED analysis. (b) The 3D structure Prediction of the SARS-CoV-2 M protein by PDB online server.

antigen. The protein was then used for further study. ProtParam's estimate of the physiochemical characteristics of the SARS-CoV-2 S protein indicated that it had 1273 amino acids and a molecular weight of 141.1495 kDa [42]. The theoretical isoelectric point (PI) was determined to be 6.52, indicating that it is positive because proteins are positively charged when the isoelectric point is over 7.0 [43]. ProtParam's instability-index (II) calculation of 32.79 classified our protein as stable. A protein with an 84.75 aliphatic index is thermo-stable throughout a high-temperature range [44]. A total number of carbon (C), hydrogen (H), nitrogen (N), Oxygen (O), and sulfur (S) were entitled by formula C6340H9775N1657O1890S53. The secondary structure of the protein was predicted by PSIPRED [6]. The 3D structure was predicted by the Swiss model (Fig. 1b) [45]. The secondary structure showed that the S protein has 29.59 % α helixes, 4.08 % β sheets, and 44.30 % random coil, extended strand 22.07 % loops (Fig. 1a). Transmembrane topology was predicted using the online program TMHMM (Fig. 2).

3.2. Prediction of linear B cell epitopes

The B-cell epitopes are vital for resistance to viral contagion. Potential B-cell epitopes have modified properties that direct B-cell recognition and activation of strong immune responses against specific viral infections. Used for the study of likely B-cell epitopes, we expected to use techniques based on amino acid screening. We predicted probable B-cell epitopes using a consensus-based technique. Six linear B cell epitopes were predicted by using the Bepipred Linear Epitopes server, with a threshold score of 0.350 (Fig. 3a). The epitopes were selected for vaccine development based on these principles, with higher scores representing the chosen epitopes (Table 1).

Kolaskar and Tongaonkar method was used for the antigenicity of epitopes [46]. Antigenicity was found to have a maximum tendency of 1.261 and a minimum tendency of 0.866 (Table 2 and Fig. 3b). show the anticipated variation in S protein as an average experiential value of 1.041. For effective B-cell epitopes, good surface accessibility is an essential factor. The surface accessibility was predicted by Emini surface accessibility tool [47], which has the lowest value 0.042 and a maximum value of 6.065 with an average observed value 1.00 (Fig. 3c). Beta turns are hydrophilic, which is essential in immune defense response. For the prediction of beta-turn in S protein Chou and Fasman beta-turn evaluating algorithm was used [48]. The values range from a minimum of 0.541 to a maximum of 1.484. 0.997 was the average value (Fig. 3d) Karplus and Schulz flexibility prediction for flexibility [40]. The calculated outcomes are 0.876 (minimum) and 1.125 (maximum). The average computed value was 0.993 (Fig. 3e). Parker's hydrophobicity was used to identify hydrophobicity [49] the values which are −7.629 (minimum) and 7.743 (maximum). The average value was 1.34 (Fig. 3f). The epitopes were further selected based on toxicity, allergenicity, and antigenicity, with a threshold value of 0.4. Non-toxic and non-allergenic Epitopes were chosen after toxic, allergic epitopes were eliminated. 6 out of 17 epitopes were assumed as effective B-cell epitopes that can evoke lymphocyte (GKQGNF, LTPGDSSSGWTAG, VRQIAPGQTGKIADVRQIAPGQTGKIAD, VITPGTNTSN, ILPDPSKPSKRS, and KNHTSPDVDLG).

3.3. T-cell epitopes identification

3.3.1. Predication of MHC Class-1 binding profile for conserved epitopes

The SMM method and *Homo sapiens* as the MHC source are chosen for the analysis of a varied collection of MHC HLA alleles in humans. This program provides an output interface for epitopes' affinity for binding to HLA molecules by the IC50 nM unit. A lower IC50 value shows more epitope-MHC Class-I molecule binding affinity. A total of 1500 epitopes were selected, with those with larger affinities predicted to relate with several MHC Class-1 alleles and IC50 values less than 200. The 84 of 800 epitopes were selected based on the highest interaction MHC Class-1 allele. 79epitopes were finalized for further study based on toxicity, allergic and antigenicity. The remaining epitopes that have less than a 0.4 antigenic score were excluded. The core epitopes FTISVTTEI and WTAGAAAY were

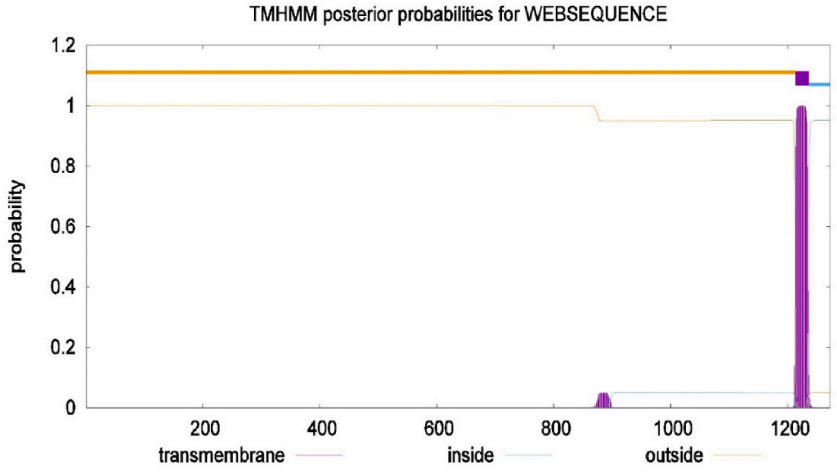


Fig. 2. Transmembrane helices prediction in proteins by TMHMM.

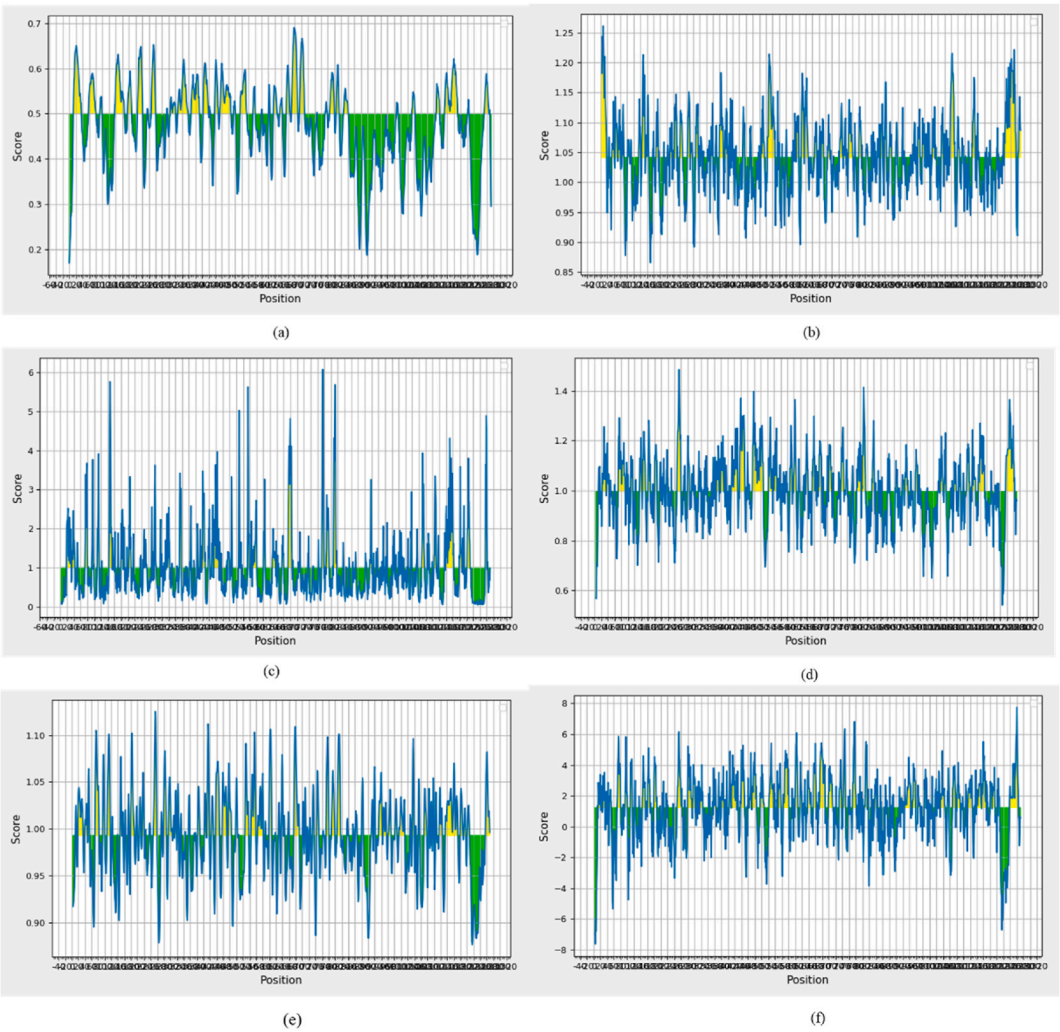


Fig. 3. **(a)** Bepipred linear epitope prediction of the S protein. **(b)** Prediction of antigenicity of S protein by Kolaskar and Tongaonkar antigenicity scale. **(c)** Emini surface accessibility prediction of S protein. **(d)** Chou and Fasman's beta turn analysis technique is used to predict beta turns in the S protein. (e) Prediction of the flexibility of the S protein by Karplus and Schulz **(f)** Prediction of Hydrophilicity of S protein by Parker HydrophilicityPrediction.

Table 1
The list of linear B cell epitopes.

Start	End	Peptide	Length	Antigenic Score	Allergenicity	Toxicity
181	186	GKQGNF	6	2.134	Non	Non
249	261	LTPGDSSSGWTAG	13	0.495	Non	Non
407	420	VRQIAPGQTGKIAD	14	1.261	Non	Non
597	606	VITPGTNTSN	10	0.4217	Non	Non
805	816	ILPDPSKPSKRS	12	0.5322	Non	Non
1157	1167	KNHTSPDVDLG	11	1.404	Non	Non

identified to be the domain binders with five alleles (HLA-A*68:02, HLA-A*02:06, HLA-A*02:03, HLA-A*02:01, HLA-B*58:01, HLA-A*26:01, HLA-A*01:01, HLA-A*30:02, HLA-A*68:01, HLA-B*35:01 (Table 3). The core epitope GKQGNF showed the highest antigenic score, 2.134(Table 1).

3.3.2. Prediction of MHC Class-II binding profile for conserved epitopes

It was found that 1300 identified conserved epitopes with an IC50 of less than 200 interacted with MHC Class-II alleles. From a total of 630 epitopes, 63 were chosen because they were associated to more than six MHC Class-II alleles. Based on allergenicity, toxicity,

Table 2
Antigenicity predicted by Kolaskar and Tongaonkar method.

Start	End	Peptide	Length
4	18	FLVLLPLVSSQCVNL	15
34	41	RGVYYPDK	8
44	51	RSSVLHST	8
53	60	DLFLPFFS	8
65	70	FHAIHV	6
81	87	NPVLPFN	7
115	121	QSLLIVN	7
125	134	NVVIKVCEFQ	10
136	146	CNDPFLGVYYH	11
168	174	FEYVSQP	7
210	216	INLVRDL	7
223	230	LEPLVDLP	8
239	248	QTLLALHRSY	10
263	270	AAYYVGYL	8
272	278	PRTFLLK	7
288	295	AVDCALDP	8
333	339	TNLCPFG	7
359	371	SNCVADYSVLYNS	13
376	385	TFKCYGVSPT	10
430	435	TGCVIA	6
485	495	GFNCYFPLQSY	11
505	527	YQPYRVVVLSFELLHAPATVCGP	23
592	599	FGGVSVIT	8
607	615	QVAVLYQGV	9
617	627	CTEVPVAIHAD	11
647	653	AGCLIGA	7
667	674	GAGICASY	8
687	693	VASQSII	7
723	733	TTEILPVSITK	11
735	741	SVDCTMY	7
750	763	SNLLLQYGSFCTQL	14
781	788	VFAQVKQI	8
803	808	SQILPD	6
837	843	YGDCLGD	7
847	853	RDLICAQ	7
858	864	LTVLPPL	7
873	880	YTSALLAG	8
959	966	LNTLVKQL	8
973	979	ISSVLND	7
1003	1011	SLQTYVTQQ	9
1030	1037	SECVLGQS	8
1057	1070	PHGVVFLHVTYVPA	14
1079	1085	PAICHDG	7
1123	1132	SGNCDVVIGI	10
1221	1256	IAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKF	36

and antigenicity, 63 epitopes were selected for further study. The core epitopes FAMQMAYRF is considered to be the top binder as it interacts with 26 alleles; (HLA-DRB1*01:01, HLA-DRB1*01:01, HLA-DRB1*01:01, HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DRB5*01:01, HLA-DRB5*01:01, HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DRB5*01:01, HLA-DRB1*07:01, HLA-DRB1*01:01, HLA-DRB1*07:01, HLA-DRB1*07:01, HLA-DRB1*07:01, HLA-DRB1*07:01, HLA-DRB3*01:01, HLA-DRB1*07:01, HLA-DRB1*07:01, HLA-DRB3*01:01, HLA-DRB1*09:01, HLA-DRB3*01:01,HLA-DRB5*01:01, HLA-DRB1*15:01) followed by LYIIKLIFL and LFLTWICLL are supposed to bind with 10 and 9 alleles, respectively (Table 4).

3.4. Construction of multi-epitope vaccine

For the multi-epitomic vaccination chimera, 6 B cell epitopes, 79 MCH Class-I and 63 MHC Class-II epitopes were used. With UniProt (ID: P9WHE3), 50S ribosomal protein L7/L12 was used as an adjuvant. The adjuvant was first linked to B cell epitopes using an EAAAK linker at the amino (N) terminus. Additionally, GPGPG and AAY linkers were used to link the B-cell, HTL and CTL epitopes. A 6× His tag was added to the vaccination sequence's C-terminus to help detect and separate the protein (Table 5). The final vaccine construct sequence had a molecular weight of 160.1596 kDa and contained 1437 amino acids the 3D structure of the construct vaccine (Fig. 4).

Table 3
Most potential non-allergen, non-allergen, no toxic 79 T-cell epitopes with interacting MHC-I alleles, IC_{50} less than 200 and percentile rank.

Peptide	Antigenic Score	Allele	Start	End	Length	Ic50	Rank	Conservancy %
AALQIPFAM	**0.7747**	HLA-B*35:01	52	60	9	26.99	0.08	100
AEIRASANL	**0.7082**	HLA-B*40:01	36	44	9	11.05	0.03	100
		HLA-B*44:03	36	44	9	71.39	0.09	
		HLA-B*44:02	36	44	9	99.35	0.14	
ALNTLVKQL	**0.5716**	HLA-A*02:03	48	56	9	50.21	0.76	100
ALQIPFAMQM	**0.962**	HLA-B*15:01	53	62	10	68.2	0.33	100
AVDCALDPL	**0.6604**	HLA-A*02:06	8	16	9	7.96	0.07	100
AYYVGYLQPR	**1.3309**	HLA-A*31:01	54	63	10	22.08	0.22	100
DLPIGINITR	**1.8171**	HLA-A*68:01	18	27	10	62.49	0.53	100
EQYIKWPWYI	**1.1122**	HLA-A*23:01	17	26	10	21.01	0.07	100
		HLA-A*24:02	17	26	10	45.33	0.08	
FAMQMAYRF	**1.0278**	HLA-B*35:01	58	66	9	5.03	0.02	100
		HLA-B*53:01	58	66	9	13.26	0.02	
		HLA-B*58:01	58	66	9	32.13	0.19	
		HLA-B*08:01	58	66	9	74.55	0.19	
FELLHAPATV	**0.5982**	HLA-A*02:03	25	34	10	8.17	0.12	100
		HLA-A*02:01	25	34	10	20.98	0.19	
		HLA-A*02:06	25	34	10	80.64	0.72	
FEYVSQPFL	**0.6324**	HLA-B*40:01	28	36	9	19.59	0.05	100
FFSNVTWFH	**0.5951**	HLA-A*33:01	58	66	9	53.66	0.21	100
FKIYSKHTPI	**1.0160**	HLA-A*02:03	61	70	10	9.13	0.13	100
FKIYSKHTPI	**0.5609**	HLA-A*02:06	61	70	10	77.23	0.69	100
FPNITNLCP	**1.6218**	HLA-B*53:01	49	57	9	39.11	0.06	100
FPREGVFVS	**0.5509**	HLA-B*35:01	39	47	9	35.75	0.11	100
FTISVTTEI	**0.8535**	HLA-A*68:02	18	26	9	3.05	0.02	100
		HLA-A*02:06	18	26	9	8.29	0.08	
		HLA-A*02:03	18	26	9	9.07	0.13	
		HLA-A*02:01	18	26	9	25.37	0.22	
		HLA-B*58:01	18	26	9	48.78	0.27	
FTISVTTEIL	**0.5849**	HLA-A*68:02	18	27	10	23.1	0.21	100
		HLA-A*02:06	18	27	10	90.6	0.8	
FVFLVLLPL	**0.8601**	HLA-A*02:06	2	10	9	19.6	0.22	100
		HLA-A*02:01	2	10	9	41.68	0.37	
		HLA-A*68:02	2	10	9	47.56	0.35	
		HLA-A*02:03	2	10	9	68.65	0.93	
FVFLVLLPLV	**0.8044**	HLA-A*02:06	2	11	10	9.2	0.09	100
		HLA-A*02:01	2	11	10	32.64	0.29	
		HLA-A*02:03	2	11	10	45.1	0.69	
		HLA-A*68:02	2	11	10	70.7	0.46	
GINASFVNI	**1.2711**	HLA-A*02:03	51	59	9	97.73	1.3	100
GLIAIVMVTI	**1.0813**	HLA-A*02:03	33	42	10	56.34	0.82	100
		HLA-A*02:01	33	42	10	80.84	0.73	
GSFCTQLNR	**0.9306**	HLA-A*31:01	57	65	9	50.93	0.55	100
		HLA-A*11:01	57	65	9	51.4	0.31	
		HLA-A*68:01	57	65	9	73.21	0.6	
GVLTESNKK	**0.8797**	HLA-A*11:01	60	68	9	64.67	0.4	100
GVVFLHVTYV	**1.4551**	HLA-A*02:03	9	18	10	39.32	0.62	100
GVYFASTEK	**0.7112**	HLA-A*11:01	19	27	9	15.17	0.06	100
		HLA-A*03:01	19	27	9	23.87	0.08	
		HLA-A*68:01	19	27	9	43	0.37	
		HLA-A*30:01	19	27	9	98.99	0.34	
GWTAGAAAYY	**0.5358**	HLA-A*30:02	47	56	10	32.48	0.06	100
		HLA-A*01:01	47	56	10	68.69	0.13	
GYLQPRTFLL	**0.7535**	HLA-A*23:01	58	67	10	32.94	0.11	100
		HLA-A*02:01	58	67	10	36.12	0.32	
HWFVTQRNFY	**0.6677**	HLA-A*30:02	51	60	10	74.74	0.2	100
IAIPTNFTI	**0.7052**	HLA-B*58:01	12	20	9	11.76	0.06	100
		HLA-A*02:06	12	20	9	54.88	0.54	
		HLA-B*53:01	12	20	9	72.25	0.08	
IGAGICASY	**0.6368**	HLA-A*30:02	36	44	9	18.04	0.02	100
INASFVNIQK	**1.2242**	HLA-A*68:01	52	61	10	38.22	0.33	100
		HLA-A*11:01	52	61	10	66.7	0.41	
IPTNFTISV	**0.8820**	HLA-B*07:02	14	22	9	95.09	0.33	100
ITPCSFGGV	**1.3871**	HLA-A*68:02	27	35	9	20.32	0.18	100
		HLA-A*02:06	27	35	9	25.07	0.29	
KEIDRLNEV	**0.53**	HLA-A*02:06	61	69	9	85.27	0.77	100
KIYSKHTPI	**0.7455**	HLA-A*02:03	62	70	9	9.39	0.14	100
		HLA-A*32:01	62	70	9	10.98	0.03	

(*continued on next page*)

Table 3 (*continued*)

Peptide	Antigenic Score	Allele	Start	End	Length	Ic50	Rank	Conservancy %
		HLA-A*02:06	62	70	9	68.18	0.63	
KLNDLCFTNV	**2.6927**	HLA-A*02:03	36	45	10	6.22	0.08	100
		HLA-A*02:01	36	45	10	15.27	0.14	
		HLA-A*02:06	36	45	10	28.21	0.31	
KNLNESLIDL	**0.7093**	HLA-A*02:03	1	10	10	85.66	1.2	100
KSNLKPFER	**0.949**	HLA-A*31:01	38	46	9	8.83	0.07	100
LLFNKVTLA	**0.615**	HLA-A*02:03	51	59	9	4.98	0.07	100
		HLA-A*02:01	51	59	9	25.41	0.22	
LPFNDGVYF	**0.5593**	HLA-B*35:01	14	22	9	9.15	0.03	100
		HLA-B*53:01	14	22	9	20.89	0.03	
LPIGINITRF	**1.3027**	HLA-B*53:01	19	28	10	25.21	0.04	100
		HLA-B*35:01	19	28	10	32.8	0.1	
LQSYGFQPT	**0.7917**	HLA-A*02:06	2	10	9	18.53	0.21	100
LQYGSFCTQL	**1.4443**	HLA-B*15:01	54	63	10	13.44	0.07	100
		HLA-A*02:06	54	63	10	24.73	0.29	
LSFELLHAPA	**0.9434**	HLA-A*02:06	23	32	10	69.11	0.64	100
NFTISVTTEI	**1.0857**	HLA-A*68:02	17	26	10	13.27	0.12	100
NKSWMESEFR	**0.6813**	HLA-A*31:01	9	18	10	83.45	0.84	100
		HLA-A*68:01	9	18	10	85.09	0.67	
NTQEVFAQVK	**0.5676**	HLA-A*68:01	7	16	10	53.44	0.48	100
NVTWFHAIHV	**0.9051**	HLA-A*68:02	61	70	10	25.28	0.23	100
PFAMQMAYR	**1.3315**	HLA-A*33:01	57	65	9	84.41	0.35	100
PFAMQMAYRF	**1.1051**	HLA-A*23:01	57	66	10	28.72	0.09	100
		HLA-A*24:02	57	66	10	38.93	0.07	
PIGAGICASY	**0.8411**	HLA-A*30:02	35	44	10	84.37	0.22	100
PYRVVVLSF	**1.0281**	HLA-A*23:01	17	25	9	30.39	0.1	100
QIPFAMQMAY	**1.2149**	HLA-B*35:01	55	64	10	23.9	0.07	100
		HLA-B*15:01	55	64	10	55.7	0.29	
QTRAGCLIGA	**1.3933**	HLA-A*68:02	14	23	10	39.78	0.32	100
RASANLAATK	**0.6339**	HLA-A*11:01	39	48	10	71.1	0.44	100
RKSNLKPFER	**0.8187**	HLA-A*31:01	37	46	10	29.47	0.31	100
RQIAPGQTGK	**1.7893**	HLA-A*03:01	58	67	10	38	0.13	100
SLIDLQELGK	**1.0275**	HLA-A*11:01	6	15	10	38.67	0.22	100
SPRRARSVA	**0.7729**	HLA-B*07:02	50	58	9	4.17	0.01	100
SPRRARSVAS	**0.5591**	HLA-B*07:02	50	59	10	18.4	0.07	100
TAGAAAYYV	**0.6397**	HLA-A*68:02	49	57	9	13.38	0.12	100
TFEYVSQPF	**0.6641**	HLA-A*23:01	27	35	9	81.44	0.24	100
TLADAGFIK	**0.5781**	HLA-A*11:01	57	65	9	48.34	0.28	100
		HLA-A*68:01	57	65	9	60.95	0.52	
TSVDCTMYI	**0.7328**	HLA-A*68:02	34	42	9	3.83	0.03	100
		HLA-A*02:06	34	42	9	81.13	0.73	
TTRTQLPPA	**1.254**	HLA-A*30:01	19	27	9	8.31	0.05	100
TYVPAQEKNF	**0.7276**	HLA-A*23:01	16	25	10	68.13	0.2	100
VLSFELLHA	**1.0776**	HLA-A*02:03	22	30	9	33.91	0.55	100
VTLADAGFIK	**0.8702**	HLA-A*11:01	56	65	10	23.29	0.12	100
VTWFHAIHV	**0.5426**	HLA-A*68:02	62	70	9	35.91	0.31	100
		HLA-A*02:06	62	70	9	74.83	0.67	
VVFLHVTYV	**1.5122**	HLA-A*02:03	10	18	9	13.02	0.21	100
		HLA-A*02:06	10	18	9	21.97	0.25	
		HLA-A*02:01	10	18	9	36.56	0.33	
		HLA-A*68:02	10	18	9	51.51	0.38	
WIFGTTLDSK	**0.6625**	HLA-A*03:01	34	43	10	62.53	0.24	100
WTAGAAAY	**0.6306**	HLA-A*26:01	48	56	9	11.63	0.02	100
		HLA-A*01:01	48	56	9	12.27	0.03	
		HLA-A*30:02	48	56	9	16.16	0.02	
		HLA-A*68:01	48	56	9	30.13	0.26	
		HLA-B*35:01	48	56	9	66.67	0.18	
WTAGAAAYYV	**0.5371**	HLA-A*68:02	48	57	10	2.8	0.02	100
		HLA-A*02:06	48	57	10	38.43	0.41	
YGSFCTQLNR	**0.9565**	HLA-A*68:01	56	65	10	20.16	0.17	100
YIWLGFIAGL	**0.5798**	HLA-A*02:01	25	34	10	20.48	0.19	100
		HLA-A*02:06	25	34	10	52.25	0.52	
		HLA-A*02:03	25	34	10	60.94	0.86	
YQPYRVVVL	**0.5964**	HLA-A*02:06	15	23	9	99.74	0.85	100
YTSALLAGT	**0.5487**	HLA-A*68:02	33	41	9	17.78	0.16	100
		HLA-A*02:06	33	41	9	78.35	0.7	
		HLA-A*02:03	33	41	9	79.29	1.2	
YYVGYLQPR	**1.4692**	HLA-A*33:01	55	63	9	21.9	0.07	100
		HLA-A*31:01	55	63	9	56.41	0.61	

Table 4
Most potential non-allergen, nontoxic, 63 T-cell epitopes with interacting MHC-II alleles, epitope conservancy score, IC_{50} less than 200 and percentile rank.

Core Peptide	Antigen score	Allele	Start	End	Length	Peptide	Ic50	Rank	Conservancy%
AALQIPFAM	0.7747	HLA-DRB4*01:01	890	904	15	AGAALQIPFAMQMAY	58.2	3	100
AGAAAYYVG	0.5762	HLA-DQA1*05:01/DQB1*03:01	256	270	15	SGWTAGAAAYYVGYL	28.4	1.1	100
		HLA-DQA1*05:01/DQB1*03:01	257	271	15	GWTAGAAAYYVGYLQ	32.1	1.4	
		HLA-DQA1*05:01/DQB1*03:01	258	272	15	WTAGAAAYYVGYLQP	39.7	2	
		HLA-DQA1*05:01/DQB1*03:01	259	273	15	TAGAAAYYVGYLQPR	77	4.8	
CLGDIAARD	1.4292	HLA-DRB1*03:01	837	851	15	YGDCLGDIAARDLIC	98.4	4.1	100
CVLGQSKRV	0.9083	HLA-DRB1*01:01	1028	1042	15	KMSECVLGQSKRVDF	29.5	10	100
		HLA-DRB1*01:01	1029	1043	15	MSECVLGQSKRVDFC	30.4	11	
		HLA-DRB1*01:01	1030	1044	15	SECVLGQSKRVDFCG	38.2	13	
		HLA-DRB1*01:01	1027	1041	15	TKMSECVLGQSKRVD	46.6	15	
		HLA-DRB1*01:01	1031	1045	15	ECVLGQSKRVDFCGK	62.7	19	
DFGGFNFSQ	1.1378	HLA-DPA1*02:01/DPB1*01:01	794	808	15	IKDFGGFNFSQILPD	97.1	1.8	100
DQLTPTWRV	0.7445	HLA-DRB1*13:02	621	635	15	PVAIHADQLTPTWRV	96.2	9.8	100
DSKTQSLLI	0.5911	HLA-DRB1*07:01	106	120	15	FGTTLDSKTQSLLIV	25.3	2.9	100
		HLA-DRB1*07:01	107	121	15	GTTLDSKTQSLLIVN	30.1	3.6	
		HLA-DRB1*07:01	105	119	15	IFGTTLDSKTQSLLI	30.3	3.6	
		HLA-DRB1*07:01	108	122	15	TTLDSKTQSLLIVNN	35.3	4.2	
		HLA-DRB1*07:01	109	123	15	TLDSKTQSLLIVNNA	44	5.3	
		HLA-DRB1*07:01	110	124	15	LDSKTQSLLIVNNAT	63.9	7.5	
FAMQMAYRF	1.0278	HLA-DRB1*01:01	895	909	15	QIPFAMQMAYRFNGI	7.3	1.6	100
		HLA-DRB1*01:01	894	908	15	LQIPFAMQMAYRFNG	8.1	2	
		HLA-DRB1*01:01	896	910	15	IPFAMQMAYRFNGIG	9.8	2.6	
		HLA-DRB1*01:01	893	907	15	ALQIPFAMQMAYRFN	10.3	2.8	
		HLA-DRB5*01:01	893	907	15	ALQIPFAMQMAYRFN	14.9	2.1	
		HLA-DRB1*01:01	897	911	15	PFAMQMAYRFNGIGV	15.8	5.1	
		HLA-DRB5*01:01	894	908	15	LQIPFAMQMAYRFNG	15.9	2.3	
		HLA-DRB5*01:01	895	909	15	QIPFAMQMAYRFNGI	17.2	2.6	
		HLA-DRB5*01:01	892	906	15	AALQIPFAMQMAYRF	17.5	2.7	
		HLA-DRB1*01:01	892	906	15	AALQIPFAMQMAYRF	18.8	6.2	
		HLA-DRB5*01:01	896	910	15	IPFAMQMAYRFNGIG	23.4	3.7	
		HLA-DRB5*01:01	897	911	15	PFAMQMAYRFNGIGV	35.9	5.4	
		HLA-DRB1*07:01	892	906	15	AALQIPFAMQMAYRF	36.2	4.3	
		HLA-DRB1*01:01	898	912	15	FAMQMAYRFNGIGVT	40.3	14	
		HLA-DRB1*07:01	893	907	15	ALQIPFAMQMAYRFN	41.7	5	
		HLA-DRB1*07:01	897	911	15	PFAMQMAYRFNGIGV	56.1	6.6	
		HLA-DRB1*07:01	895	909	15	QIPFAMQMAYRFNGI	56.3	6.7	
		HLA-DRB1*07:01	894	908	15	LQIPFAMQMAYRFNG	62.9	7.3	
		HLA-DRB3*01:01	893	907	15	ALQIPFAMQMAYRFN	66.6	2.8	
		HLA-DRB1*07:01	896	910	15	IPFAMQMAYRFNGIG	70.7	8.2	
		HLA-DRB1*07:01	898	912	15	FAMQMAYRFNGIGVT	76.8	8.8	
		HLA-DRB3*01:01	895	909	15	QIPFAMQMAYRFNGI	79.5	3.2	
		HLA-DRB1*09:01	892	906	15	AALQIPFAMQMAYRF	83.3	8.3	
		HLA-DRB3*01:01	894	908	15	LQIPFAMQMAYRFNG	88.4	3.6	
		HLA-DRB5*01:01	898	912	15	FAMQMAYRFNGIGVT	98.8	13	
		HLA-DRB1*15:01	897	911	15	PFAMQMAYRFNGIGV	99.6	7.2	
FCTQLNRAL	0.5159	HLA-DRB1*01:01	756	770	15	YGSFCTQLNRALTGI	32.1	11	100
		HLA-DRB1*01:01	757	771	15	GSFCTQLNRALTGIA	34.4	12	
		HLA-DRB1*01:01	755	769	15	QYGSFCTQLNRALTG	37.7	13	
		HLA-DRB1*01:01	754	768	15	LQYGSFCTQLNRALT	40.9	14	
		HLA-DRB1*01:01	753	767	15	LLQYGSFCTQLNRAL	60.1	19	
		HLA-DRB1*01:01	512	526	15	VLSFELLHAPATVCG	2.7	0.03	
		HLA-DRB1*01:01	513	527	15	LSFELLHAPATVCGP	2.8	0.03	
		HLA-DRB1*01:01	511	525	15	VVLSFELLHAPATVC	2.9	0.03	
		HLA-DRB1*01:01	514	528	15	SFELLHAPATVCGPK	3.3	0.09	
		HLA-DRB1*01:01	510	524	15	VVVLSFELLHAPATV	3.3	0.09	
		HLA-DRB1*01:01	515	529	15	FELLHAPATVCGPKK	5.6	0.71	
		HLA-DRB1*01:01	509	523	15	RVVVLSFELLHAPAT	6.2	0.96	
		HLA-DRB1*04:05	511	525	15	VVLSFELLHAPATVC	37.7	1.1	
		HLA-DRB1*04:05	512	526	15	VLSFELLHAPATVCG	42.4	1.3	
		HLA-DRB1*04:05	510	524	15	VVVLSFELLHAPATV	45.8	1.6	
		HLA-DRB1*04:05	513	527	15	LSFELLHAPATVCGP	53.2	2	
		HLA-DRB1*04:05	509	523	15	RVVVLSFELLHAPAT	68.3	3	
		HLA-DRB1*15:01	511	525	15	VVLSFELLHAPATVC	82.8	6.1	
		HLA-DRB1*09:01	511	525	15	VVLSFELLHAPATVC	96.7	9.6	
		HLA-DRB1*04:01	510	524	15	VVVLSFELLHAPATV	99.3	4.3	
FEYVSQPFL	0.6324	HLA-DRB1*01:01	165	179	15	NCTFEYVSQPFLMDL	20.5	6.8	100
		HLA-DRB1*01:01	166	180	15	CTFEYVSQPFLMDLE	22.5	7.6	

(*continued on next page*)

Table 4 (*continued*)

Core Peptide	Antigen score	Allele	Start	End	Length	Peptide	Ic50	Rank	Conservancy%
		HLA-DRB1*01:01	167	181	15	TFEYVSQPFLMDLEG	27.4	9.3	
		HLA-DRB1*01:01	164	178	15	NNCTFEYVSQPFLMD	32.7	11	
		HLA-DPA1*01:03/DPB1*02:01	163	177	15	ANNCTFEYVSQPFLM	43.5	1.5	
		HLA-DRB1*01:01	163	177	15	ANNCTFEYVSQPFLM	48.4	16	
		HLA-DQA1*05:01/DQB1*02:01	164	178	15	NNCTFEYVSQPFLMD	56	0.89	
		HLA-DRB1*07:01	163	177	15	ANNCTFEYVSQPFLM	56.2	6.7	
		HLA-DQA1*05:01/DQB1*02:01	165	179	15	NCTFEYVSQPFLMDL	57.9	0.95	
		HLA-DPA1*01:03/DPB1*02:01	162	176	15	SANNCTFEYVSQPFL	61.3	2.6	
		HLA-DRB1*07:01	165	179	15	NCTFEYVSQPFLMDL	62.7	7.3	
		HLA-DRB1*07:01	164	178	15	NNCTFEYVSQPFLMD	64.6	7.6	
		HLA-DQA1*05:01/DQB1*02:01	166	180	15	CTFEYVSQPFLMDLE	65.2	1.3	
		HLA-DQA1*05:01/DQB1*02:01	163	177	15	ANNCTFEYVSQPFLM	69.3	1.5	
		HLA-DQA1*05:01/DQB1*02:01	162	176	15	SANNCTFEYVSQPFL	77.9	1.7	
		HLA-DRB1*07:01	166	180	15	CTFEYVSQPFLMDLE	78.9	9	
		HLA-DQA1*05:01/DQB1*02:01	167	181	15	TFEYVSQPFLMDLEG	81	1.8	
		HLA-DRB1*07:01	162	176	15	SANNCTFEYVSQPFL	95.1	11	
FFSNVTWFH	0.5951	HLA-DRB3*02:02	55	69	15	FLPFFSNVTWFHAIH	41.9	2.1	100
		HLA-DRB1*01:01	56	70	15	LPFFSNVTWFHAIHV	48.7	16	
		HLA-DRB3*02:02	56	70	15	LPFFSNVTWFHAIHV	49	2.3	
		HLA-DRB3*02:02	54	68	15	LFLPFFSNVTWFHAI	54.6	2.6	
		HLA-DRB5*01:01	55	69	15	FLPFFSNVTWFHAIH	60	8.5	
		HLA-DRB1*04:05	54	68	15	LFLPFFSNVTWFHAI	62	2.5	
		HLA-DPA1*02:01/DPB1*01:01	55	69	15	FLPFFSNVTWFHAIH	63.6	0.76	
		HLA-DPA1*03:01/DPB1*04:02	55	69	15	FLPFFSNVTWFHAIH	66.7	2.7	
		HLA-DPA1*02:01/DPB1*01:01	54	68	15	LFLPFFSNVTWFHAI	67.6	0.91	
		HLA-DPA1*02:01/DPB1*01:01	53	67	15	DLFLPFFSNVTWFHA	69.5	0.94	
		HLA-DPA1*03:01/DPB1*04:02	56	70	15	LPFFSNVTWFHAIHV	71	3	
		HLA-DPA1*03:01/DPB1*04:02	54	68	15	LFLPFFSNVTWFHAI	72.3	3	
		HLA-DPA1*03:01/DPB1*04:02	53	67	15	DLFLPFFSNVTWFHA	72.9	3	
		HLA-DPA1*02:01/DPB1*01:01	56	70	15	LPFFSNVTWFHAIHV	73.7	1.1	
		HLA-DRB3*02:02	53	67	15	DLFLPFFSNVTWFHA	79.5	3.4	
		HLA-DRB1*04:05	55	69	15	FLPFFSNVTWFHAIH	87.2	4.3	
FGAISSVLN	0.5435	HLA-DRB1*01:01	967	981	15	SSNFGAISSVLNDIL	6.3	0.99	100
		HLA-DRB1*01:01	966	980	15	LSSNFGAISSVLNDI	7.4	1.6	
		HLA-DRB1*01:01	968	982	15	SNFGAISSVLNDILS	8	1.9	
		HLA-DRB1*01:01	965	979	15	QLSSNFGAISSVLND	9.5	2.4	
		HLA-DRB1*01:01	969	983	15	NFGAISSVLNDILSR	14.5	4.6	
		HLA-DRB1*01:01	964	978	15	KQLSSNFGAISSVLN	15.7	5	
		HLA-DRB1*01:01	970	984	15	FGAISSVLNDILSRL	35.1	12	
		HLA-DRB1*09:01	964	978	15	KQLSSNFGAISSVLN	59.4	5.6	
		HLA-DRB1*09:01	965	979	15	QLSSNFGAISSVLND	73.7	7.1	
		HLA-DRB1*09:01	966	980	15	LSSNFGAISSVLNDI	96.4	9.6	
		HLA-DRB1*04:01	967	981	15	SSNFGAISSVLNDIL	99.9	4.4	
FGGFNFSQI	1.273	HLA-DPA1*01:03/DPB1*04:01	794	808	15	IKDFGGFNFSQILPD	22.4	0.56	100
		HLA-DPA1*01:03/DPB1*04:01	793	807	15	PIKDFGGFNFSQILP	24.3	0.65	
		HLA-DPA1*01:03/DPB1*04:01	792	806	15	PPIKDFGGFNFSQIL	27.9	0.79	
		HLA-DPA1*01:03/DPB1*04:01	791	805	15	TPPIKDFGGFNFSQI	41.2	1.3	
		HLA-DPA1*01:03/DPB1*04:01	795	809	15	KDFGGFNFSQILPDP	47.1	1.5	
		HLA-DRB1*09:01	794	808	15	IKDFGGFNFSQILPD	94.8	9.5	
		HLA-DRB1*09:01	796	810	15	DFGGFNFSQILPDPS	97.9	9.8	
FKIYSKHTP	0.9886	HLA-DRB1*11:01	199	213	15	GYFKIYSKHTPINLV	8.5	0.57	100
		HLA-DRB1*11:01	198	212	15	DGYFKIYSKHTPINL	9.3	0.68	
		HLA-DRB1*11:01	197	211	15	IDGYFKIYSKHTPIN	9.9	0.73	
		HLA-DRB1*11:01	196	210	15	NIDGYFKIYSKHTPI	12.2	1.2	
		HLA-DRB1*11:01	195	209	15	KNIDGYFKIYSKHTP	19.7	2	
		HLA-DRB1*11:01	200	214	15	YFKIYSKHTPINLVR	28.7	3	
		HLA-DRB1*15:01	198	212	15	DGYFKIYSKHTPINL	92.5	6.7	
		HLA-DRB1*15:01	199	213	15	GYFKIYSKHTPINLV	93.9	6.7	
FLGVYYHKN	0.9675	HLA-DPA1*01:03/DPB1*04:01	134	148	15	QFCNDPFLGVYYHKN	27.6	0.78	100
		HLA-DPA1*01:03/DPB1*04:01	135	149	15	FCNDPFLGVYYHKNN	32.4	1	
		HLA-DPA1*01:03/DPB1*04:01	136	150	15	CNDPFLGVYYHKNNK	47.2	1.5	
		HLA-DPA1*01:03/DPB1*04:01	137	151	15	NDPFLGVYYHKNNKS	55.2	1.8	
		HLA-DPA1*01:03/DPB1*02:01	134	148	15	QFCNDPFLGVYYHKN	70.7	3	
FNATRFASV	0.5609	HLA-DPA1*01:03/DPB1*04:01	338	352	15	FGEVFNATRFASVYA	4.7	0.03	100
		HLA-DPA1*01:03/DPB1*04:01	337	351	15	PFGEVFNATRFASVY	4.8	0.03	
		HLA-DPA1*01:03/DPB1*04:01	339	353	15	GEVFNATRFASVYAW	5.5	0.06	
		HLA-DPA1*01:03/DPB1*04:01	336	350	15	CPFGEVFNATRFASV	5.9	0.06	
		HLA-DPA1*01:03/DPB1*04:01	340	354	15	EVFNATRFASVYAWN	10.7	0.2	
		HLA-DPA1*01:03/DPB1*04:01	341	355	15	VFNATRFASVYAWNR	25.8	0.75	

(*continued on next page*)

Table 4 (*continued*)

Core Peptide	Antigen score	Allele	Start	End	Length	Peptide	Ic50	Rank	Conservancy%
		HLA-DRB1*07:01	336	350	15	CPFGEVFNATRFASV	37.8	4.5	
		HLA-DRB1*07:01	337	351	15	PFGEVFNATRFASVY	40.5	4.9	
		HLA-DRB1*07:01	338	352	15	FGEVFNATRFASVYA	41.6	5	
		HLA-DRB1*07:01	341	355	15	VFNATRFASVYAWNR	44.8	5.3	
		HLA-DRB1*07:01	339	353	15	GEVFNATRFASVYAW	50.6	6	
		HLA-DRB1*07:01	340	354	15	EVFNATRFASVYAWN	57.8	6.8	
		HLA-DRB1*07:01	342	356	15	FNATRFASVYAWNRK	68.8	8.1	
		HLA-DRB1*09:01	339	353	15	GEVFNATRFASVYAW	88.6	8.9	
FNFSQILPD	0.5831	HLA-DRB1*04:05	797	811	15	FGGFNFSQILPDPSK	37.8	1.1	100
		HLA-DRB1*04:05	796	810	15	DFGGFNFSQILPDPS	48.8	1.8	
		HLA-DRB1*07:01	794	808	15	IKDFGGFNFSQILPD	77.1	8.8	
		HLA-DRB1*04:05	795	809	15	KDFGGFNFSQILPDP	94.7	4.9	
FRKSNLKPF	0.628	HLA-DRB1*07:01	451	465	15	YLYRLFRKSNLKPFE	14.5	1.4	100
		HLA-DRB1*07:01	452	466	15	LYRLFRKSNLKPFER	16.3	1.7	
		HLA-DRB1*07:01	450	464	15	NYLYRLFRKSNLKPF	17.2	1.8	
		HLA-DRB1*07:01	453	467	15	YRLFRKSNLKPFERD	20.7	2.3	
		HLA-DRB1*07:01	454	468	15	RLFRKSNLKPFERDI	32.5	3.9	
		HLA-DRB1*07:01	455	469	15	LFRKSNLKPFERDIS	57.1	6.7	
FSNVTWFHA	0.8156	HLA-DRB1*07:01	55	69	15	FLPFFSNVTWFHAIH	16.6	1.7	100
		HLA-DPA1*01:03/DPB1*04:01	54	68	15	LFLPFFSNVTWFHAI	28.1	0.8	
		HLA-DPA1*01:03/DPB1*04:01	55	69	15	FLPFFSNVTWFHAIH	28.8	0.81	
		HLA-DPA1*01:03/DPB1*04:01	53	67	15	DLFLPFFSNVTWFHA	29.1	0.84	
		HLA-DPA1*01:03/DPB1*02:01	53	67	15	DLFLPFFSNVTWFHA	33.2	1.1	
		HLA-DPA1*01:03/DPB1*02:01	54	68	15	LFLPFFSNVTWFHAI	35	1.2	
		HLA-DPA1*01:03/DPB1*04:01	56	70	15	LPFFSNVTWFHAIHV	36.6	1.2	
		HLA-DPA1*01:03/DPB1*02:01	55	69	15	FLPFFSNVTWFHAIH	50.1	2	
		HLA-DRB1*01:01	57	71	15	PFFSNVTWFHAIHVS	55.6	17	
		HLA-DPA1*01:03/DPB1*02:01	56	70	15	LPFFSNVTWFHAIHV	57.5	2.3	
		HLA-DPA1*01:03/DPB1*04:01	57	71	15	PFFSNVTWFHAIHVS	59.8	1.9	
		HLA-DRB1*04:01	55	69	15	FLPFFSNVTWFHAIH	69.1	2.7	
		HLA-DPA1*03:01/DPB1*04:02	57	71	15	PFFSNVTWFHAIHVS	73.9	3.1	
		HLA-DPA1*01:03/DPB1*02:01	57	71	15	PFFSNVTWFHAIHVS	91.3	3.9	
FTISVTTEI	0.8535	HLA-DRB1*07:01	712	726	15	IAIPTNFTISVTTEI	6.4	0.34	100
		HLA-DRB1*07:01	713	727	15	AIPTNFTISVTTEIL	6.5	0.35	
		HLA-DRB1*07:01	715	729	15	PTNFTISVTTEILPV	8	0.51	
		HLA-DRB1*07:01	716	730	15	TNFTISVTTEILPVS	8.2	0.52	
		HLA-DRB1*07:01	714	728	15	IPTNFTISVTTEILP	8.2	0.52	
		HLA-DRB1*07:01	717	731	15	NFTISVTTEILPVSI	9.6	0.69	
		HLA-DRB1*07:01	718	732	15	FTISVTTEILPVSIT	13.7	1.3	
		HLA-DRB1*09:01	713	727	15	AIPTNFTISVTTEIL	13.8	0.49	
		HLA-DRB1*09:01	715	729	15	PTNFTISVTTEILPV	14.6	0.59	
		HLA-DRB1*09:01	714	728	15	IPTNFTISVTTEILP	15	0.64	
		HLA-DRB1*09:01	716	730	15	TNFTISVTTEILPVS	19.7	1.1	
		HLA-DRB1*09:01	712	726	15	IAIPTNFTISVTTEI	20.2	1.1	
		HLA-DRB1*09:01	717	731	15	NFTISVTTEILPVSI	26.4	1.8	
		HLA-DRB1*13:02	713	727	15	AIPTNFTISVTTEIL	30.7	4.2	
		HLA-DRB1*13:02	715	729	15	PTNFTISVTTEILPV	32.2	4.3	
		HLA-DRB1*13:02	714	728	15	IPTNFTISVTTEILP	40.9	5.3	
		HLA-DRB1*13:02	712	726	15	IAIPTNFTISVTTEI	45.3	5.6	
		HLA-DRB1*01:01	715	729	15	PTNFTISVTTEILPV	46	15	
		HLA-DRB1*01:01	716	730	15	TNFTISVTTEILPVS	51.6	17	
		HLA-DRB1*01:01	714	728	15	IPTNFTISVTTEILP	62	19	
		HLA-DRB3*01:01	713	727	15	AIPTNFTISVTTEIL	64.7	2.7	
		HLA-DRB1*01:01	713	727	15	AIPTNFTISVTTEIL	68.1	20	
		HLA-DRB3*01:01	715	729	15	PTNFTISVTTEILPV	76.9	3.1	
		HLA-DRB3*01:01	714	728	15	IPTNFTISVTTEILP	78.5	3.2	
		HLA-DRB1*09:01	718	732	15	FTISVTTEILPVSIT	80.5	8	
		HLA-DRB3*02:02	715	729	15	PTNFTISVTTEILPV	86.2	3.7	
		HLA-DRB1*01:01	717	731	15	NFTISVTTEILPVSI	92.1	25	
GINITRFQT	0.6699	HLA-DRB4*01:01	229	243	15	LPIGINITRFQTLLA	57.7	3	100
GVGYQPYRV	1.3587	HLA-DRB1*01:01	500	514	15	TNGVGYQPYRVVVLS	50.7	16	100
		HLA-DRB1*01:01	499	513	15	PTNGVGYQPYRVVVL	64.5	19	
IAPGQTGKI	1.6527	HLA-DQA1*05:01/DQB1*03:01	406	420	15	EVRQIAPGQTGKIAD	95.5	6	100
		HLA-DQA1*05:01/DQB1*03:01	405	419	15	DEVRQIAPGQTGKIA	95.5	6	
		HLA-DQA1*05:01/DQB1*03:01	407	421	15	VRQIAPGQTGKIADY	95.9	6	
IAWNSNNLD	1.3597	HLA-DRB3*02:02	431	445	15	GCVIAWNSNNLDSKV	14.1	0.86	100
		HLA-DRB3*02:02	430	444	15	TGCVIAWNSNNLDSK	15	0.91	
		HLA-DRB1*13:02	431	445	15	GCVIAWNSNNLDSKV	17.2	2.4	
		HLA-DRB1*13:02	430	444	15	TGCVIAWNSNNLDSK	19	2.6	

(*continued on next page*)

Table 4 (*continued*)

Core Peptide	Antigen score	Allele	Start	End	Length	Peptide	Ic50	Rank	Conservancy%
		HLA-DRB3*02:02	429	443	15	FTGCVIAWNSNNLDS	19.2	1.2	
		HLA-DRB1*13:02	429	443	15	FTGCVIAWNSNNLDS	20.2	2.8	
		HLA-DRB1*13:02	432	446	15	CVIAWNSNNLDSKVG	21.8	3.1	
		HLA-DRB3*02:02	432	446	15	CVIAWNSNNLDSKVG	24.4	1.4	
		HLA-DRB1*13:02	428	442	15	DFTGCVIAWNSNNLD	25.3	3.5	
		HLA-DRB1*13:02	433	447	15	VIAWNSNNLDSKVGG	48.2	5.9	
		HLA-DRB1*04:01	431	445	15	GCVIAWNSNNLDSKV	90	3.7	
		HLA-DRB1*04:01	430	444	15	TGCVIAWNSNNLDSK	92.4	3.8	
		HLA-DRB1*04:01	429	443	15	FTGCVIAWNSNNLDS	96.2	4.1	
IAYTMSLGA	1.0212	HLA-DRB1*07:01	689	703	15	SQSIIAYTMSLGAEN	7.9	0.49	100
		HLA-DRB1*07:01	688	702	15	ASQSIIAYTMSLGAE	8.3	0.52	
		HLA-DRB1*07:01	687	701	15	VASQSIIAYTMSLGA	8.4	0.54	
		HLA-DRB1*07:01	690	704	15	QSIIAYTMSLGAENS	8.6	0.55	
		HLA-DRB1*07:01	691	705	15	SIIAYTMSLGAENSV	8.8	0.59	
		HLA-DRB1*07:01	692	706	15	IIAYTMSLGAENSVA	10.5	0.78	
		HLA-DRB1*07:01	693	707	15	IAYTMSLGAENSVAY	15.4	1.5	
		HLA-DRB1*15:01	689	703	15	SQSIIAYTMSLGAEN	28.4	1.6	
		HLA-DRB1*15:01	690	704	15	QSIIAYTMSLGAENS	29.9	1.8	
		HLA-DRB1*15:01	691	705	15	SIIAYTMSLGAENSV	39.1	2.6	
		HLA-DRB1*01:01	688	702	15	ASQSIIAYTMSLGAE	59.6	19	
		HLA-DRB1*09:01	688	702	15	ASQSIIAYTMSLGAE	96.6	9.6	
IGINITRFQ	1.3386	HLA-DRB1*13:02	228	242	15	DLPIGINITRFQTLL	39.9	5.2	100
		HLA-DRB1*13:02	227	241	15	VDLPIGINITRFQTL	40.9	5.3	
		HLA-DRB1*13:02	229	243	15	LPIGINITRFQTLLA	45.5	5.7	
		HLA-DRB1*13:02	225	239	15	PLVDLPIGINITRFQ	47.8	5.8	
		HLA-DRB1*13:02	230	244	15	PIGINITRFQTLLAL	53.6	6.4	
		HLA-DRB1*13:02	226	240	15	LVDLPIGINITRFQT	83.9	8.9	
IHVSGTNGT	0.8621	HLA-DRB1*07:01	64	78	15	WFHAIHVSGTNGTKR	79	9	100
		HLA-DRB1*07:01	65	79	15	FHAIHVSGTNGTKRF	79.7	9.1	
		HLA-DRB1*07:01	63	77	15	TWFHAIHVSGTNGTK	93.6	11	
IIAYTMSLG	0.8621	HLA-DRB1*15:01	688	702	15	ASQSIIAYTMSLGAE	31.2	1.9	100
		HLA-DRB1*15:01	687	701	15	VASQSIIAYTMSLGA	44.5	3	
		HLA-DRB1*15:01	686	700	15	SVASQSIIAYTMSLG	83.2	6.1	
INASFVNIQ	1.7246	HLA-DRB1*13:02	1169	1183	15	ISGINASFVNIQKEI	76.5	8.3	100
IYQTSNFRV	0.7886	HLA-DRB1*01:01	309	323	15	EKGIYQTSNFRVQPT	36.2	12	100
		HLA-DRB1*01:01	310	324	15	KGIYQTSNFRVQPTE	41.5	14	
		HLA-DRB1*01:01	308	322	15	VEKGIYQTSNFRVQP	43.7	15	
		HLA-DRB1*01:01	307	321	15	TVEKGIYQTSNFRVQ	49.8	16	
		HLA-DRB1*01:01	311	325	15	GIYQTSNFRVQPTES	79.1	22	
		HLA-DRB1*07:01	307	321	15	TVEKGIYQTSNFRVQ	97.4	11	
		HLA-DRB1*13:02	307	321	15	TVEKGIYQTSNFRVQ	99.4	10	
LGFIAGLIA	0.5797	HLA-DRB1*01:01	1214	1228	15	WYIWLGFIAGLIAIV	65	20	100
LKGVKLHYT	1.5748	HLA-DRB1*12:01	1259	1273	15	DDSEPVLKGVKLHYT	58.6	1.2	100
LLFNKVTLA	0.615	HLA-DRB1*13:02	817	831	15	FIEDLLFNKVTLADA	12.8	1.7	100
		HLA-DRB1*13:02	816	830	15	SFIEDLLFNKVTLAD	12.8	1.7	
		HLA-DRB1*13:02	815	829	15	RSFIEDLLFNKVTLA	13	1.7	
		HLA-DRB1*13:02	818	832	15	IEDLLFNKVTLADAG	14.6	1.9	
		HLA-DRB1*13:02	819	833	15	EDLLFNKVTLADAGF	22.4	3.1	
		HLA-DRB1*13:02	820	834	15	DLLFNKVTLADAGFI	30.1	4.1	
		HLA-DRB1*01:01	819	833	15	EDLLFNKVTLADAGF	52.4	17	
		HLA-DRB1*01:01	818	832	15	IEDLLFNKVTLADAG	58.5	18	
		HLA-DRB1*13:02	821	835	15	LLFNKVTLADAGFIK	65.4	7.4	
		HLA-DRB3*02:02	817	831	15	FIEDLLFNKVTLADA	70.5	3.1	
		HLA-DRB1*01:01	817	831	15	FIEDLLFNKVTLADA	73.1	21	
		HLA-DRB3*02:02	818	832	15	IEDLLFNKVTLADAG	74.2	3.3	
		HLA-DRB1*01:01	816	830	15	SFIEDLLFNKVTLAD	99.1	26	
		HLA-DRB1*01:01	815	829	15	RSFIEDLLFNKVTLA	99.2	26	
LLQYGSFCT	0.8188	HLA-DRB1*15:01	750	764	15	SNLLLQYGSFCTQLN	10.5	0.27	100
		HLA-DRB1*15:01	749	763	15	CSNLLLQYGSFCTQL	11.5	0.33	
		HLA-DRB1*15:01	751	765	15	NLLLQYGSFCTQLNR	12.1	0.36	
		HLA-DRB1*15:01	748	762	15	ECSNLLLQYGSFCTQ	15.1	0.61	
		HLA-DRB1*15:01	752	766	15	LLLQYGSFCTQLNRA	20.6	1.1	
		HLA-DRB1*15:01	747	761	15	TECSNLLLQYGSFCT	27.4	1.6	
		HLA-DRB1*01:01	751	765	15	NLLLQYGSFCTQLNR	64.6	19	
		HLA-DRB1*15:01	753	767	15	LLQYGSFCTQLNRAL	93.9	6.7	
LSSTASALG	0.9155	HLA-DRB1*01:01	935	949	15	QDSLSSTASALGKLQ	97.1	25	100
LTGIAVEQD	0.8891	HLA-DQA1*05:01/DQB1*02:01	761	775	15	TQLNRALTGIAVEQD	92.1	2.1	100
		HLA-DQA1*01:02/DQB1*06:02	762	776	15	QLNRALTGIAVEQDK	98.8	2.8	
NRALTGIAV	0.5302	HLA-DRB1*01:01	760	774	15	CTQLNRALTGIAVEQ	9.4	2.4	100

(*continued on next page*)

Table 4 (*continued*)

Core Peptide	Antigen score	Allele	Start	End	Length	Peptide	Ic50	Rank	Conservancy%
		HLA-DRB1*01:01	761	775	15	TQLNRALTGIAVEQD	9.7	2.5	
		HLA-DRB1*01:01	759	773	15	FCTQLNRALTGIAVE	10.2	2.8	
		HLA-DRB1*01:01	762	776	15	QLNRALTGIAVEQDK	11.6	3.5	
		HLA-DRB1*01:01	758	772	15	SFCTQLNRALTGIAV	14	4.4	
		HLA-DRB1*01:01	763	777	15	LNRALTGIAVEQDKN	16.5	5.4	
		HLA-DRB1*01:01	764	778	15	NRALTGIAVEQDKNT	39.1	13	
NVLYENQKL	0.6017	HLA-DRB1*15:01	911	925	15	VTQNVLYENQKLIAN	86.6	6.3	100
		HLA-DRB1*15:01	912	926	15	TQNVLYENQKLIANQ	92.7	6.7	
		HLA-DRB1*15:01	910	924	15	GVTQNVLYENQKLIA	94.3	6.8	
PFAMQMAYR	1.3315	HLA-DRB4*01:01	894	908	15	LQIPFAMQMAYRFNG	22.2	0.73	100
		HLA-DRB4*01:01	893	907	15	ALQIPFAMQMAYRFN	24.1	0.81	
		HLA-DRB4*01:01	895	909	15	QIPFAMQMAYRFNGI	24.5	0.82	
		HLA-DRB4*01:01	892	906	15	AALQIPFAMQMAYRF	26.2	0.93	
		HLA-DRB4*01:01	891	905	15	GAALQIPFAMQMAYR	37.2	1.8	
		HLA-DRB4*01:01	896	910	15	IPFAMQMAYRFNGIG	46	2.3	
PFLGVYYHK	0.9406	HLA-DPA1*03:01/DPB1*04:02	134	148	15	QFCNDPFLGVYYHKN	65.2	2.6	100
		HLA-DPA1*03:01/DPB1*04:02	135	149	15	FCNDPFLGVYYHKNN	75	3.1	
		HLA-DPA1*03:01/DPB1*04:02	137	151	15	NDPFLGVYYHKNNKS	92.6	4.1	
		HLA-DPA1*03:01/DPB1*04:02	136	150	15	CNDPFLGVYYHKNNK	96.4	4.2	
		HLA-DPA1*02:01/DPB1*01:01	134	148	15	QFCNDPFLGVYYHKN	98.4	1.8	
PIGAGICAS	0.6621	HLA-DQA1*05:01/DQB1*03:01	663	677	15	DIPIGAGICASYQTQ	30.8	1.2	100
		HLA-DQA1*05:01/DQB1*03:01	662	676	15	CDIPIGAGICASYQT	31	1.3	
		HLA-DQA1*05:01/DQB1*03:01	664	678	15	IPIGAGICASYQTQT	33.8	1.5	
		HLA-DQA1*05:01/DQB1*03:01	661	675	15	ECDIPIGAGICASYQ	36.8	1.7	
		HLA-DQA1*05:01/DQB1*03:01	660	674	15	YECDIPIGAGICASY	42.4	2.2	
		HLA-DQA1*05:01/DQB1*03:01	665	679	15	PIGAGICASYQTQTN	46.5	2.5	
PLVDLPIGI	1.0525	HLA-DRB1*13:02	222	236	15	ALEPLVDLPIGINIT	49.1	5.9	100
		HLA-DRB1*13:02	223	237	15	LEPLVDLPIGINITR	61.9	7.1	
		HLA-DRB1*13:02	221	235	15	SALEPLVDLPIGINI	63	7.2	
		HLA-DRB1*13:02	220	234	15	FSALEPLVDLPIGIN	87.5	9.2	
		HLA-DRB1*13:02	219	233	15	GFSALEPLVDLPIGI	95.1	9.7	
PQTLEILDI	1.6404	HLA-DPA1*03:01/DPB1*04:02	576	590	15	VRDPQTLEILDITPC	67.5	2.7	100
		HLA-DPA1*03:01/DPB1*04:02	577	591	15	RDPQTLEILDITPCS	69	2.8	
		HLA-DPA1*03:01/DPB1*04:02	575	589	15	AVRDPQTLEILDITP	73.1	3	
RALTGIAVE	0.5744	HLA-DQA1*04:01/DQB1*04:02	761	775	15	TQLNRALTGIAVEQD	94.9	0.23	100
SALLAGTIT	0.6456	HLA-DRB1*01:01	873	887	15	YTSALLAGTITSGWT	98.1	25	100
SVLHSTQDL	0.5654	HLA-DRB1*01:01	42	56	15	VFRSSVLHSTQDLFL	69.1	20	100
		HLA-DRB1*01:01	43	57	15	FRSSVLHSTQDLFLP	94.1	25	
TAGAAAYYV	0.6397	HLA-DQA1*05:01/DQB1*03:01	255	269	15	SSGWTAGAAAYYVGY	25.3	0.86	100
		HLA-DRB1*01:01	256	270	15	SGWTAGAAAYYVGYL	36.2	12	
		HLA-DRB1*01:01	255	269	15	SSGWTAGAAAYYVGY	37.1	13	
		HLA-DQA1*01:02/DQB1*06:02	255	269	15	SSGWTAGAAAYYVGY	58.9	1.3	
		HLA-DRB1*01:01	257	271	15	GWTAGAAAYYVGYLQ	61.6	19	
		HLA-DQA1*01:02/DQB1*06:02	256	270	15	SGWTAGAAAYYVGYL	67.2	1.5	
		HLA-DQA1*01:02/DQB1*06:02	254	268	15	SSSGWTAGAAAYYVG	72.7	1.7	
		HLA-DQA1*01:02/DQB1*06:02	257	271	15	GWTAGAAAYYVGYLQ	88.7	2.4	
VGYLQPRTF	1.2441	HLA-DRB1*01:01	262	276	15	AAAYYVGYLQPRTFL	23.7	8	100
		HLA-DRB5*01:01	264	278	15	AYYVGYLQPRTFLLK	31.5	4.8	
		HLA-DRB1*09:01	263	277	15	AAYYVGYLQPRTFLL	40.3	3.3	
		HLA-DRB1*09:01	264	278	15	AYYVGYLQPRTFLLK	46.9	4.1	
		HLA-DRB1*09:01	262	276	15	AAAYYVGYLQPRTFL	48.8	4.4	
		HLA-DRB1*09:01	261	275	15	GAAAYYVGYLQPRTF	61.8	5.8	
		HLA-DRB1*09:01	265	279	15	YYVGYLQPRTFLLKY	62.9	6	
		HLA-DRB1*09:01	266	280	15	YVGYLQPRTFLLKYN	85.5	8.6	
VLSFELLHA	1.0776	HLA-DPA1*03:01/DPB1*04:02	508	522	15	YRVVVLSFELLHAPA	11.3	0.05	100
		HLA-DPA1*03:01/DPB1*04:02	509	523	15	RVVVLSFELLHAPAT	16.6	0.12	
		HLA-DPA1*03:01/DPB1*04:02	510	524	15	VVVLSFELLHAPATV	20.9	0.27	
		HLA-DPA1*03:01/DPB1*04:02	511	525	15	VVLSFELLHAPATVC	41.2	1.3	
		HLA-DRB1*15:01	507	521	15	PYRVVVLSFELLHAP	60.2	4.3	
		HLA-DPA1*01:03/DPB1*02:01	509	523	15	RVVVLSFELLHAPAT	68.3	2.8	
		HLA-DRB1*15:01	508	522	15	YRVVVLSFELLHAPA	68.8	5	
		HLA-DRB1*15:01	509	523	15	RVVVLSFELLHAPAT	77.6	5.7	
		HLA-DRB1*15:01	510	524	15	VVVLSFELLHAPATV	85.7	6.2	
VNQNAQALN	0.5164	HLA-DRB3*02:02	948	962	15	LQDVVNQNAQALNTL	30.2	1.6	100
		HLA-DRB3*02:02	949	963	15	QDVVNQNAQALNTLV	49.2	2.4	
		HLA-DRB3*02:02	950	964	15	DVVNQNAQALNTLVK	74.3	3.3	
VTLADAGFI	1.2653	HLA-DRB1*01:01	822	836	15	LFNKVTLADAGFIKQ	57.6	18	100
		HLA-DRB1*01:01	823	837	15	FNKVTLADAGFIKQY	65.9	20	
		HLA-DRB1*01:01	824	838	15	NKVTLADAGFIKQYG	89.6	24	

(*continued on next page*)

Table 4 (*continued*)

Core Peptide	Antigen score	Allele	Start	End	Length	Peptide	Ic50	Rank	Conservancy%
VTWFHAIHV	0.5426	HLA-DRB1*07:01	56	70	15	LPFFSNVTWFHAIHV	12.7	1.1	100
		HLA-DRB1*07:01	57	71	15	PFFSNVTWFHAIHVS	14.8	1.5	
		HLA-DRB1*07:01	58	72	15	FFSNVTWFHAIHVSG	18.1	1.9	
		HLA-DRB1*07:01	59	73	15	FSNVTWFHAIHVSGT	22.1	2.5	
		HLA-DRB1*07:01	62	76	15	VTWFHAIHVSGTNGT	38.4	4.6	
		HLA-DRB1*07:01	60	74	15	SNVTWFHAIHVSGTN	48.2	5.6	
		HLA-DRB1*07:01	61	75	15	NVTWFHAIHVSGTNG	55.7	6.5	
		HLA-DRB1*15:01	57	71	15	PFFSNVTWFHAIHVS	82.1	6.1	
		HLA-DRB1*01:01	59	73	15	FSNVTWFHAIHVSGT	83.6	23	
		HLA-DRB1*01:01	58	72	15	FFSNVTWFHAIHVSG	88.4	24	
		HLA-DRB1*15:01	58	72	15	FFSNVTWFHAIHVSG	97.8	7.1	
VVLSFELLH	1.409	HLA-DPA1*02:01/DPB1*01:01	506	520	15	QPYRVVVLSFELLHA	63.8	0.77	100
		HLA-DPA1*02:01/DPB1*01:01	507	521	15	PYRVVVLSFELLHAP	70.7	0.97	
		HLA-DPA1*02:01/DPB1*01:01	508	522	15	YRVVVLSFELLHAPA	70.8	0.97	
		HLA-DPA1*02:01/DPB1*01:01	509	523	15	RVVVLSFELLHAPAT	97.9	1.8	
VVVLSFELL	1.0909	HLA-DPA1*01:03/DPB1*02:01	507	521	15	PYRVVVLSFELLHAP	36.8	1.2	100
		HLA-DPA1*01:03/DPB1*02:01	506	520	15	QPYRVVVLSFELLHA	39.3	1.3	
		HLA-DPA1*01:03/DPB1*02:01	504	518	15	GYQPYRVVVLSFELL	41.8	1.5	
		HLA-DPA1*01:03/DPB1*02:01	508	522	15	YRVVVLSFELLHAPA	42.2	1.5	
		HLA-DPA1*01:03/DPB1*02:01	505	519	15	YQPYRVVVLSFELLH	46	1.6	
WTAGAAAYY	0.6306	HLA-DRB1*09:01	253	267	15	DSSSGWTAGAAAYYV	10.9	0.3	100
		HLA-DRB1*09:01	254	268	15	SSSGWTAGAAAYYVG	11	0.3	
		HLA-DRB1*09:01	255	269	15	SSGWTAGAAAYYVGY	11.3	0.3	
		HLA-DRB1*09:01	256	270	15	SGWTAGAAAYYVGYL	13.2	0.45	
		HLA-DRB1*09:01	252	266	15	GDSSSGWTAGAAAYY	14.7	0.6	
		HLA-DRB1*09:01	257	271	15	GWTAGAAAYYVGYLQ	20.6	1.1	
		HLA-DQA1*05:01/DQB1*03:01	254	268	15	SSSGWTAGAAAYYVG	28.1	1.1	
		HLA-DQA1*05:01/DQB1*03:01	253	267	15	DSSSGWTAGAAAYYV	35.4	1.6	
		HLA-DQA1*05:01/DQB1*03:01	252	266	15	GDSSSGWTAGAAAYY	48.2	2.6	
		HLA-DRB1*01:01	254	268	15	SSSGWTAGAAAYYVG	57.4	18	
		HLA-DRB1*09:01	258	272	15	WTAGAAAYYVGYLQP	68.7	6.6	
		HLA-DRB1*09:01	880	894	15	GTITSGWTFGAGAAL	35.9	2.7	
		HLA-DRB1*09:01	881	895	15	TITSGWTFGAGAALQ	37.2	2.9	
		HLA-DQA1*05:01/DQB1*03:01	881	895	15	TITSGWTFGAGAALQ	38.4	1.8	
		HLA-DRB1*01:01	881	895	15	TITSGWTFGAGAALQ	47.8	16	
		HLA-DRB1*01:01	880	894	15	GTITSGWTFGAGAAL	81.4	23	
YAWNRKRIS	0.8209	HLA-DRB1*11:01	346	360	15	RFASVYAWNRKRISN	12.6	1.2	100
		HLA-DRB1*11:01	345	359	15	TRFASVYAWNRKRIS	15	1.4	
		HLA-DRB1*11:01	348	362	15	ASVYAWNRKRISNCV	15.1	1.4	
		HLA-DRB1*11:01	347	361	15	FASVYAWNRKRISNC	15.1	1.4	
		HLA-DRB1*11:01	349	363	15	SVYAWNRKRISNCVA	15.5	1.4	
		HLA-DRB1*11:01	350	364	15	VYAWNRKRISNCVAD	52.1	5.1	
		HLA-DRB3*02:02	346	360	15	RFASVYAWNRKRISN	78.7	3.4	
		HLA-DRB3*02:02	347	361	15	FASVYAWNRKRISNC	94.5	3.9	
YGSFCTQLN	0.9142	HLA-DRB1*04:05	751	765	15	NLLLQYGSFCTQLNR	40.4	1.2	100
		HLA-DRB1*04:01	754	768	15	LQYGSFCTQLNRALT	43.6	1.3	
		HLA-DRB1*04:05	750	764	15	SNLLLQYGSFCTQLN	47.4	1.7	
		HLA-DRB1*04:01	753	767	15	LLQYGSFCTQLNRAL	49.3	1.6	
		HLA-DRB1*04:05	752	766	15	LLLQYGSFCTQLNRA	50.3	1.8	
		HLA-DRB1*04:01	755	769	15	QYGSFCTQLNRALTG	51.6	1.8	
		HLA-DRB1*04:05	753	767	15	LLQYGSFCTQLNRAL	58.3	2.3	
		HLA-DRB1*04:05	754	768	15	LQYGSFCTQLNRALT	65.9	2.8	
		HLA-DRB1*04:01	752	766	15	LLLQYGSFCTQLNRA	78.9	3.1	
YIKWPWYIW	0.9673	HLA-DPA1*01:03/DPB1*02:01	1203	1217	15	LGKYEQYIKWPWYIW	35.9	1.2	100
		HLA-DRB1*15:01	1205	1219	15	KYEQYIKWPWYIWLG	93.2	6.7	
YIWLGFIAG	0.7254	HLA-DPA1*01:03/DPB1*02:01	1210	1224	15	IKWPWYIWLGFIAGL	26.9	0.69	100
		HLA-DPA1*01:03/DPB1*02:01	1209	1223	15	YIKWPWYIWLGFIAG	28.1	0.73	
		HLA-DPA1*01:03/DPB1*02:01	1211	1225	15	KWPWYIWLGFIAGLI	33.8	1.1	
		HLA-DPA1*01:03/DPB1*02:01	1212	1226	15	WPWYIWLGFIAGLIA	39.2	1.3	
		HLA-DPA1*01:03/DPB1*02:01	1213	1227	15	PWYIWLGFIAGLIAI	44.4	1.6	
YNENGTITD	0.7650	HLA-DRB3*02:02	276	290	15	LLKYNENGTITDAVD	54.7	2.6	100
YQPYRVVVL	0.5964	HLA-DRB1*01:01	501	515	15	NGVGYQPYRVVVLSF	50.5	16	100
		HLA-DRB1*01:01	502	516	15	GVGYQPYRVVVLSFE	78.6	22	
YTSALLAGT	0.5487	HLA-DRB1*01:01	870	884	15	IAQYTSALLAGTITS	24.7	8.4	100
		HLA-DRB1*01:01	869	883	15	MIAQYTSALLAGTIT	25.1	8.6	
		HLA-DRB1*01:01	871	885	15	AQYTSALLAGTITSG	29.4	9.9	
		HLA-DQA1*05:01/DQB1*03:01	870	884	15	IAQYTSALLAGTITS	38.3	1.8	
		HLA-DQA1*05:01/DQB1*03:01	869	883	15	MIAQYTSALLAGTIT	43.2	2.3	
		HLA-DRB1*01:01	872	886	15	QYTSALLAGTITSGW	44.7	15	

(*continued on next page*)

Table 4 (*continued*)

Core Peptide	Antigen score	Allele	Start	End	Length	Peptide	Ic50	Rank	Conservancy%
		HLA-DQA1*05:01/DQB1*03:01	868	882	15	EMIAQYTSALLAGTI	64	3.9	
		HLA-DRB1*09:01	868	882	15	EMIAQYTSALLAGTI	99.1	9.8	
YVGYLQPRT	1.3483	HLA-DRB1*15:01	262	276	15	AAAYYVGYLQPRTFL	22.8	1.3	100
		HLA-DRB1*15:01	263	277	15	AAYYVGYLQPRTFLL	27.1	1.5	
		HLA-DRB1*15:01	264	278	15	AYYVGYLQPRTFLLK	30	1.8	
		HLA-DRB1*15:01	265	279	15	YYVGYLQPRTFLLKY	30.8	1.9	
		HLA-DRB1*04:05	263	277	15	AAYYVGYLQPRTFLL	56.3	2.2	
		HLA-DRB1*04:05	264	278	15	AYYVGYLQPRTFLLK	59.3	2.3	
		HLA-DRB1*04:05	262	276	15	AAAYYVGYLQPRTFL	78.2	3.6	
		HLA-DRB1*01:01	261	275	15	GAAAYYVGYLQPRTF	81.6	23	
		HLA-DRB1*04:05	261	275	15	GAAAYYVGYLQPRTF	95.6	4.9	
		HLA-DRB5*01:01	263	277	15	AAYYVGYLQPRTFLL	40	6	
		HLA-DRB5*01:01	262	276	15	AAAYYVGYLQPRTFL	63.2	8.9	

3.5. Population coverage and epitope conservancy analysis

A population coverage analysis was performed to evaluate how well MHC Class-I and MHC-II allele interaction epitopes are covered globally. The established candidate epitopes for each coverage methodology have been classified using the IEDB population coverage analysis tool. The MHC HLA allele distribution varies across a diversity of environmental positions. Therefore, the design of a possible effective vaccination must incorporate population coverage. The (80.79 %) world population coverage of the MHC Class-II allele was computed. The highest population exposure for the MHC Class-II allele was computed in Europe (85.67 %), which was followed by Northeast Asia, North Africa and East Africa, Central Africa, West Africa, South America, Central America, Southeast Asia, Oceania, Southwest Asia, East Asia and West Africa with population coverage of 58.76 %, 74.75 %, 68.30 %, 62.54 %, 64.01 %, 43.75 %,28.74 %, 56.57 %, 58.81 %, 43.77 %, 78.82 %, 64.01 % respectively (Fig. 6). However, Central America has the lowest population coverage (28.74 %) in MHC Class-I. The maximum world population coverage for the MHC Class-I allele was calculated to be (98.25 %). At the same time, the highest population coverage for MHC Class-I allele was calculated in Europe (99.56 %) which was followed by East Asia, North Africa, Oceania, West Africa, Northeast Asia, South Asia, Southwest Asia, East Africa, Central Africa, Southeast Asia with population coverage of 97.63 %, 95.31 %, 94.60, 94.93 %, 93.99 %, 93.77 %, 90.93 %, 89.94 %, 85.22 %, 90.1 % respectively (Fig. 5).

The feature of this calculation is the population coverage outcomes of the study for high binders to MHC Class-I and MHC Class-II alleles in mutual method, which display a significant coverage of 97.63 % and 78.82 %, respectively (Figs. 5 and 6). The IEDB conservancy analysis tool was used to determine identified epitope conservancies (Tables 3 and 4).

3.6. Antigenicity and allergnicity evaluation of vaccine protein

Using the VaxiJen 2.0 web server, the antigenicity of the Construct vaccine was identified to be 0.7809. The results of AllerTOP v2 show that the vaccine is not allergic, whether the adjuvant is connected or not. Toxinpred tested the protein's toxicity with or without an adjuvant, and it was shown to be non-toxic.

3.7. Physiochemical and solubility analysis of construct vaccine

ExPASY ProtParam was used to determine physicochemical properties, and the results presented several properties key to the nature of the protein. The multi-epitope element has a molecular weight (MW) of 160.1596 kDa. Protein pI was 9.39, according to calculations. Protein is essential according to this value. There were 28.29 in the instability index (II). Based on calculations of the protein's instability index, it was determined that the protein was stable because a value larger than 40 showed that the protein was thermostable. The aliphatic index values, determined to be 96.61, indicated that it was a thermostable protein [50], and GRAVY index was 0.283. According to the positive value, the protein is not hydrophilic [51]. With a score of 0.106, our vaccine design indicated a greater solubility rate as determined by the SOLpro server.

3.8. Secondary structure analysis

The online SOPMA tool (https://npsa-prabi.ibcp.fr) for identification of the secondary structure of the protein. The results determined the protein structure to have 28.11 % coils, 31.87%alpha-helix and 10.09 % beta strands (Fig. 7). Secondary structure of construct vaccine by PISPRED (Fig. 8)

3.9. Tertiary structure analysis

The tertiary structure of the chimeric vaccination contract was predicated by the online RCSB PDB database (https://www.rcsb.org/). The models based on high coverage values of the top ten threading templates were predicated. The present study's model with a high coverage score was chosen for method refinement (Fig. 9) [52].

Table 5
The sequence of epitopes merged to form final vaccine construct.

Vaccine Construct	
Adjevent	50S ribosomal protein L7/L12
Linker	**EAAAK**
B-cell Epitops	GKQGNF
	LTPGDSSSGWTAG
	VRQIAPGQTGKIADVRQIAPGQTGKIAD
	VITPGTNTSN
	ILPDPSKPSKRS
	KNHTSPDVDLG
Linker	**GPGPG**
MHC-I Epitope	AALQIPFAM
	AEIRASANL
	ALNTLVKQL
	ALQIPFAMQM
	AVDCALDPL
	AYYVGYLQPR
	DLPIGINITR
	EQYIKWPWYI
	FAMQMAYRF
	FELLHAPATV
	FEYVSQPFL
	FFSNVTWFH
	FKIYSKHTPI
	FKIYSKHTPI
	FPNITNLCP
	FPREGVFVS
	FTISVTTEI
	FTISVTTEIL
	FTISVTTEIL

	FVFLVLLPL
	FVFLVLLPLV
	GINASFVNI
	GLIAIVMVTI
	GVLTESNKK
	GVVFLHVTYV
	GVYFASTEK
	GWTAGAAAYY
	GYLQPRTFLL
	HWFVTQRNFY
	IAIPTNFTI
	IGAGICASY
	INASFVNIQK
	IPTNFTISV
	ITPCSFGGV
	KEIDRLNEV
	KIYSKHTPI
	KLNDLCFTNV
	KNLNESLIDL
	KSNLKPFER
	LLFNKVTLA
	LPFNDGVYF
	LPIGINITRF
	LQSYGFQPT
	LQYGSFCTQL
	LSFELLHAPA
	NFTISVTTEI
	NKSWMESEFR
	NTQEVFAQVK
	NVTWFHAIHV

	PFAMQMAYR
	PFAMQMAYRF
	PIGAGICASY
	PYRVVVLSF
	QIPFAMQMAY
	QTRAGCLIGA
	RASANLAATK
	RKSNLKPFER
	RQIAPGQTGK
	SLIDLQELGK
	SPRRARSVA
	SPRRARSVAS
	TAGAAAYYV
	TFEYVSQPF
	TLADAGFIK
	TSVDCTMYI
	TTRTQLPPA
	TYVPAQEKNF
	VLSFELLHA
	VTLADAGFIK
	VTWFHAIHV
	VVFLHVTYV
	WIFGTTLDSK
	WTAGAAAY
	WTAGAAAYYV
	YGSFCTQLNR
	YIWLGFIAGL
	YQPYRVVVL
	YTSALLAGT
	YYVGYLQPR

Linker	AAY
MHC-II Epitope	AALQIPFAM
	AGAAAYYVG
	CLGDIAARD
	CVLGQSKRV
	DFGGFNFSQ
	DQLTPTWRV
	DSKTQSLLI
	FAMQMAYRF
	FCTQLNRAL
	FEYVSQPFL
	FFSNVTWFH
	FGAISSVLN
	FGGFNFSQI
	FKIYSKHTP
	FLGVYYHKN
	FNATRFASV
	FNFSQILPD
	FRKSNLKPF
	FSNVTWFHA
	FTISVTTEI
	GINITRFQT
	GVGYQPYRV
	IAPGQTGKI
	IAWNSNNLD
	IAYTMSLGA
	IGINITRFQ
	IHVSGTNGT
	IIAYTMSLG
	INASFVNIQ

Linker	AAY
MHC-II Epitope	AALQIPFAM
	AGAAAYYVG
	CLGDIAARD
	CVLGQSKRV
	DFGGFNFSQ
	DQLTPTWRV
	DSKTQSLLI
	FAMQMAYRF
	FCTQLNRAL
	FEYVSQPFL
	FFSNVTWFH
	FGAISSVLN
	FGGFNFSQI
	FKIYSKHTP
	FLGVYYHKN
	FNATRFASV
	FNFSQILPD
	FRKSNLKPF
	FSNVTWFHA
	FTISVTTEI
	GINITRFQT
	GVGYQPYRV
	IAPGQTGKI
	IAWNSNNLD
	IAYTMSLGA
	IGINITRFQ
	IHVSGTNGT
	IIAYTMSLG
	INASFVNIQ
	YNENGTITD
	YQPYRVVVL
	YTSALLAGT
	YVGYLQPRT
6x His tag	HHHHHH

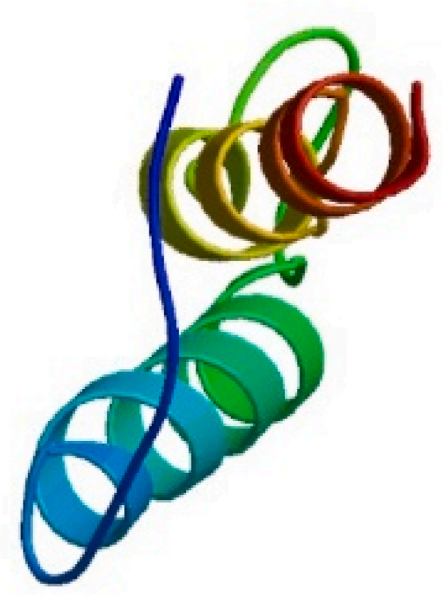


Fig. 4. Shows 3D structure of construct vaccine.

3.10. Refinement of tertiary structure

The GalaxyRefine tool refined the vaccine chimera, generating five models. A few variables were considered all over the refining process, with GDT-HA (0.8885), RMSD (0.568), and MolProbity (1.573). The computed crash score was 7.4, the weak rotamer score was 0, the refinement energy was determined to be −4851.05, and the Ramachandran score was predicted to be 97.1 %. Since Model 3 was the most authentic, it was selected for further analysis.

3.11. Validation of 3D structure

The RAMPAGE server verified the modified tertiary structure. After analysis, a Ramachandran plot for the protein structure was created (Fig. 10). RAMPAGE achieved the best results and, later, the refinement. These findings show that 87.6 % of residues were in the preferred zone, 11.3 % were in the additional authorized region, and 0.6 % was in a generously specified region.

3.12. Protein-protein docking

Using the online protein-protein docking service Cluspro2.0, molecular docking was done to calculate the interaction between the modified vaccination model and the immune receptor TLR3's ligand-binding domain [53]. Using docking, we may analyze numerous models. Model number 8 was found to be the best-docked model after a review of all 30 docked stances. Following docking, the 0 Cluster was selected because it had the lowest energy level (−956.2), as seen in (Fig. 11).

3.13. Molecular dynamics simulations

IModS performed a critical structure study by adjusting the force field at several time intervals. Less distortion is seen in the final model at each level of residue capacity. The eigon value is 6.218649e-05. Better relationships between the different residues were demonstrated by heat maps with low RMSD and highly linked regions (Fig. 12). The findings of the IModS molecular dynamics simulation are shown in detail in the image. Fig. 12a displays the MNA mobility in the protein structure, and Fig. 12b represents deformability, which exhibits minor deformation at all residues. The B-factor is indicated in Fig. 12c. The eigon values in Fig. 12d are indicated as 6.218649e-05, and the difference described in Fig. 12e is shown in green and red. The elastic network and co-variance of the complex are also shown in Fig. 12(f) and (g).

3.14. Codon optimization of designed vaccine peptide for expression analysis

Codon optimization was done using the Java Codon Adaptation Tool (JCat) to generate the highest quality protein expression. The optimized sequence has a codon adaptation index (CAI) of 0.9698, a GC content of 51.82 % and an optimum codon length of 4300 nucleotides. These numbers show stable vector expression in *Escherichia coli* (*E. coli*) because the ideal range for GC content is between 30 and 70 % (Fig. 13a). The optimized sequence was amplified by *in-silico* PCR using SnapGene software, and the recombinant plasmid

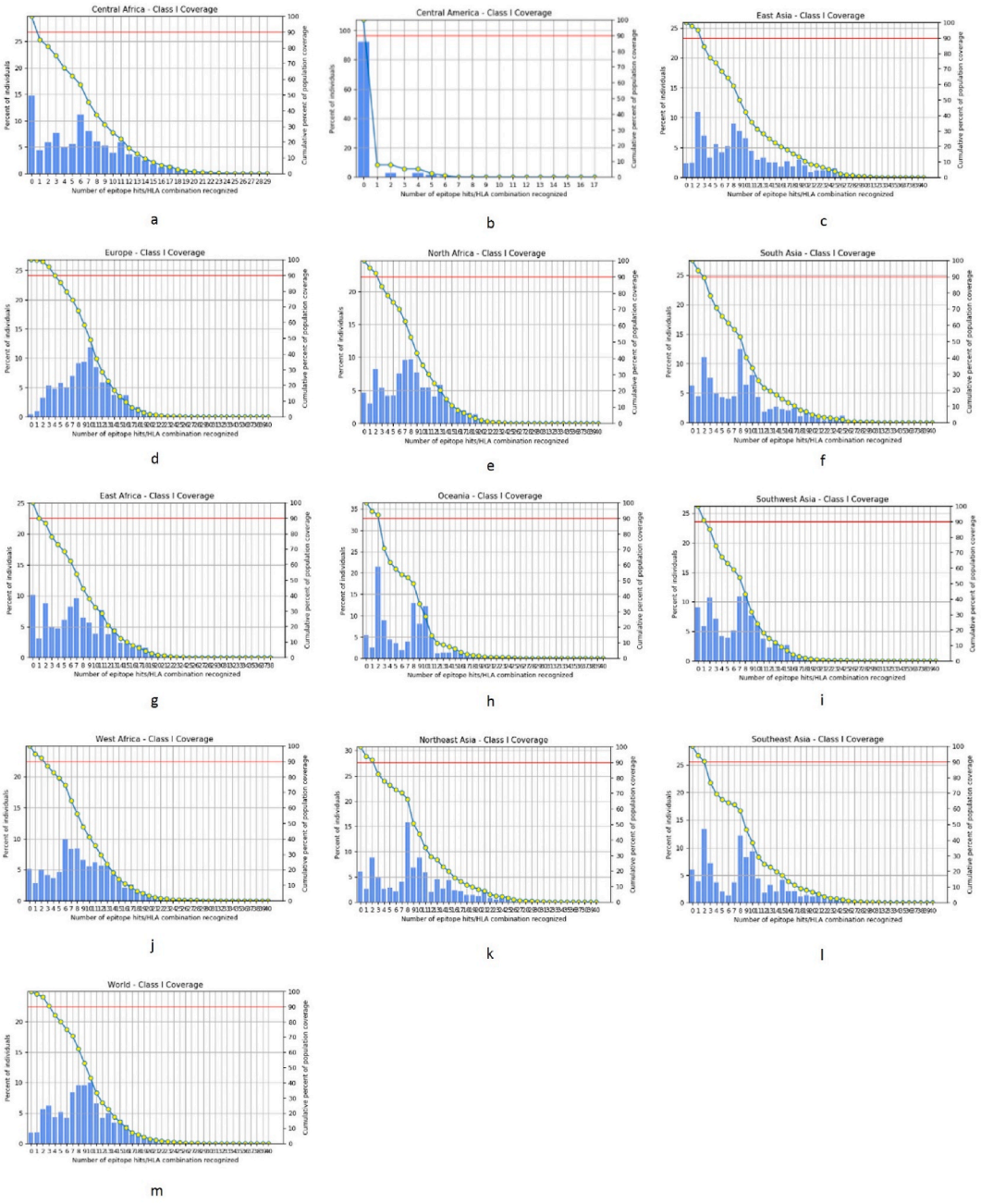
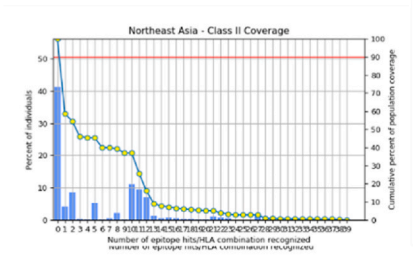
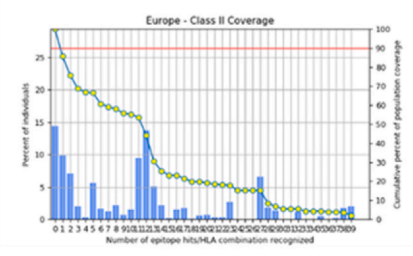


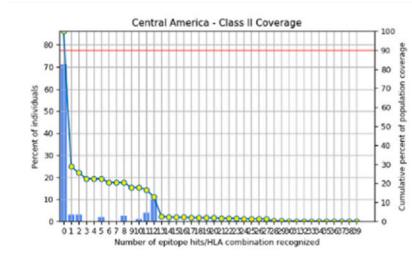
Fig. 5. Based on MHC-I restriction data, population coverage was designed. The whole world, as well as all continents, have been chosen to evaluate the population coverage of specified epitopes. (a) Show 85.22 % population coverage of specified epitopes in Central Africa. (b) Show 28.74 % population coverage of specified epitopes in Central America. (c) Show 97.63 % population coverage of specified epitopes in East Asia. (d) Show 99.56 % population coverage of specified epitopes in Europe. (e) Show 95.31 % population coverage of specified epitopes in North Africa. (f) Show 93.77 % population coverage of specified epitopes in South Asia. (g) Show 89.94 % population coverage of specified epitopes in East Africa. (h) Show 94.60 % population coverage of specified epitopes in Oceania. (i) Show 90.93 % population coverage of specified epitopes in Southwest Asia. (j) Show 94.93 % population coverage of specified epitopes in West Africa. (k) Show 93.99 % population coverage of specified epitopes in Northeast Asia. (l) Show 90.1 % population coverage of specified epitopes in Southeast Asia. (m) Show 98.25 % population coverage of specified epitopes in World.



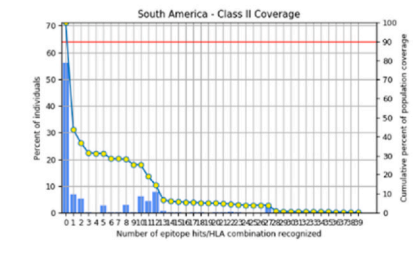
a



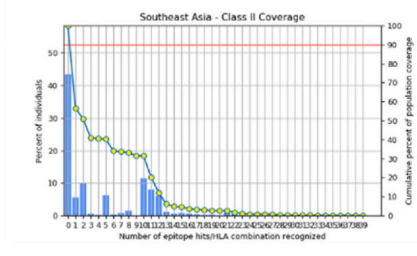
b



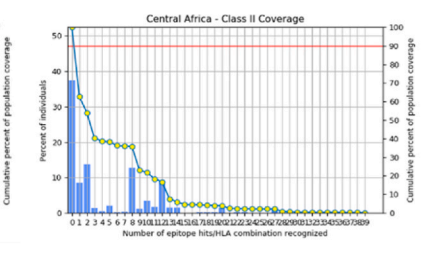
c



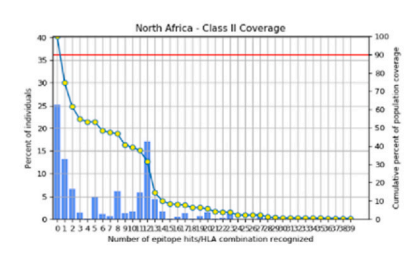
d



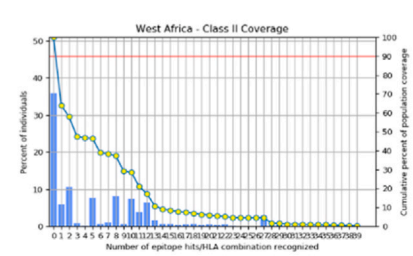
e



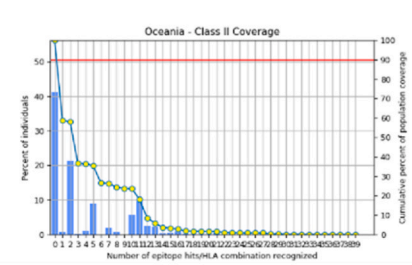
f



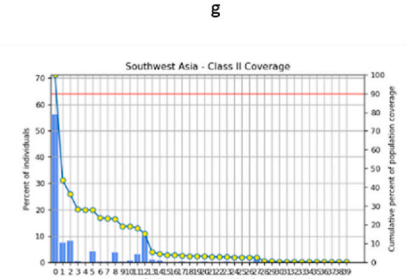
g



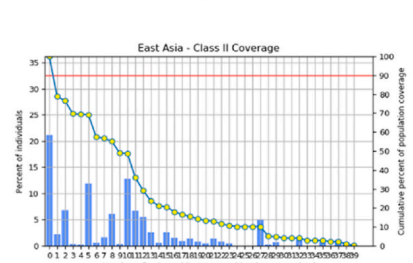
h



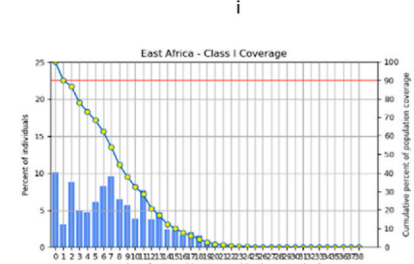
i



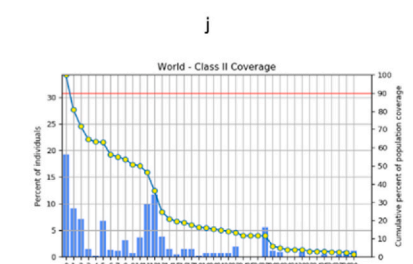
j



k



l



m

(caption on next page)

Fig. 6. Based on MHC-II restriction data, population coverage was designed. The whole world, as well as all continents, have been chosen to evaluate the population coverage of specified epitopes. (a) Show 58.76 % population coverage of specified epitopes in Northeast Asia. (b) Show 85.67 % population coverage of specified epitopes in Europe. (c) Show 28.74 % population coverage of specified epitopes Central America. (d) Show 43.75 % population coverage of specified epitopes in South America. (e) Show 56.57 % population coverage of specified epitopes in Southeast Asia. (f) Show 62.54 % population coverage of specified epitopes in Central Africa. (g) Show 74.75 % population coverage of specified epitopes in North Africa. (h) Show 64.01 % population coverage of specified epitopes in West Africa. (i) Show 58.81 % population coverage of specified epitopes in Oceania. (j) Show 43.77 % population coverage of specified epitopes in Southwest Asia. (k) Show 78.82 % population coverage of specified epitopes in East Asia. (l) Show 68.30 % population coverage of specified epitopes in East Africa. (m) Show 80.79 % population coverage of specified epitopes in world.

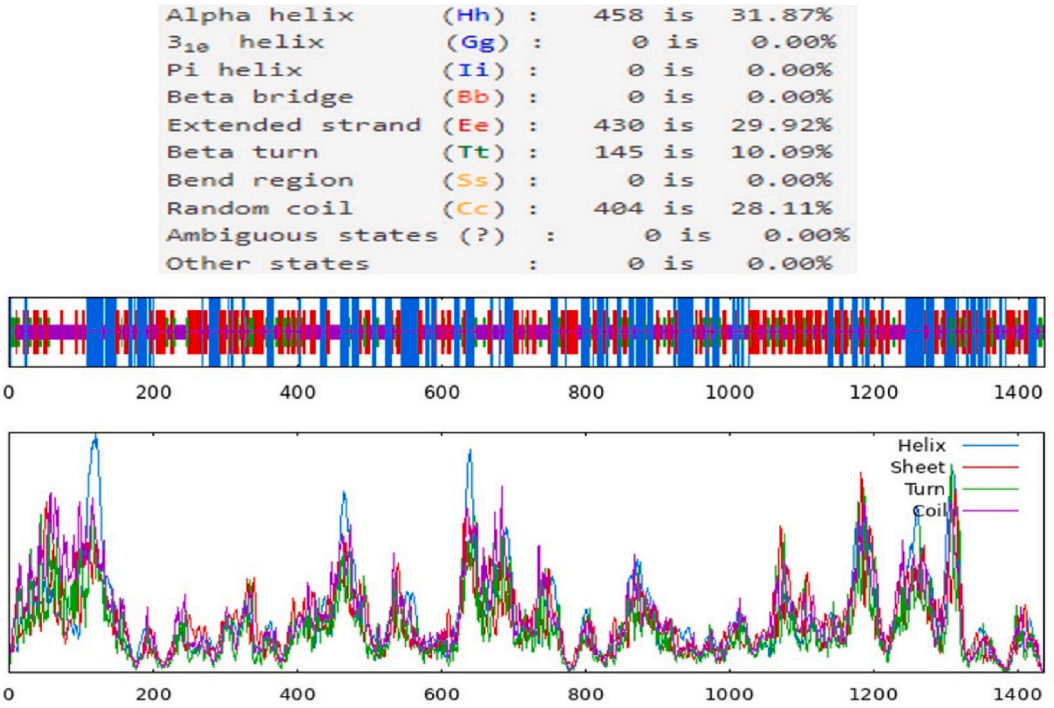


Fig. 7. Secondary structure of construct vaccine predicted by SOPMA tool.

was developed by cloning in the Pet28 (+) vector (Fig. 13b-c-d).

4. Discussion

The urgent necessity of the hour is to move forward in the search for a treatment for this unique disease, given the rise in worldwide disasters of this fatal virus. With the advancement of computationally aided sequence-based technology, bioinformatics tools provide a vital approach to peptide-based vaccine designing. Other viruses with peptide-based vaccines, including dengue, rhino, and SLE viruses, have previously shown a potential channel of viral targeting [54]. The high level of mutation in the genome of the RNA-based SARS-CoV-2 causes multiple resistance. This is because the primary focus of this research is the membrane protein or outer protein (S Protein) that covers the virus [55]. However, the targeted protein's physicochemical and secondary structure research suggests it is a highly antigenic helical candidate for vaccine construction. *In-silico* methods are used in this study to construct a multi-epitope.

B-cells were used as a basis in the past to develop a latent vaccine [56]. The antigen may become removed from the antibody memory response over time due to antigenic drift. While T-cell immunity produces a long-lasting immune response [57], developing an epitope into a vaccine is focused on a number of secure rations. We first form a database of all potential S protein epitopes. Five strategies have been selected to predict the antigenicity of B-cell epitopes in the IEDB database. The accessibility, polarity of the exposed surface, flexibility, hydrophobicity, and susceptibility to the antigenicity of polypeptide chains are all linked to the epitope's site. The Karplus and Schulz, Emini surface, Kolaskar and Tongaonkar, Chou and Fasman beta-turn, and Bepipred linear epitope calculations indicate through visuals each residue's capacity to contribute to epitope era [58]. However, IEDB is equally useful for evaluating T-cell binding and processing predictions [59]. Here, we identified that T cell epitopes were active against their targeted allele and had an IC50 value of less than 200.

The consensus technique has been used to separate the epitopes interacting with more than five MHC Class-II molecules for further screening. The fight against epitopes increases after being proven to be the best among a few criteria. The capacity of T and B cells to

Fig. 8. secondary structure of constructed vaccine by PISPRED tool.

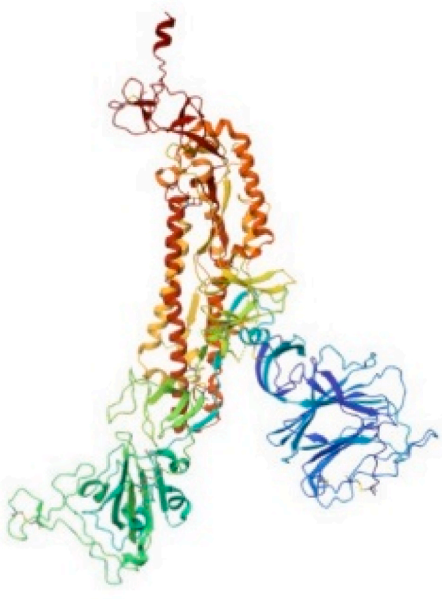


Fig. 9. 3D Structure of assemble vaccine protein.

detect antigenic features is essential. The antigenic epitope must stay nontoxic in its ability to produce an immune response properly. Toxinpred is based on a support vector machine (SVM) used to estimate peptide toxicity. Allergenicity is the great difficulty in constructing vaccinations [60]. Many vaccinations used nowadays include side effects to boost the immune system. AllerTop v.2.0 is a tool that helps calculate the score value for the chance of an allergy. This score value is calculated using Quantitative Structural Activity

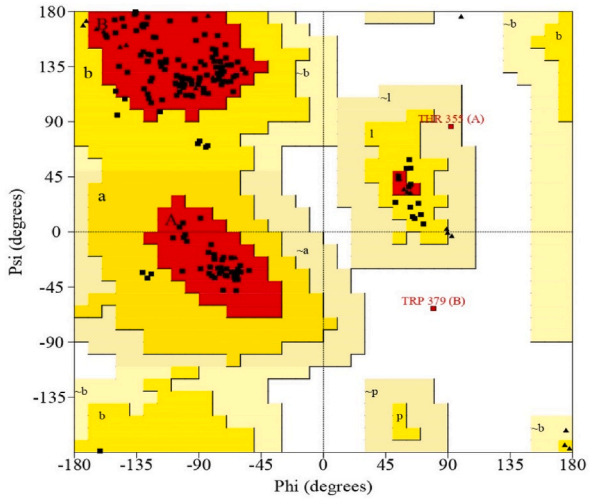


Fig. 10. Validation of refine structure shows 87.6 % residues in favored regions 11.3 % residues in allowed region and 0.6 % residues in outer region.

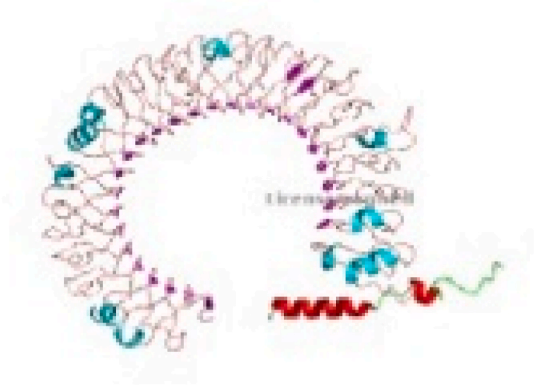


Fig. 11. Docking complex of TLR-3 and the design vaccine.

Relationship (QSAR). The approach was developed with auto cross-covariance (ACC), which indicated epitopes screened out as non-allergens with no allergenicity. According to FAO/WHO, any peptide comprising at least six contiguous amino acids can be considered allergic [30]. Because our selected epitopes were unable to meet the standards set for FAO/WHO evaluation method of allergenicity prediction, they were categorized as non-allergens by AllerTop v.2.0. The predicted antigenic epitopes that are free of toxicity and allergens result in immunoreactive peptides for our further research [61].

The IEDB's screened-out epitope exhibits good protein sequence conservation in a portion of the sequences, and this conservation determines the step of identification between strains. The IEDB server is used for population coverage analysis [40]. Many T-cell peptides with varying HLA bindings were selected and studied for this goal. Both MHC classes show exemplary conservation globally. As a result, a more significant portion of the peptide-based vaccine-vulnerable CoV patient group will be covered. After satisfying all criteria, six B-cell epitopes, 79 MHC Class-I T-cell epitopes, and 63 MHC Class-II T-cell epitopes were selected as subunits for the vaccine-building process.

It was chosen as an adjuvant to boost the immunoreactive feature after a study on 50S ribosomal protein L7/L12 showed the importance of increasing the vaccination reaction in viral identification and immune system stimulation [53]. The multi-epitope subunit vaccine design was first developed by combining selected B and T-cell epitopes with appropriate linkers. Spacer sequences are regarded as essential to vaccine development processes due to their optimal effect. GPGPG and AAY linkers were inserted among identified epitopes from previous studies to construct a possible vaccination with the maximum antigenicity. An EAAAK linker was added to the sequence design to link the adjuvant with the first predicted B cell epitope. This linker's entanglement was exploited to produce functional peptides enhancing the attached protein. The 6× His tag, also known as a polyhistidine tag, is a motif formed up of at minimum 6 histidine residues attached to the sequence's carboxyl (C-) terminus. The sequence is made simple to function in a buffer environment by binding histidine residues to immobilized ions [9].

The immunological and bioinformatics analysis showed that the developed protein sequence is free of allergenic and harmful

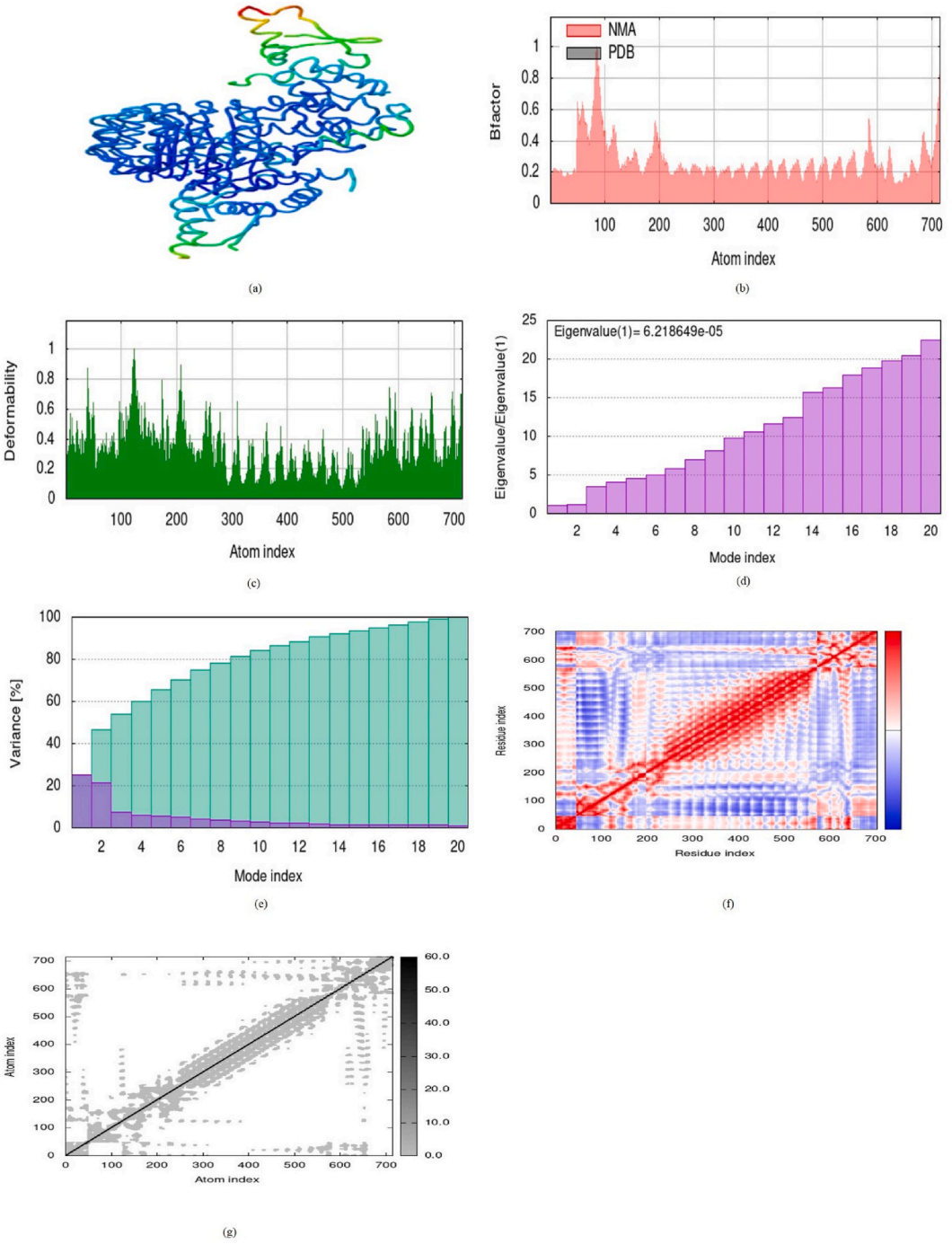
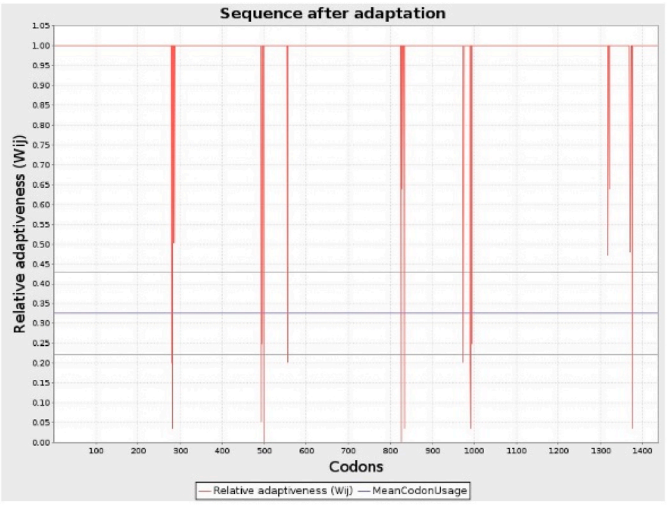


Fig. 12. The results of the IModS molecular dynamics simulation are shown in detail in the (a) represents the MNA mobility in the protein structure as it is, and figure (b) displays deformability, which exhibits limited deformation at all residues. The B-factor is indicated in (c). The eigon values in (d) are indicated as 6.218649e-05, and the difference given in (e) is shown in both red and green color. The elastic network and co-variance of the complex are also shown in the other (f) and (g) figures.

features. A small number of studies reveal the poor antigenicity of vaccine constructs. However, this artificial vaccination chimera displayed an effective antigenic score, whether or not it was linked to an adjuvant. The molecular weight of the constructed protein was 160.1596 kDa, and its solubility was studied considering its improved antigenicity. The pI of the vaccine is 9.39, showing the basic properties of the vaccine protein. With a low instability score of 28.29, the vaccine's recommended vaccine protein is stable and can be considered to function as a vaccine model when expressed. Secondary and tertiary structures are essential for designing a vaccine.

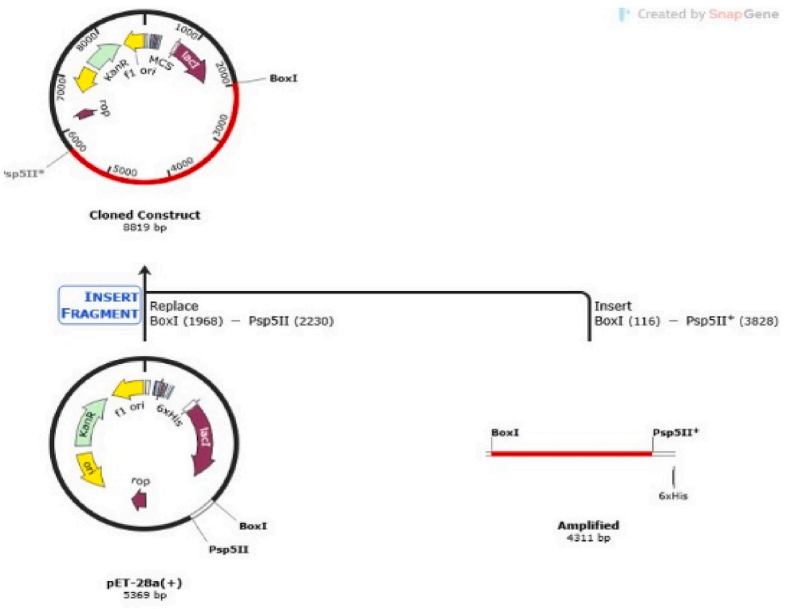


a

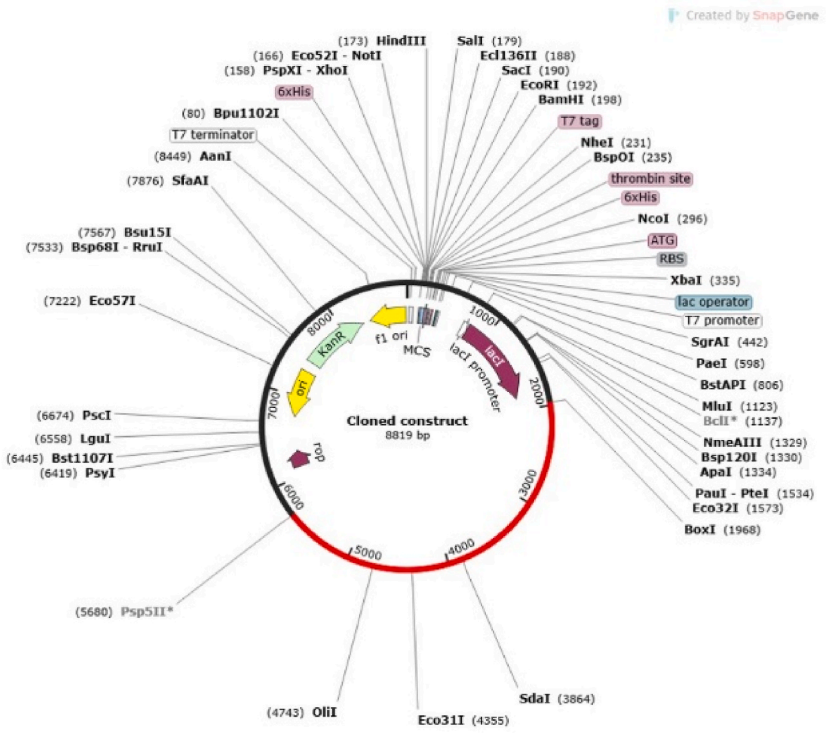
GAAGCTGCTGCTAAAGGTAAACAGGGTAACTTCCTGACCCCGGGTGACTCTTCTTCTGGTTGGACCGCTGGTGTTCGTCAGATCGCTCCGGGTCAGACC
G GTAAAATCGCTGACGTTCGTCAGATCGCTCCGGGTCAGACCGGTAAAATCGCTGACGTTATCACCCCGGGTACCAACACCTCTAACATCCTGCCGGA
CCCGTCTAAACCGTCTAAACGTTCTATCCTGCCGGACCCGTCTAAACCGTCTAAACGTTCTAAAAACCACACCTCTCCGGACGTTGACCTGGGTGGTCC
GGGTCCGGGTGCTGCTCTGCAGATCCCGTTCGCTATGGCTGAAATCCGTGCTTC TGCTAACCTGGCTCTGAACACCCTGGTTAAACAGCTGGCTCTGCA
GATCCCGTTCGCTATGCAGATGGCTGTTGACTGCGCTCTGGACCCGCTGGCTTACTACGTTGGTTACCTGCAGCCGCGTGACCTGCCGATCGGTATCAA
CATCACCCGTGAACAGTACATCAAATGGCCGTGGTACATCTTCGCTATGCAGATGGCTTACCGTTTCTTCGAACTGCTGCACGCTCCGGCTACCGTTTTC
GAATACGTTTCTCAGCCGTTCCTGTTCTTCTCTAACGTTACCTGGTTCCACTTCAAAATCTACTCTAAACACACCCCGATCTTCAAAATCTACTCTAAAC
ACACCCCGATCTTCCCGAACATCACCAACCTGTGCCCGTTCCCGCGTGAAGGTGTTTTCGTTTCTTTCACCATCTCTGTTACCACCGAAATCTTCACCAT
CTCTGTTACCACCGAAATCCTGTTCACCATCTCTGTTACCACC**GAG**ATC**CTC**TTC**GTA**TTCCTG**GTA**CTGCTGCCGCTGTTCGTTTTCCTGGTTCTGCTG
CCGCTGGTTGGTATCAACGCTTCTTTCGTTAACATCGGTCTGATCGCTATCGTTATGGTTACCATCGGTTCTTTCTGCACCCAGCTGAACCGTGGTGTTCT
GACCGAATCTAACAAAAAAGGTGTTGTTTTCCTGCACGTTACCTACGTTGGTGTTTACTTCGCTTCTACCGAAAAAGGTTGGACCGCTGGTGCTGCTGCT
TACTACGGTTACCTGCAGCCGCGTACCTTCCTGCTGCACTGGTTCGTTACCCAGCGTAACTTCTACATCGCTATCCCGACCAACTTCACCATCATCGGTG
CTGGTATCTGCGCTTCTTACATCAACGCTTCTTTCGTTAACATCCAGAAAATCCCGACCAACTTCACCATCTCTGTTATCACCCCGTGCTCTTTCGGTGGT
GTTAAAGAAATCGACCGTCTGAACGAAGTTAAAATCTACTCTAAACACACCCCGATCAAACTGAACGACCTGTGCTTCACCAACGTTAAAAACCTGAA
CGAATCTCTGATCGACCTGAAATCTAACCTGAAACCGTTCGAACGTCTGCTGTTCAACAAAGTTACCCTGGCTCTGCCGTTC**AATGAT**GGT**GTG**TACTT
C**CTA**CCG ATCGGTATCAACATCACCCGTTTCCTGCAGTCTTACGGTTTCCAGCCGACCCTGCAGTACGGTTCTTTCTGCACCCAGCTGCTGTCTTTCGA
ACTGCTGCACGCTCCGGCTAACTTCACCATCTCTGTTACCACCGAAATCAACAAATCTTGGATGGAATCTGAA**TTT**CGTAACACCCAGGAAGTTTTCGC
TCAGGTTAAAAACGTTACCTGGTTCCACGCTATCCACGTTCCGTTCGCTATGCAGATGGCTTACCGTCCGTTCGCTATGCAGATGGCTTACCGTTTCCCG
ATCGGTGCTGGTATCTGCGCTTCTTACCCGTACCGTGTTGTTGTTCTGTCTTTCCAGATCCCGTTCGCTATGCAGATGGCTTACCAGACCCGTGCTGGTTG
CCTGATCGGTGCTCGTGCTTCTGCTAACCTGGCTGCTACCAAACGTAAATCTAACCTGAAACCGTTCGAACGTCGTCAGATCGCTCCGGGTCAGACCGG
TAAATCTCTGATCGACCTGCAGGAACTGGGTAAATCTCCGCGTCGTGCTCGTTCTGTTGCTTCTCCGCGTCGTGCTCGTTCTGTTGCTTCTACCGCTGGT
GCTGCTGCTTACTACGTTACCTTCGAATACGTTTCTCAGCCGTTCACCCTGGCTGACGCTGGTTTCATCAAAACCTCTGTTGACTGCACCATGTACATCA
CCACCCGTACCCAGCTGCCGCCGGCTACCTACGTTCCGGCTCAGGAAAAAAACTTCGTTCTGTCTTTCGAACTGCTGCACGCTGTTACCCTGGCTGACG
CTGGTTTCAT CAAAGTTACCTGGTTCCACGCTATCCACGTTGTTGTTTTCCTGCACGTTACCTACGTTTGGATCTTCGGTACCACCCTGGACTCTAAATG
GACCGCTGGTGCTGCTGCTTACTGGACCGCTGGTGCTGCTGCTTACTACGTTTACGGTTCTTTCTGCACCCAGCTGAACCGTTAC**ATA**TGGCTG**GGC**TTC
ATCGCTGGT**CTC**TACCAGCCGTACCGTGTTGTTGTTCTGTACACCTCTGCTCTGCTGGCTGGTACCTACTACGTTGGTTACCTGCAGCCGCGTGCTGCTT
ACGCTGCTCTGCAGATCCCGTTCGCTATGGCTGGTGCTGCTGCTTACTACGTTGGTTGCCTGGGTGACATCGCTGCTCGTGACTGCGTTCTGGGTCAGTC
TAAACGTGTTGACTTCGGTGGTTTCAACTTCTCTCAGGACCAGCTGACCCCGACCTGGCGTGTTGACTCTAAAACCCAGTCTCTGCTGATCTTCGCTATG
CAGATGGCTTACCGTTTCTTCTGCACCCAGCTGAACCGTGCTCTGTTCGAATACGTTTCTCAGCCGTTCCTGTTCTTCTCTAACGTTACCTGGTTCCACTT
CGGTGCTATCTCTTCTGTTCTGAAC**TTTGGCGGC**TTCAACTTCTCTCAGATCTTCAAAATCTACTCTAAACACACCCCG**TTTCTC**GGT**GTG**TACTACCA
CAAAAACTTCAACGCTACCCGTTTCGCTTCTGTTTTCAACTTCTCTCAGATCCTGCCGGACTTCCGTAAATCTAACCTGAAACCGTTCTTCTCTAACGTTA
CCTGGTTCCACGCTTTCACCATCTCTGTTACCACCGAAATCGGTATCAACATCACCCGTTTCCAGACCGGTGTTGGTTACCAGCCGTACCGTGTTATCGC
TCCGGGTCAGACCGGTAAAATCATCGCTTGGAACTCTAACAACCTGGACATCGCTTACACCATGTCTCTGGGTGCTATCGGTATCAACATCACCCGTTT
CCAGATCCACGTTTCTGGTACCAACGGTACCATCATCGCTTACACCATGTCTCTGGGTATCAACGCTTCTTTCGTTAACATCCAGATCTACCAGACCTCT
AACTTCCGTGTTCTGGGTTTCATCGCTGGTCTGATCGCTCTGAAAGGTGTTAAACTGCACTACACCCTGCTGTTCAACAAAGTTACCCTGGCTCTGCTGC
AGTACGGTTCTTTCTGCACCCTGTCTTCTACCGCTTCTGCTCTGGGTCTGACCGGTATCGCTGTTGAACAGGACAACCGTGCTCTGACCGGTATCGCTGT
TAACGTTCTGTACGAAAACCAGAAACTGCCGTTCGCTATGCAGATGGCTTACCGTCCGTTCCTGGGTGTTTACTACCACAAACCGATCGGTGCTGGTAT
CTGCGCTTCTCCGCTGGTTGACCTGCCGATCGGTATCCCGCAGACCCTGGAAATCCTGGACATCCGTGCTCTGACCGGTATCGCTGTTGAATCTGCTCTG
CTGGCTGGTACCATCACCTCTGTTCTGCACTCTACCCAGGACCTGACCGCTGGTGCTGCTGCTTACTACGTTGTTGGTTACCTGCAGCCGCGTACCTTCG
TTCTGTCTTTCGAACTGCTGCACGCTGTTAACCAGAACGCTCAGGCTCTGAACGTTACCCTG**GCG**GACGCT**GGC**TTCATCGTTACCTGGTTCCACGCTA
TCCACGTTGTTGTTCTGTCTTTCGAACTGCTGCACGTTGTTGTTCTGTCTTTCGAACTGCTGTGGACCGCTGGTGCTGCTGCTTACTACTACGCTTGGAAC
CGTAAACGTATCTCTTAC**GGCAGC**TTCTGCACCCAG**CTC**AACTACATCAAATGGCCGTGGTACATCTGGTACATCTGGCTGGGTTTCATCGCTGGTTAC
AACGAAAACGGTACCATCACCGACTACCAGCCGTACCGTGTTGTTGTTCTGTACACCTCTGCTCTGCTGGCTGGTACCTACGTTGGTTACCTGCAGCCG
CGTACCCACCACCACCACCACCAC

b

Fig. 13. a. Codon optimization of designed vaccine construct. The result showed that optimized codon have a CAI score of 0.9698 and GC content is 51.82 %. b. Final DNA sequence of the vaccine after codon optimization. c. *In silico* PCR amplification of vaccine construct followed by addition of restriction sites and cloning in PET28 (+) vector. d. Recombinant plasmid obtained after cloning of peptide in vector. Vaccine construct is showed in red color and black line is representing the vector backbone.



c



d

Fig. 13. (*continued*).

According to the secondary structure of chimeric vaccine construct, it contains 31.8 % helix, 10.09 % beta turns, and 28.11 % coil. To refine 3D structure of the construct vaccine Galax Refine was used. The Ramachandran Plot, or RAMPAGE, indicates the critical characteristics of a potential vaccination. Results show that only a minimum quantity of residues was present in the outlier area, whereas the common residues were found in favored areas. It indicates the suitable model quality that was intended.

The next phase in the vaccine's development is vital to achieving a positive result. It has been acknowledged that

immunoinformatic, which uses computational techniques such as molecular docking, is a powerful tool for predicting protein-protein interaction. A study shows that TLR3 triggers an immunological reaction in SARS-CoV2 [62]. Based on docking using Clustpro2.0, the vaccine-receptor complex model 9 interacts favorably with the reference structure. The minimum energy value of the docked complex displayed a stable connection and a lower RMSD [63]. A single picture of the involved physiological movement was provided by docking. As a result, it is necessary to study protein-protein interaction in an adjustable situation [64]. Based on the highest eigenvalue, the resultant complex is stable and displays a decreased chance of deformation during an immunological response. Covariance matrix analysis revealed the immunological simulation of the proposed construct, and its results matched the immune reactions [65].

Immune simulation was used to produce an average immune response. Frequently increasing antigen exposure increases the immunological response. The memory B cells have developed and maintained their consistency for several months. Along with helper T-cell simulation, memory T-cell development also occurred. After the initial injection, an elevated IL-2 level was indicated. Studies have shown that structural proteins stimulate a stronger T-cell response [66]. It was evaluated for immunoreactivity [67], which was done by expressing it in *E. coli* [68]. The codon was modified according to the host for maximum expression, yielding a CAI of 0.9698 and a GC content of 51.82 %.

No vaccine has been effectively constructed using the S-protein of the new coronavirus, vaccine cloning was not used when evolving previous vaccinations, and several potential vaccine candidates have limited coverage despite extensive *in-silico* studies on various potential vaccine candidates. If in vitro and in vivo research is performed, it would be a successful potential vaccine candidate.

5. Conclusion

In this study, computational methods have been used to develop a vaccine effective against SARS-CoV-2 infections. Because there are few risks of failure, the illustration of a promising vaccine design technique recommends it as a time- and money-saving strategy. A practical choice for scientific trials is an in vitro vaccine through a positive immune response and high population coverage. The immunological response was shown in an *in-silico* immunological simulation in line with antigen clearance. Snapgene-assisted computational cloning in the PET28a (+) plasmid revealed practical protein expression. However, experimental validation is essential for confirming the vaccine construct's efficacy against COVID-19. Peptide vaccines have shown satisfactory results with a more robust immune response in several trials; thus, this vaccine construction could be considered. By developing a potent vaccination, this study would help remove the worldwide risk presented by COVID-19.

6. Limitation

This *in-silico* study shows the path towards formulating a new drug to cure and treat COVID-19 infections. *In vivo* trials are needed to justify the computational behavior in animal/human pathology.

7. Future perspective

This bioinformatics study may help design a new vaccine/therapeutic target to treat lethal infections such as COVID-19 better.

Ethic approved

Not applicable because there are no human and animals in this study.

Funding

This work was funded by the Researcher Supporting Project number (RSP2024R123), King Saud University, Riyadh, Saudi Arabia.

CRediT authorship contribution statement

Sarmad Frogh Arshad: Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Rehana Rehana:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Muhammad Asif Saleem:** Writing – review & editing, Data curation. **Muhammad Usman:** Methodology, Formal analysis, Data curation. **Hasan Junaid Arshad:** Methodology, Investigation. **Rizwana Rizwana:** Software. **Shakeela Shakeela:** Software. **Asma Shah Rukh:** Software. **Imran Ahmad Khan:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Conceptualization. **M. Ali Hayssam:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Conceptualization. **Muhammad Anwar:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors would like to extend their sincere appreciation to theResearchers Supporting Project Number (RSP2024R123), King Saud University, Riyadh, Saudi Arabia.

References

[1] L. Ren, H. Cui, Y. Wang, F. Ju, Y. Cai, X. Gang, G. Wang, The role of lipotoxicity in kidney disease: from molecular mechanisms to therapeutic prospects, Biomed. Pharmacother. 161 (2023) 114465, https://doi.org/10.1016/j.biopha.2023.114465.
[2] W.S. Al-Qahtani, L.M. Alneghery, A.Q.S. Alqahtani, M.D. Alkahtani, S. Alkahtani, A review of comparison study between corona viruses (Sars-cov, mers-cov) and novel corona virus (COVID-19), Rev. Mex. Ing. Quim. 19 (2020) 201–212, https://doi.org/10.24275/rmiq/Bio1692.
[3] A.A. Rabaan, S.H. Al-Ahmed, S. Haque, R. Sah, R. Tiwari, Y.S. Malik, K. Dhama, M.I. Yatoo, D.K. Bonilla-Aldana, A.J. Rodriguez-Morales, SARS-CoV-2, SARS-CoV, and MERS-CoV: a comparative overview, Infez. Med. 28 (2020) 174–184.
[4] R. Tiwari, K. Dhama, K. Sharun, M. Iqbal Yatoo, Y.S. Malik, R. Singh, I. Michalak, R. Sah, D.K. Bonilla-Aldana, A.J. Rodriguez-Morales, COVID-19: animals, veterinary and zoonotic links, Vet. Q. 40 (2020) 169–182, https://doi.org/10.1080/01652176.2020.1766725.
[5] D.T. Chu, S.M. Vu Ngoc, H. Vu Thi, Y.V. Nguyen Thi, T.T. Ho, V.T. Hoang, V. Singh, J.A. Al-Tawfiq, COVID-19 in Southeast Asia: current status and perspectives, Bioengineered 13 (2022) 3797–3809, https://doi.org/10.1080/21655979.2022.2031417.
[6] I.A. Asante, S.N. Hsu, L. Boatemaa, L. Kwasah, M. Adusei-Poku, J.K. Odoom, Y. Awuku-Larbi, B.H. Foulkes, J. Oliver-Commey, E.K. Asiedu, M.D. Parker, O. Mitja, R.M. Eggo, L. de Oliveira-Martins, F. Asiedu-Bekoe, D.O. Laryea, P. Kuma-Aboagye, M. Marks, T.I. de Silva, W.K. Ampofo, Repurposing an integrated national influenza platform for genomic surveillance of SARS-CoV-2 in Ghana: a molecular epidemiological analysis, Lancet Global Health 11 (2023), https://doi.org/10.1016/S2214-109X(23)00189-4 e1075–e1085.
[7] S. Rehman, F.S. Hayat, S. Norin, A. Aziz, S.U. Rahman, i l n e n s r i t i t r e l c i l n e n s r i t 2 (2022) 1–8.
[8] M.A. Nugent, The Future of the COVID-19 Pandemic : How Good (or Bad) Can the SARS-CoV2 Spike Protein Get ?, 2022.
[9] M. Naveed, S. Tehreem, S. Arshad, S.A. Bukhari, M.A. Shabbir, R. Essa, N. Ali, S. Zaib, A. Khan, A. Al-Harrasi, I. Khan, Design of a novel multiple epitope-based vaccine: an immunoinformatics approach to combat SARS-CoV-2 strains, J. Infect. Public Health. 14 (2021) 938–946, https://doi.org/10.1016/j.jiph.2021.04.010.
[10] K.A. Madurani, S. Suprapto, M.Y. Syahputra, I. Puspita, A. Masudi, H.D. Rizqi, A.M. Hatta, J. Juniastuti, M.I. Lusida, F. Kurniawan, Review—recent development of detection methods for controlling COVID-19 outbreak, J. Electrochem. Soc. 168 (2021) 037511, https://doi.org/10.1149/1945-7111/abe9cc.
[11] M. Shirzad, M. Nourigorji, A. Sajedi, M. Ranjbar, F. Rasti, Z. Sourani, M. Moradi, S. Mostafa Mir, M.Y. Memar, Targeted therapy in Coronavirus disease 2019 (COVID-19): implication from cell and gene therapy to immunotherapy and vaccine, Int. Immunopharm. 111 (2022) 109161, https://doi.org/10.1016/j.intimp.2022.109161.
[12] H.M. Dabbous, M.H. El-Sayed, G. El Assal, H. Elghazaly, F.F.S. Ebeid, A.F. Sherief, M. Elgaafary, E. Fawzy, S.M. Hassany, A.R. Riad, M.A. TagelDin, Safety and efficacy of favipiravir versus hydroxychloroquine in management of COVID-19: a randomised controlled trial, Sci. Rep. 11 (2021) 1–7, https://doi.org/10.1038/s41598-021-85227-0.
[13] A.K. Singh, A. Singh, A. Shaikh, R. Singh, A. Misra, Chloroquine and hydroxychloroquine in the treatment of COVID-19 with or without diabetes: a systematic search and a narrative review with a special reference to India and other developing countries, Diabetes Metab. Syndr. Clin. Res. Rev. 14 (2020) 241–246, https://doi.org/10.1016/j.dsx.2020.03.011.
[14] L. Huang, Y. Rong, Q. Pan, K. Yi, X. Tang, Q. Zhang, W. Wang, J. Wu, F. Wang, SARS-CoV-2 vaccine research and development: conventional vaccines and biomimetic nanotechnology strategies, Asian J. Pharm. Sci. 16 (2021) 136–146, https://doi.org/10.1016/j.ajps.2020.08.001.
[15] F.M. Alshabrmi, F. Alrumaihi, S.F. Alrasheedi, W.A.I. Al-Megrin, A. Almatroudi, K.S. Allemailem, An In-Silico Investigation to Design a Multi-Epitopes Vaccine against Multi-Drug Resistant Hafnia Alvei, vol. 10, 2022, https://doi.org/10.3390/vaccines10071127. Vaccines.
[16] S. Gautam, Identification of Alpha 1B Glycoprotein and Haptoglobin 1 as Probable Biomarkers for Ear Ly Diagnosis of Leprosy, 2023, pp. 1–18.
[17] S.K. G, A. Joshi, V. Kaushik, T cell epitope designing for dengue peptide vaccine using docking and molecular simulation studies, Mol. Simulat. 46 (2020) 787–795, https://doi.org/10.1080/08927022.2020.1772970.
[18] S. Kalajirao, S.A. Ramakrishnan, V.R. Palanimuthu, Predicting, designing, characterization and evaluation of a new novel anticancer peptide SSVAM-9 against the lung carcinoma, an insilico approach, J. Med. Pharm. Allied Sci. 12 (2023) 5706–5712, https://doi.org/10.55522/jmpas.V12I2.4768.
[19] M. Alam, A. Saim, M. Saikat, Bioinformatics Approaches for Structural and Functional Annotation of an Uncharacterized Protein of Helicobacter pylori †, 2023, pp. 1–5.
[20] M. Naveed, U. Ali, T. Aziz, M.J. Rasool, A. Ijaz, M. Alharbi, M.E. Alharbi, A. Alshammari, A.F. Alasmari, A reverse vaccinology approach to design an mRNA-based vaccine to provoke a robust immune response against HIV-1, Acta Biochim. Pol. 70 (2023), https://doi.org/10.18388/abp.2020_6696.
[21] A. Ras-Carmona, H.F. Pelaez-Prestel, E.M. Lafuente, P.A. Reche, Bceps: a web server to predict linear b cell epitopes with enhanced immunogenicity and cross-reactivity, Cells 10 (2021), https://doi.org/10.3390/cells10102744.
[22] N. Kafee, F. Sefid, In Silico Identification and Validation of OSPA 3D Structure and its Topology as a Candidate for Lyme Disease Vaccine in Silico Identification and Validation of OSPA 3D Structure and its ABSTRACT, 2016.
[23] M. Kamthania, S. Srivastava, M. Desai, A. Jain, A. Shrivastav, D.K. Sharma, Immunoinformatics approach to design T-cell epitope-based vaccine against hendra virus, Int. J. Pept. Res. Therapeut. 25 (2019) 1627–1637, https://doi.org/10.1007/s10989-018-09805-z.
[24] A.S.S. Girija, Prediction of antigenic vaccine peptide candidates from BfmRS associated with biofilm formation in acinetobacter baumannii, Cureus 15 (2023), https://doi.org/10.7759/cureus.47804.
[25] M.Y. Mohammed, M. Abdelrhman, M. Khaier, N. Agabna, S. Shaddad, Vaccine Design for Marburg Virus Using VP35 Protein 10 (2022) 85–93, https://doi.org/10.11648/j.ijgg.20221004.11.
[26] A. Ullah, F.A. Shahid, M.U. Haq, M. Tahir ul Qamar, M. Irfan, B. Shaker, S. Ahmad, F. Alrumaihi, K.S. Allemailem, A. Almatroudi, An integrative reverse vaccinology, immunoinformatic, docking and simulation approaches towards designing of multi-epitopes based vaccine against monkeypox virus, J. Biomol. Struct. Dyn. 41 (2023) 7821–7834, https://doi.org/10.1080/07391102.2022.2125441.
[27] G. Samata, P. Sudhakar, G. Jyothsna, In silico Analysis of Spike Protein Glycoprotein A of Omicron variant and identification of variant specific peptide based Vaccine, Res. J. Biotechnol. 18 (2023) 1–6, https://doi.org/10.25303/1807rjbt01006.
[28] M. Validi, A. Karkhah, V. Kumar, H. Reza, Immuno-informatics based approaches to design a novel multi epitope-based vaccine for immune response reinforcement against Leptospirosis, Mol. Immunol. 104 (2018) 128–138, https://doi.org/10.1016/j.molimm.2018.11.005.
[29] S. Akter, M. Shahab, M.H. Sarkar, C. Hayat, T.A. Banu, B. Goswami, I. Jahan, E. Osman, M.S. Uzzaman, A. Habib, Immunoinformatics approach to epitope - based vaccine design against the SARS - CoV - 2 in Bangladeshi patients, J. Genet. Eng. Biotechnol. 8 (2022), https://doi.org/10.1186/s43141-022-00410-8.
[30] M. Bhattacharjee, M. Banerjee, A. Mukherjee, In silico designing of a novel polyvalent multi-subunit peptide vaccine leveraging cross-immunity against human visceral and cutaneous leishmaniasis: an immunoinformatics-based approach, J. Mol. Model. 29 (2023), https://doi.org/10.1007/s00894-023-05503-w.
[31] M. Tourani, S.S. Tehrani, A. Movahedpour, S.R. Arablouydareh, A. Maleksabet, A. Savardashtaki, H.G. Berenji, M. Taheri-anganeh, Design and evaluation of a multi-epitope vaccine for COVID- 19 : an in silico approach 2 (2023) 180–204.
[32] Z.M. Aqsha, M.A. Dharmawan, V.D. Kharisma, A. Nur, M. Ansori, N.I. Sumantri, Reverse Vaccinology Analysis of B-cell Epitope against Nipah Virus using Fusion Protein 16 (2023) 499–507.
[33] M. Liang, W. Hong, J. Shao, Bioinformatics-based prediction of character of envelope glycoprotein and analysis of epitopes of B- and T-cell of gp120, Asian J. Complement. Altern. Med. 11 (2023) 27–39, https://doi.org/10.53043/2347-3894.acam11006.

[34] S.A.H. Mohammadzadeh Hosseini Moghri, G. Mahmoodi Chalbatani, M. Ranjbar, C. Raposo, A. Abbasian, CD171 Multi-epitope peptide design based on immuno-informatics approach as a cancer vaccine candidate for glioblastoma, J. Biomol. Struct. Dyn. 41 (2023) 1028–1040, https://doi.org/10.1080/07391102.2021.2020166.
[35] M. Yousaf, S. Ismail, A. Ullah, S. Bibi, Immuno-informatics profiling of monkeypox virus cell surface binding protein for designing a next generation multi-valent peptide-based vaccine, Front. Immunol. 13 (2022) 1–20, https://doi.org/10.3389/fimmu.2022.1035924.
[36] M.A. Shabbir, M.H. Javed, A. Waqar, M. Huraira, Computational design and analysis of a novel inhibitor for FERMT1 gene : a novel treatment strategy for the, Kindler Syndrome 1 (2023) 55–68.
[37] H. Ur Rashid, V. Bolzani, K. Khan, L.A. Dutra, N. Ahmad, A. Wadood, Homology modeling of alpha-glucosidase from Candida albicans: sequence analysis and structural validation studies in silico, J. Braz. Chem. Soc. (2023), https://doi.org/10.21577/0103-5053.20230123.
[38] M.N.Z. Zakaria, A.F. Aththar, M. Fai, M.A. Hamami, V.D. Kharisma, A. Affan, A. Murtadlo, A. Nur, M. Ansori, T.H. Sucipto, R. Zainul, A Novel Multi-Epitope Vaccine Design Targeting E1/E2 Envelope Glycoprotein of Chikungunya Virus : an Immunoinformatics Approach A Novel Multi-Epitope Vaccine Design Targeting E1/E2 Envelope Glycoprotein of Chikungunya Virus, An Immunoinformatics Appr, 2024, https://doi.org/10.26655/JMCHEMSCI.2024.2.6.
[39] E. Aktaş, İ. Saygılı, E. Kahveci, Z. Tekbıyık, N.Ö. Özgentürk, Bioinformatic investigation of Nipah virus surface protein mutations: molecular docking with Ephrin B2 receptor, molecular dynamics simulation, and structural impact analysis, Microbiol. Immunol. (2023), https://doi.org/10.1111/1348-0421.13098.
[40] N.A. Khamjan, M. Lohani, M.F. Khan, S. Khan, A. Algaissi, Immunoinformatics strategy to develop a novel universal multiple epitope-based COVID-19 vaccine, Vaccines 11 (2023) 1090, https://doi.org/10.3390/vaccines11061090.
[41] S. Aiman, A. Ahmad, A.K. Id, Y.A. Id, A. Malik, Y.A. Id, Multi-Epitope-Based Vaccine Engineering against Multidrug Resistant Shigella Sonnei : Immunoinformatics and Chemoinformatics Approaches, 2023, pp. 1–25, https://doi.org/10.1371/journal.pone.0289773.
[42] A. Gupta, N. Sahu, V.K. Singh, R.P. Sinha, Evolutionary aspects of mutation in functional motif and post-translational modifications in SARS-CoV-2 3CLpro (Mpro): an in-silico study, J. Protein Proteonomics 14 (2023) 99–109, https://doi.org/10.1007/s42485-023-00105-9.
[43] J. Du, G. Wu, Q. Chen, C. Yu, G. Xu, A. Liu, Since January 2020 Elsevier Has Created a COVID-19 Resource Centre with Free Information in English and Mandarin on the Novel Coronavirus COVID- 19 . The COVID-19 Resource Centre Is Hosted on, Elsevier Connect , the company ' s public news and information, 2020.
[44] N.H.N. Bazlan, N.M. Yunus, M.F.Z.R. Yahya, In Silico analysis and homology modelling of carbohydrate metabolic enzymes from Corynebacterium Pseudotuberculosis, AIP Conf. Proc. 2720 (2023), https://doi.org/10.1063/5.0137771.
[45] A. Jamil, A.R.A. Bakar, N.H. Khairuddin, Structural elucidation of the Masp1 and Masp2 protein from Nephila pilipes web via bioinformatics approaches, J. Adv. Res. Appl. Sci. Eng. Technol. 30 (2023) 132–144, https://doi.org/10.37934/araset.30.1.132144.
[46] T. Soni, V. Prabha, Advances in in Silico Technologies and Their Applications : Special Attention to Drug Development , Vaccine Design , and Advances in in Silico Technologies and Their Applications : Special Attention to Drug Development , Vaccine Design , and Molecular Mim, 2023, https://doi.org/10.9734/bpi/rhst/v6/6740A.
[47] J.J. Hiu, J.K.Y. Fung, H.S. Tan, M.K.K. Yap, Unveiling the functional epitopes of cobra venom cytotoxin by immunoinformatics and epitope-omic analyses, Sci. Rep. 13 (2023) 1–14, https://doi.org/10.1038/s41598-023-39222-2.
[48] L. Li, Z. Zhao, X. Yang, Z. Su, W. Li, S. Chen, L. Wang, T. Sun, C. Du, Z. Li, Z. Yang, M. Li, T. Wang, Y. Wang, Y. Fan, H. Wang, J. Zhang, A newly identified spike protein targeted linear B-cell epitope based dissolvable microneedle array successfully eliciting neutralizing activities against SARS-CoV-2 wild-type strain in mice, Adv. Sci. 2207474 (2023), https://doi.org/10.1002/advs.202207474, 1–13.
[49] K.M.K. Kibria, M.O. Faruque, M.S. bin Islam, H. Ullah, S. Mahmud, M. Miah, A.A. Saleh, A conserved subunit vaccine designed against SARS-CoV-2 variants showed evidence in neutralizing the virus, Appl. Microbiol. Biotechnol. 106 (2022) 4091–4114, https://doi.org/10.1007/s00253-022-11988-x.
[50] Y. Han, Z. Lin, J. Zhou, G. Yun, R. Guo, J.J. Richardson, F. Caruso, Polyphenol-mediated assembly of proteins for engineering functional materials, Angew. Chem. Int. Ed. 59 (2020) 15618–15625, https://doi.org/10.1002/anie.202002089.
[51] A.S.S. Girija, S. Gunasekaran, S. Habib, M. Aljeldah, B.R. Al Shammari, A.A. Alshehri, A.S.S. Alwashmi, S.A. Turkistani, A. Alawfi, A. Alshengeti, Prediction of Putative Epitope Peptides against BaeR Associated with TCS Adaptation in Acinetobacter Baumannii Using an in Silico Approach, 2023.
[52] S. Saha, S. Vashishtha, B. Kundu, M. Ghosh, In - silico design of an immunoinformatics based multi - epitope vaccine against Leishmania donovani, BMC Bioinf. (2022) 1–28, https://doi.org/10.1186/s12859-022-04816-6.
[53] S. Huang, C. Zhang, J. Li, Z. Dai, J. Huang, F. Deng, X. Wang, X. Yue, X. Hu, Y. Li, Y. Deng, Y. Wang, W. Zhao, Z. Zhong, Y. Wang, Designing a multi-epitope vaccine against coxsackievirus B based on immunoinformatics approaches, Front. Immunol. 13 (2022) 1–19, https://doi.org/10.3389/fimmu.2022.933594.
[54] I. Ahmad, S.S. Ali, B. Zafar, H.F. Hashmi, I. Shah, S. Khan, M. Suleman, M. Khan, S. Ullah, S. Ali, J. Khan, M. Ali, A. Khan, D.Q. Wei, Development of multi-epitope subunit vaccine for protection against the norovirus' infections based on computational vaccinology, J. Biomol. Struct. Dyn. 40 (2022) 3098–3109, https://doi.org/10.1080/07391102.2020.1845799.
[55] W. Yan, Y. Zheng, X. Zeng, B. He, W. Cheng, Structural biology of SARS-CoV-2: open the door for novel therapies, Signal Transduct. Targeted Ther. 7 (2022), https://doi.org/10.1038/s41392-022-00884-5.
[56] W.G. dos Santos, Impact of virus genetic variability and host immunity for the success of COVID-19 vaccines, Biomed. Pharmacother. 136 (2021), https://doi.org/10.1016/j.biopha.2021.111272.
[57] C. Agrati, C. Castilletti, D. Goletti, A. Sacchi, V. Bordoni, D. Mariotti, S. Notari, G. Matusali, S. Meschi, L. Petrone, A. Aiello, S. Najafi Fard, C. Farroni, F. Colavita, D. Lapa, S. Leone, A. Agresta, M. Capobianchi, G. Ippolito, F. Vaia, V. Puro, E. Cimini, E. Tartaglia, R. Casetti, G. Grassi, F. Cristofanelli, A. Capri, A. Santoro, N. Orchi, A. Bettini, M. Francalancia, E. Specchiarello, Persistent Spike-specific T cell immunity despite antibody reduction after 3 months from SARS-CoV-2 BNT162b2-mRNA vaccine, Sci. Rep. 12 (2022) 1–5, https://doi.org/10.1038/s41598-022-07741-z.
[58] A.N.M. Ansori, M.K.J. Kusala, I. Normalina, S. Indrasari, M.Y. Alamudi, R.V. Nidom, K.P. Santoso, K. Rachmawati, C.A. Nidom, Immunoinformatic investigation of three structural protein genes in Indonesian SARS-CoV-2 isolates, Sys. Rev. Pharm. 11 (2020) 422–434, https://doi.org/10.31838/srp.2020.7.62.
[59] J. Sidney, B. Peters, A. Sette, Epitope prediction and identification- adaptive T cell responses in humans, Semin, Immunol. 50 (2020) 101418, https://doi.org/10.1016/j.smim.2020.101418.
[60] J. Dey, S.R. Mahapatra, P. Singh, S. Patro, G.S. Kushwaha, N. Misra, M. Suar, B and T cell epitope-based peptides predicted from clumping factor protein of Staphylococcus aureus as vaccine targets, Microb. Pathog. 160 (2021) 105171, https://doi.org/10.1016/j.micpath.2021.105171.
[61] M.T. Khan, A. Mahmud, M. Hasan, K.F. Azim, M.K. Begum, M.H. Rolin, A. Akter, S.I. Mondal, Proteome exploration of Legionella pneumophila to identify novel therapeutics: a hierarchical subtractive genomics and reverse vaccinology approach, Microbiol. Spectr. 10 (2022), https://doi.org/10.1128/spectrum.00373-22.
[62] K. Rawal, R. Sinha, B.A. Abbasi, A. Chaudhary, S.K. Nath, P. Kumari, P. Preeti, D. Saraf, S. Singh, K. Mishra, P. Gupta, A. Mishra, T. Sharma, S. Gupta, P. Singh, S. Sood, P. Subramani, A.K. Dubey, U. Strych, P.J. Hotez, M.E. Bottazzi, Identification of vaccine targets in pathogens and design of a vaccine using computational approaches, Sci. Rep. 11 (2021) 1–25, https://doi.org/10.1038/s41598-021-96863-x.
[63] M. Hammed-Akanmu, M. Mim, A.Y. Osman, A.M. Sheikh, E. Behmard, A.A. Rabaan, R. Suppain, K. Hajissa, Designing a multi-epitope vaccine against toxoplasma gondii: an immunoinformatics approach, Vaccines 10 (2022), https://doi.org/10.3390/vaccines10091389.
[64] R. Khan, Computational and in Out, 2023.
[65] A. Moodley, A. Fatoba, M. Okpeku, T. Emmanuel Chiliza, M. Blessing Cedric Simelane, O.J. Pooe, Reverse vaccinology approach to design a multi-epitope vaccine construct based on the Mycobacterium tuberculosis biomarker PE_PGRS17, Immunol. Res. 70 (2022) 501–517, https://doi.org/10.1007/s12026-022-09284-x.
[66] N. Dixit, C. Fanton, J.L. Langowski, Y. Kirksey, P. Kirk, T. Chang, J. Cetz, V. Dixit, G. Kim, P. Kuo, M. Maiti, Y. Tang, L.A. VanderVeen, P. Zhang, M. Lee, J. Ritz, Y. Kamihara, C. Ji, W. Rubas, T.D. Sweeney, S.K. Doberstein, J. Zalevsky, NKTR-358: a novel regulatory T-cell stimulator that selectively stimulates expansion

and suppressive function of regulatory T cells for the treatment of autoimmune and inflammatory diseases, J. Transl. Autoimmun. 4 (2021) 100103, https://doi.org/10.1016/j.jtauto.2021.100103.

[67] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics 16 (2000) 404–405, https://doi.org/10.1093/bioinformatics/16.4.404.

[68] K.S. Gunnarsen, L.S. Høydahl, R.S. Neumann, K. Bjerregaard-Andersen, N.R. Nilssen, L.M. Sollid, I. Sandlie, G.Å. Løset, Soluble T-cell receptor design influences functional yield in an E. Coli chaperone-assisted expression system, PLoS One 13 (2018) 1–15, https://doi.org/10.1371/journal.pone.0195868.