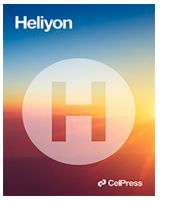
Heliyon 10 (2024) e35129

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Bioinformatic elucidation of conserved epitopes to design a potential vaccine candidate against existing and emerging SARS-CoV-2 variants of concern

Amber Rastogi [a,b], Sakshi Gautam [a,b], Manoj Kumar [a,b,*]

[a] *Virology Unit and Bioinformatics Centre, Institute of Microbial Technology, Council of Scientific and Industrial Research (CSIR), Sector 39A, Chandigarh, 160036, India*
[b] *Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, 201002, India*

ARTICLE INFO

Keywords:
SARS-CoV-2
COVID-19
Variants of concern
Immunogenic epitope
Multi-epitope
Vaccine design

ABSTRACT

The COVID-19 pandemic caused by SARS-CoV-2 poses a significant adverse effects on health and economy globally. Due to mutations in genome, COVID-19 vaccine efficacy decreases. We used immuno-informatics to design a Multi epitope vaccine (MEV) candidate for SARS-CoV-2 variants of concern (VOCs). Hence, we predicted binders/epitopes MHC-I, $CD8^+$, MHC-II, $CD4^+$, and CTLs from spike, membrane and envelope proteins of VOCs. In addition, we assessed the conservation of these binders and epitopes across different VOCs. Subsequently, we designed MEV by combining the predicted CTL and $CD4^+$ epitopes from spike protein, peptide linkers, and an adjuvant. Further, we evaluated the binding of MEV candidate against immune receptors namely HLA class I histocompatibility antigen, HLA class II histocompatibility antigen, and TLR4, achieving binding scores of −1265.3, −1330.7, and −1337.9. Molecular dynamics and normal mode analysis revealed stable docking complexes. Moreover, immune simulation suggested MEV candidate elicits both innate and adaptive immune response. We anticipate that this conserved MEV candidate will provide protection from VOCs and emerging strains.

1. Introduction

The COVID-19 pandemic caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has adversely affected humans and the economy globally. SARS-CoV-2 is a (+) sense-enveloped RNA virus belonging to the family *Coronaviridae* and genus Beta CoV [1]. Genomic analyses indicated immunologically important SARS-CoV-2 genes subjected to increased selective pressure because of a change in the host environment [2]. SARS-CoV-2 is continuously evolving with various variants worldwide, which are classified as variant of interest (VOI), variant of concern (VOC), and variant under monitoring (VUM) [3]. In addition, variants with more mutations were emerging compared to previously circulating variants [4]. Consequently, with the emergence of new variants of SARS-CoV-2, the population's infection rate is rapidly increasing [5]. Due to a large number of mutations, new variants are emerging, leading to enhanced transmissibility or severity along with plausible depletion in the vaccine's efficacy [6]. Various mutations were associated within the spike protein (RBD region), responsible for increased transmissibility, reduced sensitivity for neutralizing

* Corresponding author. Virology Unit and Bioinformatics Centre, Institute of Microbial Technology, Council of Scientific and Industrial Research (CSIR), Sector 39A, Chandigarh, 160036, India.
E-mail addresses: amberrastogi@imtech.res.in (A. Rastogi), manojk@imtech.res.in (M. Kumar).

https://doi.org/10.1016/j.heliyon.2024.e35129
Received 27 September 2023; Received in revised form 22 July 2024; Accepted 23 July 2024
Available online 23 July 2024

antibodies, and resistance against immunization [6–8]. According to researchers, a rapid increase in the SARS-CoV-2 variants can evade neutralization by antibodies [9,10]. Current scenarios suggest that crucial mutations in SARS-CoV-2 variants may impact their overall structure-function. Therefore, it is also harmful for immunization programs when various mutations in proteins cause protein structure changes.

In this context, various vaccines developed by PfIzer-BioNTech, Moderna, J&J Janssen, which largely controlled SARS-CoV-2 globally [11]. However, Jamie et al. reported 74.5 % and 67.0 % effectiveness of the ChAdOx1 nCoV-19 vaccine among persons infected with alpha and delta variants, respectively [12]. Zhou et al. suggested no vaccine can be applied to all situations or cases due to the emergence of new variants therefore suggested to upgrade and develop vaccines promptly [13]. Henceforth, there is a requirement to identify conserved epitopes and design a multiple epitopes vaccines (MEV) against existing and emerging variant of concerns of SARS-CoV-2.

Therefore, we adopted an integrated immune-informatics approach to predict putative immunogenic epitopes derived from various proteins, *viz.,* spike, membrane, envelope proteins from different SARS-CoV-2 VOCs. In addition, we also assessed the conservancy of these binders and epitopes for different VOCs. Eventually, 23 conserved CTL epitopes and 7 HTL ($CD4^+$) epitopes derived from S-protein were used to design the MEV candidate that might be capable of inducing innate and adaptive immune responses. The advantage of MEV containing conserved epitopes is that they may protect against SARS-CoV-2 omicron and other emerging VOCs.

2. Materials and methods

2.1. Data collection

The present work aims to predict immunogenic epitopes among the SARS-CoV-2 variants of concern and design a MEV candidate. We utilized distinct SARS-CoV-2 VOCs for this work mentioned in Supplementary Table S1.

2.2. T-cell epitope prediction

MHC-I and MHC-II binders, CTL epitopes, $CD8^+$ and $CD4^+$ immunogenic T-cell epitopes were predicted from spike, membrane and envelope protein sequence. For the prediction of MHC-I and MHC-II, we have utilized MHC-I binding prediction tools available on the IEDB epitope analysis web page using the consensus method [14,15]. Similarly, MHC-II binders were predicted utilizing the IEDB recommended 2.2 method, species/locus (Human), Full HLA reference set, and default length [14]. Further, Binder peptides were sorted based on low-affinity value (IC50 $<$ 50 nM) and percentile rank [16]. The schematic representation of MHC-I and MHC-II binder's prediction from proteins of SARS-CoV-2 VOCs are shown in Supplementary Figs. S1–S2.

The IEDB Class I Immunogenicity tool with default parameters was utilized to predict the $CD8^+$ immunogenic peptides [17]. This tool predicts the immunogenicity based on amino acid position in the peptide sequences and amino acid properties. The potential MHC-I epitopes were used as input sequences to predict $CD8^+$ immunogenic peptides. It gives output in the tabular form consisting of a peptide, length, and score. The higher the score, the more likely it is induce an immunological response [17].

$CD4^+$ T cell immunogenicity tool on the IEDB web page used for immunogenicity prediction of MHC-II restricted epitopes [18]. The IEDB-recommended (combined) approach was adopted to predict the $CD4^+$ immunogenic epitopes, which gives results by combining the output of the 7-allele and immunogenicity methods [18,19]. The potential MHC-II epitopes were used as input sequences to predict $CD4^+$ T cell immunogenic epitopes. The low values indicate more immunogenic peptides and vice versa [20]. An illustration depicting the immunogenic prediction of $CD8^+$ and $CD4^+$ from proteins of SARS-CoV-2 variants of VOCs is given in Supplementary Fig. S3.

CTL epitopes were predicted for all HLA supertypes using NetCTLpan [21]. Predicting CTL epitopes considers multiple sequential processes like proteasome cleavage, TAP binding, and MHC-I binding [21]. CTL epitopes were predicted at default parameters (Peptide length 9 mers, weight on C terminal cleavage 0.225, and weight on TAP transport efficiency 0.025). The methodology and flowchart of CTL epitope prediction is given in Supplementary Fig. S4.

2.3. B-cell linear epitope

The B-cell epitope tool based on antigen sequence properties available at IEDB was used to predict linear B-cell epitopes [22]. BepiPred 2.0 method adopted for B-cells epitope prediction. It relies on the random forest algorithm. It considers epitope amino acids and non-epitope amino acids from crystal structure of antigen-antibody complex [22]. A group of amino acid residues with a score higher than the standard cut-off value 0.5 is called an epitope [22]. The schematic representation of B-cell linear epitope prediction from proteins of VOCs is shown in Supplementary Fig. S5.

2.4. Antigenicity prediction

VaxiJen v2.0 server [23] was utilized to accomplish antigenicity prediction of filtered CTL epitopes, B-cell epitopes, MHC-I, and MHC-II binders. VaxiJen employs an alignment-independent approach based on peptide sequences ACC transformation [23]. To predict the antigenicity of these peptides, antigenicity prediction was done using the standard cut-off of 0.4 for viral antigens [23].

2.5. Toxicity prediction

ToxinPred webserver was used to predict the toxicity of antigenic epitope (VaxiJen score >0.4). This web server relies on a quantitative matrix, SVM, and numerous peptide properties [24]. Here, we adopted the SVM (Swiss-Prot) method for prediction at default parameters except for SVM threshold of 0.4.

2.6. Interferon-gamma (IFN-γ)

Interferon-gamma (IFN-γ) secreted from T-helper cells plays a very crucial role in the clearance of viruses from host cells [25]. For the prediction of IFN-γ-inducing peptides, the IFN-epitope web server was utilized at default parameters except the SVM model [20].

2.7. Allergenicity prediction

AllerTOP v. 2.0 was used to predict the allergenicity of non-toxic epitopes. This method involves transforming protein sequences into vectors of uniform length through auto cross-covariance to differentiate allergens from non-allergens [26].

2.8. Population coverage analysis of MEV candidate

Different populations possess polymorphic HLAs of varying frequencies, which can lead to biased coverage of epitopes on some HLAs in populations [27]. Therefore, it is necessary to consider population coverage when designing vaccines to avoid decreasing their applicability in some populations [28]. IEDB population coverage tool was used to analyze the predicted epitopes and the associated HLA alleles [28]. At least one T-cell epitope from the collection will likely trigger an immune response in individuals in a population [30].

2.9. Designing of MEV candidate

We designed a MEV candidate that contains 23 conserved CTL epitopes with 7 $CD4^+$ immunogenic epitopes from S-protein of VOCs. Various linkers help in epitope separation and assist epitope presentation to MHC class I and II receptors [29]. Two different types of linkers, AAY and GPGPG, were used to join various CTLs and $CD4^+$ immunogenic epitopes for immune responses [29,30]. The AAY linker efficiently separates CTL epitopes within cells, reducing immunogenicity at the junctions [31]. Additionally, the GPGPG linker is used with $CD4^+$ immunogenic epitopes to induce an HTL immune response and restore each epitope's immunogenicity post-processing [32]. The bi-lysine (KK) linker ensures that the epitopes maintain their autonomous immunogenic functions within the vaccine. These linkers were thus crucial in developing our MEV candidate [33]. Further, CTxB adjuvant linked to the N-terminal region of a MEV candidate using EAAAK linker to enhance the immune response [34]. Eventually, to boost the immunogenicity inside the host, thirteen amino acids PADRE sequence (AKFVAAWTLKAAA) was added at C-terminal end.

2.10. Physiochemical properties of MEV candidate

ExPasy ProtParam tool was utilized to evaluate the putative MEV candidate physiochemical properties [35]. The number of amino acids, aliphatic index, isoelectric point (pI), total atoms number, molecular weight, and projected half-life of the relevant sequence were calculated.

2.11. Secondary and tertiary prediction of MEV candidate

The secondary structure was investigated utilizing the SOPMA [36] and PSIPRED 4.0 [37] self-optimized prediction methods. Transmembrane helices, globular areas, random coil, and coiled-coil regions, allare shown in a graphical fashion using these servers. Additionally, Protein–Sol was used to predict the solubility of putative MEV candidate [38]. Protein–Sol server developed to predict the solubility of proteins. It utilizes available data on *E. coli* protein solubility in a cell-free expression system to calculate properties derived from the protein sequence. Tertiary protein structures provide vital information for understanding molecular processes at the atomic level. Raptor-X [39], a template-based protein structure modeling server, was applied to predict the tertiary structure of a putative MEV candidate. Further, the predicted structure was refined using GalaxyWEB Server [40].

2.12. Tertiary structure validation of MEV candidate

The tertiary structure of putative MEV candidate protein model was validated using primarily three web-based servers: ProSA [41], ERRAT [42], and PROCHECK [43]. The ProSA server confirmed the protein's tertiary structure [39]. Further, model protein non-bonded interaction was statistically analyzed using the ERRAT tool [42]. Finally, for Ramachandran plot analysis, we used the PROCHECK [43].

2.13. Di-sulfide modifications in MEV candidate

Disulfide by Design 2 (DbD2) tool [44] utilized to decreased conformational entropy and enhanced protein stability. This tool measure uncertainty or unpredictability that affects 3D conformations of modeled vaccine construct by adding new disulfide bonds [44].

2.14. Molecular docking of MEV candidate

To analyze putative MEV candidate binding affinity for various antigenic receptors that exist on the immune cell surface like Toll-like receptor-4 (PDB ID: 3FXI), HLA class I histocompatibility antigen, B-8 alpha chain (PDB ID: 6P2F), and HLA class II histocompatibility antigen, DR alpha chain (PDB ID: 3C5J). We used the ClusPro server to perform molecular docking. The models were generated using ClusPro based on electrostatic interactions and desolvation energy [45]. Subsequently, the binding of putative MEV candidate to various antigenic molecules were validated via. HDOCK server [46].

2.15. Molecular dynamic simulation

We performed a molecular dynamics simulation of a putative MEV candidate and HLA class II histocompatibility antigen, DR alpha chain. Molecular dynamics simulation of HLA class II and putative MEV candidate docked complex was performed by WEBGRO server [47] to examine the binding stability of the complex [48]. Water molecules were used to solvate the whole complex. In addition, the system was neutralized, and 0.15 M NaCl salt was added using GROMOS96 43a1 force field parameters. An energy decrease of 5000 steps resulted from the sharpest descending strategy. Equilibration was done using constant quantity, volume, temperature, and pressure (NVT/NPT). A simulation was conducted at 300K temperature and pressure of 1.0 bar for a simulation time of 50ns with 1000 frames per simulation. Simulation parameters were calculated as the Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF).

2.16. Normal mode analysis (NMA) of MEV candidate

The collective functional movements of biological macromolecules are naturally reproduced by NMA based on dihedral co-ordinates [49]. NMA based on internal coordinates to display the cooperative movement of the MEV candidate was performed using the iMODS webserver [49].

2.17. Immune simulation profiling of MEV candidate

The C-ImmSim web server [50] was used to characterize our putative MEV candidate's immune responses. In this study, some parameters were adjusted at 12345, 1000, and 25 for random seed, simulation step, and simulation volume. We used three injections at different time interval (0, 30, and 60 days) to accomplish the immunological simulation in this model. The default settings were used for the remaining parameters [51].

2.18. In silico cloning of MEV candidate in pET28a (+) vector

The protein sequence was reverse translated and optimized using the Java Codon Adaptation Tool server [52] for continual putative MEV candidate expression in the (*E. coli*) system. The restriction sites *Xho*I and *Bam*HI were attached at the N and C-terminal of the sequence. Further, the sequence was incorporated into the pET28a (+) vector between *Xho*I and *Bam*HI restriction sites. Snap gene version 6.0 was used for the in-silico cloning experiment.

2.19. Conservancy of CTLs epitopes

The conservancy of 23 CTL epitope were examined in major sublineages of VOCs using MEGA software. Further, MSA was performed using a muscle alignment algorithm.

3. Results

3.1. Prediction for MHC-I binders and $CD8^+$ immunogenic epitope

12 potential MHC-I binders were predicted from S-protein among all SARS-CoV-2 VOCs. ALQIPFAMQMAYR and RSYLTPGDSSSGW binders were found across all VOCs and the Wuhan strain. The binder FSTFKCYGVSPTK was identified in five VOCs; it was absent in the omicron variant due to the substitution of FSTFK to VKLHY. Additionally, three novel binders found exclusively in omicron, namely SVLYNLAPFFTFK, YSVLYNLAPFFTF, and VLYNLAPFFTFKC, as a result of the mutations indicated by S372L, S374P, and S376F respectively. In the context of the M protein, 10 MHC-I were predicted, out of which six epitopes were common across all VOCs, whereas only three were observed in five VOCs. Specifically, the IATAMACLVGLMW sequence was present in delta variant. The detailed list of MHC-I binders predicted from M proteins of VOCs was given in Supplementary Table S2.

Using the "IEDB Class I Immunogenicity Tool" with the default parameters, we identified $CD8^+$ immunogenic epitopes for different proteins restricted to MHC-I. S protein has 12 $CD8^+$ immunogenic epitopes, as mentioned in Table 1. For the S protein, all variants of concern VOCs contained four $CD8^+$ epitopes. The epitope FSTFKCYGVSPTK was found in five of the VOCs, with the exception being omicron. Conversely, seven epitopes were unique to a single VOC. Similarly, for the M protein, six epitopes were common across all VOCs, while three specific epitopes were identified in five of the VOCs. In a distinct case, the IATAMACLVGLMW epitope was exclusively observed in the delta variant. However, the E protein does not contain any $CD8^+$ immunogenic epitopes Supplementary Table S3.

3.2. Prediction for MHC-II and $CD4^+$ immunogenic epitope

41 potential binders from S protein were predicted from VOCs, as shown in Supplementary Table S4. Regarding the S protein, there were 16 binders common across all VOCs. Another set of nine binders were detected in five different VOCs. The binder LQYGSFCTQLNRALT was observed in four VOCs, while FLVLLPLVSSQCVNF was noted in two VOCs. Additionally, 14 binders were novel to a single variant. For the E protein, a total of nine binders were found to be conserved across all VOCs. In addition, 24 binders were identified from the M-protein, with LVIGFLFLTWICLLQ and LVGLMWLSYFIASFR being present in all VOCs. While twelve binders were identified in five VOCs, and ten binders were specific to just one VOC. In addition, the detailed list of MHC-II binders predicted from E and M proteins of SARS-CoV-2 VOCs was given in Supplementary Table S4.

Using default parameters in the "IEDB Class II Immunogenicity Tool" we also predict $CD4^+$ immunogenic epitopes for different proteins restricted to MHC-II. For the S protein, 8 $CD4^+$ immunogenic epitopes were predicted in all VOCs Table 2. The epitopes AYYVGYLQPRTFLLK and RFASVYAWNRKRISN were identified across all variants of concern (VOCs). Additionally, there were three epitopes present in four different VOCs, and the epitope FLVLLPLVSSQCVNF was found in two VOCs. Meanwhile, two novel epitopes were detected in one VOC. For the E-protein, it was observed that all nine of its epitopes were conserved across all VOCs. On the other hand, the M-protein exhibited 12 immunogenic epitopes, with the epitope LVGLMWLSYFIASFR being common among all VOCs. Furthermore, six epitopes were common to five VOCs, and five epitopes was present in one VOC as given in Supplementary Table S5.

3.3. Prediction for CTL epitope

Using NetCTLpan, we predicted CTL epitopes from proteins among all VOCs. The S-protein revealed 41 CTL epitopes, including 23 common epitopes across all VOCs as detailed in Table 3. Additionally, four epitopes were shared among five variants, while SPRRARSVA epitope identified in four variants. Moreover, twelve epitopes were present in only one variant. Similarly, for the E protein, a total of six epitopes were predicted, with five epitopes common across all variants. The SFYVYSRVK epitope was found in five variants. In the case of the M protein, 14 epitopes were predicted, with 11 epitopes conserved across all variants. Conversely, the LWPVTLACF epitope was present in five variants, while three epitopes were specific to individual variants. In addition, CTL epitopes form E and M protein were given in Supplementary Table S6.

3.4. Linear B-cell epitopes

Using the "IEDB B-cell epitope prediction tool" linear B-cell epitopes were predicted from proteins of VOCs. In the S protein, various epitopes show overlapping protein sequences in all strains, as shown in Supplementary Table S7. The result of B-cell epitope prediction for membrane and envelope proteins is also provided in Supplementary Tables S8–9.

3.5. Designing of MEV candidate

We design an MEV that contains 23 (4 novel) conserved CTL epitopes and 7 (5 novel) HTL ($CD4^+$) immunogenic epitopes from

Table 1
Predicted $CD8^+$ Epitopes from S Protein of SARS-CoV-2 Omicron and Other VOCs; higher Scores Indicate increased immunogenicity.

Allele	Peptide	Score	Wuhan	Alpha	Beta	Gamma	Delta	omicron
HLA-A*11:01/HLA-A*03:01	SVLYNLAPFFTFK	0.32824	–	–	–	–	–	+
HLA-A*03:01	VLYNLAPFFTFKC	0.24613	–	–	–	–	–	+
HLA-A*23:01	YSVLYNLAPFFTF	0.24052	–	–	–	–	–	+
HLA-A*01:01	VNFTTRTQLPPAY	0.08252	–	–	+	–	–	–
HLA-A*11:01	**SVTTEILPVSMTK**	**0.01611**	+	+	+	+	+	+
HLA-A*23:01	**AAYYVGYLQPRTF**	**−0.003**	+	+	+	+	+	+
HLA-A*31:01	FKIYSKHTPIIVR	−0.0114	–	–	–	–	–	+
HLA-A*01:01	VNFTNRTQLPSAY	−0.09184	–	–	–	+	–	–
HLA-A*31:01	**ALQIPFAMQMAYR**	**−0.18062**	+	+	+	+	+	+
HLA-A*11:01	FSTFKCYGVSPTK	−0.20491	+	+	+	+	+	–
HLA-A*31:01	SYQTQTKSHRRAR	−0.29788	–	–	–	–	–	+
HLA-B*57:01	**RSYLTPGDSSSGW**	**−0.36741**	+	+	+	+	+	+

(+) sign Indicate the presence of $CD8^+$ epitope while (−) Indicate the absence of $CD8^+$ epitope. Conserved $CD8^+$ epitopes in all SARS-CoV-2 VOCs are shown in bold.

Table 2
The table presents predicted $CD4^+$ immunogenic epitopes derived from the S protein of SARS-CoV-2 Omicron variant and other VOCs. The alleles, peptides, combined scores, and their presence (+) or absence (−) with different variants are listed.

Allele	Peptide	Combined _score	Wuhan	Alpha	Beta	Gamma	Delta	Omicron
HLA-DRB1*15:01 HLA-DRB5*01:01 HLA-DRB1*01:01	FQTLLALHRSYLTPG	40.50464	+	+	–	+	+	+
HLA-DRB1*11:01	GNYNYRYRLFRKSNL	42.80992	–	–	–	–	+	–
HLA-DRB5*01:01 HLA-DRB1*01:01	GINITRFQTLLALHR	43.57824	+	+	–	+	+	+
HLA-DRB5*01:01 HLA-DRB1*11:01	**RFASVYAWNRKRISN**	**45.3264**	+	+	+	+	+	+
HLA-DRB4*01:01	IGINITRFQTLHRSY	45.57656	–	–	+	–	–	–
HLA-DRB1*01:01	FLVLLPLVSSQCVNF	45.75596	–	–	+	+	–	–
HLA-DRB5*01:01 HLA-DRB1*11:01	RAAEIRASANLAATK	45.87236	+	+	+	–	+	+
HLA-DQA1*05:01 HLA-DQB1*03:01	**AYYVGYLQPRTFLLK**	**47.8952**	+	+	+	+	+	+

(+) Indicate the presence of $CD4^+$ epitope; (−) Indicate the absence of $CD4^+$ epitope; Conserved $CD4^+$ epitope in all SARS-CoV-2 VOCs are shown in bold. Low combined score indicate more immunogenic epitope.

Table 3
CTLs immunogenic epitopes predicted from S protein of SARS-CoV-2 omicron and other VOC.

Allele	Peptide	Wuhan	Alpha	Beta	Gamma	Delta	Omicron
HLA-A*24:02	YYHKNNKSW	+	–	–	–	–	–
HLA-A*01:01, HLA-A*26:01, HLA-B*58:01	**WTAGAAAYY**	+	+	+	+	+	+
HLA-A*01:01	**LGAENSVAY**	+	+	-	+	+	+
HLA-A*01:01	**STECSNLLL**	+	+	+	+	+	+
HLA-A*02:01	**FVFLVLLPL**	+	+	+	+	+	+
HLA-A*02:01,HLA-A*26:01	**FTISVTTEI**	+	+	+	+	+	+
HLA-A*02:01	**LLFNKVTLA**	+	+	+	+	+	+
HLA-A*02:01	**VVFLHVTYV**	+	+	+	+	+	+
HLA-A*03:01	GVYFASTEK	+	+	+	+	+	–
HLA-A*03:01	**TLADAGFIK**	+	+	+	+	+	+
HLA-A*03:01	**VLKGVKLHY**	+	+	+	+	+	+
HLA-A*24:02	VYHKNNKSW		+	–	–	–	–
HLA-A*24:02	**TFEYVSQPF**	+	+	+	+	+	+
HLA-A*24:02	**PYRVVVLSF**	+	+	+	+	+	+
HLA-A*26:01, HLA-B*08:01	**ESNKKFLPF**	+	+	+	+	+	+
HLA-A*26:01, HLA-B*39:01	**WTFGAGAAL**	+	+	+	+	+	+
HLA-B*07:02	**LPFNDGVYF**	+	+	+	+	+	+
HLA-B*08:01	**KIYSKHTPI**	+	+	+	+	+	+
HLA-B*08:01	**FNATRFASV**	+	+	+	+	+	+
HLA-B*08:01	**FRKSNLKPF**	+	+	+	+	+	+
HLA-B*08:01	NSHRRARSV		+	–	–	–	–
HLA-B*27:05	**VRFPNITNL**	+	+	+	+	+	+
HLA-B*39:01	**YQPYRVVVL**	+	+	+	+	+	+
HLA-B*40:01	**FEYVSQPFL**	+	+	+	+	+	+
HLA-B*40:01	**AEIRASANL**	+	+	+	+	+	+
HLA-B*40:01	**KEIDRLNEV**	+	+	+	+	+	+
HLA-B*58:01	**TSVDCTMYI**	+	+	+	+	+	+
HLA-B*58:01	**FAMQMAYRF**	+	+	+	+	+	+
HLA-B*58:01	**YIKWPWYIW**	+	+	+	+	+	+
HLA-A*24:02, HLA-B*58:01	IAIPTNFTI	+	–	+	+	+	+
HLA-B*07:02	SPRRARSVA	+	–	+	+	–	–
HLA-B*07:02	IPTNFTISV	+	–	+	+	+	+
HLA-A*03:01	GTIADYNYK	–	–	–	+	–	–
HLA-A*01:01, HLA-A*03:01, HLA-B*58:01, HLA-B*15:01	RSYSFRPTY	–	–	–	–	–	+
HLA-A*03:01	GVYFASIEK	–	–	–	–	–	+
HLA-A*03:01	NLAPFFTFK	–	–	–	–	–	+
HLA-A*03:01	VLKGVKLHY	–	–	–	–	–	+
HLA-A*24:02	IYSKHTPII	–	–	–	–	–	+
HLA-A*24:02	YNLAPFFTF	–	–	–	–	–	+
HLA-B*08:01	QLSSKFGAI	–	–	–	–	–	+
HLA-B*15:01	SVLYNLAPF	–	–	–	–	–	+

(+) Indicate the presence of CTLs epitope; (−) Indicate the absence of CTLs epitope; Conserved CTLs epitope in all SARS-CoV-2 VOCs are shown in bold.

SARS-CoV-2 spike protein, as shown in Fig. 1.

3.6. Population coverage of a putative MEV candidate

We used the IEDB tool to analyze population coverage of MHC-I and MHC-II alleles based on epitopes. MHC-I and MHC-II epitopes obtained from all VOCs showed promising results from the population coverage analysis, as shown in Supplementary Fig. S6.

3.7. Physiochemical properties of a putative MEV candidate

Several physicochemical parameters were calculated using the ProtParam tool. The putative MEV candidate with molecular weight ~63.20 kDa and theoretical isoelectric (PI) ~ 9.71 shows an aliphatic index of 87.92, indicating the protein is thermostable. Moreover, the protein was classified as stable based on the instability index (II) value of 25.38. In addition, a putative MEV candidate estimated half-life in different organisms, *viz.,* >10 h in *E.coli*, >20 h in yeast, and 30 h in mammalian reticulocytes. The putative MEV candidate shows a suitable physicochemical landscape for a stable, safer MEV candidate.

3.8. Secondary and tertiary structure prediction of a putative MEV candidate

The putative MEV candidate secondary structure was predicted using the SOPMA secondary structure prediction method. The putative MEV candidate has an alpha helix of 49.83 %, an extended strand of 20.63 %, a beta-turn of 7.34 %, and a random coil of 22.20 %, as shown in Supplementary Figs. S7(A–B). Further, the putative MEV candidate solubility was predicted via a protein-sol server. The putative MEV candidate showed a scaled solubility value of 0.503) Supplementary Fig. S7(C). The putative MEV candidate raw tertiary and refined tertiary structure were shown in Supplementary Figs. S7(D–E). Furthermore, PSIPRED server 4.0 also showed a suitable secondary structure of a putative MEV candidate as shown in Supplementary Figs. S8(A–B).

3.9. Tertiary structure validation of a putative MEV candidate

The putative MEV candidate refined structure was validated using PROSA, ERRAT, and PROCHECK web servers. ERRAT server reveals an overall quality factor 86.28 % of a putative MEV candidate, as shown in Supplementary Fig. S9(A). The PROSA server output suggests putative MEV candidate have a good protein structure, having a z-score value of −7.17 Supplementary Figs. S9(B–C). The Ramachandran plot obtained from the PROCHECK server showed that 93 % of residues belonged to the most favoured region, 4.9 % in the additional allowed area, 1 % in the generously allowed region, as shown in Supplementary Fig. S9(D) and only 0.6 % lied in the disallowed region as given in Table 4.

3.10. Di-sulfide modifications in a putative MEV candidate

For disulfide modification, the DbD2 server showed 63 pairs of residues capable of making disulfide bonds. We have used only residues with less than 2.5 values for disulfide bond formation in a putative MEV candidate. Five disulfide bonds were formed, satisfying the above criteria, as shown in Supplementary Fig. S10.

3.11. Molecular docking of a putative MEV candidate

Molecular docking of a putative MEV candidate was carried out to evaluate the binding with Toll-like receptor-4 (TLR-4 PDB ID.3FXI), HLA molecules, namely HLA class I histocompatibility antigen, B-8 alpha chain (PDB ID. 6P2F), and HLA class II histocompatibility antigen, DR alpha chain (PDB ID. 3C5J). With a putative MEV candidate, the ClusPro server predicted 29 models each for TLR-4, HLA class I histocompatibility antigen, and HLA class II histocompatibility antigen. Minimum binding scores of −1337.9, −1265.3, and −1330.7 were achieved in the case of TLR-4, HLA class I histocompatibility antigen, B-8 alpha chain, and HLA class II histocompatibility antigen respectively. Furthermore, the binding of a putative MEV candidate among different antigenic receptors

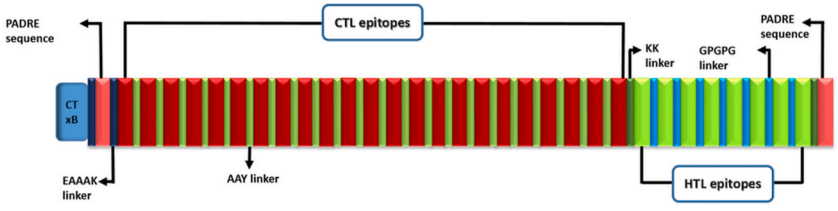


Fig. 1. Schematic representation of the developed multi-epitope vaccine (MEV) candidate. The N-terminal region consists of CTxB, which is linked to the PADRE sequence using an EAAAK linker. Moreover, two different types of linkers, AAY and GPGPG, were used to join various CTLs and $CD4^+$ immunogenic epitopes. Additionally, a KK linker was used to join CTLs and $CD4^+$ immunogenic epitopes. The C-terminal region has a PADRE sequence which is linked to $CD4^+$ immunogenic epitopes via a KK linker.

Table 4
Ramachandran plot illustrating the amino acid residue distribution of a vaccine candidate.

Amino acid positions	Residue number	Percentage %	Total
In most favoured region	477	93.5	477
Addition allowed region	25	4.9	25
Generously allowed region	5	1	5
Disallowed region	3	0.6	3
End-residues	NA		2
Glycine residues			33
Proline residues			27
Total residues			572

NA: Not Available.

was validated using the HDOCK server. It provided the best models for TLR-4, HLA class I histocompatibility antigen, and HLA class II histocompatibility antigen, having minimum binding energy of −386.28, −329.64, and −286.44, respectively Fig. 2(A–F).

3.12. Molecular dynamics simulation analysis of a putative MEV candidate

The RMSD value obtained from the WEBGRO server ranges from 0.4 to 1.25 nm on 50 ns runs and has stable confirmation. The stability of the complex is higher due to the presence of many stable bonds in HLA class II histocompatibility antigen and a putative MEV candidate. The molecular dynamic simulation of HLA class II histocompatibility antigen, and putative MEV candidate complex RMSD, RMS, is shown in Supplementary Figs. S11(A–B).

3.13. Normal mode analysis (NMA) of a putative MEV candidate

Normal mode analysis (NMA) is a powerful computational method. A variety of new algorithms have been developed in recent years, especially the coarse-grained NMA, which enables simulations of supramolecular complexes at unprecedented levels of resolution and length scales [53]. The normal mode analysis of HLA class II histocompatibility antigen with putative MEV candidate revealed various properties like deformability, B-factor, elastic network, Eigenvalue, variance, and covariance map. The Eigenvalue was 9.593398e-05, which indicated the motion stiffness of molecules. The complex elastic map, as shown in Fig. 3(A–D), represents the relationship between the atoms of big molecules. Moreover, normal mode analysis performed for the HLA class I histocompatibility antigen, and TLR-4 molecules showed a similar pattern with TLR-4 and MEV candidate Supplementary Figure S12(A-D)-13(A-D).

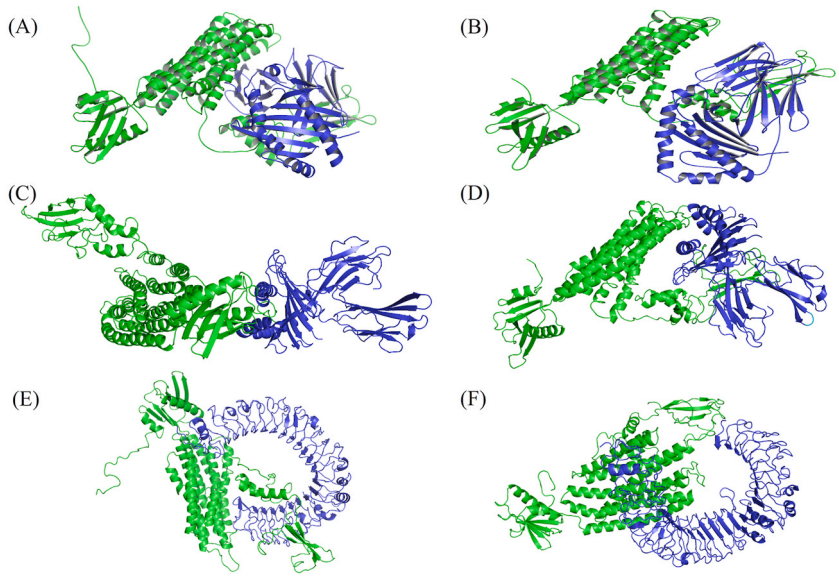


Fig. 2. Docked poses of a putative MEV candidate with different immunological receptors molecules. (A) Docked complex of putative MEV candidate (green colour) with HLA class I histocompatibility antigen, B-8 alpha chain (blue colour) using ClusPro server. (B) Docked complex of putative MEV candidate (green colour) with HLA class I histocompatibility antigen, B-8 alpha chain (blue colour) using HDOCK server (C) Docked complex of putative MEV candidate (green colour) with HLA class II histocompatibility antigen, DR alpha chain (blue colour) using ClusPro server (D) Docked complex of putative MEV candidate (green colour) with HLA class II histocompatibility antigen, DR alpha chain (blue colour) using HDOCK server (E) Docked complex of putative MEV candidate (green colour) with TLR-4 (Blue colour) using ClusPro server (F) Docked complex of putative MEV candidate (green colour) with TLR-4 (Blue colour) using HDOCK server.

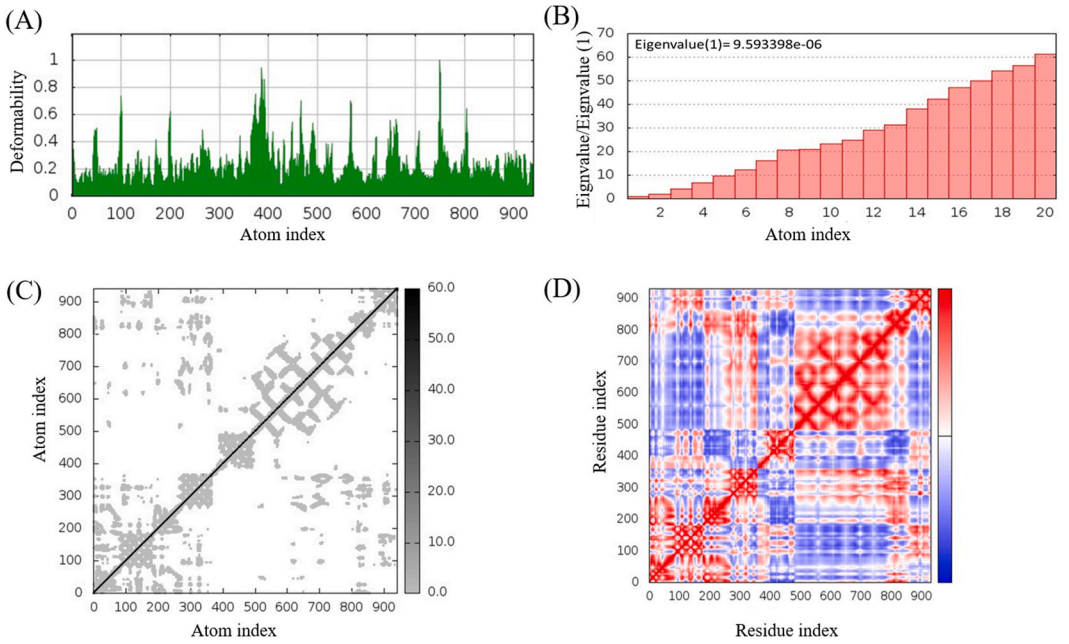


Fig. 3. Normal mode analysis of docked complex of putative MEV candidate with HLA class II histocompatibility antigen, DR alpha chain to show their stability. (A) Deformability plot of putative MEV candidate with HLA class II histocompatibility antigen, DR alpha chain docked complex reveals that maximum deformability in atom index ~750. (B) The putative MEV candidate with HLA class II histocompatibility antigen, DR alpha chain docked complex associated to each normal mode represents the motion stiffness. The higher Eignvalue of docked complex indicate large amount of energy required to deform structure. (C) The elasticity network model of putative MEV candidate with HLA class II histocompatibility antigen, DR alpha chain docked complex represent the connections between pairs of atoms using springs. Each node in the graph represents a spring connecting two atoms, with the node's color indicating the spring's stiffness. Darker nodes represent stiffer springs, while lighter nodes indicate softer springs. (D) The covariance matrix of putative MEV candidate with HLA class II histocompatibility antigen, DR alpha chain docked complex demonstrates the interdependence between pairs of residues, indicating whether they exhibit correlated (red), uncorrelated (white), or anti-correlated (blue) anti-correlated (blue colour) in motions.

3.14. Immune simulation profiling of putative MEV candidate

C-IMMSIM web server was utilized to evaluate the ability of a putative MEV candidate to induce various types of immune cells, namely T-cells, B-cells, Natural killer cells, etc. After the putative MEV candidate administration, antibody concentration rapidly increased, thereby indicating an immune response. Additionally, the B-cell population was expanded and showed memory B-cells after administration. Similarly, the number of Cytotoxic T-lymphocytes elevated after 8 days of putative MEV candidate administration and declined after ~20 days, as shown in Fig. 4(A–L).

Consequently, the TH-cell population surged after the putative MEV candidate administration. Moreover, the number of TH-cells remains at ~1000 after 350 days of administration. In correspondence with T-cells and B-cells, the number of natural killer cells and dendritic cells also increased after putative MEV candidate administration. Apart from various immune cells, cytokines like IFN-g, TGF-b, and IL-10 were also elevated. The level of the IFN-g showed a high concentration of about ~1×10^6 ng/ml after the first dose of a MEV candidate.

3.15. In silico cloning of MEV candidate in pET28a (+) vector

In the course of codon adaptation, the Java Codon Adaptation Tool server was used, and the protein sequence was reverse transcribed into nucleotide sequences and optimized for cloning and expression in *E. coli* (K12 strain), as shown in Supplementary Fig. S14. Consequently, optimized sequences showed CAI-value ~0.912 and GC% ~51.22 %.

3.16. Conservancy of MEV epitopes in SARS-CoV-2 VOCs and omicron sublineages

In designing our putative MEV candidate, we utilized 23 CTL epitopes, which were conserved across the Wuhan, alpha, beta, gamma, delta, and omicron variants as provided in Supplementary Fig. S15. These CTL epitopes are 100 % conserved in the aforementioned reference sequences of VOCs. In case of CTLs epitopes, 207 amino acid residues are predicted as a part of the epitopes; however, 04 amino acid residues were changed due to point mutations. Additionally, we evaluated the conservancy of these CTL epitopes within the sublineages of the Omicron variant, finding that the overall conservancy rate is approximately 93.66 %. Furthermore, our putative MEV design incorporates 7 HTL epitopes, with 2 epitopes being conserved across the Wuhan, alpha, beta,

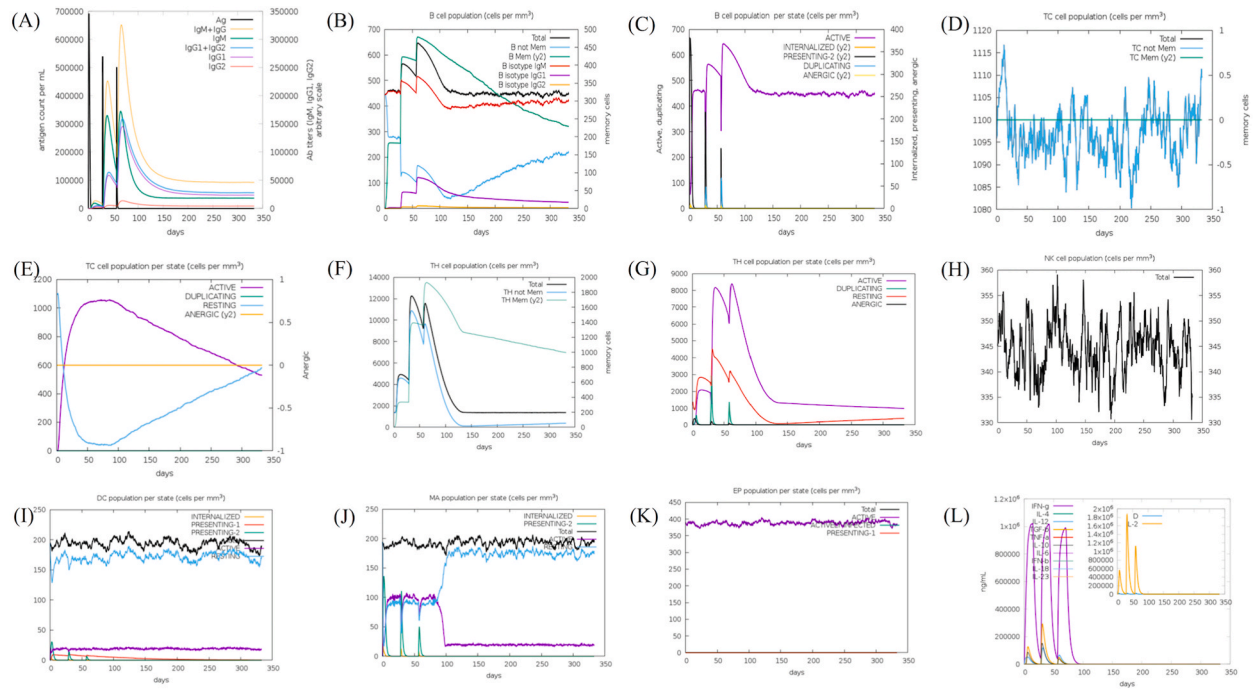


Fig. 4. In-silico immune simulation of putative MEV candidate using C-IMMSIM server using 3 injection at different time interval (0^{th}, 30th and 60th day). (A) initially on 0^{th} day MEV candidate (~700000 antigen count/mL) the level of immunoglobulin (IgM, IgG1, IgG2, IgM + IgG, and IgG1 + IgG2) Ab titers were low at arbitrary scale, later on 30th and 60th days the level of immunoglobulin (IgM, IgG1, IgG2 IgM + IgG, and IgG1 + IgG2) Ab titers increases significantly about ~200000 and ~320000 respectively at arbitrary scale. (B) In case of B lymphocyte population, the number of B-Memory cells (y2) increases after the each injection while the number of B-not Memory cells decline. Additionally the number of B-isotype IgM cells population increases. (C) In total population of different types of B-cell (cells per mm3); the number of active and duplicating B-cells population increases after each injection. (D) The population of TC lymphocyte count increases ~1115 after the 0th day injection and comes to normal level (E) The number of the active TC-cells increases after the 0^{th} day injection and remained high during the 30th and 60th day injection while the number of resting TC-cells declined after 0^{th}, 30th and 60th days injection. (F) The population of T-helper (TH) lymphocyte after 0^{th} day injection the number of TH not Mem increases more compare to the TH mem (Y2) cell type while after 30th and 60th days injection the number of TH mem (Y2) remain high in compare to TH not Mem cell type. (G) Total number of TH cells: after 0th day injection the number of active and duplicating TH cells increases while the number of resting TH cells declined after 30th day injection. (H) Total population of Natural killer cells (cells per mm^3), after 0^{th}, 30th and 60th day injection the population of Natural killer cells increases (I) Dc cells population per state mm^3 (internalized, presentating-1, presentating-2, active, resting). (J) MA population (cells per stage mm^3) after the each injection the population of active and resting MA cells population increases. (K) In case of EP population (cells per state mm3) there is no much variation was observed in active EP population. (L) Concentrations of cytokines and interleukins after three injections (0^{th}, 30th and 60th day); After the each injection the IFN-G concentration ~1×10^6 ng/mL while the TGF-b concentration about 2×10^5 ng/mL. Additionally, the concentration of the concentration of IL-2 ~1.8×10^6 after 30th day injection.

gamma, delta, and omicron, as well as omicron sublineages. Subsequently, 126 amino acid residues are predicted as a part of the epitopes. However, 08 residues were alerted due to point mutation and deletion. In addition, the total percentage of amino acid residues conservancy in putative MEV candidate epitopes is about 97 %.

3.17. Comparison of the putative MEV candidate with other candidate

Our proposed putative MEV candidate shows a suitable landscape to provide protection against the omicron and other SARS-CoV-2 VOCs. Further, we compare the various properties of our designed putative MEV candidate with the existing studies, as shown in Supplementary Table S10.

4. Discussion

With the increased number of SARS-CoV-2 variants, there is a need to identify conserved epitopes against existing and emerging SARS-CoV-2 variants. Using bioinformatics tools to design peptide-based vaccines has become increasingly important due to the advancement of computational-aided sequence-based technology. A potential pipeline of viral targeting tools has already been demonstrated with peptide-based vaccines, such as those developed for dengue, chikungunya, and rhinovirus [54]. Several mutations in the genome of SARS-CoV-2 result in multiple resistance against various vaccines. Therefore, this study focuses on designing a MEV

candidate that contains conserved epitopes in different SARS-CoV-2 VOCs.

In this study, 12 MHC-I binders were predicted from S-protein. From 12 MHC binders, we found 2 binders, namely AAYYVGYLQPRTF and SVTTEILPVSMTK, have previously been experimentally reported in IEDB [55]. Moreover, the FSTFKCYGVSPTK and SYQTQTKSHRRAR binders were identified through computational analysis [56,57]. Additionally, the remaining eight binders were not previously reported and are considered novel binders. In the case of M protein, RLFARTRSMWSFN binder was experimentally validated in IEDB. Furthermore, the remaining nine binders have not been reported prior to this study and are therefore identified as novel.

In the case of $CD8^+$ immunogenic epitopes, we found 12 immunogenic epitopes from S-protein that included eight novel immunogenic epitopes, and the remaining four epitopes were reported in various studies. SVTTEILPVSMTK and AAYYVGYLQPRTF immunogenic epitopes were reported experimentally in IEDB [55] while two epitopes, FSTFKCYGVSPTK and SYQTQTKSHRRAR computationally reported [56,57]. Additionally, eight binders had not been reported before and are consider as novel. The details of $CD8^+$ immunogenic epitopes from E and M proteins are in Supplementary Table S3.

For MHC-II, 41 binders were anticipated from S-protein, from which eight were novel, and 33 were reported in IEDB [58]. We found five of the eight novel binders exclusively in the delta variant. The FLVLLPLVSSQCVNF binder was present in two VOCs, while the IGINITRFQTLHRSY binder was detected solely in the beta VOC. Moreover, the QYGSFCTQLKRALTG novel binder was uniquely present in the omicron VOC. Similarly, 12 new binders for the M-protein were identified, among which the LVIGFLFLTWICLLQ binder was found in all VOCs. The EQWNLVIGFLFLTWI binder was present in five VOCs. Conversely, the ITGGIATAMACLVGL binder was exclusively detected in the delta variant. Furthermore, Nine MHC-II binding epitopes were novel and exclusively found in the Omicron variant. The binders for M-proteins are listed in Supplementary Table S4.

In the case of $CD4^+$ immunogenic epitopes, a total of 8 immunogenic epitopes were predicted from S-protein that include three novel immunogenic epitopes viz. IGINITRFQTLHRSY, and GNYNYRYRLFRKSNL and FLVLLPLVSSQCVNF. However, five $CD4^+$ immunogenic epitopes were experimentally reported in IEDB. Additionally, we identified nine epitopes within the E-protein, which have been reported in multiple studies [59–62]. In case of M protein, 12 $CD4^+$ immunogenic epitopes were predicted out of which, seven epitopes were reported. Five novel epitopes TCFVLAAVYRINWIT, TLTCFVLAAVYRINW, LTCFVLAAVYRINWI, PVTLTCFVLAAVYRI, and IKLIFLWLLWPVTLT were present only in omicron.

Moreover, 41 CTL epitopes were predicted from S-protein, out of which eight epitopes were novel, and the remaining were reported in the literature and IEDB [63–65]. From the 8 epitopes, two novel epitopes were common in all VOCs and Wuhan strain: three epitopes in omicron, 1 epitope in Alpha, one epitope in Wuhan and alpha, one epitope in gamma. The RSYSFRPTY epitope was found in only omicron due to the substitution of Q494R, G497S, and Q499R. Moreover, the GVYFASIEK peptide epitope was identified due to the mutation of T96I. In addition, six and 14 CTL epitopes were predicted from the E and M protein respectively.in case of M-protein novel LWPVTLTCF epitope present in omicron VOC. Remaining all the CTLs epitopes were experimentally validated. In addition, the details of E and M protein CTL epitopes were provided in Supplementary Table S6.

In addition, S-protein CTL epitopes and HTL immunogenic epitopes were used to design the putative MEV candidate against all VOCs and the Wuhan strain. Our designed putative MEV candidate included common CTL epitopes (23) and 7 HTL ($CD4^+$) immunogenic epitopes derived from all VOCs as well as the Wuhan strain along with various adjuvants (CTxB, PADRE sequence) and peptide linkers. Moreover, EAAAK linkers were used to join the CTxB and the PADRE sequence to the N-terminal, followed by an epitopes. The PADRE sequence also plays a significant role in boosting CTL reaction against various antigens [34]. Furthermore, multiple linkers like AAY, GPGPG, and KK were used to fuse epitopes.

ExPasy ProtParam server output reveals that the putative MEV candidate was stable according to the aliphatic index of 87.92; since the aliphatic index was higher, this implies greater thermostability of the MEV candidate [66]. In addition, the instability index of a putative MEV candidate is 25.38 based on an instability index of less than 40; it represents the stable character of the MEV candidate [66]. Moreover, the Prot-Proam result indicates that the design putative MEV candidate physiochemical properties are better than the previously reported studies [67,68]. Protein-sol reveals that the putative MEV candidate is water-soluble, comparable to previously reported similar results [34].

The SOPMA and PSIPRED reveal that putative MEV candidate form a good-quality secondary structure. Additionally, our proposed putative MEV candidate shows a better ERRAT score (86.28) and Z-score (−7.17) than previously reported studies [34,67,69–71]. Furthermore, the Ramachandran plot showed a preferable structure. The molecular docking analysis reveals that the putative MEV candidate effectively binds with three molecules: Toll-like receptor-4, HLA class I histocompatibility antigen, B-8 alpha chain, and HLA class II histocompatibility antigen, DR alpha chain. TLR-4 significantly contributes to the induction of pro-inflammatory response, and both MHC-I and MHC-II help in antigen presentation for immune cells [72]. The ClusPro model showed the lowest binding score of −1337.9, −1265.3, and 1330.7 for TLR-4, HLA class I histocompatibility antigen, B-8 alpha chain, and HLA class II histocompatibility antigen, DR alpha chain, respectively. The HDOCK server also provided models with the lowest docking scores, −386.28, −329.64, and 286.44 for TLR-4, HLA class I histocompatibility antigen, B-8 alpha chain, and HLA class II histocompatibility antigen, DR alpha chain. The results of clusPro (−1337.9) and HDOCK (−312.89) for TLR-4 were better than the previously reported lowest binding score [34,71,73].

The NMA provides properties like Eigenvalue, deformability, B-factor, covariance, and elastic network. In each NMA, an eigenvalue is assigned to represent the stiffness of the motion. Eigenvalues are directly related to the energy needed to deform a structure. Deformation becomes easier with a lower eigenvalue. Moreover, molecular dynamics and normal mode analysis revealed a stable docking complex.

Interestingly, our putative MEV candidate can induce an immune response against all the SARS-CoV-2 VOCs. Additionally, the putative MEV candidate induced a significant secondary response compared to the primary response. Even after 300 days, the

population of memory T-cells and B-cells remains ~1000 and ~200 (cells per mm3), respectively. The C-IMM-SIM results indicated the immune response of the putative MEV candidate and compared this result with the previously reported study [34]. Furthermore, we have checked the putative MEV candidate epitopes amino acid residue conservancy in omicron variant sublineages. The overall conservancy of amino acid residues in a MEV candidate is about 97 % in omicron sublineages. The major advantage of the study is that we have predicted binders/epitopes from spike, membrane and envelope proteins and analyze their conservancy among SARS-CoV-2 VOCs. Our predicted novel binders/epitopes increase the repertoire against all VOCs of SARS-CoV-2. Furthermore, 23 conserved CTL epitopes and 7 HTL ($CD4^+$) epitopes derived from spike protein were used to design a MEV candidate that might be capable of inducing innate and adaptive immune responses. The advantage of our designed putative MEV candidate containing conserved epitopes is that they may protect against SARS-CoV-2 omicron and other emerging VOCs to mitigate the COVID-19 pandemic.

5. Conclusion

In this study, we have predicted various peptides/epitopes from various proteins of SARS-CoV-2 VOCs. We compared our predicted binders/epitopes in the Immune Epitope Database (IEDB) and literature. We found experimental evidences of about 79 out of 208 epitopes/binders in IEDB. This demonstrate the utility of our approach for predicting epitopes/binders. Furthermore, in this study there are 71 novel epitopes/binders are identified. These include novel 17 MHC-I, 20 MHC-II, 17 $CD8^+$, 8 $CD4^+$, and 9 CTLs epitopes/binders. In addition, we have also developed a putative MEV candidate against recently emerged VOCs and the Wuhan strain. Our designed putative MEV candidate has antigenic, non-allergic, and relevant physicochemical properties that fulfill vaccine requirements. Additionally, the validation of its structural integrity reveals a stable structure. Moreover, molecular docking and molecular dynamics analysis reveal that putative MEV candidate shows binding affinity and stable interactions with various immunogenic receptors. The immune simulation analysis has shown that putative MEV candidate can elicit a good immune response. Subsequently, the overall conservancy of amino acid residues in a MEV candidate is about 97 % in sublineages. However, MEV candidate must be validated through in vitro and in vivo tests on animals and human clinical trials to ensure its efficacy. A conserved putative MEV candidate may effectively combat present VOCs and future emerging strains of SARS-CoV-2.

Funding

This work was supported by the Council of Scientific and Industrial Research (CSIR), Government of India (STS-038 and OLP0192).

CRediT authorship contribution statement

Amber Rastogi: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Sakshi Gautam:** Methodology, Formal analysis. **Manoj Kumar:** Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, 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.

Acknowledgments

We wish to acknowledge the efforts of GISAID (http://www.gisaid.org/) in sharing hCoV-19 open data. All authors from originating and submitting labs are also to be acknowledged for their contribution to these genetic sequence data. Mr. Amber Rastogi and Ms. Sakshi Gautam are thankful to the Council of Scientific and Industrial Research (CSIR) for providing fellowships.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e35129.

References

[1] A.K. Gupta, M.S. Khan, S. Choudhury, A. Mukhopadhyay, Sakshi, A. Rastogi, A. Thakur, P. Kumari, M. Kaur, Shalu, C. Saini, V. Sapehia, Barkha, P.K. Patel, K. T. Bhamare, M. Kumar, CoronaVR: a computational resource and analysis of epitopes and therapeutics for severe acute respiratory syndrome coronavirus-2, Front. Microbiol. 11 (2020), https://doi.org/10.3389/FMICB.2020.01858.

[2] R.A. Bender Ignacio, A.E. Shapiro, R.M. Nance, B.M. Whitney, J.A.C. Delaney, L. Bamford, D. Wooten, M.Y. Karris, W.C. Mathews, H.N. Kim, J. Keruly, G. Burkholder, S. Napravnik, K.H. Mayer, J. Jacobson, M. Saag, R.D. Moore, J.J. Eron, A.L. Willig, K.A. Christopoulos, J. Martin, P.W. Hunt, H.M. Crane, M. M. Kitahata, E.R. Cachay, Racial and ethnic disparities in coronavirus disease 2019 disease incidence independent of comorbidities, among people with HIV in the United States, AIDS 36 (2022) 1095–1103, https://doi.org/10.1097/QAD.0000000000003223.

[3] Tracking SARS-CoV-2 variants, (n.d.). https://www.who.int/activities/tracking-SARS-CoV-2-variants (accessed April 26, 2023).
[4] W.T. Harvey, A.M. Carabelli, B. Jackson, R.K. Gupta, E.C. Thomson, E.M. Harrison, C. Ludden, R. Reeve, A. Rambaut, S.J. Peacock, D.L. Robertson, SARS-CoV-2 variants, spike mutations and immune escape, Nat. Rev. Microbiol. 19 (2021) 409–424, https://doi.org/10.1038/S41579-021-00573-0.
[5] T. Kirby, New variant of SARS-CoV-2 in UK causes surge of COVID-19, Lancet Respir. Med. 9 (2021) e20–e21, https://doi.org/10.1016/S2213-2600(21)00005-9.
[6] J. Lopez Bernal, N. Andrews, C. Gower, E. Gallagher, R. Simmons, S. Thelwall, J. Stowe, E. Tessier, N. Groves, G. Dabrera, R. Myers, C.N.J. Campbell, G. Amirthalingam, M. Edmunds, M. Zambon, K.E. Brown, S. Hopkins, M. Chand, M. Ramsay, Effectiveness of Covid-19 vaccines against the B.1.617.2 (delta) variant, N. Engl. J. Med. 385 (2021) 585–594, https://doi.org/10.1056/NEJMOA2108891.
[7] Y. Cao, J. Wang, F. Jian, T. Xiao, W. Song, A. Yisimayi, W. Huang, Q. Li, P. Wang, R. An, J. Wang, Y. Wang, X. Niu, S. Yang, H. Liang, H. Sun, T. Li, Y. Yu, Q. Cui, S. Liu, X. Yang, S. Du, Z. Zhang, X. Hao, F. Shao, R. Jin, X. Wang, J. Xiao, Y. Wang, X.S. Xie, Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies, Nature 602 (2022) 657–663, https://doi.org/10.1038/S41586-021-04385-3.
[8] A.J. Greaney, A.N. Loes, K.H.D. Crawford, T.N. Starr, K.D. Malone, H.Y. Chu, J.D. Bloom, Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies, Cell Host Microbe 29 (2021) 463–476.e6, https://doi.org/10.1016/J.CHOM.2021.02.003.
[9] Z. Sun, T. Wu, H. Xie, Y. Li, J. Zhang, X. Su, H. Qi, The role of cellular immunity in the protective efficacy of the SARS-CoV-2 vaccines, Vaccines 10 (2022), https://doi.org/10.3390/VACCINES10071103.
[10] J. Ai, H. Zhang, Y. Zhang, K. Lin, Y. Zhang, J. Wu, Y. Wan, Y. Huang, J. Song, Z. Fu, H. Wang, J. Guo, N. Jiang, M. Fan, Y. Zhou, Y. Zhao, Q. Zhang, Q. Liu, J. Lv, P. Li, C. Qiu, W. Zhang, Omicron variant showed lower neutralizing sensitivity than other SARS-CoV-2 variants to immune sera elicited by vaccines after boost, Emerg. Microb. Infect. 11 (2022) 337–343, https://doi.org/10.1080/22221751.2021.2022440.
[11] W.H. Self, M.W. Tenforde, J.P. Rhoads, M. Gaglani, A.A. Ginde, D.J. Douin, S.M. Olson, H.K. Talbot, J.D. Casey, N.M. Mohr, A. Zepeski, T. McNeal, S. Ghamande, K.W. Gibbs, D.C. Files, D.N. Hager, A. Shehu, M.E. Prekker, H.L. Erickson, M.N. Gong, A. Mohamed, D.J. Henning, J.S. Steingrub, I.D. Peltan, S. M. Brown, E.T. Martin, A.S. Monto, A. Khan, C.L. Hough, L.W. Busse, C.C. ten Lohuis, A. Duggal, J.G. Wilson, A.J. Gordon, N. Qadir, S.Y. Chang, C. Mallow, C. Rivas, H.M. Babcock, J.H. Kwon, M.C. Exline, N. Halasa, J.D. Chappell, A.S. Lauring, C.G. Grijalva, T.W. Rice, I.D. Jones, W.B. Stubblefield, A. Baughman, K. N. Womack, C.J. Lindsell, K.W. Hart, Y. Zhu, L. Mills, S.N. Lester, M.M. Stumpf, E.A. Naioti, M. Kobayashi, J.R. Verani, N.J. Thornburg, M.M. Patel, N. Calhoun, K. Murthy, J. Herrick, A. McKillop, E. Hoffman, M. Zayed, M. Smith, N. Seattle, J. Ettlinger, E. Priest, J. Thomas, A. Arroliga, M. Beeram, R. Kindle, L.-A. Kozikowski, L. De Souza, S. Ouellette, S. Thornton-Thompson, O. Mehkri, K. Ashok, S. Gole, A. King, B. Poynter, N. Stanley, A. Hendrickson, E. Maruggi, T. Scharber, J. Jorgensen, R. Bowers, J. King, V. Aston, B. Armbruster, R.E. Rothman, R. Nair, J.-T.T. Chen, S. Karow, E. Robart, P.N. Maldonado, M. Khan, P. So, J. Levitt, C. Perez, A. Visweswaran, J. Roque, A. Rivera, L. Angeles, T. Frankel, L. Angeles, J. Goff, D. Huynh, M. Howell, J. Friedel, M. Tozier, C. Driver, M. Carricato, A. Foster, P. Nassar, L. Stout, Z. Sibenaller, A. Walter, J. Mares, L. Olson, B. Clinansmith, C. Rivas, H. Gershengorn, E. McSpadden, R. Truscon, A. Kaniclides, L. Thomas, R. Bielak, W.D. Valvano, R. Fong, W.J. Fitzsimmons, C. Blair, A.L. Valesano, J. Gilbert, C.D. Crider, K.A. Steinbock, T.C. Paulson, L. A. Anderson, C. Kampe, J. Johnson, R. McHenry, M. Blair, D. Conway, M. LaRose, L. Landreth, M. Hicks, L. Parks, J. Bongu, D. McDonald, C. Cass, S. Seiler, D. Park, T. Hink, M. Wallace, C.-A. Burnham, O.G. Arter, Comparative effectiveness of Moderna, pfizer-BioNTech, and janssen (johnson & johnson) vaccines in preventing COVID-19 hospitalizations among adults without immunocompromising conditions - United States, March-August 2021, MMWR Morb. Mortal. Wkly. Rep. 70 (2021) 1337–1343, https://doi.org/10.15585/MMWR.MM7038E1.
[12] J. Lopez Bernal, N. Andrews, C. Gower, E. Gallagher, R. Simmons, S. Thelwall, J. Stowe, E. Tessier, N. Groves, G. Dabrera, R. Myers, C.N.J. Campbell, G. Amirthalingam, M. Edmunds, M. Zambon, K.E. Brown, S. Hopkins, M. Chand, M. Ramsay, Effectiveness of Covid-19 vaccines against the B.1.617.2 (delta) variant, N. Engl. J. Med. 385 (2021) 585–594, https://doi.org/10.1056/NEJMOA2108891.
[13] W. Zhou, W. Wang, Fast-spreading SARS-CoV-2 variants: challenges to and new design strategies of COVID-19 vaccines, Signal Transduct. Targeted Ther. 6 (2021), https://doi.org/10.1038/S41392-021-00644-X.
[14] P. Wang, J. Sidney, Y. Kim, A. Sette, O. Lund, M. Nielsen, B. Peters, Peptide binding predictions for HLA DR, DP and DQ molecules, BMC Bioinf. 11 (2010), https://doi.org/10.1186/1471-2105-11-568.
[15] M. Moutaftsi, B. Peters, V. Pasquetto, D.C. Tscharke, J. Sidney, H.H. Bui, H. Grey, A. Sette, A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus, Nat. Biotechnol. 24 (2006) 817–819, https://doi.org/10.1038/NBT1215.
[16] A.K. Gupta, M.S. Khan, S. Choudhury, A. Mukhopadhyay, Sakshi, A. Rastogi, A. Thakur, P. Kumari, M. Kaur, Shalu, C. Saini, V. Sapehia, Barkha, P.K. Patel, K. T. Bhamare, M. Kumar, CoronaVR: a computational resource and analysis of epitopes and therapeutics for severe acute respiratory syndrome coronavirus-2, Front. Microbiol. 11 (2020), https://doi.org/10.3389/FMICB.2020.01858.
[17] J.J.A. Calis, M. Maybeno, J.A. Greenbaum, D. Weiskopf, A.D. De Silva, A. Sette, C. Keşmir, B. Peters, Properties of MHC class I presented peptides that enhance immunogenicity, PLoS Comput. Biol. 9 (2013), https://doi.org/10.1371/JOURNAL.PCBI.1003266.
[18] S. Paul, C.S. Lindestam Arlehamn, T.J. Scriba, M.B.C. Dillon, C. Oseroff, D. Hinz, D.M. McKinney, S. Carrasco Pro, J. Sidney, B. Peters, A. Sette, Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes, J. Immunol. Methods 422 (2015) 28–34, https://doi.org/10.1016/J.JIM.2015.03.022.
[19] S.K. Dhanda, E. Karosiene, L. Edwards, A. Grifoni, S. Paul, M. Andreatta, D. Weiskopf, J. Sidney, M. Nielsen, B. Peters, A. Sette, Predicting HLA CD4 immunogenicity in human populations, Front. Immunol. 9 (2018), https://doi.org/10.3389/FIMMU.2018.01369.
[20] S.K. Dhanda, P. Vir, G.P.S. Raghava, Designing of interferon-gamma inducing MHC class-II binders, Biol. Direct 8 (2013), https://doi.org/10.1186/1745-6150-8-30.
[21] T. Stranzl, M.V. Larsen, C. Lundegaard, M. Nielsen, NetCTLpan: pan-specific MHC class I pathway epitope predictions, Immunogenetics 62 (2010) 357–368, https://doi.org/10.1007/S00251-010-0441-4.
[22] M.C. Jespersen, B. Peters, M. Nielsen, P. Marcatili, BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes, Nucleic Acids Res. 45 (2017) W24–W29, https://doi.org/10.1093/NAR/GKX346.
[23] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinf. 8 (2007), https://doi.org/10.1186/1471-2105-8-4.
[24] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, In silico approach for predicting toxicity of peptides and proteins, PLoS One 8 (2013), https://doi.org/10.1371/JOURNAL.PONE.0073957.
[25] D.A. Chesler, C.S. Reiss, The role of IFN-γ in immune responses to viral infections of the central nervous system, Cytokine Growth Factor Rev. 13 (2002) 441–454, https://doi.org/10.1016/S1359-6101(02)00044-8.
[26] I. Dimitrov, I. Bangov, D.R. Flower, I. Doytchinova, AllerTOP v.2–a server for in silico prediction of allergens, J. Mol. Model. 20 (2014), https://doi.org/10.1007/S00894-014-2278-5.
[27] J. Sidney, A. Steen, C. Moore, S. Ngo, J. Chung, B. Peters, A. Sette, Divergent motifs but overlapping binding repertoires of six HLA-DQ molecules frequently expressed in the worldwide human population, J. Immunol. 185 (2010) 4189–4198, https://doi.org/10.4049/JIMMUNOL.1001006.
[28] H.H. Bui, J. Sidney, K. Dinh, S. Southwood, M.J. Newman, A. Sette, Predicting population coverage of T-cell epitope-based diagnostics and vaccines, BMC Bioinf. 7 (2006), https://doi.org/10.1186/1471-2105-7-153.
[29] M. Kavoosi, A.L. Creagh, D.G. Kilburn, C.A. Haynes, Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli, Biotechnol. Bioeng. 98 (2007) 599–610, https://doi.org/10.1002/BIT.21396.
[30] M. Saadi, A. Karkhah, H.R. Nouri, Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches, Infect. Genet. Evol. 51 (2017) 227–234, https://doi.org/10.1016/J.MEEGID.2017.04.009.
[31] M. Kavoosi, A.L. Creagh, D.G. Kilburn, C.A. Haynes, Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli, Biotechnol. Bioeng. 98 (2007) 599–610, https://doi.org/10.1002/BIT.21396.
[32] B. Livingston, C. Crimi, M. Newman, Y. Higashimoto, E. Appella, J. Sidney, A. Sette, A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes, J. Immunol. 168 (2002) 5499–5506, https://doi.org/10.4049/JIMMUNOL.168.11.5499.

[33] Y. Gu, X. Sun, B. Li, J. Huang, B. Zhan, X. Zhu, Vaccination with a paramyosin-based multi-epitope vaccine elicits significant protective immunity against Trichinella spiralis infection in mice, Front. Microbiol. 8 (2017), https://doi.org/10.3389/FMICB.2017.01475.

[34] M. Bhattacharya, A.R. Sharma, P. Ghosh, S.S. Lee, C. Chakraborty, A next-generation vaccine candidate using alternative epitopes to protect against wuhan and all significant mutant variants of SARS-CoV-2: an immunoinformatics approach, Aging Dis 12 (2021) 2173–2195, https://doi.org/10.14336/AD.2021.0518.

[35] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Protein identification and analysis tools in the ExPASy server, Methods Mol. Biol. 112 (1999) 531–552, https://doi.org/10.1385/1-59259-584-7:531.

[36] C. Geourjon, G. Deléage, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, Comput. Appl. Biosci. 11 (1995) 681–684, https://doi.org/10.1093/BIOINFORMATICS/11.6.681.

[37] 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.

[38] M. Hebditch, M.A. Carballo-Amador, S. Charonis, R. Curtis, J. Warwicker, Protein-Sol: a web tool for predicting protein solubility from sequence, Bioinformatics 33 (2017) 3098–3100, https://doi.org/10.1093/BIOINFORMATICS/BTX345.

[39] J. Peng, J. Xu, RaptorX: exploiting structure information for protein alignment by statistical inference, Proteins 79 (Suppl 10) (2011) 161–171, https://doi.org/10.1002/PROT.23175.

[40] J. Ko, H. Park, L. Heo, C. Seok, GalaxyWEB server for protein structure prediction and refinement, Nucleic Acids Res. 40 (2012), https://doi.org/10.1093/NAR/GKS493.

[41] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007), https://doi.org/10.1093/NAR/GKM290.

[42] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519, https://doi.org/10.1002/PRO.5560020916.

[43] R.A. Laskowski, J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, J. Biomol. NMR 8 (1996) 477–486, https://doi.org/10.1007/BF00228148.

[44] D.B. Craig, A.A. Dombkowski, Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins, BMC Bioinf. 14 (2013), https://doi.org/10.1186/1471-2105-14-346.

[45] S. Vajda, C. Yueh, D. Beglov, T. Bohnuud, S.E. Mottarella, B. Xia, D.R. Hall, D. Kozakov, New additions to the ClusPro server motivated by CAPRI, Proteins 85 (2017) 435–444, https://doi.org/10.1002/PROT.25219.

[46] Y. Yan, D. Zhang, P. Zhou, B. Li, S.Y. Huang, HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy, Nucleic Acids Res. 45 (2017) W365–W373, https://doi.org/10.1093/NAR/GKX407.

[47] WebGro | UAMS, (n.d.). https://simlab.uams.edu/(accessed June 5, 2024).

[48] E. Lindahl, P. Bjelkmar, P. Larsson, M.A. Cuendet, B. Hess, Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models, J. Chem. Theor. Comput. 6 (2010) 459–466, https://doi.org/10.1021/CT900549R.

[49] J.R. López-Blanco, J.I. Aliaga, E.S. Quintana-Ortí, P. Chacón, iMODS: internal coordinates normal mode analysis server, Nucleic Acids Res. 42 (2014), https://doi.org/10.1093/NAR/GKU339.

[50] N. Rapin, O. Lund, M. Bernaschi, F. Castiglione, Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system, PLoS One 5 (2010), https://doi.org/10.1371/JOURNAL.PONE.0009862.

[51] F. Castiglione, F. Mantile, P. De Berardinis, A. Prisco, How the interval between prime and boost injection affects the immune response in a computational model of the immune system, Comput. Math. Methods Med. 2012 (2012), https://doi.org/10.1155/2012/842329.

[52] A. Grote, K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D.C. Hempel, D. Jahn, JCat: a novel tool to adapt codon usage of a target gene to its potential expression host, Nucleic Acids Res. 33 (2005), https://doi.org/10.1093/NAR/GKI376.

[53] P. Doruker, R.L. Jernigan, I. Bahar, Dynamics of large proteins through hierarchical levels of coarse-grained structures, J. Comput. Chem. 23 (2002) 119–127, https://doi.org/10.1002/JCC.1160.

[54] 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.

[55] S. Lu, X. xiu Xie, L. Zhao, B. Wang, J. Zhu, T. rui Yang, G. wen Yang, M. Ji, C. ping Lv, J. Xue, E. hei Dai, X. ming Fu, D. qun Liu, L. Zhang, S. jie Hou, X. lin Yu, Y. ling Wang, H. xia Gao, X. han Shi, C. wen Ke, B. xia Ke, C. guo Jiang, R. tian Liu, The immunodominant and neutralization linear epitopes for SARS-CoV-2, Cell Rep. 34 (2021), https://doi.org/10.1016/J.CELREP.2020.108666.

[56] P. Cihan, Z.B. Ozger, A new approach for determining SARS-CoV-2 epitopes using machine learning-based in silico methods, Comput. Biol. Chem. 98 (2022), https://doi.org/10.1016/J.COMPBIOLCHEM.2022.107688.

[57] A. Jorkesh, S. Rothenberger, L. Baldassar, B. Grybaite, P. Kavaliauskas, V. Mickevicius, M. Dettin, F. Vascon, L. Cendron, A. Pasquato, Screening of small-molecule libraries using SARS-CoV-2-derived sequences identifies novel furin inhibitors, Int. J. Mol. Sci. 25 (2024), https://doi.org/10.3390/IJMS25105079.

[58] A. Tarke, J. Sidney, C.K. Kidd, J.M. Dan, S.I. Ramirez, E.D. Yu, J. Mateus, R. da Silva Antunes, E. Moore, P. Rubiro, N. Methot, E. Phillips, S. Mallal, A. Frazier, S. A. Rawlings, J.A. Greenbaum, B. Peters, D.M. Smith, S. Crotty, D. Weiskopf, A. Grifoni, A. Sette, Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases, Cell Reports, Méd. 2 (2021), https://doi.org/10.1016/J.XCRM.2021.100204.

[59] S.K. Hotop, S. Reimering, A. Shekhar, E. Asgari, U. Beutling, C. Dahlke, A. Fathi, F. Khan, M. Lütgehetmann, R. Ballmann, A. Gerstner, W. Tegge, L. Cicin-Sain, U. Bilitewski, A.C. McHardy, M. Brönstrup, Peptide microarrays coupled to machine learning reveal individual epitopes from human antibody responses with neutralizing capabilities against SARS-CoV-2, Emerg. Microb. Infect. 11 (2022) 1037–1048, https://doi.org/10.1080/22221751.2022.2057874.

[60] F.J. Obermair, F. Renoux, S. Heer, C.H. Lee, N. Cereghetti, M. Loi, G. Maestri, Y. Haldner, R. Wuigk, O. Iosefson, P. Patel, K. Triebel, M. Kopf, J. Swain, J. Kisielow, High-resolution profiling of MHC II peptide presentation capacity reveals SARS-CoV-2 CD4 T cell targets and mechanisms of immune escape, Sci. Adv. 8 (2022), https://doi.org/10.1126/SCIADV.ABL5394.

[61] R.P. dos Santos Alves, J. Timis, R. Miller, K. Valentine, P.B.A. Pinto, A. Gonzalez, J.A. Regla-Nava, E. Maule, M.N. Nguyen, N. Shafee, S. Landeras-Bueno, E. Olmedillas, B. Laffey, K. Dobaczewska, Z. Mikulski, S. McArdle, S.R. Leist, K. Kim, R.S. Baric, E. Ollmann Saphire, A. Elong Ngono, S. Shresta, Human coronavirus OC43-elicited CD4+ T cells protect against SARS-CoV-2 in HLA transgenic mice, Nat. Commun. 15 (2024), https://doi.org/10.1038/S41467-024-45043-2.

[62] J. Heide, S. Schulte, M. Kohsar, T.T. Brehm, M. Herrmann, H. Karsten, M. Marget, S. Peine, A.M. Johansson, A. Sette, M. Lütgehetmann, W.W. Kwok, J. Sidney, J.S. zur Wiesch, Broadly directed SARS-CoV-2-specific CD4+ T cell response includes frequently detected peptide specificities within the membrane and nucleoprotein in patients with acute and resolved COVID-19, PLoS Pathog. 17 (2021), https://doi.org/10.1371/JOURNAL.PPAT.1009842.

[63] C. Hu, M. Shen, X. Han, Q. Chen, L. Li, S. Chen, J. Zhang, F. Gao, W. Wang, Y. Wang, T. Li, S. Li, J. Huang, J. Wang, J. Zhu, D. Chen, Q. Wu, K. Tao, D. Pang, A. Jin, Identification of cross-reactive CD8+ T cell receptors with high functional avidity to a SARS-CoV-2 immunodominant epitope and its natural mutant variants, Genes Dis 9 (2022) 216–229, https://doi.org/10.1016/J.GENDIS.2021.05.006.

[64] S. Prakash, R. Srivastava, P.-G. Coulon, N.R. Dhanushkodi, A.A. Chentoufi, D.F. Tifrea, R.A. Edwards, C.J. Figueroa, S.D. Schubl, L. Hsieh, M.J. Buchmeier, M. Bouziane, A.B. Nesburn, B.D. Kuppermann, L. BenMohamed, Genome-wide B cell, CD4+, and CD8+ T cell epitopes that are highly conserved between human and animal coronaviruses, identified from SARS-CoV-2 as targets for preemptive pan-coronavirus vaccines, J. Immunol. 206 (2021) 2566–2582, https://doi.org/10.4049/jimmunol.2001438.

[65] J. Zhao, L. Wang, M. Schank, X. Dang, Z. Lu, D. Cao, S. Khanal, L.N. Nguyen, L.N.T. Nguyen, J. Zhang, Y. Zhang, J.L. Adkins, E.M. Baird, X.Y. Wu, S. Ning, M. El Gazzar, J.P. Moorman, Z.Q. Yao, SARS-CoV-2 specific memory T cell epitopes identified in COVID-19-recovered subjects, Virus Res. 304 (2021), https://doi.org/10.1016/J.VIRUSRES.2021.198508.

[66] M.O. Rafi, K. Al-Khafaji, M.T. Sarker, T. Taskin-Tok, A.S. Rana, M.S. Rahman, Design of a multi-epitope vaccine against SARS-CoV-2: immunoinformatic and computational methods, RSC Adv. 12 (2022) 4288–4310, https://doi.org/10.1039/D1RA06532G.

[67] K. Khan, S.A. Khan, K. Jalal, Z. Ul-Haq, R. Uddin, Immunoinformatic approach for the construction of multi-epitopes vaccine against omicron COVID-19 variant, Virology 572 (2022) 28–43, https://doi.org/10.1016/J.VIROL.2022.05.001.
[68] V. Li, C. Lee, D.A. Yoo, S. Cho, H. Kim, In silico SARS-CoV-2 vaccine development for Omicron strain using reverse vaccinology, Genes Genomics 44 (2022) 937–944, https://doi.org/10.1007/S13258-022-01255-8.
[69] K. Jalal, K. Khan, Z. Basharat, M.N. Abbas, R. Uddin, F. Ali, S.A. Khan, S.S. ul Hassan, Reverse vaccinology approach for multi-epitope centered vaccine design against delta variant of the SARS-CoV-2, Environ. Sci. Pollut. Res. Int. 29 (2022) 60035–60053, https://doi.org/10.1007/S11356-022-19979-1.
[70] P.R. Sirohi, J. Gupta, P. Somvanshi, V.K. Prajapati, A. Grover, Multiple epitope-based vaccine prediction against SARS-CoV-2 spike glycoprotein, J. Biomol. Struct. Dyn. 40 (2022) 3347–3358, https://doi.org/10.1080/07391102.2020.1846626.
[71] S. Srinivasan, G. Selvaraj, V. Gopalan, P. Padmanabhan, K. Ramesh, K. Govindan, A. Chandran, P. Dhandapani, K. Krishnasamy, S. Kitambi, Epitope identification and designing a potent multi-epitope vaccine construct against SARS-CoV-2 including the emerging variants, J. Global Infect. Dis. 14 (2022) 24–30, https://doi.org/10.4103/JGID.JGID_96_21.
[72] M. Molteni, S. Gemma, C. Rossetti, The role of toll-like receptor 4 in infectious and noninfectious inflammation, Mediat. Inflamm. 2016 (2016), https://doi.org/10.1155/2016/6978936.
[73] S. Singh Jyotisha, I.A. Qureshi, Multi-epitope vaccine against SARS-CoV-2 applying immunoinformatics and molecular dynamics simulation approaches, J. Biomol. Struct. Dyn. 40 (2022) 2917–2933, https://doi.org/10.1080/07391102.2020.1844060.