



Research Article

Epitope-Based Vaccine Design against Novel Coronavirus SARS-CoV-2 Envelope Protein

 Hanuman Singh Dagur,¹  Saurabh Singh Dhakar,²  Apporva Gupta³

¹Department of Biotechnology, Jaipur National University, Jaipur, India

²Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India

³Department of Chemical Sciences, Indian Institute of Science Education and Research, Bhopal, Madhya Pradesh, India

Abstract

Objectives: The outbreak of the 2019 novel coronavirus (SARS-CoV-2) has infected millions of people, with a large number of deaths across 200 countries. The sudden emergence of the virus leads to limited available therapies for SARS-CoV-2. Therefore, vaccines and antiviral medicines are in desperate need.

Methods: This study took immune-informatics approaches to identify B- and T-cell epitopes for the envelope membrane protein (E) of SARS-CoV-2, followed by estimating their antigenicity and interactions with the human leukocyte antigen (HLA) alleles.

Results: We identified three B cell epitopes (“NSSRVPD”, “SRVKNLNSSRV”, “SRVPDLLV”), two MHC class-I (“FLLVTLAIL”, “VSLVKPSFY”) and one MHC class-II binding T-cell epitopes (“LLFLAFVVFLVTLA”), which showed highly antigenic features. Allergenicity, toxicity and physicochemical properties analysis confirmed the specificity and selectivity of epitopes. The stability and safety of epitopes were confirmed by digestion analysis.

Conclusion: Epitopes were thus identified and some of them can be potential candidates for vaccine development.

Keywords: Envelope Protein, SARS-CoV-2, Vaccine

Cite This Article: Dagur HS, Dhakar SS, Gupta A. Epitope-Based Vaccine Design against Novel Coronavirus SARS-CoV-2 Envelope Protein. EJMO 2020;4(3):201–208.

The SARS-CoV-2 has recently emerged as a human pathogen infected 23,59,346 confirmed cases globally and 161,975 deaths by April 19, 2020.^[1] SARS-CoV-2 is an enveloped, 29,811 bp positive-sense RNA genome virus that belongs to the genus betacoronavirus and contains both structural and nonstructural proteins. It is transmitted through airborne droplets and spreads rapidly through person-to-person, as well as from contact with infected surfaces and objects.^[2–6] Structural proteins play a crucial role in fusion, virion assembly, morphogenesis and entry into the host cell. Because of these properties structural proteins are important targets for vaccine and antiviral drug development.^[7] The E protein is a structural protein

which acts as a viroporin and self-assembles in host membranes forming pentameric protein-lipid pores that allow ion transport and plays a role in the induction of apoptosis. The critical function of the E protein can be a breakthrough in vaccine design and development.^[8, 9]

Pandemic SARS-CoV-2 vaccines are not currently available for treatment, developing an effective treatment for SARS-CoV-2 is therefore a research priority. The design of novel vaccines against viruses are time-consuming and expensive processes using kits and antibodies.^[10] Thus, we chose immune-informatics methods, which are much efficient and more applicable for deep analysis of viral antigens, prediction of B- and T- cell linear epitope and

Address for correspondence: Hanuman Singh Dagur, MS. Department of Biotechnology, Jaipur National University, Jaipur, India

Phone: +918349572145 **E-mail:** hanumansinghdagur1996@gmail.com

Submitted Date: April 20, 2020 **Accepted Date:** May 14, 2020 **Available Online Date:** May 15, 2020

©Copyright 2020 by Eurasian Journal of Medicine and Oncology - Available online at www.ejmo.org

OPEN ACCESS This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.



evaluation of immunogenicity and virulence of pathogens. B-cell can recognize and activate defense responses against viral infection, while T-cell and antibody reactions may recover severe respiratory infection. In this article, we applied an Immuno-informatics approach to identify potential B- and T-cell epitopes based on the E protein of SARS-CoV-2. The antigenicity of the selected epitopes was estimated and their interactions with the human leukocyte antigen (HLA) alleles were evaluated for MHC class-I epitopes. Allergenicity, toxicity, stability and physicochemical properties were also investigated for exploring the antigenicity, stability and safety of the identified epitopes. Some of these identified epitopes can be used as promising vaccine candidates.

Methods

Data Retrieval and Structural Analysis

The amino acid sequence of SARS-CoV-2 envelope protein was taken from the NCBI database using accession number NC_045512.^[11] The 3D structure of E protein was obtained by applying the homology swiss-model available at online.^[12] The E protein sequence was examined by ProtParam for physicochemical properties, including grand average of hydropathicity (GRAVY), half-life, stability index, molecular weight and amino acid atomic composition.^[13] The transmembrane topology of E protein was examined using TMHMM v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) and the secondary structure was analyzed by PSIPRED.^[14] An online tool DIANNA v1.1 using a trained neural system to predict was used to examine the presence of disulfide-bonds.^[15] The antigenicity of full-length E protein was evaluated by Vaxijen v2.0.^[16]

B-Cell Epitope Prediction

IEDB (Immune-Epitope-Database And Analysis-Resource)^[17] with default parameter settings was used to predict 'B-cell epitopes'. Epitopes predicted by linear epitope prediction of Bepipred and Bepipred 2.0, Kolaskar and Tongaonkar antigenicity, Parker hydrophilicity, Chou and Fasman beta turn, and Karplus and Schulz flexibility. BcePred^[18] was also used to predict B-cell epitopes using accessibility, antigenic propensity, exposed surface, flexibility, hydrophilicity, polarity and turns.

B-cell epitopes predicted by IEDB and BcePred were combined to the B-cell epitope candidate list. Based on transmembrane topology of E protein predicted by TMHMM v2.0, only epitopes on the outer surface were left, and other intracellular epitopes were eliminated VaxiJen 2.0^[16] was applied to evaluate the antigenicity of the remaining epitopes. A stringent criteria was used to have an antigenic-

ity score of 0.4946 and 7 residue-length epitopes deemed adequate to initiate a defensive immune reaction.

T-Cell Epitope Prediction

Cytotoxic T-lymphocyte epitopes are important in vaccine development. The Peptide_binding_to_MHC_class_I_molecules tool of IEDB and HLA class I set^[19] was employed to predict MHC class I binding T-cell epitopes for E protein. The Peptide_binding_to_MHC_class_II_molecules tool of IEDB and HLA class II set^[20] was used to predict T-cell epitopes for E protein. Percentile rank with a threshold of 1% for MHC class I binding epitopes and 10% for MHC class II binding epitopes were used to filter out peptide-allele with weak binding affinity. The antigenicity score of each epitope was calculated by Vaxijen v2.0. A high stringent standard was used to filter peptides with an antigenicity score equal to or greater than 0.6, with the number of binding alleles greater than or equal to 3 for MHC class I binding epitopes and equal to 5 for MHC class II binding epitopes.

Characterization of Selected B-Cell and T-Cell Epitopes

All selected B-cell and MHC class I and II binding T cell epitopes were examined for allergenicity, hydrochemical and physicochemical features, toxicity and digestion. The allergenicity of B-cell and T-cell epitopes was assessed by Allergen FP 1.0 (<http://ddg-pharmfac.net/AllergenFP/>). The toxicity of B-cell and T-cell epitopes along with hydrophobicity, hydropathicity, hydrophilicity, molecular weight, steric hindrance, and charge were evaluated by ToxinPred (<https://webs.iitd.edu.in/raghava/toxinpred/index.html>). The peptides digested by fewer enzymes are more stable and favorable vaccine candidates. The protein digest server ('<http://db.systemsbio.net:8080/proteomicsToolkit/proteinDigest.html>') was used for the digestion of B- and T-cell epitopes by 13 enzymes including Trypsin, Chymotrypsin, Clostripain, Cyanogen Bromide, Iodosobenzoate, Proline Endopept, Staph Protease, Trypsin K, Trypsin R, AspN, Chymotrypsin (modified), Elastase, and Elastase/Trypsin/Chymotrypsin.

Protein-Epitope Interaction Evaluation

The 3D structure of human HLA-B35:01 (PDB ID: 1A9E) at a resolution of 2.5 Å, HLA-B*51:01 (PDB ID: 1E27) at a resolution of 2.2 Å and HLA-B*53:01 (PDB ID: 1A10) at a resolution of 2.3 Å were downloaded from the protein data bank (RCSB PDB) and used to evaluate their interactions with selected epitopes. Protein-peptide interactions were performed by PepSite^[21] with the top prediction selected from a total of 10 epitope-protein interaction reports.

Conservation Analysis of Selected B- and T-Cell Epitopes

The E protein sequences were taken from an open-access database <http://covdb.popgenetics.net/v2/> and NCBI. Multiple sequence alignment was performed by online clustal omega [ref] against the E protein collected from different locations. The selected epitopes were examined for their variability and conservation using an online PSV web server.^[22]

Results

Structural Analysis of Envelope Protein of SARS-CoV-2

The ProtParam software demonstrates that E protein contains 75 amino acids (AA) with a molecular weight of 8.36 kDa. Of the 75 residues, 3 (AA) were negatively charged and 5 (AA) were positively charged. The theoretical isoelectric point (PI) of the E protein was 8.57, representing its positive nature. The instability-index (II) was calculated as 38.68, which classified the E protein as stable. Aliphatic-index was 144.00 with a grand average of hydropathicity (GRAVY) value of 1.128, indicating high hydrophobic nature inside the chain. The half-life of the S protein was estimated as 30 hours for mammalian reticulocytes (in vitro), 20 hours for yeast (in vivo), 10 hours for Escherichia coli (in vivo), which determine the total time taken for its disappear after it has been synthesized in the cell. The Carbon, Hydrogen, Nitrogen, Oxygen and Sulfur numbers of a total of 1213 atoms were formulated as C390H625N91O103S4. The Secondary structure of the E protein was generated by PSIPRED^[14] which showed that Helixes (77.33%) and coil (22.66%) are present in structure (Fig. 1). Disulfide (S-S) bonds were identified by DiANNA (Fig. 2). The antigenicity of the full-length E protein was measured by Vaxijen, which confirmed that the E protein (0.6025) had an antigenicity score greater than the S protein antigenicity (0.4646).^[16] The transmembrane protein topology predicted by TMHMM showed that residues 1 to 11 were inside the transmembrane region, and residues 12 to 34 were within the core-region, and resi-

Disulfide bond scores			
Cysteine sequence position	Distance	Bond	Score
40 - 43	3	TALRLCAYCCN-RLCAYCCNIVN	0.02912
40 - 44	4	TALRLCAYCCN-LCAYCCNIVNV	0.01037
43 - 44	1	RLCAYCCNIVN-LCAYCCNIVNV	0.01037

Step 5: Weighted matching	
Predicted bonds	
40 - 43	TALRLCAYCCN - RLCAYCCNIVN
Predicted connectivity	
1-2	

Figure 2. Identified disulfide bonds by DiANNA.

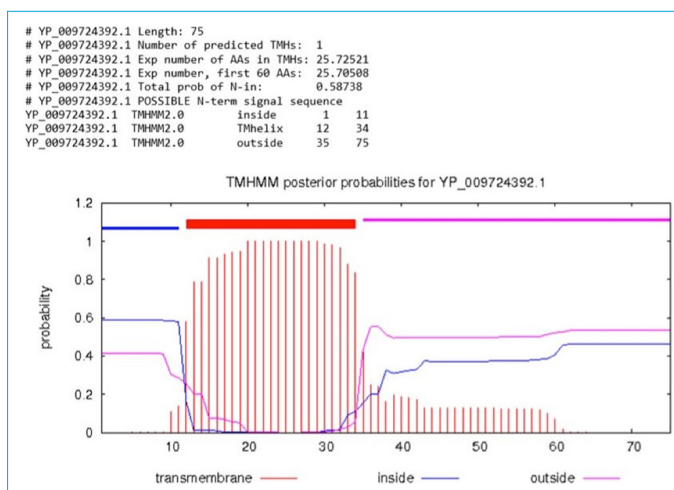


Figure 3. Envelope protein transmembrane topology predicted by TMHMM.

dues 35 to 75 were exposed on the surface of the E protein (Fig. 3).

Identification of B-Cell Epitopes

B-cell epitopes can direct B-cell to recognize and operate defense mechanisms against viral infection. Recognition of B-cell epitopes depended on predictions of linear epitopes, antigenicity, hydrophilicity, surface accessibility, beta-turn and flexibility.^[23] B-cell epitopes of E protein were predicted using IEDB.^[17] A total of 4 linear epitopes (two epitopes each) were identified by Bepipred and Bepipred2.0. Kolaskar and Tongaonkar antigenicity of E protein was

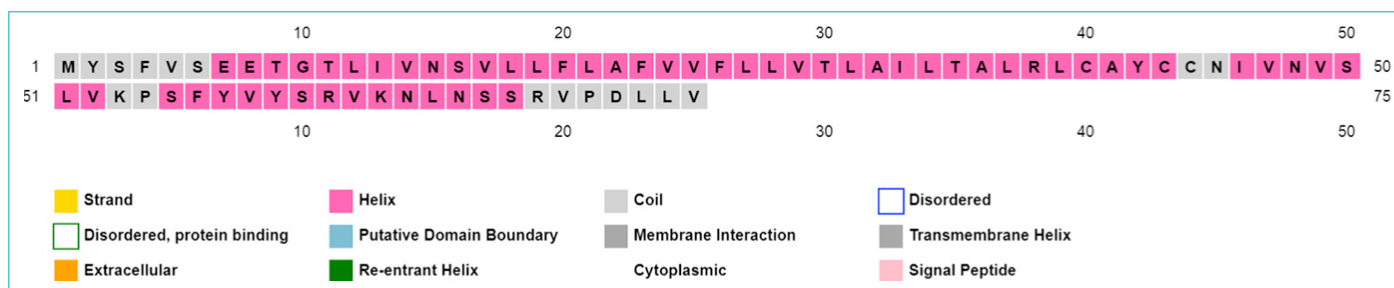


Figure 1. Secondary structure of E protein obtained by PSIPRED.

analyzed with default parameter settings to assess the physicochemical properties. The antigenic affinity value of E protein was 1.119 (average), 0.947 (minimum) and 1.262 (maximum) (Fig. 4a). Hydrophilic regions play a fundamental role in initiating an immune response, which is generally uncovered on the surface of a protein. The Parker hydrophilicity of E protein was estimated as -0.911 (average), -6.843 (minimum) and 4.929 (maximum) (Fig. 4b). To find the surface availability of B-cell epitopes, Emini surface accessibility was predicted for E protein with an average of 1.00, a minimum of 0.088, and a maximum of 4.316 (Fig. 4c). The surface exposure and hydrophilic nature of beta constitute a vital structure in the initial defense response. Therefore, we predicted that the Chou and Fasman beta-turn would result in 0.883 (average), 0.554 (minimum) and 1.264 (maximum) in E protein (Fig. 4d). We estimated the Karplus and Schulz flexibility of E protein and the result was 0.965 (average), 0.894 (minimum), and 1.081 (maximum) (Fig. 4e). B-cell epitopes were selected based on the combination of results. BcePred^[18] was used to predict B-cell epitopes using accessibility, antigenic propensity, exposed surface, flexibility, hydrophilicity, polarity and turns. Overall, we obtained a totally 11 B-cell epitopes (Table 1).

The antigenicity of predicted B-cell epitopes was further evaluated by Vaxijen v2.0 with a high stringent threshold of 0.4649. Based on the transmembrane topology of E protein predicted by TMHMM v2.0, intracellular epitopes

were eliminated. B-cell epitopes with residue length from 7 to 16 were obtained, including "NSSRVPD", "SRVPDLLV", "LCAYCCNIV" and "SRVKNLNSSRV" (Table 1).

Allergenicity, Toxicity and Stability Analysis of B-Cell Epitopes

The allergenicity of B-cell epitopes was determined by Allergen FP 1.0, indicating that all four B-cell epitopes were predicted to be non-allergenic. Toxicity, steric hindrance, hydrophobicity, hydrophobicity, hydrophobicity, amphipathicity, molecular weight, net hydrogen, isoelectric point and charge of B-cell epitopes were evaluated by ToxinPred, a support vector machine (SVM) based method (Fig. 5). The result indicates that three of the four B-cell epitopes were non-toxic. The stability of B-cell epitopes was evaluated through the number of peptides digesting enzymes by the protein digest server. All B-cell epitopes were found to have multiple non-digesting enzymes.

Identification of T-Cell Epitopes

The peptide binding to MHC class I molecules tool of IEDB and HLA class I set^[19] was used to predict T-cell epitopes for E protein. A threshold percentile of 1% was applied to filter the peptide-allele with weak binding affinity. The antigenicity score of each peptide was determined by VaxiJen v2.0 to calculate its antigenicity. A peptide having both a high antigenicity score and capacity to bind with a larger

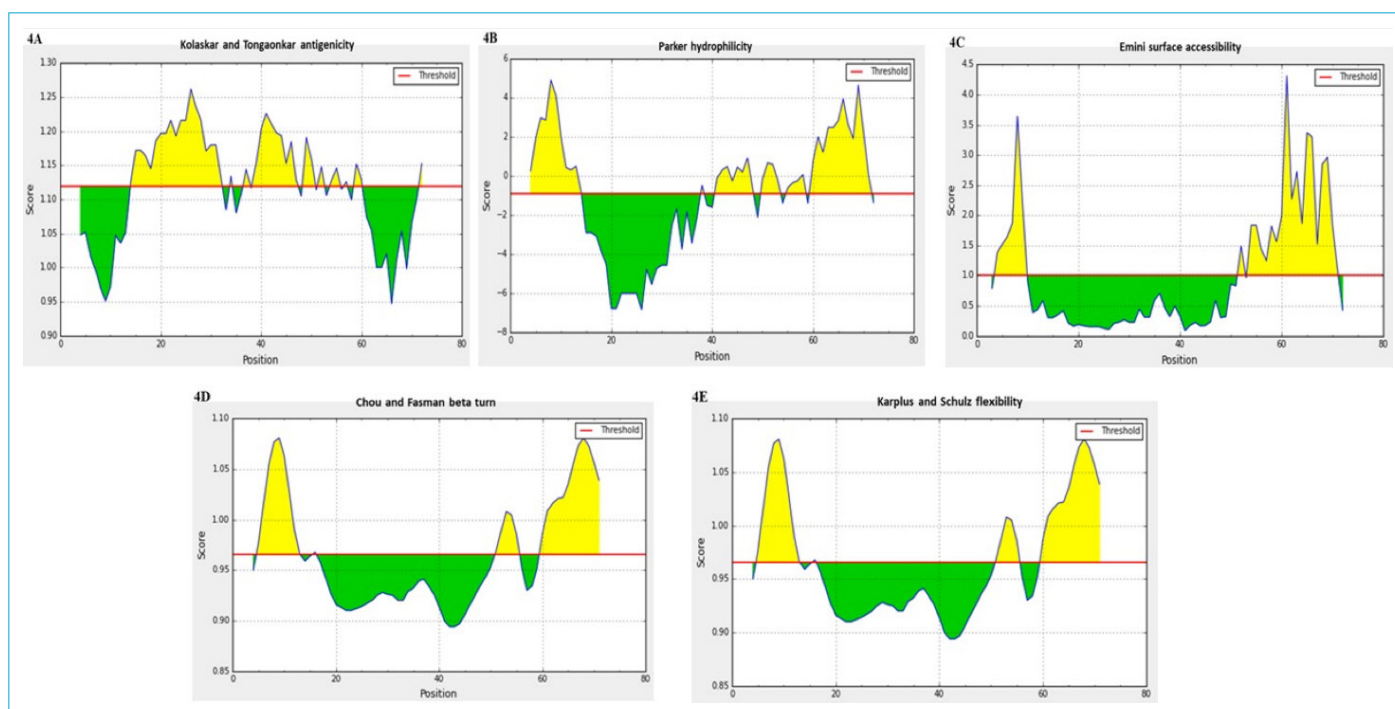


Figure 4. Linear epitope identification of IEDB resource analysis. (a) Antigenicitybased prediction by Kolaskar and Tongaonkar scale. (b) Hydrophilicity based prediction by Parker Hydrophilicity software. (c) Surface accessibility based by Emini surface accessibility software. (d) Beta turn based prediction using Chou & Fasman beta turn program. (e) Flexibility based prediction by Karplus & Schulz scale.

Table 1. Predicted epitopes with their sequence, length, antigenicity, allergenicity, toxicity, stability and type of epitope

Epitope Sequence	Length	Antigenicity	Allergenicity	Toxicity	Stability (Yes/No)	Type of Epitope
NSSRVPD	7	0.4934	Non-Allergen	Non-Toxic	Y	B-cell
SRVKNLNSSRV	11	0.6816	Non-Allergen	Non-Toxic	Y	B-cell
YVYSRVKLNSSRVPD	16	0.5457	Allergen	Non-Toxic	Y	B-cell
YCCNIVNLSLVKPSFYVYSRVK	22	0.8817	Non-Allergen	Non-Toxic	N	B-cell
SRVPDLLV	8	0.5064	Non-Allergen	Non-Toxic	Y	B-cell
LCAYCCNIV	9	0.7349	Non-Allergen	Toxic	Y	B-cell
PSFYVYSRVKLNSSRVP	18	0.5796	Non-Allergen	Non-Toxic	Y	B-cell
NIVNLSLVKPSFYVYSRVKLNSSRVPD	28	0.7457	Non-Allergen	Non-Toxic	N	B-cell
LAFVFLV	9	0.7976	Non-Allergen	Non-Toxic	Y	MHC I T-cell
FLLVTLAIL	9	0.9694	Non-Allergen	Non-Toxic	Y	MHC I T-cell
FVFLVTL	9	0.7403	Allergen	Non-Toxic	N	MHC I T-cell
VSLVKPSFY	9	0.7476	Non-Allergen	Non-Toxic	Y	MHC I T-cell
LLFLAFVFLVTLA	15	0.8122	Non-Allergen	Non-Toxic	Y	MHC II T-cell
TLAILTALRLCAYCC	15	0.7304	Non-Allergen	Toxic	Y	MHC II T-cell

number of alleles is considered to have high potential to initiate a strong defensive response. A high stringent criteria were used to filter peptides with antigenicity scores greater than or equal to 0.7 and the number of binding alleles greater than or equal to 6. Using the evaluation method on the peptide 'FLLVTLAIL', the highest antigenicity score is 0.9694 and binds with more than six alleles including A*68:01, B*35:01, A*57:01, B*58:01, B*40:01, B*44:02 and A*33:01. The peptide 'VSLVKPSFY' has the highest number of binding MHC class-I alleles (25 alleles), including B*35:01, B*53:01, A*23:01, B*58:01, A*24:02, and B*08:01 with an antigenicity score of 0.7476. The peptide binding to MHC class II molecules tool of IEDB and HLA class II set[20] was used to identify T-cell epitopes for E protein. A

threshold percentile of 10% was applied to filter peptide-allele with weak binding affinity. The antigenicity score of each peptide was determined by VaxiJen v2.0 to calculate its antigenicity. A high stringent standard was applied to filter peptides with antigenicity score greater than or equal to 0.7 and the number of binding alleles greater than or equal to 5. The peptide 'LLFLAFVFLVTLA' has the highest antigenicity score of 0.8122.

Allergenicity, Toxicity and Stability Analysis of T-Cell Epitopes

The allergenicity of T-cell epitopes was determined by Allergen FP 1.0, indicating that five of the six T-cell epitopes were predicted to be non-allergenic. The toxicity, steric hindrance, hydrophobicity, hydrophobicity, hydrophilicity, am-

Query Peptides													
Peptide ID	Peptide Sequence	SVM Score	Prediction	Hydrophobicity	Steric hindrance	Sidebulk	Hydrophobicity	Amphipathicity	Hydrophilicity	Net Hydrogen	Charge	pi	Mol wt
1	NSSRVPD	-0.53	Non-Toxin	-0.45	0.62	0.62	-1.50	0.35	0.78	1.29	0.00	6.19	773.88
2	SRVKNLNSSRV	-0.61	Non-Toxin	-0.46	0.64	0.64	-0.92	0.78	0.50	1.55	3.00	12.01	1259.58
3	YVYSRVKLNSSRVPD	-0.90	Non-Toxin	-0.33	0.64	0.64	-0.85	0.54	0.15	1.25	2.00	9.72	1897.34
4	YCCNIVNLSLVKPSFYVYSRVK	-0.04	Non-Toxin	-0.06	0.65	0.65	0.45	0.44	-0.55	0.82	3.00	9.27	2582.39
5	SRVPDLLV	-0.90	Non-Toxin	-0.08	0.60	0.60	0.70	0.31	-0.04	0.75	0.00	6.19	898.18
6	LCAYCCNIV	1.39	Toxin	0.17	0.64	0.64	1.89	0.00	-1.19	0.33	0.00	5.79	1001.38
7	PSFYVYSRVKLNSSRVP	-1.16	Non-Toxin	-0.24	0.62	0.62	-0.54	0.48	-0.16	1.11	3.00	10.29	2113.65
8	LAFVFLV	-1.33	Non-Toxin	0.52	0.62	0.62	3.49	0.00	-1.71	0.00	0.00	5.88	1020.48
9	FLLVTLAIL	-1.43	Non-Toxin	0.45	0.59	0.59	3.09	0.00	-1.54	0.11	0.00	5.88	1002.45
10	FVFLVTL	-1.37	Non-Toxin	0.47	0.62	0.62	3.21	0.00	-1.70	0.11	0.00	5.88	1050.49
11	VSLVKPSFY	-1.20	Non-Toxin	0.06	0.60	0.60	0.73	0.41	-0.67	0.56	1.00	6.94	1039.36
12	LLFLAFVFLVTLA	-1.54	Non-Toxin	0.46	0.60	0.60	3.11	0.00	-1.61	0.07	0.00	5.88	1679.40
13	TLAILTALRLCAYCC	0.23	Toxin	0.11	0.58	0.58	1.99	0.16	-0.91	0.47	1.00	8.26	1628.25
14	SEE	-0.79	Non-Toxin	-0.50	0.63	0.63	-2.80	0.85	2.10	1.00	-2.00	3.80	383.35
15	LNSSR	-0.73	Non-Toxin	-0.48	0.61	0.61	-1.16	0.49	0.40	1.80	1.00	10.11	575.68
16	VSEETGT	-0.71	Non-Toxin	-0.17	0.62	0.62	-0.77	0.36	0.57	0.71	-2.00	3.80	721.81
17	NIVNLSLVKPSFYVYSRVKLNSSRVPD	-1.24	Non-Toxin	-0.17	0.64	0.64	-0.14	0.44	-0.15	1.00	3.00	10.00	3198.08

Figure 5. Epitopes toxicity results obtained by ToxinPred.

phipathicity, molecular weight, net hydrogen, isoelectric point and charge of the T-cell epitopes were evaluated by a ToxinPred, support vector machine (SVM) based method. (Fig. 5). The result indicates that four out of the five B-cell epitopes were non-toxic, while 'TLAILTALRLCAYCC' was a toxic epitope. The stability of T-cell epitopes was evaluated through the number of peptides digesting enzymes by the protein digest server. All T-cell epitopes were found to have multiple non-digesting enzymes.

Interaction of T-Cell Epitopes with HLA Alleles

Protein-peptide interactions are critical in cellular signaling pathways. The two MHC class-I binding epitopes, 'FLLVTLAIL' and 'VSLVKPSFY', were considered to be non-allergic and non-toxic. Both epitopes were predicted to bind to HLA-B35:01, HLA-B*51:01, and HLA-B*53:01. The 3D structure of human HLA-B35:01 (PDB ID: 1A9E),^[24] HLA-B*51:01 (PDB ID: 1E27)^[25] and HLA-B*53:01 (PDB ID: 1A10)^[26] proteins were accessible with co-crystallized peptides in the PDB database. Protein-peptide interactions were performed by PepSite,^[21] 10 epitope-protein interactions were reported.

The epitope 'FLLVTLAIL' was predicted to bind significantly on the surface of HLA-B35:01 (PDB ID: 1A9E) via six hydrogen bonds with Phe-1, Thr-5, Leu-6, Ala-7, Ile-8. Epitope 'FLLVTLAIL' bond moderately significantly to HLA-B35:01 (PDB ID: 1A9E) via six hydrogen bonds with Phe-1, Leu-3, Ala-7. Similarly, both epitope 'FLLVTLAIL' and 'VSLVKPSFY' show strong and stable binding with HLA-B*51:01 (1E27) residues and HLA-B*53:01 (PDB ID: 1A10) residues, respectively (Fig. 6).

Conservation of B- and T-Cell Epitopes

The amino acid sequences of E protein were obtained and MSA (<https://www.ebi.ac.uk/Tools/msa/muscle/>) was performed on sequences combined with predicted B-cell and T-cell epitopes. No mutations were observed in all selected epitopes confirming that all selected epitopes were conserved in all sequences used in the analysis (Fig. 7).

Discussion

The emergence of SARS-CoV-2 is a daunting situation for the whole society, so there is a vital need for medications and preventive measures. The SARS-CoV-2 survives in the

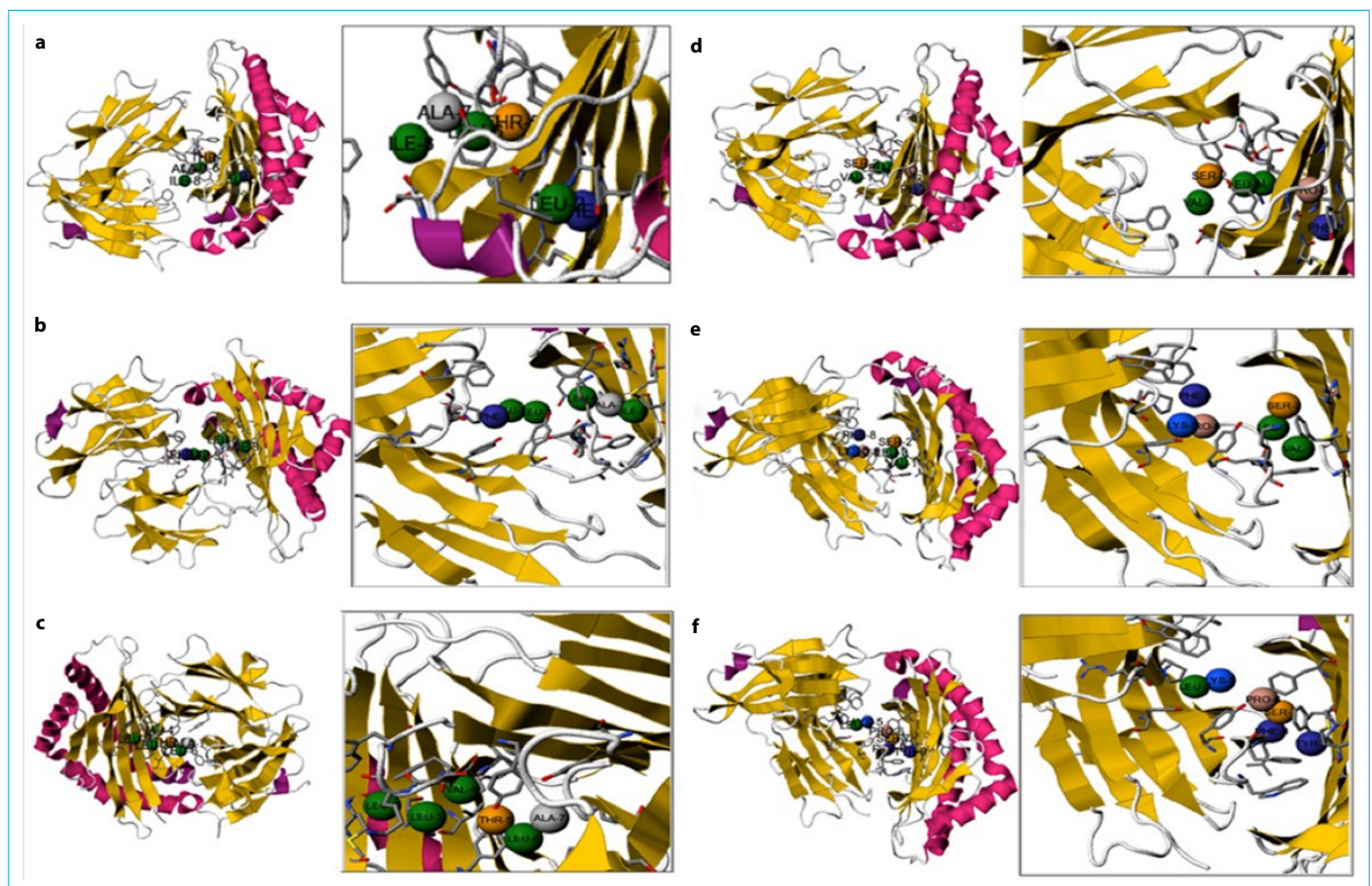


Figure 6. Results obtained by PepSite showing interaction between human leucocyte antigen and E protein linear epitope. (a) 1A10 and "FLLVTLAIL" (b) 1A9E and "FLLVTLAIL" (c) 1E27 and "FLLVTLAIL" (d) 1A10 and "VSLVKPSFY" (e) 1A9E and "VSLVKPSFY" (f) 1E27 and "VSLVKPSFY".

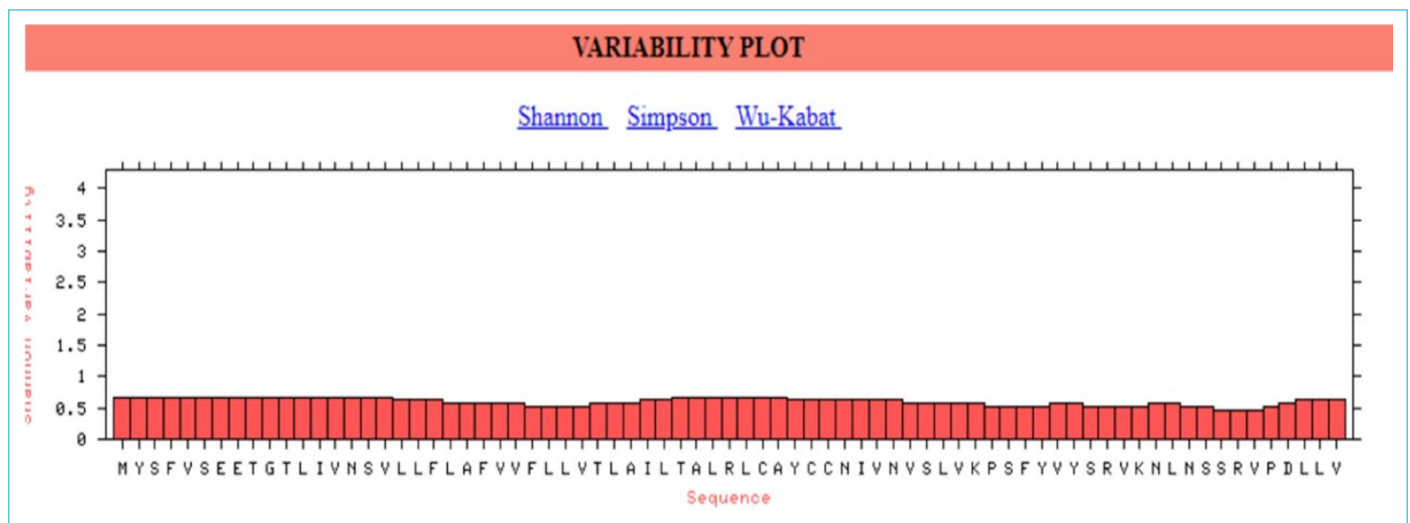


Figure 7. Conservation and variability plot obtained from PSV web server. Selected epitopes are conserved.

lungs and causes symptoms such as fever, cough, and dyspnea. According to the WHO (World Health Organization), symptoms of SARS-CoV-2 can occur within 2 to 24 days which can be transmitted through human to human or through contact with infected surfaces and objects.^[5-7] It is essential to quickly identify immune epitopes. The E protein is crucial in the virion assembly and inward spread of the virus,^[8] as a result it is a primary target for neutralizing antibodies.

The specificity of 'epitope-based vaccines' can be enhanced by selecting the E protein fragments exposed on the surface of the membrane.^[27] Medical biotechnology is consistently working to develop vaccines against SARS-CoV-2. Nonetheless, in-silico immune-informatics can improve time and economic efficiency, therefore, it is also an essential method in immunogenic analysis and vaccine development.

In this study, we characterized the physicochemical characteristics of the Covid-19 viral genome for epitopes and adopted an 'In-silico' method with highly stringent criteria to identify E protein targeted B- and T-cell epitopes which may potentially advance immune response inside the host cell. The antigenicity, flexibility, solvent accessibility, disulfide bonds of predicted candidates were evaluated, yielding three potential B-cell epitope and vaccine candidates. Allergenicity and toxicity analysis confirmed that the three B-cell epitopes were non-allergen and non-toxic. Stability analysis revealed that they cannot be digested by multiple enzymes. In addition, two MHC class-I and one MHC class-II binding T-cell epitopes were predicted to interact with various HLA alleles and to be highly antigenic in nature. Allergenicity, toxicity, and physicochemical properties of T-cell epitopes were determined to enhance specificity and selectivity. The stability and safety were verified by diges-

tion analysis. The B- and T-cell epitopes identified (Table 1) here may help the development of potent peptide-based vaccines to address the pandemic challenges. However, antiviral vaccines are essential to be developed before the predicted epitopes become potentially outdated. In addition, our immune-informatics-based methods also provide a framework to identify B- and T-cell epitopes having therapeutic potential with magnificent scope for SARS-CoV-2, but not limited to specific viruses.

Acknowledgement

We would like to thank IIT Bombay and JNU Jaipur for providing a platform for the study. We sincerely thank our family. We also thank Mr. Prem Pal Singh (PhD electronics at JNU Jaipur), Mr. Chandra Pal Uikey and Ms. Bhawna Upadhaya for their support.

Disclosures

Peer-review: Externally peer-reviewed.

Conflict of Interest: None declared.

Authorship Contributions: Concept – H.S.D.; Design – H.S.D.; Supervision – H.S.D.; Data collection &/or processing – H.S.D., S.S.D., A.G.; Analysis and/or interpretation – H.S.D., S.S.D.; Literature search – S.S.D., A.G.; Writing – S.S.D., A.G.; Critical review – H.S.D.

References

1. Worldometer. COVID-19 CORONAVIRUS PANDEMIC. Available at: www.worldometers.info/coronavirus/ Accessed Sep 18, 2020.
2. Dagur HS, Dhakar SS. Genome Organization of Covid-19 and Emerging Severe Acute Respiratory Syndrome Covid-19 Outbreak: A Pandemic. *Eurasian J Med Oncol* 2020; 4:107–15.
3. Liu W, Zhang Q, Chen J, Xiang R, Song H, Shu S, et al. Detection of Covid-19 in Children in Early January 2020 in Wuhan, China. *N Engl J Med* 2020;382:1370–1.
4. Chan JFW, Yuan S, Kok KH, To KKW, Chu H, Yang J, et al. A famil-

- ial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020;395:514–23.
5. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020;395:507–13.
 6. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 2020;382:1199–207.
 7. Lindenbach BD, Rice CM. Molecular biology of flaviviruses. *Adv Virus Res* 2003;59:23–61.
 8. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. *Virology* 2019;16:69.
 9. Bianchi M, Benvenuto D, Giovanetti M, Angeletti S, Ciccozzi M, Pascarella S. Sars-CoV-2 Envelope and Membrane Proteins: Structural Differences Linked to Virus Characteristics? *Biomed Res Int* 2020;2020:4389089.
 10. Tahir Ul Qamar M, Bari A, Adeel MM, Maryam A, Ashfaq UA, Du X, et al. Peptide vaccine against chikungunya virus: immunoinformatics combined with molecular docking approach. *J Transl Med* 2018;16:298.
 11. Mishra S. T Cell Epitope-Based Vaccine Design for Pandemic Novel Coronavirus 2019-nCoV. Available at: https://chemrxiv.org/articles/T_Cell_Epitope-Based_Vaccine_Design_for_Pandemic_Novel_Coronavirus_2019-nCoV/12029523 Accessed Sep 19, 2020.
 12. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;46:W296–W303.
 13. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003;31:3784–8.
 14. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res* 2013;41:W349–57.
 15. Ferrè F, Clote P. DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Res* 2006;34:W182–5.
 16. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007;8:4.
 17. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, et al. The immune epitope database and analysis resource: from vision to blueprint. *PLoS Biol* 2005;3:e91.
 18. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 2006;65:40–8.
 19. Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. *Proc Natl Acad Sci U S A* 2013;110:E2046–53.
 20. Greenbaum J, Sidney J, Chung J, Brander C, Peters B, Sette A. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics* 2011;63:325–35.
 21. Trabuco LG, Lise S, Petsalaki E, Russell RB. PepSite: prediction of peptide-binding sites from protein surfaces. *Nucleic Acids Res* 2012;40:W423–7.
 22. Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, Reche PA. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res* 2008;36:W35–41.
 23. Fieser TM, Tainer JA, Geysen HM, Houghten RA, Lerner RA. Influence of protein flexibility and peptide conformation on reactivity of monoclonal anti-peptide antibodies with a protein alpha-helix. *Proc Natl Acad Sci U S A* 1987;84:8568–72.
 24. Menssen R, Orth P, Ziegler A, Saenger W. Decamer-like conformation of a nona-peptide bound to HLA-B*3501 due to non-standard positioning of the C terminus. *J Mol Biol* 1999;285:645–53.
 25. Maenaka K, Maenaka T, Tomiyama H, Takiguchi M, Stuart DI, Jones EY. Nonstandard peptide binding revealed by crystal structures of HLA-B*5101 complexed with HIV immunodominant epitopes. *J Immunol* 2000;165:3260–7.
 26. Smith KJ, Reid SW, Harlos K, McMichael AJ, Stuart DI, Bell JI, et al. Bound water structure and polymorphic amino acids act together to allow the binding of different peptides to MHC class I HLA-B53. *Immunity* 1996;4:215–28.
 27. Ghanavi J, Farnia P, Ghomi H. The Effectiveness of Cold Atmospheric Plasma by Inhaling Anesthetic Mask or through Bronchoscopy against COVID-19. *Biomed Biotechnol Res J* 2017;50:553–8.