Received: 22 February 2020 | Accepted: 26 February 2020

DOI: 10.1002/jmv.25736

RESEARCH ARTICLE

JOURNAL OF MEDICAL VIROLOGY WILEY

Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach

Manojit Bhattacharya PhD[1,2] | Ashish R. Sharma PhD[1] | Prasanta Patra MSc[2] | Pratik Ghosh MSc[2] | Garima Sharma PhD[3] | Bidhan C. Patra PhD[2] | Sang-Soo Lee MD, PhD[1] | Chiranjib Chakraborty PhD[1,4]

[1]Institute for Skeletal Aging and Orthopedic Surgery, Hallym University-Chuncheon Sacred Heart Hospital, Chuncheon-si, Gangwon-do, Republic of Korea

[2]Department of Zoology, Vidyasagar University, Midnapore, West Bengal, India

[3]Neuropsychopharmacology and Toxicology Program, College of Pharmacy, Kangwon National University, Chuncheon, Republic of Korea

[4]Department of Biotechnology, School of Life Science and Biotechnology, Adamas University, Kolkata, West Bengal, India

Correspondence
Sang-Soo Lee, MD, PhD, Institute for Skeletal Aging and Orthopedic Surgery, Hallym University-Chuncheon Sacred Heart Hospital, Chuncheon-si, Gangwon-do 24252, Korea.
Email: 123sslee@gmail.com

Chiranjib Chakraborty, PhD, Department of Biotechnology, School of Life Science and Biotechnology, Adamas University, Barasat-Barrackpore Rd, Jagannathpur, Kolkata, West Bengal 700126, India.
Email: drchiranjib@yahoo.com

Disclosures: None.

Funding information
Hallym University Research Fund; National Research Foundation of Korea (NRF), Grant/Award Number: NRF-2017R1A2B4012944

Abstract

Recently, a novel coronavirus (SARS-COV-2) emerged which is responsible for the recent outbreak in Wuhan, China. Genetically, it is closely related to SARS-CoV and MERS-CoV. The situation is getting worse and worse, therefore, there is an urgent need for designing a suitable peptide vaccine component against the SARS-COV-2. Here, we characterized spike glycoprotein to obtain immunogenic epitopes. Next, we chose 13 Major Histocompatibility Complex-(MHC) I and 3 MHC-II epitopes, having antigenic properties. These epitopes are usually linked to specific linkers to build vaccine components and molecularly dock on toll-like receptor-5 to get binding affinity. Therefore, to provide a fast immunogenic profile of these epitopes, we performed immunoinformatics analysis so that the rapid development of the vaccine might bring this disastrous situation to the end earlier.

KEYWORDS

epitopes, immunoinformatics, SARS-COV-2, vaccine

1 | INTRODUCTION

At the end of 2019, a novel coronavirus (SARS-COV-2) was identified as the cause of a cluster of pneumonia cases in Wuhan, a city in the Hubei province of China.[1] It has a positive-sense single-stranded RNA as their genetic component and shares genome similarity with SARS-CoV and bat coronavirus,[2,3] 79.5% and 96% respectively. Phylogenetically, it belongs to the family Coronaviridae, order Nidovirales and is a β-coronavirus of 2B group.[4]

Regarding epidemiology, human-to-human transmission of the virus through the sneezes, cough, and respiratory droplets has been confirmed, yet the zoonotic nature has not been confirmed.[5-7] Epidemiologic investigation in Wuhan, China identified an initial association with a seafood market where most patients had worked or visited.[4] However, as the outbreak progressed, several confirmed cases were reported sporadically all over the world, showing the pandemic nature of the disease named as COVID-19. At last, on 30 January 2020, the World Health Organization (WHO) declared this outbreak a public health emergency of international concern.[8]

Manojit Bhattacharya and Ashish R. Sharma contributed equally to the study.

TABLE 1 List of linear B-cell epitopes along with their sequence, position, and length

Serial no.	Start	End	Sequence	Length
1	22	46	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	68	90	FSNVTWFHAIHVSGTNGTKRFDN	23
3	106	107	KS	2
4	147	163	DPFLGVYYHKNNKSWME	17
5	186	198	MDLEGKQGNFKNL	13
6	215	230	KHTPINLVRDLPQGFS	16
7	259	269	TPGDSSSGWTA	11
8	302	305	LDPL	4
9	313	331	KSFTVEKGIYQTSNFRVQP	19
10	338	372	FPNITNLCPFGEVFNATRFASVYAWNRKRISNCVA	35
11	378	402	YNSASFSTFKCYGVSPTKLNDLCFT	25
12	413	435	GDEVRQIAPGQTGKIADYNYKLP	23
13	449	510	NLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTN	62
14	525	545	ELLHAPATVCGPKKSTNLVKN	21
15	564	571	SNKKFLPF	8
16	589	592	QTLE	4
17	611	615	TNTSN	5
18	625	641	NCTEVPVAIHADQLTPT	17
19	643	653	RVYSTGSNVFQ	11
20	665	675	VNNSYECDIPI	11
21	681	699	ASYQTQTNSPRRARSVASQ	19
22	704	719	YTMSLGAENSVAYSNN	16
23	757	757	E	1
24	782	788	EQDKNTQ	7
25	795	809	KQIYKTPPIKDFGGF	15
26	816	823	PDPSKPSK	8
27	837	851	LADAGFIKQYGDCLG	15
28	997	1001	EAEVQ	5
29	1044	1052	GQSKRVDFC	9
30	1116	1127	RNFYEPQIITTD	12
31	1142	1181	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI	40
32	1212	1215	LGKY	4
33	1261	1276	SCCKFDEDDSEPVLKG	16
34	1278	1278	K	1

According to the situation report 35 (reported by 24 February 2020) of WHO, in China, 77 262 confirmed cases were reported, of which 2595 cases were with deaths. Moreover, outside of China, 2069 confirmed cases were reported in 29 other countries (https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200224-sitrep-35-covid-19.pdf?sfvrsn=1ac4218d_2).

Therefore, as the situation was getting worse and worse, the need for designing a suitable peptide vaccine component against the SARS-COV-2 was growing. Our work was to find suitable epitopes, which can generate enough immune response against the SARS-COV-2 infection. Using immunoinformatics, we could recognize and characterize potential B and T-cell epitopes for the generation of the

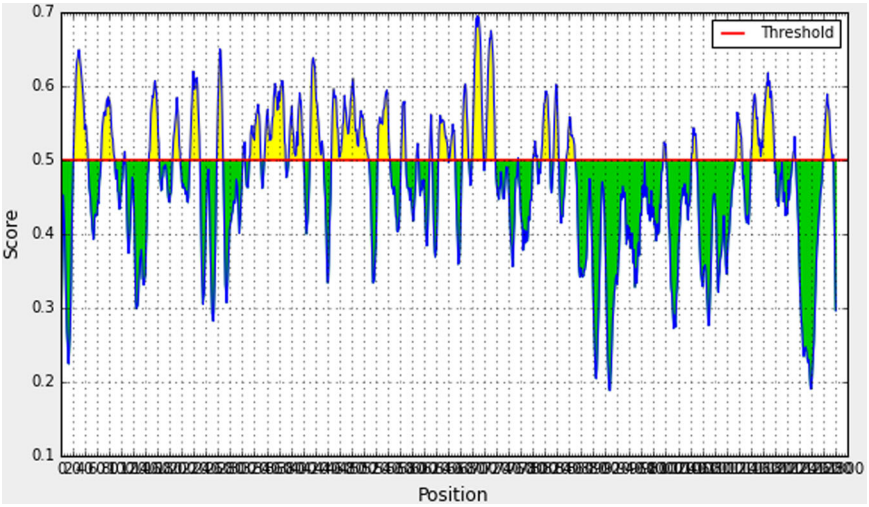

FIGURE 1 Graphical representation of linear B-cell epitopes within the spike glycoprotein of SARS-COV-2

epitopic vaccine against SARS-COV-2.[9] Specifically, the spike glycoprotein of SARS-COV-2 is considered as the target because it forms a characteristic crown of the virus and protrudes from the viral envelope.[10] So, the protein sequence of spike glycoprotein was explored thoroughly using multiple immunoinformatic-based servers and software, to identify various epitopes for an effective vaccine.

2 | MATERIALS AND METHODS

2.1 | Collection of targeted protein sequence

The amino acid sequence of the targeted protein on SARS-COV-2 was collected from the National Centre for Biotechnological Information (NCBI) database.[11] The protein sequence is very crucial for identifying the potential epitopes of the targeted protein.

2.2 | Identification of B-cell epitopes

In this subsection, we used the Immune Epitope Database (IEDB) to identify linear B-cell epitopes using the incorporated BepiPred 2.0 prediction module.[12,13] We provided the FASTA sequence of the targeted protein as an input considering all default parameters.

2.3 | Identification of T-cell epitopes and antigenicity analysis

T-cell epitopes having the binding affinity towards MHC-I and MHC-II alleles were selected to boost up both cytotoxic T-cell and helper T-cell mediated immune response. We adopted two servers which are ProPred-I and ProPred server to the selection of MHC-I and MHC-II binding epitopes respectively within preidentified B-cell epitopic region.[14,15] The selected epitopes were submitted to the VaxiJen v.2.0 server applying a virus as a target field with the given threshold value of 0.4 for analyzing the antigenic propensity.[16]

2.4 | Vaccine construction, modeling, and validation

With the help of a specific peptide linker, we fused the antigenic epitopes to construct an effectual vaccine component. Later, the vaccine component was modeled in the SPARKS-X server.[17] An adjuvant was also added with the vaccine component to accelerate the adaptive immune responses. The vaccine model passed through two different servers ProSA-web and PROCHECK—in a subsequent manner for evaluating the structural accuracy of the model.[18,19]

2.5 | Molecular docking analysis

Molecular docking is the most promising part of the modern drug-discovery method. Here, in this study, we adopted PatchDock (Beta 1.3 Version) docking server to receptor-ligand docking.[20] PatchDock server analyzes the molecular docking between the vaccine component and the toll-like receptor (TLR)-5. The generated Protein Data Bank (PDB) file of the protein-peptide docking complex was visualized in PyMOL software v.2.3.[21]

3 | RESULT

3.1 | Collection of targeted protein sequence

Spike glycoprotein of SARS-COV-2, retrieved from the NCBI has the GenBank accession ID: QHR63290.1. This spike glycoprotein has 1282-long amino acid sequences and this sequence was downloaded in a FASTA format to carry out the further process.

TABLE 2 List of epitopes with encountering MHC-I alleles, positional value, and VaxiJen antigenic score

Serial no.	Epitopic sequence	MHC-I alleles	Position	Antigenicity
1	SQCVNLTTR	HLA-A*3101 HLA-A*3302 HLA-A68.1 HLA-A20 Cattle HLA-B*2705 MHC-Db revised	22-30	1.5476 (Probable Antigen).
2	YTNSFTRGV	HLA-A2 HLA-A*0201 HLA-A2.1 HLA-B*5301 HLA-B*5401 HLA-B*51 HLA-B*5801 HLA-B61	37-45	−0.6177 (Probable Nonantigen).
3	GVYYHKNNK	HLA-A*1101 HLA-A3 HLA-A*3101 HLA-A68.1 HLA-B*2705	151-159	0.8264 (Probable Antigen).
4	GKQGNFKNL	HLA-A2 HLA-A20 Cattle HLA-B*3902 HLA-Cw*0301 MHC-Db MHC-Db revised MHC-Dd MHC-Kb	190-198	1.0607 (Probable Antigen).
5	TPINLVRDL	HLA-A24 HLA-B14 HLA-B*3501 HLA-B*3801 HLA-B*3901 HLA-B*3902 HLA-B40 HLA-B*4403 HLA-B*5101 HLA-B*5102 HLA-B*5103 HLA-B*5301 HLA-B*5401 HLA-B*51 HLA-B60 HLA-B7 HLA-B*0702 HLA-B8 HLA-Cw*0301 HLA-Cw*0401 HLA-Cw*0602 HLA-Cw*0702 MHC-Kd MHC-Ld	217-225	0.3862 (Probable Nonantigen).

TABLE 2 (Continued)

Serial no.	Epitopic sequence	MHC-I alleles	Position	Antigenicity
6	GIYQTSNFR	HLA-A*1101 HLA-A3 HLA-A*3101 HLA-A*3302 HLA-A68.1 HLA-A20 Cattle HLA-B*2705	320-328	0.5380 (Probable Antigen).
7	NLCPFGEVF	HLA-A1 HLA-A3 HLA-A2.1 HLA-B*2702 HLA-B*5201 HLA-B*5801 HLA-B62 MHC-Ld	343-351	0.1999 (Probable Nonantigen).
8	FASVYAWNR	HLA-A*3101 HLA-A*3102 HLA-A68.1 HLA-A20 Cattle HLA-B*5301 HLA-B*5401	356-364	0.0713 (Probable Nonantigen).
9	ASFSTFKCY	HLA-A1 HLA-B*2702 HLA-B*3501 HLA-B*4403 HLA-B*5401 HLA-B*51 HLA-B*5801 HLA-Cw*0702 MHC-Ld	381-389	0.2795 (Probable Nonantigen).
10	VSPTKLNDL	HLA-A24 HLA-A2.1 HLA-B*3501 HLA-B*3902 HLA-B*51 HLA-B*5801 HLA-B60 HLA-B7 HLA-B8 HLA-Cw*0401 HLA-Cw*0602 MHC-Dd MHC-Kb MHC-Ld	391-399	1.4610 (Probable Antigen).
11	KIADYNYKL	HLA-A2 HLA-A*0201 HLA-A*0205 HLA-A24 HLA-A3 HLA-A*3101 HLA-A2.1 HLA-B*2705	426-434	1.6639 (Probable Antigen).

(Continues)

TABLE 2 (Continued)

Serial no.	Epitopic sequence	MHC-I alleles	Position	Antigenicity
		HLA-B*3501 HLA-B*3801 HLA-B*3902 HLA-B7 HLA-Cw*0401		
12	KVGGNYNYL	HLA-A*0201 HLA-A*0205 HLA-A24 HLA-A68.1 HLA-B*2705 HLA-B*3501 HLA-B*3801 HLA-B*3902 HLA-B7 HLA-B*0702 HLA-Cw*0301 MHC-Db MHC-Db revised MHC-Kb	453-461	0.5994 (Probable Antigen).
13	RLFRKSNLK	HLA-A2 HLA-A*1101 HLA-A3 HLA-A*3101 HLA-A68.1 HLA-A20 Cattle HLA-B*2705	463-471	−0.2829 (Probable Nonantigen).
14	FERDISTEI	HLA-B*3701 HLA-B40 HLA-B*4403 HLA-B*5301 HLA-B*5401 HLA-B*51 HLA-B60 HLA-B61 MHC-Kk	473-481	−0.7442 (Probable Nonantigen).
15	EGFNCYFPL	HLA-A2 HLA-B14 HLA-B*3902 HLA-B40 HLA-B*5101 HLA-B*5102 HLA-B*5103 HLA-B*5401 HLA-B60 HLA-B7 HLA-Cw*0301 MHC-Dd MHC-Kb	493-501	0.5453 (Probable Antigen).
16	ELLHAPATV	HLA-A2 HLA-A*0201 HLA-A2.1 HLA-B*5103 HLA-B62	525-533	0.2109 (Probable Nonantigen).

TABLE 2 (Continued)

Serial no.	Epitopic sequence	MHC-I alleles	Position	Antigenicity
17	GPKKSTNLV	HLA-B*3501 HLA-B*5101 HLA-B*5102 HLA-B*5103 HLA-B*5301 HLA-B*5401 HLA-B*51 HLA-B61 HLA-B7 HLA-B*0702 HLA-B8 HLA-Cw*0401 MHC-Ld	535-543	0.6828 (Probable Antigen).
18	TEVPVAIHA	HLA-B*3701 HLA-B40 HLA-B*4403 HLA-B60 HLA-B61 MHC-Kk	627-635	0.2687 (Probable Nonantigen).
19	RVYSTGSNV	HLA-A2 HLA-A*0201 HLA-A*0205 HLA-A2.1 HLA-B*2702 HLA-B*2705 HLA-B*5102 HLA-B*5103 HLA-B*5201 HLA-B*5401 HLA-B*0702	643-651	0.2636 (Probable Nonantigen).
20	NSYECDIPI	HLA-B*2702 HLA-B*3501 HLA-B*5101 HLA-B*5102 HLA-B*5103 HLA-B*5401 HLA-B*5801 MHC-Db revised MHC-Kk	667-675	0.2216 (Probable Nonantigen).
21	SPRRARSVA	HLA-B*3501 HLA-B*5101 HLA-B*5301 HLA-B*5401 HLA-B*51 HLA-B7 HLA-B*0702 HLA-B8 MHC-Ld	689-697	0.7729 (Probable Antigen).
22	LGAENSVAY	HLA-B*3501 HLA-B*4403 HLA-B*51 HLA-B62	707-715	0.4173 (Probable Antigen).

(Continues)

TABLE 2 (Continued)

Serial no.	Epitopic sequence	MHC-I alleles	Position	Antigenicity
		HLA-Cw*0702 MHC-Dd		
23	KQIYKTPPI	HLA-A2 HLA-A*0201 HLA-A*0205 HLA-B*2702 HLA-B*2705 HLA-B*5102 HLA-B*5201 HLA-B62 HLA-B*0702 MHC-Dd MHC-Kd	795-803	0.2705 (Probable Nonantigen).
24	FIKQYGDCL	HLA-A2.1 HLA-B*3501 HLA-B*5301 HLA-B*5401 HLA-B*51 HLA-B7 HLA-B8	842-850	−0.4436 (Probable Nonantigen).
25	RNFYEPQII	HLA-B*2702 HLA-B*2705 HLA-B*5102 HLA-B*5201 HLA-B*5401	1116-1124	0.3282 (Probable Nonantigen).
26	VNNTVYDPL	HLA-A24 HLA-B*3701 HLA-B*3902 HLA-B*5301 HLA-B*51 HLA-B60 HLA-B7 HLA-Cw*0301 MHC-Kb	1142-1150	0.2397 (Probable Nonantigen).
27	ELDSFKEEL	HLA-A2 HLA-A3 HLA-A2.1 HLA-B*3801 HLA-B*3902 HLA-Cw*0401 HLA-Cw*0602	1153-1161	−0.6805 (Probable Nonantigen).
28	FKNHTSPDV	HLA-A2 HLA-A20 Cattle HLA-A2.1 HLA-B*5301 HLA-B*5401 HLA-B*51	1165-1173	0.4846 (Probable Antigen).
29	DEDDSEPVL	HLA-B*3701 HLA-B40 HLA-B*4403 HLA-B60 HLA-B61 MHC-Kk	1266-1274	0.5104 (Probable Antigen).

TABLE 3 List showing the epitopes with encountering MHC-II alleles, positional value and VaxiJen antigenic score

Serial no.	Sequence	Alleles	Position	VaxiJen score
1	IHVSGTNGT	DRB1_0306 DRB1_0307 DRB1_0308 DRB1_0311 DRB1_0401 DRB1_0404 DRB1_0410 DRB1_0421 DRB1_0423 DRB1_0426	77-85	0.8621 (Probable Antigen).
2	VYYHKNNKS	DRB1_0306 DRB1_0307 DRB1_0308 DRB1_0311 DRB1_0401 DRB1_0402 DRB1_0404 DRB1_0405 DRB1_0408 DRB1_0410 DRB1_0421 DRB1_0423 DRB1_0426 DRB1_1102 DRB1_1114 DRB1_1120 DRB1_1121 DRB1_1322 DRB1_1323 DRB1_1327 DRB1_1328 DRB1_1501 DRB1_1506	152-160	0.4510 (Probable Antigen).
3	LVRDLPQGF	DRB1_0301 DRB1_0305 DRB1_0306 DRB1_0307 DRB1_0308 DRB1_0309 DRB1_0311 DRB1_0421 DRB1_0426 DRB1_1107	221-229	0.1234 (Probable Nonantigen).
4	VFNATRFAS	DRB1_0301 DRB1_0305 DRB1_0309 DRB1_0802 DRB1_0804 DRB1_0813 DRB1_1101 DRB1_1102 DRB1_1104	350-358	0.1739 (Probable Nonantigen).

(Continues)

TABLE 3 (Continued)

Serial no.	Sequence	Alleles	Position	VaxiJen score
		DRB1_1106 DRB1_1107 DRB1_1114 DRB1_1120 DRB1_1121 DRB1_1301 DRB1_1302 DRB1_1304 DRB1_1307 DRB1_1311 DRB1_1322 DRB1_1323 DRB1_1327 DRB1_1328 DRB1_1501 DRB1_1506		
5	YRLFRKSNL	DRB1_0101 DRB1_0305 DRB1_0309 DRB1_0405 DRB1_0408 DRB1_0701 DRB1_0703 DRB1_0801 DRB1_0802 DRB1_0804 DRB1_0806 DRB1_0813 DRB1_0817 DRB1_1101 DRB1_1102 DRB1_1114 DRB1_1120 DRB1_1121 DRB1_1128 DRB1_1301 DRB1_1302 DRB1_1304 DRB1_1305 DRB1_1307 DRB1_1321 DRB1_1322 DRB1_1323 DRB1_1327 DRB1_1328 DRB1_1501 DRB1_1502 DRB1_1506	462-470	0.0522 (Probable Nonantigen).
6	FERDISTEI	DRB1_0305 DRB1_0401 DRB1_0426 DRB1_0309	473-481	−0.7442 (Probable Nonantigen).

TABLE 3 (Continued)

Serial no.	Sequence	Alleles	Position	VaxiJen score
		DRB1_0421 DRB1_0701 DRB1_0703		
7	YQTQTNSPR	DRB1_0421 DRB1_0401 DRB1_0405 DRB1_0408 DRB1_0426	683-691	−0.1787 (Probable Nonantigen).
8	FKNHTSPDV	DRB1_0101 DRB1_0309 DRB1_0401 DRB1_0405 DRB1_0421 DRB1_0426 DRB1_0701 DRB1_0703 DRB1_1114 DRB1_1120 DRB1_1302 DRB1_1323 DRB1_1502	1166-1174	0.4846 (Probable Antigen).

3.2 | Identification of B-cell epitopes

We obtained a total of 34 sequential linear B-cell epitopes of varying lengths from the IEDB server within spike glycoprotein of SARS-COV-2. Those B-cell epitopes were placed into Table 1 based on their positional value, sequence, and length. In Figure 1 the yellow-colored peaks represent the epitopic region, while the green-colored slopes, represent the nonepitopic region.

3.3 | Identification of T-cell epitopes and antigenicity analysis

We identified 29 MHC-I epitopes and 8 MHC-II epitopes, which fall within the preidentified B-cell epitopic region. Among them, 13 MHC-I epitopes and 3 MHC-II epitopes had the antigenic propensity, according to the VaxiJen v.2.0 server analysis. The MHC-I and MHC-II epitopes are listed in Tables 2 and 3 with encountering MHC alleles and antigenic scores.

3.4 | Vaccine construction, modeling, and validation

In this study, we linked the 13 MHC-I and 3 MHC-II antigenic epitopes with $(EAAAK)_3$ linker peptide to construct a vaccine component. This linker peptide was easily fused with the virus coat protein and increased stability as well as folding of the vaccine component.[22] The predicted structure of the vaccine component is shown in Figure 2. It has 90.0%, 7.1%, 1.6%, and 1.3% residues in most favored, additionally allowed, generously allowed and disallowed regions respectively within PROCHECK as the validation server to generate Ramachandran plot. Using the ProSA server, the "*Z*" score was −3.82 and most of the residues had negative energy value as shown in Figure 3. Results from both servers indicate the model is in a good quality.[23,24]

3.5 | Molecular docking analysis

The PatchDock server provided 20 docking complexes, and among them, we selected only the docking complex with the highest negative atomic contact energy (ACE) value for analysis. The ACE value of

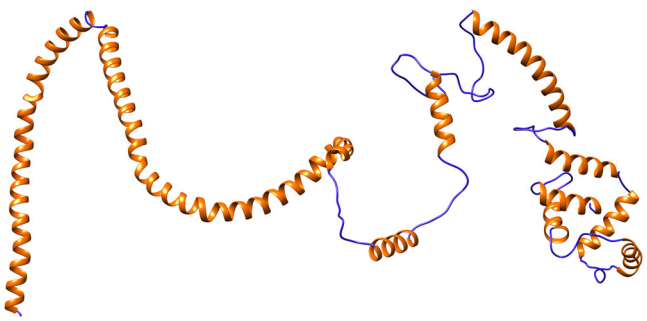

FIGURE 2 Tertiary structural model of construct vaccine component

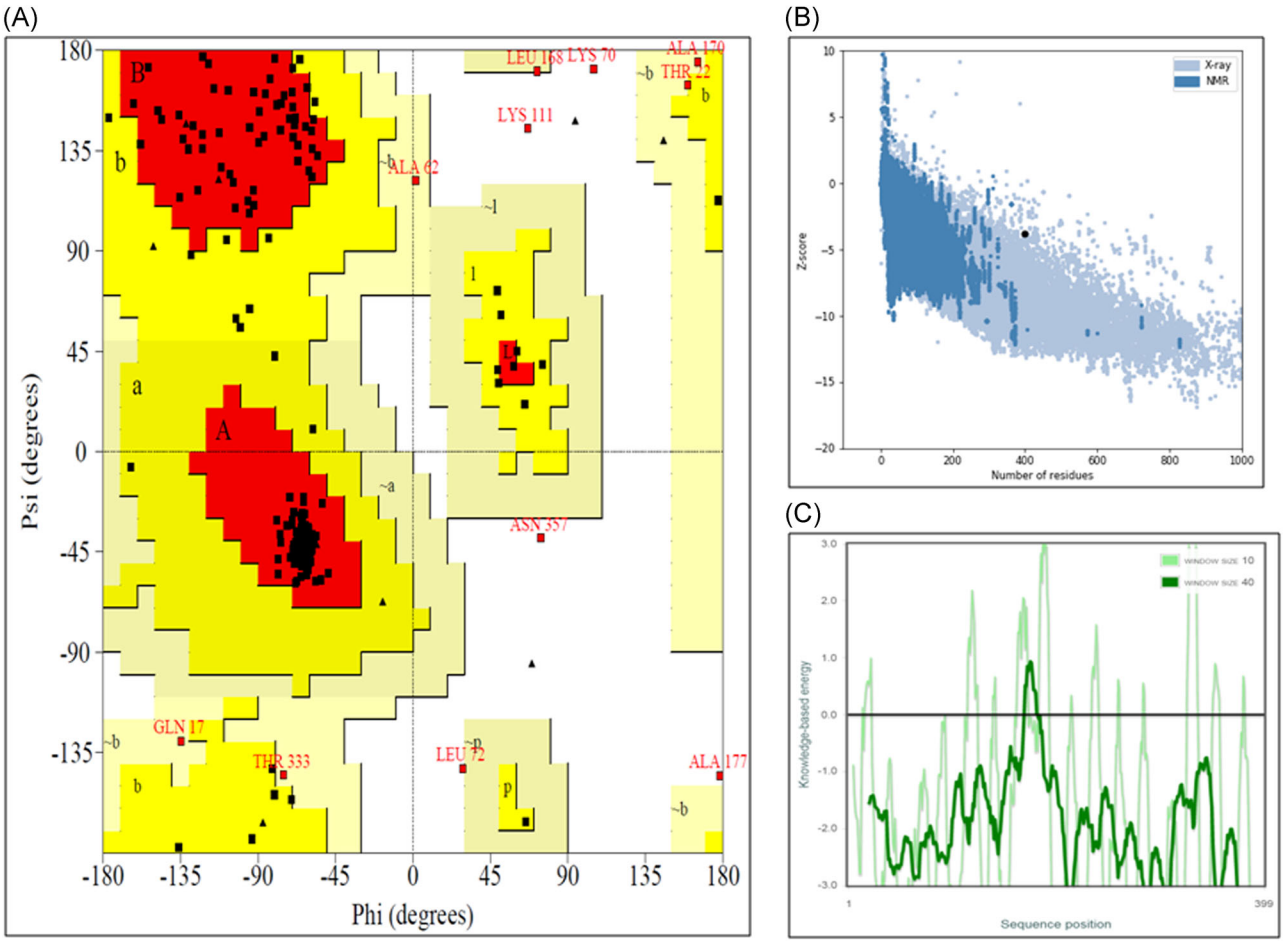

FIGURE 3 Different molecular characterization of vaccine model. (a) All atoms at Ramachandran plot, (b) "*Z*" score plot of vaccine model in ProSA server, and (c) all residue energy plot

the docking complex was −259.62, which indicates spontaneous reactivity between the vaccine component and TLR-5.[25] As proper protein-protein docking regulates the cellular functions, the docking between the vaccine component and TLR-5 will activate immune cascades for destroying the viral antigens.[26] The selected docking complex is shown in Figures 4 and 5, along with molecular surface interaction as well as some bonding interactions.

4 | DISCUSSION

The SARS-COV-2, the causative pathogen for respiratory distress syndrome, led more than 10 000 people to infection all over the world, even several to death. After first identified in Wuhan, Hubei province of China, the COVID-19 disease spread unchecked which finally became a global threat. Scientists from all over the world are struggling to find a solution to this evil outbreak.

In our present study, we attempted to find out various B-cell and T-cell epitopes against SARS-COV-2, using the immunoinformatics, as quick identification of B-cell and T-cell epitopes is crucial for designing of vaccine component against this disease. The spike

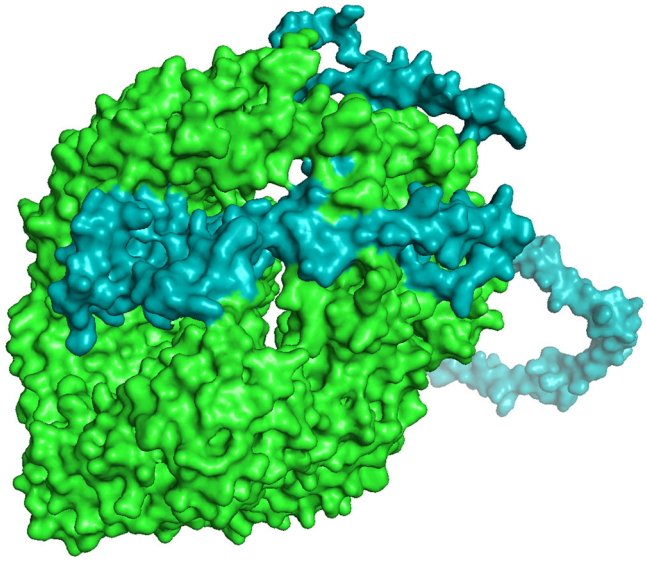

FIGURE 4 Docking complex exhibiting the surface interaction between vaccine component (cyan color) and toll-like receptor-5 (green color)

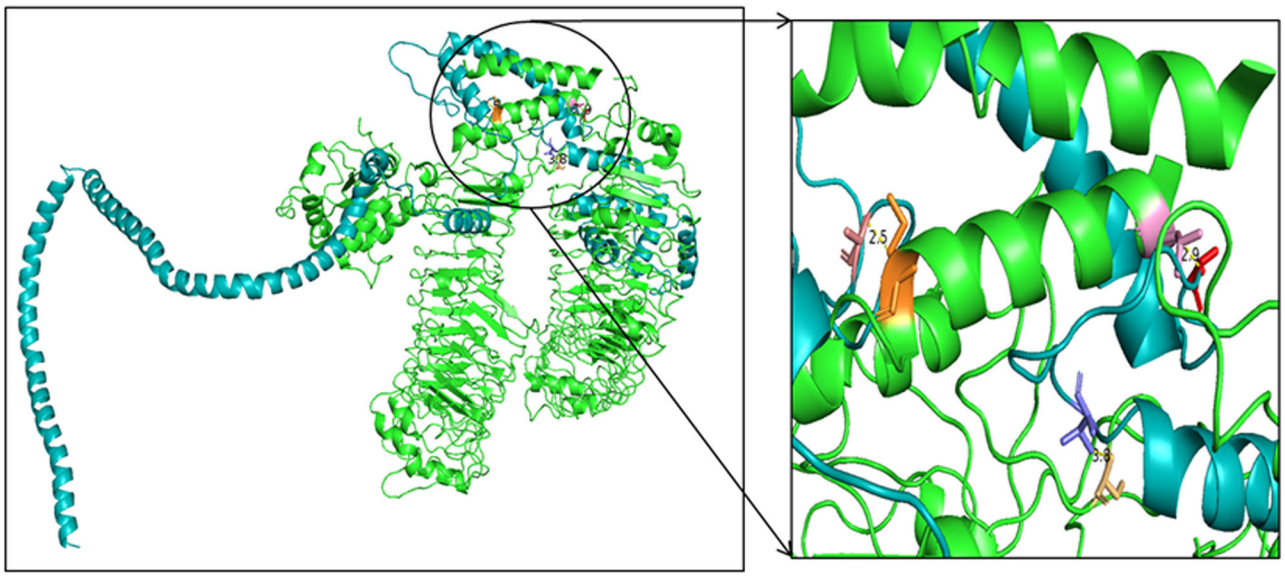

FIGURE 5 Docking complex exhibiting the bonding interaction between vaccine component and toll-like receptor-5

glycoprotein was analyzed for B-cell epitope identification in the IEDB server, and 34 linear B-cell epitopes were identified as a result. Subsequently, the sequence was also analyzed in ProPred-I and ProPred servers for the identification of the T-cell epitope that can combine with MHC-I and MHC-II molecules. Fortunately, we found 29 epitopes against MHC-I and 8 epitopes against MHC-II that can be possibly used for vaccine. Unfortunately, antigenic characterization in VaxiJen v.2.0 discarded 16 MHC-I epitopes out of 29 and 5 out of 8 MHC-II epitopes as these seemed to be nonantigenic in nature. Nevertheless, we converted the antigenic epitopes into a single vaccine component, using $(EAAAK)_3$ peptide linker.

Later, the vaccine component was modeled in the SPARK-X server and validated in PROCHECK and ProSA. A total of 90% of nonglycine and nonproline residues presented within the most favored region, while the "*Z*" score of the model was −3.82. These results from both servers indicate the model is in good quality. Molecular docking between vaccine component and TLR-5 showed significant ACE value, which indicates spontaneous reactivity within the receptor-ligand complex.

All the observations of our present work depict the effectiveness of selected epitopes within the spike glycoprotein of SARS-COV-2. These epitopes can be used to make an immunogenic multi-epitopic peptide vaccine against SARS-COV-2.

5 | CONCLUSION

Present immunoinformatic analysis pointed out 13 MHC-I and 3 MHC-II epitopes within the spike glycoprotein of SARS-COV-2. These epitopes are the ideal candidate to formulate a multi-epitopic peptide vaccine, not only because of being selected from the linear B-cell epitopic region but also because of their antigenic property was confirmed. Moreover, the molecular docking of vaccine components with the TLR-5 proves the significance and effectiveness of these epitopes as an ideal vaccine candidate against SARS-COV-2. However, these immunoinformatic analyses require several in vitro and in vivo validations before formulating the vaccine to resist COVID-19.

ACKNOWLEDGMENTS

This study was supported by Hallym University Research Fund and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1A2B4012944).

ORCID

Manojit Bhattacharya http://orcid.org/0000-0001-9669-1835
Ashish R. Sharma http://orcid.org/0000-0003-3973-6755
Prasanta Patra http://orcid.org/0000-0003-3901-8569
Pratik Ghosh http://orcid.org/0000-0001-8117-7681
Garima Sharma http://orcid.org/0000-0003-1763-0730
Bidhan C. Patra http://orcid.org/0000-0002-9188-7457
Sang-Soo Lee http://orcid.org/0000-0001-5074-7581
Chiranjib Chakraborty http://orcid.org/0000-0002-3958-239X

REFERENCES

1. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *The Lancet*. 2020;395:470-473.
2. Zhou P, Yang X-L, Wang X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020:1-4.
3. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020;382:727-733.
4. Hui DS, I Azhar E, Madani TA, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—the latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis*. 2020;91:264-266.
5. Chan JF-W, Yuan S, Kok K-H, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating

person-to-person transmission: a study of a family cluster. *The Lancet*. 2020;395:514-523.
6. Carlos WG, Dela Cruz CS, Cao B, Pasnick S, Jamil S. Novel Wuhan (2019-nCoV) coronavirus. *Am J Respir Crit Care Med*. 2020;201: P7-P8.
7. Perlman S. Another decade, another coronavirus. *N Engl J Med*. 2020; 382(8):760-762.
8. Phelan AL, Katz R, Gostin LO. The Novel coronavirus originating in Wuhan, China: challenges for global health governance. *JAMA*. 2020; 323:709.
9. Yin D, Li L, Song X, et al. A novel multi-epitope recombined protein for diagnosis of human brucellosis. *BMC Infect Dis*. 2016;16(1):219.
10. Chung M, Bernheim A, Mei X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology*. 2020:200230.
11. NCBI Resource Coordinators. Database resources of the National Center For Biotechnology Information. *Nucleic Acids Res*. 2016; 44(database issue):D7.
12. Kim Y, Ponomarenko J, Zhu Z, et al. Immune epitope database analysis resource. *Nucleic Acids Res*. 2012;40(W1):W525-W530.
13. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45(W1): W24-W29.
14. Singh H, Raghava G. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics*. 2003;19(8):1009-1014.
15. Singh H, Raghava G. ProPred: prediction of HLA-DR binding sites. *Bioinformatics*. 2001;17(12):1236-1237.
16. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8(1):4.
17. Yang Y, Faraggi E, Zhao H, Zhou Y. Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics*. 2011;27(15): 2076-2082.
18. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35(suppl_2):W407-W410.
19. Laskowski R, MacArthur M, Thornton J PROCHECK: validation of protein-structure coordinates. 2006.
20. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*. 2005;33(suppl_2):W363-W367.
21. DeLano WL. Pymol: an open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography*. 2002;40(1):82-92.
22. Chen X, Zaro JL, Shen W-C. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev*. 2013;65(10):1357-1369.
23. Al-Moubarak E, Simons C. A homology model for *Clostridium difficile* methionyl tRNA synthetase: active site analysis and docking interactions. *J Mol Model*. 2011;17(7):1679-1693.
24. Patra P, Ghosh P, Patra BC, Bhattacharya M. Biocomputational analysis and in silico characterization of an angiogenic protein (RNase5) in zebrafish (*Danio rerio*). *Int J Pept Res Ther*. 2019:1-11.
25. Ramanathan K, Shanthi V, Sethumadhavan R. In silico identification of catalytic residues in azobenzene reductase from *Bacillus subtilis* and its docking studies with azo dyes. *Interdiscip Sci: Comput Life Sci*. 2009; 1(4):290-297.
26. Lavi A, Ngan CH, Movshovitz-Attias D, et al. Detection of peptide-binding sites on protein surfaces: The first step toward the modeling and targeting of peptide-mediated interactions. *Proteins: Struct, Funct, Bioinf*. 2013;81(12):2096-2105.

How to cite this article: Bhattacharya M, Sharma AR, Patra P, et al. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. *J Med Virol*. 2020;1–14.
https://doi.org/10.1002/jmv.25736