Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Immunobiology 226 (2021) 152134

Contents lists available at ScienceDirect

Immunobiology

journal homepage: www.elsevier.com/locate/imbio

In-silico immunoinformatic analysis of SARS-CoV-2 virus for the development of putative vaccine construct

Abhishek Sharma [a], Surinder Pal [a], Anil Panwar [a], Suresh Kumar [b], Ashok Kumar [a,*]

[a] Centre for Systems Biology and Bioinformatics, Panjab University, Chandigarh 160014, India
[b] University School of Biotechnology, Guru Gobind Singh Indraprastha University, Dwarka, Sector 16C, New Delhi 110075, India

ARTICLE INFO

Keywords:
COVID-19
Immunoinformatics
Vaccine construct
Docking
Molecular simulations

ABSTRACT

COVID-19 (CoronaVirus disease 2019) is caused by the SARS-CoV-2 virus (severe acute respiratory syndrome corona virus 2). SARS-CoV-2 virus is highly contagious and affects the human respiratory tract resulting in symptoms such as high fever, body ache, cough, dysfunctions of tastebuds and smelling sense of body. The objective of the present study involves immunoinformatic analysis to predict COVID-19 protein for vaccine construct based on the genomic information SARS-CoV-2 virus. At present, as per WHO estimates, around 133 COVID-19 novel vaccines under development. Three amino acid sequences of SARS-CoV-2 were retrieved from the NCBI database for the analysis of vaccine construct. This study involves computational and immunoinformatic methods. The Immunoinformatic tools used in the present study are NetCTL server, IFN epitope server, Toxin PRED, BCPred, CTL + HTL + ADJUVANTS + LINKERS, AlgPredserver, VaxiJenserver, ProtParam to predict vaccine construct. The secondary and tertiary structure prediction is done by PSIPRED, I-TASSER, Galaxy refine, prosA + Ramachandran. Finally, docking of the vaccine constructs and ligand was done with the help of Cluspro 2.0.

C-ImmSimm webserver to simulate the potential vaccine construct. The present study demonstrated three potential Vaccine constructs for the SARS-CoV-2 virus, which were docked with TLR8 (Toll-likereceptor8). Interestingly from these, all constructs one having a high potential for the inhibition effect of the SARS-CoV-2virus. Immunological simulation data shows significant elevated amount of memory B cell; also, the high response was seen in TH(Helper) and TC(cytotoxic) cell population from the vaccine construct proposed in the current study. Hence, these constructs are suitable vaccine candidates that might be useful in developing a novel vaccine.

1. Introduction

Coronavirus disease 19 (COVID-19) is caused by the SARS-CoV-2 virus responsible for pandemic across the world. The symptoms associated with COVID-19 disease are fever, body ache, high body temperature, cough, and breathlessness (Struyf et al., 2020). The severity of COVID-19 causes accumulation of mucous in the respiratory tract resulting in respiratory failure leading to the death of the patient (Wang, 2020). The primary risk associated with SARS-CoV-2 virus infection is that it is highly contagious and spread very fast from human to human contact (Sanche et al., 2020) and there are some studies which describe that virus could affect other animals too, some are pet animals (Salajegheh Tazerji, 2020).The proneness of SARS-CoV-2 through ACE2 receptor in chicken is very less, while horse shows strong interactions (Kumar et al., 2020). The genome of the SARS-CoV-2 virus consists of around 30 K base pairs. The 5′ end of the genome comprises more than two-thirds of the total genome. It code for ORF1 ab poly protein. It is the largest coding region made up of around 21 K base pairs. The remaining one-third of the genome comprises 3′ end. The 3′ end comprises gene encoding structural proteins, including surface proteins, envelope proteins, membrane proteins, and nucleocapsid proteins. Gene coding for envelope proteins is made up of 228 base pairs. The membrane protein gene is 669 nucleotides long. Nucleocapsid N is the biggest structural gene having 908 nucleotides (Khailany et al., 2020). The viral outer layer is formed by S-protein formed; then protein helps in the viral replication, genome construction, and host cellular response; envelope shape is determined by the M-protein and maturation of the SARS-CoV-2 virus is done by the E-protein (Astuti and Ysrafil, 2020). Tilocca, B., *et al*

* Corresponding author.
E-mail addresses: sureshkumar@ipu.ac.in (S. Kumar), ashokkumar@pu.ac.in (A. Kumar).

https://doi.org/10.1016/j.imbio.2021.152134
Received 12 January 2021; Received in revised form 26 May 2021; Accepted 24 August 2021
Available online 28 August 2021

describe the significance of predicting B cell epitopes and T cell epitopes as major proteins, which play important role in the immunogenicity towards SARS-CoV-2 envelop protein, membrane protein, spike protein and nucleocapsid protein (Tilocca et al., 2020).

Based on this genomic information, a new generation of vaccines can be developed to stimulate both humoral and cellular responses. The various factors associated with the successful development of vaccines include the production of long term T and B memory cells, immunity mediated by neutralizing antibodies, and the stable nature of antigens. Cytotoxic T lymphocyte (CTL) response plays a significant role in identifying the virus infected cells by recognizing specific epitopes of viral proteins. Epitopes are particular regions of proteins with short amino acid sequences that can impart direct and indirect immune responses (Ju et al., 2020; Braun, 2020).

The patients who recover from COVID-19 disease has an evoked T-cells immune response made by T-cell epitopes of SARS-CoV-2 spike protein, and most of these immunogenic epitopes were localized to the S protein of the virus (Astuti and Ysrafil, 2020). The size of T-cell epitopes can vary from 8 to 20 amino acids, whereas the B-cell epitopes can be larger than T-cell epitopes (Dermime et al., 2004; Meloen et al., 2001).

T cell produces cellular immunity, which is vital for the functioning of a vaccine. T lymphocytes are having crucial participation in cellular and antibody-mediated immunity. CD3+ T cells, CD4+ T cells, CD8+ T cells, natural killer cells, B cells, and T lymphocytes all identify the antigen, a peptide in nature (XU et al., 2020). T helper cells response having inter connection with the production of neutralizing antibodies (Janice Oh et al., 2012). Earlier SARS-COV research demonstrated that acquired immune response against spike glycoprotein is defensive (Zhu et al., 2007). So, CD4+ T Cells are beneficial in the production and maturation of antibodies for SARS-CoV-2 (Ju et al., 2020). B cells are also a major component of the immune system against any viral infection and play an essential role in the generation of antibodies and provide immune protection. The immunological role of B cells in COVID-19 is still not much explored, so more studies could be conducted for the same (Creed, 2020).

Multi epitope peptide vaccine including T cell and B cell epitopes in a combined form that can generate a robust immune response. The significant advantages are easy manufacturing, physiochemical stability, and fewer adverse effects than conventional vaccines. This ideal characteristic makes peptide and epitope vaccine most preferable for vaccine construct (Patronov and Doytchinova, 2013). In the se multi-epitope peptide vaccine construction, prediction of B cell, cytolytic T cell, helper T cell epitopes are playing an important role (Chiarella, 2009). The whole protein sequence is not responsible for stimulating the immunogenic response, whereas only antigenic epitopes are responsible for the same, and these epitopes are easy to predict (Zheng et al., 2017). Tilocca, B., *et al* studied the *Mycobacterium avium Paratuberculosis* proteins which causes paratuberculosis, finding of the epitope binding sites which further helps in the early detection and the vaccine construct for the disease (Tilocca et al., 2020). Sarkar, B., *et al* has taken four different proteins from SARS-CoV-2 virus, and made the three possible construct for the vaccines (Sarkar et al., 2020). In the Nucleocapsid protein, at three different places i.e. 229 to 268, 349 to 399 and 405 to 419, these all amino acids having B cell epitopes and MHC II binding epitopes (Dai et al., 2020). Humoral and cell mediated immunity generated against virus by using peptide based multi epitope vaccine (Rahman et al., 2020). In spike glycoprotein there are 13 epitopes in MHC-I and 3 epitopes in MHC-II. Docking interactions with TLR 5 demonstrate these possibilities (Bhattacharya et al., 2020). Vaccine construct for these epitopes in spike glycoproteins are made with help of *in-silico* method (Bhattacharya, 2020). Some phytochemicals like sarsasapogenin, ursonic acid, curcumin, ajmalicine, novobiocin, silymarin and aranotin, piperine, gingerol, rosmarinic acid, and alpha terpinyl acetate also helpful in inhibition of viral infection by acting on Non structural protein 15 (Kumar, 2021).

The objective of the present study involves *in-silico* peptide vaccine construct based on the genomic information of the SARS-CoV-2 virus. Peptide based vaccines are biologically safe as they don't need in vitro culture, and selectivity activates immune responses. B-cell and T-cell epitopes that are immunodominant and can induce specific immune responses are constructed using immunoinformatic tools. The present study aimed to construct a peptide-based vaccine against SARS-CoV-2 virus using computational and immunoinformatic tools.

2. Material and methods

2.1. Retrieval of sequences

Three amino acid sequences of SARS-CoV-2 globally known as COVID-19 were retrieved from NCBI in FASTA format with accession number YP_009724390.1, YP_009724393.1 andYP_009724392.1.

2.2. CTL (cytotoxic T-cell) epitope prediction

NetCTL1.2 web server is used to decipher CTL epitopes in all protein sequences taken for study (Larsen et al., 2007). MHC-I binding peptide prediction, proteasomal C terminal cleavage, and the transportation efficiency Transporter Associated with antigen Processing (TAP) is the foremost essential part of this analysis. Proteasomal C terminal cleavage and MHC binding are predicted through artificial neural networks; on the other hand, the efficiency of TAP transporter is done through the weight matrix method. For finding the CTL epitopes, 0.75 was the adjusted threshold.

2.3. HTL (helper T-cell) epitope prediction

IEDB online server was used to predict epitopes for proteins of the SARS-CoV-2virus (Zhu et al., 2007). Server predicts Five epitopes of 15-mer length for human alleles (HLADRB1*01:01, HLA-DRB1*01:02, HLA-DRB1*01:03, HLA-DRB1*01:04, HLA-DRB1*01:05). IC50 score is used for peptide affinity for receptors.

2.4. Toxicity prediction

ToxinPred online server is used to find out the toxicity of selected epitopes. Only non-toxic epitopes are taken further studies, physiochemical assets are the critical point in the toxicity prediction (Gupta et al., 2013).

2.5. B cell epitope prediction

B cell has a crucial job in making host antibodies. BCPred server is used to predict the B cell epitopes (EL-Manzalawy et al., 2008). This web portal consists of a machine learning kernel algorithm that is very helpful in finding B Cell epitopes. BCPred makes predictions based on SVM with the deployment of Amino acid pair antigenicity (Patronov and Doytchinova, 2013).

2.6. Preparation of multi-epitope vaccine constructs

From earlier results of CTL, HTL, and B cell epitopes, the multi-epitope vaccine sequence is made by using linkers. The different epitopes were linked together using AAY, GPGPG, and KK linkers. To improve the immunogenicity of the vaccines, three different adjuvants were chosen and retrieved from NCBI protein Database WP_003403353. 50S ribosomal protein L7/L12 Mycobacterium (Lee et al., 2014), AAQ97601.1 human beta-defensin-3 Mature peptide, partial synthetic construct (Mei et al., 2012), AGV15514. 1heparin-binding hemagglutinin Mycobacterium avium sub spp. Paratuberculosis (Rana and Akhter, 2016) and were added at the N-terminal through an EAAAK linker.

Table 1
Predicted CTL epitopes for surface glycoproteins. Finally, the selection is made w.r.t to c-terminal, TAPscores antigenicity and toxicity.

Residue number	Peptide sequence	MHC Binding affinity	Rescale binding affinity	C terminal cleavage affinity	Transport affinity	Prediction score	Antigenic	Toxicity
865	LTDEMIAQY	0.7953	3.3768	0.9723	2.779	3.6616	Non antigenic	Non–Toxin
258	WTAGAAAYY	0.6735	2.8596	0.7339	2.863	3.1128	Antigenic	Non toxin
604	TSNQVAVLY	0.6559	2.7847	0.944	2.991	3.0758	Antigenic	Non toxin
361	CVADYSVLY	0.5348	2.2705	0.9764	3.18	2.5759	Non antigenic	Non toxin
733	KTSVDCTMY	0.4908	2.084	0.9649	3.016	2.3795	Antigenic	toxin
746	STECSNLLL	0.5136	2.1808	0.8879	0.703	2.3492	Antigenic	Non toxin
652	GAEHVNNSY	0.4042	1.7163	0.9769	2.663	1.996	Non antigenic	Non toxin
196	NIDGYFKIY	0.3921	1.6649	0.9664	3.015	1.9606	Non antigenic	Non toxin
160	YSSANNCTF	0.3975	1.6878	0.9032	2.598	1.9531	Non antigenic	Non toxin
152	WMESEFRVY	0.3902	1.6569	0.7993	2.929	1.9232	Non antigenic	Nontoxin
162	SANNCTFEY	0.3737	1.5865	0.9196	2.99	1.8739	Non antigenic	Non toxin
687	VASQSIIAY	0.3529	1.4986	0.9656	3.089	1.7978	Non antigenic	non toxin
30	NSFTRGVYY	0.3389	1.4389	0.6421	3.124	1.6915	Non antigenic	Non toxin
136	CNDPFLGVY	0.2613	1.1095	0.69	2.45	1.3355	antigenic	Non toxin
392	FTNVYADSF	0.2704	1.148	0.38	2.317	1.3208	Non antigenic	Non toxin
261	GAAAYYVGY	0.2253	0.9568	0.7608	2.969	1.2194	Antigenic	Non toxin
357	RISNCVADY	0.2106	0.8941	0.9292	3.394	1.2032	Non antigenic	Non toxin
465	ERDISTEIY	0.2097	0.8903	0.9744	2.646	1.1687	Non antigenic	Non toxin
285	ITDAVDCAL	0.235	0.9979	0.8708	0.79	1.168	Antigenic	Non toxin
1039	RVDFCGKGY	0.2036	0.8644	0.7618	3.232	1.1403	Non antigenic	Non toxin
343	NATRFASVY	0.1955	0.83	0.9342	2.873	1.1138	Non antigenic	Non toxin
1237	MTSCCSCLK	0.226	0.9595	0.7525	0.479	1.0963	Antigenic	Toxin
50	STQDLFLPF	0.1974	0.8383	0.553	2.511	1.0468	Antigenic	Non toxin
1096	VSNGTHWFV	0.2012	0.8544	0.6143	0.218	0.9574	Non antigenic	Non toxin
880	GTITSGWTF	0.1656	0.7031	0.7489	2.557	0.9433	Non antigenic	Non toxin
815	RSFIEDLLF	0.1421	0.6035	0.5938	3.032	0.8441	Non antigenic	Non toxin
1264	VLKGVKLHY	0.1262	0.5356	0.9783	2.859	0.8253	Antigenic	Non toxin
748	ECSNLLLQY	0.1413	0.6	0.5316	2.747	0.8171	Non antigenic	Non toxin
370	NSASFSTFK	0.1671	0.7093	0.5456	0.507	0.8165	Non antigenic	Non toxin
372	ASFSTFKCY	0.118	0.501	0.9587	3.275	0.8085	Non antigenic	Non toxin
628	QLTPTWRVY	0.1189	0.5047	0.9661	2.782	0.7887	Antigenic	Non toxin
296	LSETKCTLK	0.1515	0.6432	0.8919	0.22	0.7879	Antigenic	Non toxin
192	FVFKNIDGY	0.1358	0.5767	0.4093	2.913	0.7837	Non antigenic	Non toxin
445	VGGNYNYLY	0.1164	0.4941	0.9518	2.658	0.7698	Antigenic	Non toxin
83	VLPFNDGVY	0.113	0.4797	0.9703	2.846	0.7675	Antigenic	Non toxin
1095	FVSNGTHWF	0.1232	0.5231	0.7203	2.621	0.7622	Non antigenic	Non toxin
612	YQDVNCTEV	0.1531	0.6501	0.587	0.242	0.7502	Antigenic	Non toxin

2.7. Allergenicity prediction of the vaccine

AlgPred server issued to predict the allergenicity of the vaccine construct (Saha and Raghava, 2006). The most accurate results of this server are because of its six various methods in prediction; 85% accuracy is obtained with a 0.4 threshold value.

2.8. Antigenicity prediction of the vaccine

VaxiJen server isused to decipher the antigenicity of the vaccine construct (Doytchinova and Flower, 2007).Results are predicted on the basis of physiochemical assets of amino acid sequence. The virus is set as the model organism with a threshold value of 0.4 to get optimum results.

2.9. Secondary structure prediction

All physiochemical properties like nature of amino acids, half-life, PI, hydropathy plot, the molecular weight of sequence, instability index, in vitro and *in vivo* half-life, and grand average of hydropathicity are calculated with the help of online web server ProtParam (Wilkins, 1999).

Table 2
Immunogenicity results of the non-toxic + antigenic epitopes. The first three based on top scores were taken.

Peptide	Length	Scores
QLTPTWRVY	9	0.31555
WTAGAAAYY	9	0.15259
CNDPFLGVY	9	0.15232
VLPFNDG	7	0.12878
GAAAYYVGY	9	0.09963
ITDAVDCAL	9	0.08501
YQDVNCTEV	9	0.08295
STQDLFLPF	9	0.06828
GAEHVNNSY	9	−0.00296
TSNQVAVLY	9	−0.01327
VGGNYNYLY	9	−0.0148
KTSVDCTMY	9	−0.11115
LSETKCTLK	9	−0.16291
VLKGVKLHY	9	−0.18916
STECSNLLL	9	−0.20478
MTSCCSCLK	9	−0.36816

2.10. Tertiary structure prediction

I-TASSER (Iterative Threading Assembly Refinement) server was used to do homology modeling of the final constructs of vaccine peptides (Yang and Zhang, 2015). The 3D structure was generated from the FASTA format of amino acid sequence. This model is developed through multiple alignments and built simulated iterative structures.

2.11. Tertiary structure refinement and validation

Galaxy refine webserver was used in the refinement of the tertiary structure of the vaccine peptide Construct (Heo et al., 2013). Validation of the tertiary structure is done by pros A + Ramachandran. Ramachandran plot is also helping validate the 3D structure of a protein by plotting a graph that deciphers the presence of all nature and structure of the amino acid sequence. ProsA analyzed the 3D structure of the protein, whether it is showing relatedness to the native protein structure or having some mistakes in the structure (Wiederstein and Sippl, 2007).

2.12. Docking with TLRL (Toll-like receptor ligand)

TLR Docked with three vaccine constructs (A, B, C) using ClusPro protein-protein docking server. Balanced docking method was chosen using PIPER algorithm. PIPER algorithm represents the interaction energy between two proteins using an expression of the form E = w1Erep + w2Eattr + w3Eelec + w4EDARS, where Erep and Eattr denote the repulsive and attractive contributions to the van der Waals interaction energy, and Eelec is an electrostatic energy term. EDARS is a pairwise structure-based potential constructed by the Decoys as the Reference State (DARS) approach. The coefficients w1, w2, w3, and w4 define the weights of the corresponding terms. The best-Docked Structures with Coefficient Weights and Cluster Scores were retrieved and taken for further reference.

Table 4
Immunogenicity results of the non-toxic + antigenic epitopes. Epitopes the basis of top scores were taken.

Peptide	Length	Scores
AGDSGFAAY	9	0.03981
LVGLMWLSY	9	−0.06867
ATSRTLSYY	9	−0.11604
VATSRTLSY	9	−0.17295

Table 5
Predicted T-cell epitopes for Envelope protein. Finally, selection is done w.r.t to c-terminal, TAP scores antigenicity and toxicity.

Peptide	Length	Scores
LVKPSFYVY	9	−0.11106
VSLVKPSFY	9	−0.25372

2.13. Immune simulation

CImmSim server was used to characterize the immunogenicity and immune response profile of the chimeric peptide. This server also conducts *in-silico* immune simulations (Doytchinova and Flower, 2007). A position-specific scoring matrix (PSSM) used by this server for the prediction of immune epitope and immune interactions are calculated by machine learning techniques.

3. Results

3.1. Collection of proteins sequences for B and T-cell epitopes prediction

COVID-19 proteins (surface, envelope, and the membrane) amino acid sequences were retrieved from NCBI. The amino acid sequences were used to predict the B and T cell epitopes for designing the multi-epitope sub unit vaccine.

3.2. Cytotoxic T lymphocytes (CTL) epitopes prediction

Prediction of CTL epitopes was made by using an online tool, NetCTL1.2. Total 37 CTL epitopes of length 9-mer were predicted. From37epitopes by analyzing toxicity prediction and immunogenicity prediction, three epitopes were selected, which were non-toxic, antigenic (Table 1), and highly immunogenic (Table 2). For membrane protein, a total of eight CTL epitopes were predicted. Out of eight, only one epitope is selected on the basis of non-toxic, antigenic (Table 3) and highly immunogenic (Table 4). For envelope protein total of three

Table 3
Predicted T-cell epitopes for membrane glycoprotein. Finally, the selection is made w.r.t to c-terminal, TAP scores antigenicity and toxicity.

Residue number	Peptide sequence	MHC Binding affinity	Rescale binding affinity	C terminal cleavage affinity	Transport affinity	Prediction score	Antigenic	Toxicity
182	SSDNIALLV	0.6531	2.7729	0.9682	0.286	2.9325	Non–antigen	TOXIN
140	ATSRTLSYY	0.5463	2.3195	0.9375	3.09	2.6146	ANTIGEN	NON – TOXIN
165	YSRYRIGNY	0.3214	1.3648	0.9345	3.148	1.6623	NON–ANTIGEN	TOXIN
8	YANRNRFLY	0.3305	1.4031	0.4099	3.017	1.6155	NON–ANTIGEN	TOXIN
139	VATSRTLSY	0.2752	1.1684	0.9679	3.013	1.4642	ANTIGEN	NON –TOXIN
56	LVGLMWLSY	0.2694	1.144	0.724	2.897	1.3974	ANTIGEN	NON –TOXIN
157	AGDSGFAAY	0.1341	0.5695	0.9652	2.673	0.848	ANTIGEN	NON – TOXIN
181	SSSDNIALL	0.1487	0.6313	0.9639	1.098	0.8308	NON–ANTIGEN	TOXIN

Table 6
Immunogenicity results of the non-toxic + antigenic epitopes. Epitopes the basis of top scores were taken.

Residue number	Peptide sequence	MHC	Rescale binding affinity	C terminal cleavage affinity	Transport affinity	Prediction score	Antigenic	Toxicity
		Binding affinity						
34	LTALRLCAY	0.5594	2.3751	0.6272	2.933	2.6158	Non - Antigen	Toxin
49	VSLVKPSFY	0.3533	1.4999	0.3714	3.186	1.7149	Antigen	Non toxin
51	LVKPSFYVY	0.1343	0.5702	0.9767	3.119	0.8726	Antigen	Non toxin

Table 7
Helper T-Cell epitopes for surface protein of COVID-19 using IEDB MHC-II module, AlgPred, ToxinPred and VaxiJen server for identification of non-allergens non-toxic antigenic HTL epitopes.

Allele	start	end	length	Non Allergens peptide sequence	Percentile rank	Alegpred score	antigenic	toxicity	If n+/-
HLA-DRB5*01:01 1	235	249	15	ITRFQTLLALHRSYL	0.26	−0.41	Non Antigenic	Toxic	
HLA-DRB5*01:01 1	234	248	15	NITRFQTLLALHRSY	0.32	−0.45	Non Antigenic	Toxic	
HLA-DRB5*01:01 1	232	246	15	GINITRFQTLLALHR	0.52	−0.48	Antigenic	Non Toxic	Ifn+
HLA-DRB5*01:01 1	233	247	15	INITRFQTLLALHRS	0.32	−0.46	Antigenic	Non Toxic	
HLA-DRB3*01:01 1	209	223	15	PINLVRDLPQGFSAL	0.49	−0.78	Antigenic	Non Toxic	
HLA-DRB5*01:01 2	64	78	15	ATRFASVYAWNRKRI	0.49	−0.43	Non Antigenic	Toxic	
HLA-DRB5*01:01 2	65	79	15	TRFASVYAWNRKRIS	0.52	−0.55	Non Antigenic	Toxic	
HLA-DRB3*01:01 1	207	221	15	HTPINLVRDLPQGFS	0.51	−0.98	Non Antigenic	Toxic	
HLA-DRB3*01:01 1	210	224	15	INLVRDLPQGFSALE	0.51	−0.48	Antigenic	Non Toxic	
Non Overlapping Sequences									
GINITRFQTLLALHR									
PINLVRDLPQGFSAL									

Table 8
Helper T-Cell epitopes for Membrane protein of COVID-19using IEDB MHC-II module, Algpred, Toxinpred and Vaxijen server for identification of non-allergens non-toxic antigenic HTL epitopes.

Allele	Start	End	Length	Non Allergens peptide sequence	Percentile rank	antigenic	toxicity	Ifn+/-
HLA-DRB1*07:01	166	180	15	KEITVATSRTLSYYK	2.1	Antigenic	Non Toxin	
HLA-DRB3*02:02	175	189	15	TLSYYKLGASQRVAG	2.1	Antigenic	Non Toxin	
HLA-DRB1*07:01	164	178	15	LPKEITVATSRTLSY	2.2	Antigenic	Non Toxin	
HLA-DRB1*07:01	88	102	15	VGLMLSYFIASFRL	3.4	Antigenic	Non Toxin	
HLA-DRB5*01:01	174	188	15	RTLSYYKLGASQRVA	6.4	Antigenic	Non Toxin	Ifn+
Non Overlapping Sequences								
VGLMWLSYFIASFRL								
RTLSYYKLGASQRVA								

Table 9
Helper T-Cell epitopes for Envelope protein of COVID-19 using IEDB MHC-II module, Algpred, Toxinpred and Vaxijen server for identification of non-allergens non-toxic antigenic HTL epitopes.

Allele	Start	End	Length	Non Allergens peptide sequence	Percentile rank	Alegpred	antigenic	toxicity	Ifn+/-
HLA-DRB1*15:01	9	23	15	TGTLIVNSVLLFLAF	0.3354	Non allergen	Antigenic	Non Toxin	
HLA-DRB1*03:01	10	24	15	GTLIVNSVLLFLAFV	0.3383	Non allergen	Antigenic	Non Toxin	Ifn+
Non Overlapping Sequences									
GTLIVNSVLLFLAFV									

epitopes were predicted out of three; two were further selected for the experiment based on non-toxic, antigenic (Table 5) and Immunogenicity scores (Table 6). Epitopes with non-antigenic + non-toxic, toxin + antigenic and toxin + non antigenic were discarded, while epitopes with non-toxic + antigenic and highly immunogenic were accepted for further investigations.

Table 10
B-cell epitope prediction of surface protein. Top antigenic scores were taken for further investigations.

Epitope	ANTIGENCITY	Score	TOXIN SCORE	TOXIN/ NON TOXIN
GVSVITPGTNTSNQVA	Probable ANTIGEN	0.4651	−1.58	Non-Toxin
GWTAGAAAYYVGYLQP	Probable ANTIGEN	0.0621	−1.29	Non-Toxin
HRSYLTPGDSSSGWTA	Probable ANTIGEN	0.6017	−0.7	Non-Toxin
TVEKGIYQTSNFRVQ	Probable ANTIGEN	0.438	−1.78	Non-Toxin

3.3. Helper T cell epitope prediction

HTL epitopes for human alleles were predicted by the MHC II prediction module of IEDB. For surface protein HLA DRB5*01:01 with position 232–246, HLA DRB 3*01:01209-223 with a length of 15-mer were selected on the basis of antigenic, non-toxic and non-overlapping epitopes (Table 7). In the case of membrane protein HLA DRB1*07:01 with position 88–102 and HLA DRB 5*01:01 with position 174–188 with 15-mer length based on Antigenicity and non-toxicity both were selected (Table 8). For envelope protein, HLA DRB1*15:01 start 9 ends with 23, and HLA DRB1*03:01 start 10 and end with 24 were predicted, based on parameters mentioned above, only HLA DRB1*03:01 selected for further investigation (Table 9). Antigenic + Non-allergenic + Non-Toxic +

Table 11
B-cell epitope prediction of membrane protein. Top antigenic scores were taken for further investigations.

Epitope	ANTIGENCITY	Score	Score toxicity	Toxicity
RSMWSFNPETNILLNV	Probable ANTIGEN	0.4451	−1.2	Non-Toxin
SFRLFARTRSMWSFNP	Probable ANTIGEN	0.951	−0.84	Non-Toxin
RFLYIIKLIFLWLLWP	Probable ANTIGEN	0.4532	−0.56	Non-Toxin
GDSGFAAYSRYRIGNY	Probable ANTIGEN	0.898	−1.19	Non-Toxin
RCDIKDLPKEITVATS	Probable ANTIGEN	0.5606	−1.1	Non-Toxin
RINWITGGIAIAMACL	Probable ANTIGEN	1.2392	−0.71	Non-Toxin
PKEITVATSRTLSYYK	Probable ANTIGEN	0.5935	−1.26	Non-Toxin
GIAIAMACLVGLMWLS	Probable ANTIGEN	0.9132	−0.93	Non-Toxin
LVIGFLFLTWICLLQF	Probable ANTIGEN	0.9967	0.15	Toxin

Table 12
B-cell epitope prediction of Envelope protein. Top antigenic scores were taken for further investigations.

PEPTIDE	ANTIGENCITY	Score	TOXIN SCORE	TOXIN/ NON TOXIN
TLAILTALRLCAYCCN	Probable ANTIGEN	0.6628	0.54	Toxin
NVSLVKPSFYVYSRVK	Probable ANTIGEN	0.7865	−1.44	Non-Toxin
YVYSRVKNLNSSRVPD	Probable ANTIGEN	0.5457	−0.9	Non-Toxin
LCAYCCNIVNVSLVKP	Probable ANTIGEN	0.7286	0.92	Toxin

interferon-gamma positive epitopes were accepted, and others were rejected for further analysis.

3.4. B cell epitope prediction

BCPred online server was used to predict the b cell epitopes for three mentioned proteins for each protein non-toxicity and antigenicity parameter was used to qualify epitope for further analysis. For surface protein, two epitopes with the highest score of 0.6210 and 0.6017 were selected (Table 10). In the case of membrane protein, from nine epitopes, two were chosen further, with the highest score of 0.9132 and 0.9510. One epitope is discarded as it shows toxicity (Table 11). For envelope protein, from four epitopes, two were predicted as toxin, and two were predicted as non-toxin. Two non-toxin, having a score of 0.66 and 0.54, were selected (Table 12).

3.5. Construction of multi-subunit vaccine

Total17epitopes were selected from three protein shaving CTL, HTL and B-cell epitopes. For joining of adjuvant with CTL epitopes, EAAAK linker was used. AAY linker was used to connect each CTL epitope. GPGPG linker was used to join CTL epitope with HTL epitope and to connect each HTL epitope. KK linker was used to join HTL with B cell epitope and to connect each B-cell epitope (Figs. 1a–1c). Allergenicity of vaccine construct was checked by ALgPred, and antigenicity was reviewed by Vaxijen Server (Table 13). All three vaccine constructs qualify the parameters of non-allergen, and antigenic three final vaccine constructs were created by using three different adjuvants (Table 13).

Physicochemical analysis of vaccine constructs:

The molecular weight of the three constructs ranges from 35 to 52 K Da. Construct three shows maximum weight. PI value indicates that three constructs are basic in nature. The instability index is a parameter that is used to calculate the stability inside the test tube. Instability index for the final constructs predicted less than 40, which is considered as stable in nature. GRAVY value for three constructs indicates that vaccine construct 1 has a hydrophobic score while vaccine construct 3 has a hydrophilic score suggesting its hydrophilic nature (Table 14).

Homology modelling and tertiary structure refinement:

For tertiary structure prediction, all the three constructs sequences were submitted to-TASSER online web server. I-TASSER resulted in five models for each vaccine construct. Thereafter each model was submitted to Galaxy Refine server form model refinement of structure, which results back with total of 75 models for each I-TASSER vaccine constructs model. RMSD score, clash score and rama favoured best models for three constructs were selected for further analysis. For vaccine construct 1, model no. 2, having 88.0% rama favourable regions was selected (Table 15). Whether in the case of vaccine construct 2, model no. 5, having 83.4% rama favourable regions was selected (Table 16). For Vaccine construct 3, least poor rotamers, lowest clash scores and most rama favoured region model was selected (Table 17). Three vaccine

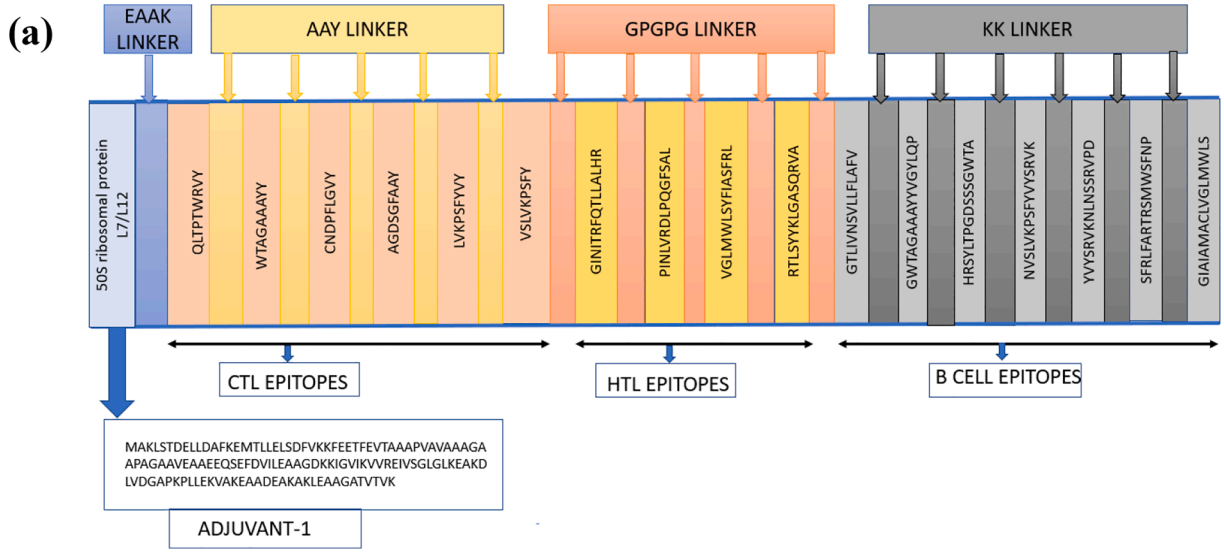

Fig. 1a. Final Multi-epitope vaccine structure 1.

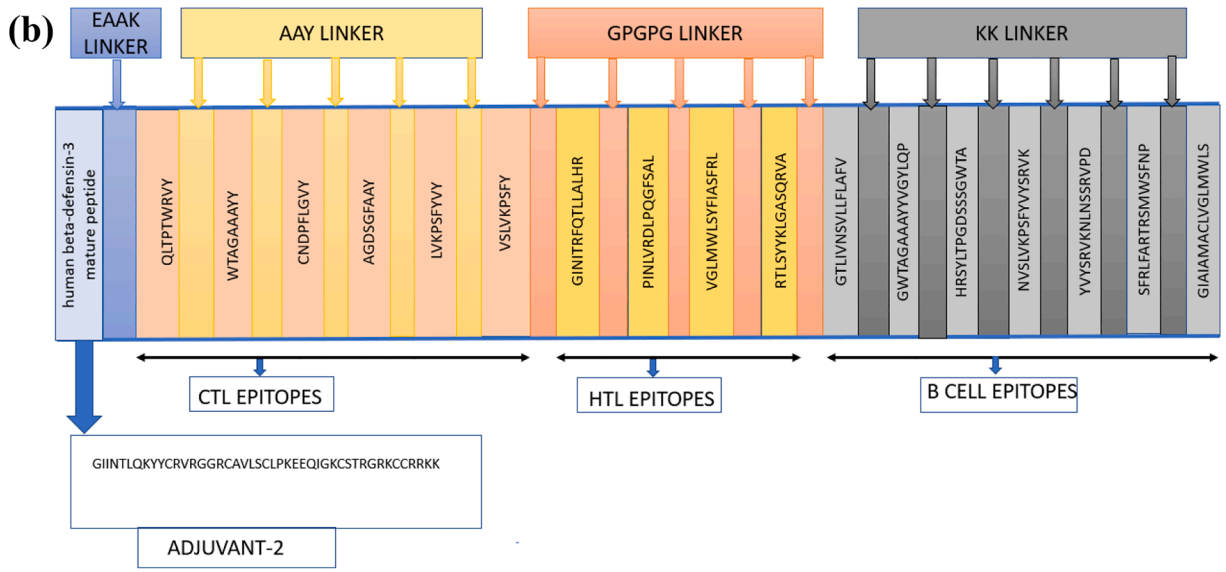

Fig. 1b. Final Multi-epitope vaccine structure 2.

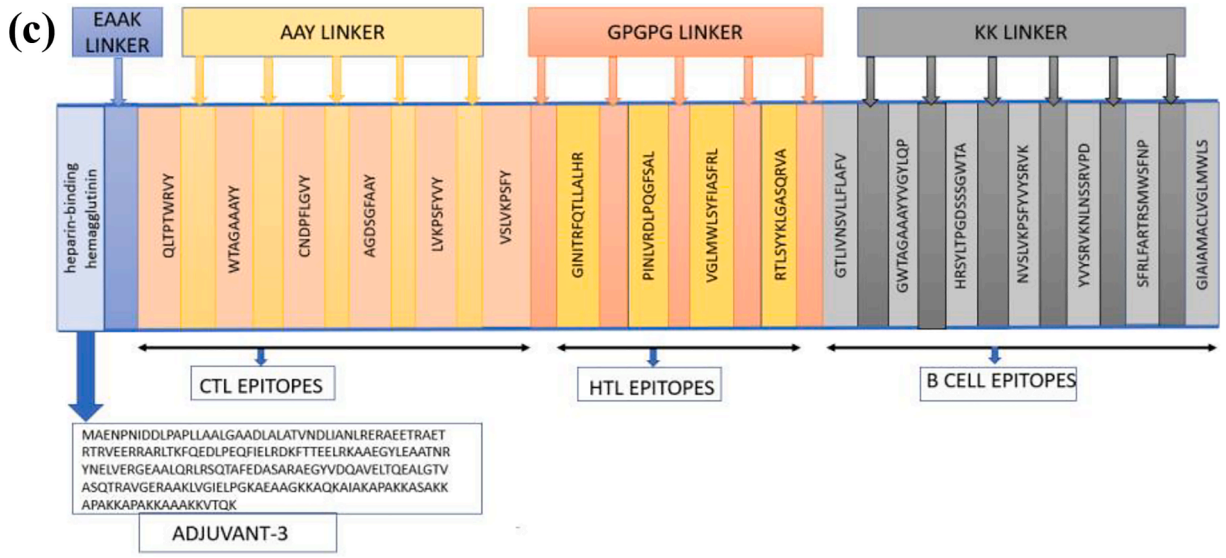

Fig. 1c. Final Multi-epitope vaccine structure 3.

structures predicted by I-TASSER was shown in (Fig. 2).

Secondary structures of three vaccine constructs:

Secondary structures of vaccine construct peptides obtained from PSIPRED are shown in (Figs. 3a, 3b, 3c (see Supplementary Figures)). Every vaccine constructs has alpha helix, beta-sheet, and beta-turn.

Tertiary structure validation:

PROSA server was used to verify the structures. Z-score determines the overall model quality and is represented in the plot. In the graphical representation, X-ray and NMR both are differentiated with different colours. From the server, it was investigating that all the vaccine constructs show homology with X-ray structures (Figs. 4a, 4b, 4c (see Supplementary Figures)) (Figs. 5–8).

3.6. Comparative results of docking of three vaccine constructs

The TLR 8 protein structure was docked to three vaccine construct with the Clus Pro server. Balanced docking method was chosen among available four methods: Balanced, Electrostatic-favored, Hyrdophobic-favored and VdW + Elec. Coefficient Weight were calculated by using algorithm E = 0.40 Erep + −0.40 Eatt + 600 Eelec + 1.00 EDARS. The best orientation PDB files with Coefficient Weights and Cluster Scores were retrieved and taken for protein-ligand interactions studies. The cluster score of all three docked complexes is given in Table 18. Interaction studies were done by Dimplot and Ligplot + .Hydrogen bond interactions vary from 21 to 34, as shown in Tables 19–21. Vaccine Construct 2 with 21 Hydrogen Bonds showed least interaction with TLR8, whereas Vaccine Construct3 interacted with 34 Hydrogen Bonds showed maximum interaction with TLR8. Interacting residues of TLR8 with all three Vaccine Construct ligand along with name and position are given in Tables 19–21. Based on the Coefficient Weights score of the docked complex obtained from ClusPro Protein-Protein Docking Server, along with Hydrogen Bonds and AAs interactions between TLR 8 and vaccine construct predicted by Dimplot, Vaccine Construct 3 seems to be promising to fight against the COVID-19 pandemic.

Immunological simulations of final potential construct:

The TLR 8 protein structure was docked to three vaccine construct with the ClusPro server. Balanced docking method was chosen among available four methods: Balanced, Electrostatic-favored, Hyrdophobic –

Table 13
Three Final multi-epitope vaccine construct sequences further predicted for allergen toxicity and antigenicity.

Construct No.	Vaccine Sequence	Allergenic	Toxicity	Antigenicity
1	MAKLSTDELLDAFKEMTLLELSDFVKKFEETFEVTAAAPV AVAAAGAAPAGAAVEAAEEQSEFDVILEAAGDKKIGVIKV VREIVSGLGLKEAKDLVDGAPKPLLEKVAKEAADEAKAK LEAAGATVTVKEAAAKQLTPTWRVYAAYWTAGAAAYY AAYCNDPFLGVYAAYAGDSGFAAYAAYLVKPSFYVYAA YVSLVKPSFYGPGPGGINITRFQTLLALHRGPGPGPINLVRD LPQGFSALGPGPGVGLMWLSYFIASFRLGPGPGRTLSYYK LGASQRVAGPGPGGTLIVNSVLLFLAFVKKGWTAGAAAY YVGYLQPKKHRSYLTPGDSSSGWTAKKNVSLVKPSFYVY SRVKKKYVYSRVKNLNSSRVPDKKSFRLFARTRSMWSFN PKKGIAIAMACLVGLMWLS	NO	NO	YES
2	GIINTLQKYYCRVRGGRCAVLSCLPKEEQIG KCSTRGRKCCRRKKEAAAKQLTPTWRVYAAY WTAGAAAYYAAYCNDPFLGVYAAYAGDSGFA AYAAYLVKPSFYVYAAYVSLVKPSFYGPGPG GINITRFQTLLALHRGPGPGPINLVRDLPQG FSALGPGPGVGLMWLSYFIASFRLGPGPGRTLSYYKLGASQRVAGPGPGGTLIVNSVLLFLA FVKKGWTAGAAAYYVGYLQPKKHRSYLTPGD SSSGWTAKKNVSLVKPSFYVYSRVKKKYVYS RVKNLNSSRVPDKKSFRLFARTRSMWSFNPK KGIAIAMACLVGLMWLS	NO	NO	YES
3	MAENPNIDDLPAPLLAALGAADLALATVNDL IANLRERAEETRAETRTRVEERRARLTKFQE DLPEQFIELRDKFTTEELRKAAEGYLEAATN RYNELVERGEAALQRLRSQTAFEDASARAEG YVDQAVELTQEALGTVASQTRAVGERAAKLV GIELPGKAEAAGKKAQKAIAKAPAKKASAKK APAKKAPAKKAAAKKVTQKEAAAKQLTPTWR VYAAYWTAGAAAYYAAYCNDPFLGVYAAYAG DSGFAAYAAYLVKPSFYVYAAYVSLVKPSFY GPGPGGINITRFQTLLALHRGPGPGPINLVR DLPQGFSALGPGPGVGLMWLSYFIASFRLGP GPGRTLSYYKLGASQRVAGPGPGGTLIVNSV LLFLAFVKKGWTAGAAAYYVGYLQPKKHRSY LTPGDSSSGWTAKKNVSLVKPSFYVYSRVKK KYVYSRVKNLNSSRVPDKKSFRLFARTRSMW SFNPKKGIAIAMACLVGLMWLS	NO	NO	YES

Table 14
Physiochemical properties of three constructs:

S.No	Vaccine Construct	Molecular weight (kDa)	Total Amino Acid	PI	GRAVY	Instability Index	Aliphatic Index
1.	C1	44.033	412	9.56	0.097	28.64	88.20
2.	C2	35.753	327	10.13	−0.045	35.48	79.11
3.	C3	52.768	487	9.89	−0.215	36.05	81.54

Table 15
Galaxy refine results showing different scores for vaccine construct Model 1.

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1	0	3.059	4.4	18.7	66.1
MODEL 1	0.9284	0.466	2.326	13.3	1.3	87.6
MODEL 2	0.9302	0.468	2.342	14.3	1.3	88
MODEL 3	0.9175	0.506	2.487	14.3	1.9	87.6
MODEL 4	0.9272	0.472	2.307	14.9	0.6	86.6
MODEL 5	0.9278	0.473	2.285	14.3	0.6	86.8

Table 16
Galaxy refine results showing different scores for vaccine construct Model 2.

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1	0	3.54	14.2	18	57.8
MODEL 1	0.9205	0.487	2.673	25.9	1.2	81.2
MODEL 2	0.9251	0.478	2.709	28.5	1.2	81.5
MODEL 3	0.9159	0.505	2.883	28.9	2	81.5
MODEL 4	0.9159	0.501	3.011	26.2	3.1	80.6
MODEL 5	0.9235	0.483	2.81	25.9	2	83.4

favored and VdW + Elec. Coefficient Weights were calculated by using algorithm E = 0.40 Erep +−0.40 Eatt + 600Eelec + 1.00EDARS. The best orientation PDB files with Coefficient Weights and Cluster Scores were retrieved and taken for protein-ligand interactions studies. The cluster score of all three docked Complexes is given in Table 18. Interaction studies were done by Dimplot and Ligplot + Hydrogen bond interactions vary from 21 to 34, as shown in Tables 19–21. Vaccine Construct 2 with 21 Hydrogen Bonds showed least interaction with TLR8, whereas Vaccine Construct 3 interacted with 34 Hydrogen Bonds showed maximum interaction with TLR 8. Interacting residues of TLR 8 with all three Vaccine Construct ligand along with name and position are given in Tables 19–21.

Based on the Coefficient Weights score of the docked complex obtained from ClusPro Protein – Protein Docking Server, along with Hydrogen Bonds and AAs interactions between TLR8 and vaccine construct predicted by Dimplot, Vaccine Construct 3 seems to be promising to fight against the COVID-19 pandemic.

Table 17
Galaxy refine results showing different scores for vaccine construct Model 3.

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1	0	3.057	6.8	19.1	81.9
MODEL 1	0.9302	0.478	1.905	13.5	0.3	96.1
MODEL 2	0.9394	0.452	2.055	14.4	1.6	96.5
MODEL 3	0.9317	0.466	1.911	14	1.1	96.5
MODEL 4	0.9379	0.462	1.917	14.5	0.8	96.3
MODEL 5	0.9343	0.459	1.913	14.4	0.8	96.3

3.7. *Immunological simulations of final potential construct:*

When antigen was initially exposed to the system, a very high level of IgM noticed. After the secondary and tertiary response, there was an

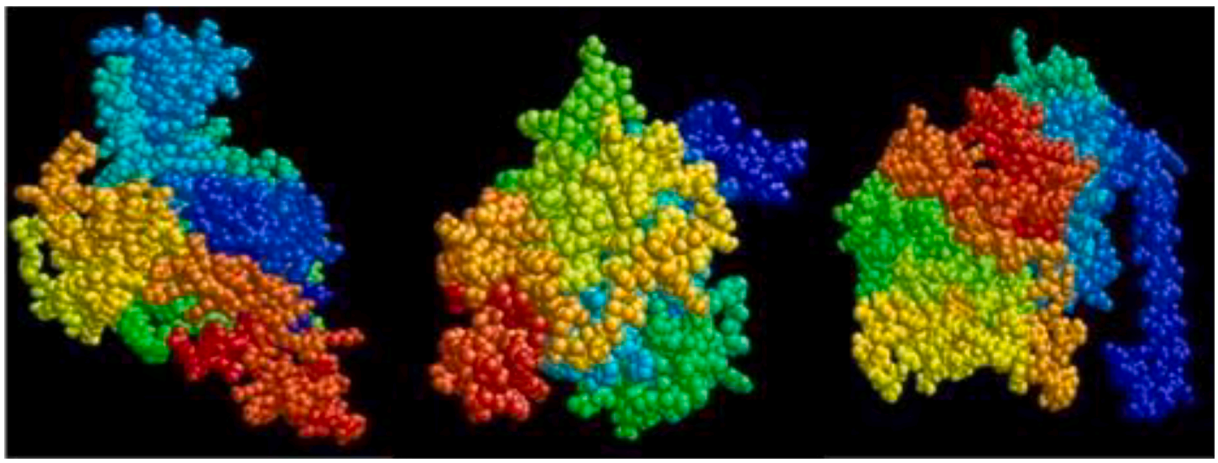

Fig. 2. 3D model structures of vaccine peptides predicted by I-TASSER.

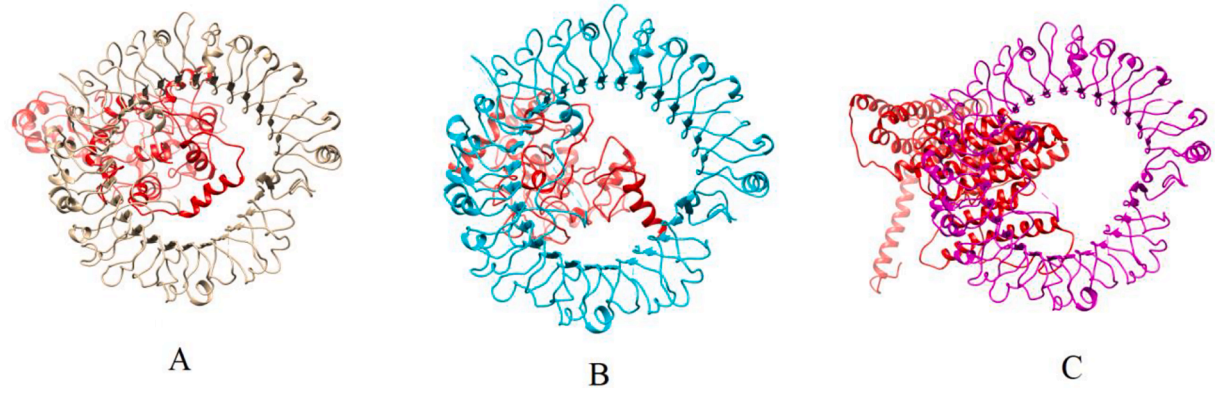

Fig. 5. Most stable orientation of TLR8 Docked with three vaccines constructs (A, B, C). Docked Complexes were obtained from the ClusPro.

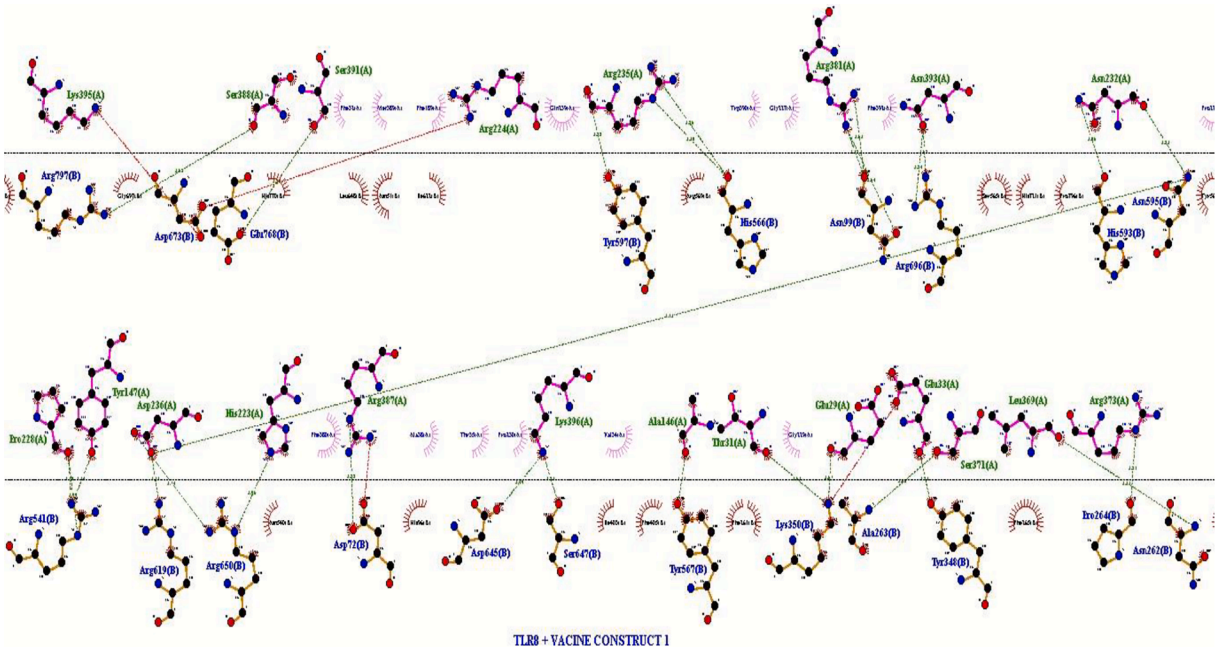

Fig. 6. H Bond interactions as plotted by Dimplot Between TLR8 and vaccine construct 1. Dashed lines Hydrogen bonds, arcs hydrophobic interactions.

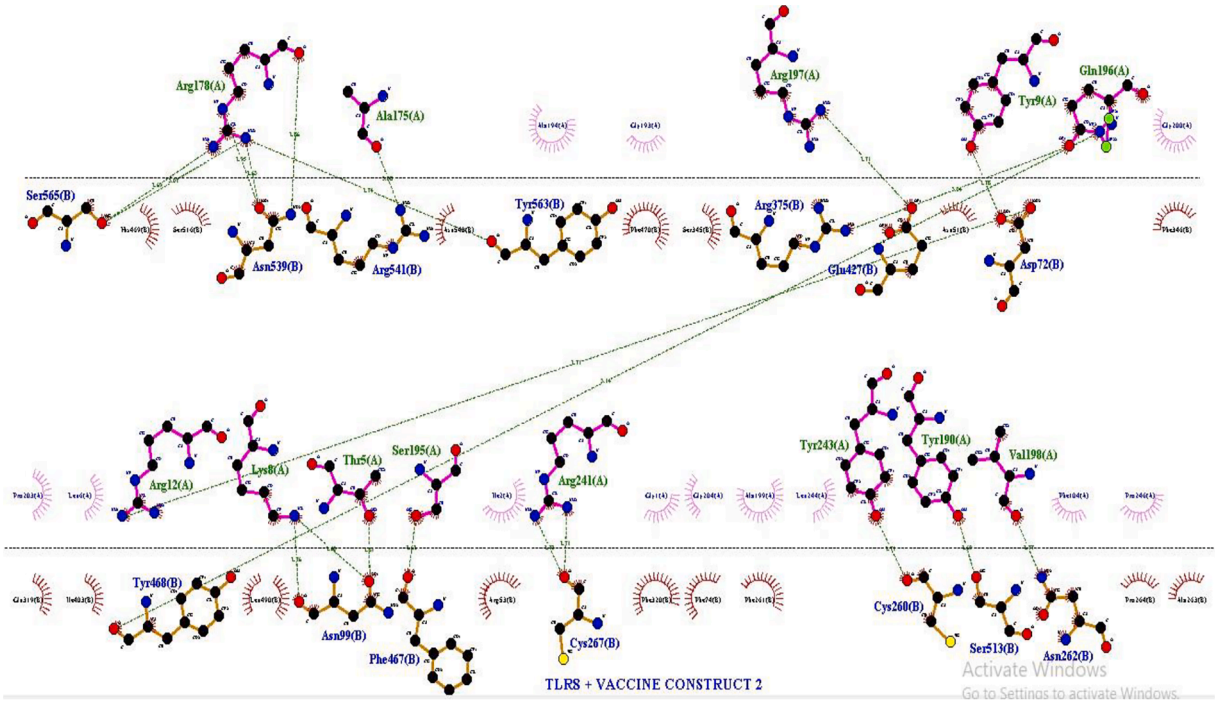

Fig. 7. H Bond interactions as plotted by Dimplot Between TLR8 and vaccine construct 2. Dashed lines Hydrogen bonds, arcs hydrophobic interactions.

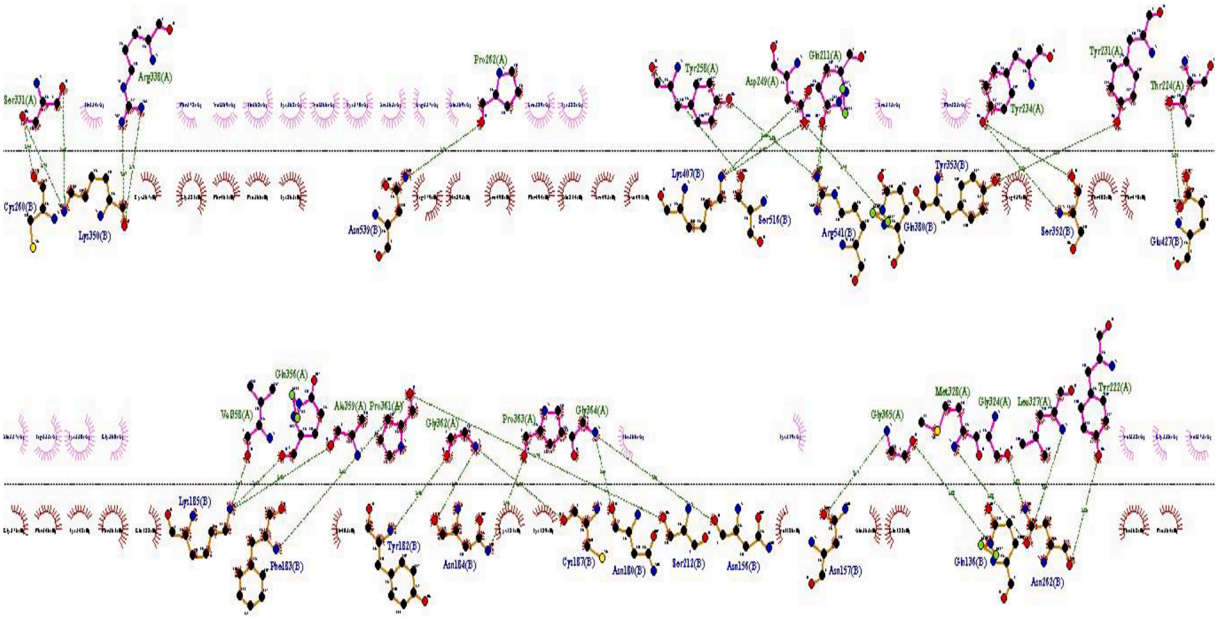

Fig. 8. H Bond interactions as plotted by Dimplot Between TLR8 and vaccine construct 3. Dashed lines Hydrogen bonds, arcs hydrophobic interactions.

Table 18
Weighted Score of TLR8 and ligand (Vaccine Construct 1–3) obtained from the ClusPro.

S.No.	Docked Complex	Weighted Score	
		Center Energy	Lowest energy
1.	TLR8 + Vaccine Construct 1	−999.3	−1382.7
2.	TLR8 + Vaccine Construct 2	−837.0	−1059.4
3.	TLR8 + Vaccine Construct 3	−1175.7	−1478.0

increase in the concentration of IgM + IgG, Ig1 and Ig2 with the decrease in the concentration of antigen (Fig. 9A). Total B cell concentration from three exposures results in increasing while memory B cell concentration remains high for several months (Fig. 9B). Similarly, a high response was seen in TH (Helper) and TC (cytotoxic) cell population (Fig. 9C,9D). Interferon – Gamma and Il-2 concentration rose with exposure to antigen. They remained at a high level with exposure to the antigen, which indicates that it leads to an increase in TC, B cell and immunoglobins productions.

Table 19
H Bond interactions between TLR8 and ligand as Predicted by Dimplot (Vaccine Construct 1).

No.	Hydrogen Bonds	Interacting AAs of TLR8		Interacting AAs of ligand (Vaccine Construct 1)	
1	33	ASN	262	ARG	387
		ALA	263	ARG	381
		TYR	348	ARG	373
		LYS	350	ARG	235
		ARG	541	ASN	232
		TYR	567	LYS	396
		ASN	595	SER	391
		TYR	597	LEU	369
		ARG	619	SER	371
		ARG	650	GLU	33
		ARG	696	GLU	29
		ARG	797	THR	31
		ASP	72	PRO	228
		ASP	673	TYR	147
		ASN	99	ALA	146
		PRO	264	ASP	236
		HIS	566	HIS	223
		HIS	593	ASN	393
		ASP	645	SER	388
		SER	647	ARG	224
		GLU	768	LYS	395

Table 20
H Bond interactions between TLR8 and ligand as Predicted by Dimplot (Vaccine Construct 2).

No.	Hydrogen Bonds	Interacting AAs of TLR8		Interacting AAs of ligand (Vaccine Construct 2)	
1	21	ASN	262	TYR	9
		ARG	375	ARG	12
		SER	513	THR	5
		ASN	539	LYS	8
		ARG	541	TYR	243
		ASP	72	ARG	241
		ASN	99	ARG	197
		CYS	260	SER	195
		CYS	267	GLN	196
		GLU	427	ARG	178
		PHE	467	VAL	198
		TYR	468	TYR	190
		TYR	563	ALA	175
		SER	565		

Table 21
H Bond interactions between TLR8 and ligand as Predicted by Dimplot (Vaccine Construct 3).

No.	Hydrogen Bonds	Interacting AAs of TLR8		Interacting AAs of ligand (Vaccine Construct 2)	
1	34	GLN	136	GLY	364
		TYR	182	GLY	365
		PHE	183	GLY	362
		ASN	184	SER	331
		LYS	185	LEU	327
		SER	212	MET	328
		ASN	262	TYR	222
		LYS	350	ARG	338
		SER	352	TYR	234
		GLN	380	TYR	231
		LYS	407	THR	224
		SER	516	PRO	361
		ASN	539	PRO	363
		ARG	541	GLN	356
		ASN	156	VAL	358
		ASN	157	ALA	359
		ASN	180	GLY	324
		CYS	187	ASP	249
		CYS	260	TYR	258
		TYR	353	PRO	262
		GLU	427	GLN	211

4. Discussion

Vaccines are developed with the purpose to generate immune responses that protect human being from several viral diseases. Recently, epitope vaccines were developed by engineering CTL, HTL and B-cell epitopes. The advantages associated with epitope vaccines as compared to conventional vaccines are their non-toxicity, highly immunogenic and Non-allergen. The present study based on *in-silico* approach for designing and developing multi-epitope vaccine construct and its immune simulation to identify predicted immune responses. The *in-silico* approach is used in the present scenario because it saves a lot of time, does not require microbial culture and safe to develop effective new generation vaccine using modern computational and immunoinformatic tools (Scarselli et al., 2005). Recent studies have shown that advanced *in-silico* tools not only helping in the prediction of immune response but also paying the way for designing and developing new generation vaccines (Groot and Rappuoli, 2004; Korber et al., 2006; Purcell et al., 2007). In the present study, SARS-CoV-2 three proteins antigenic epitope prediction, the interaction of MHC alleles with epitopes have been identified for generation of multi-epitope vaccine constructs. Initially, these proteins were checked in the BlastP online program of NCBI to ensure that they don't show any similarity with human proteome. Further B-cell and T cell epitopes were retrieved from online servers. B-cells are associated with the generation of humoral response by secreting antibodies from plasma cells and memory B cells, providing life long immunity (Silva et al., 2016).

The purpose of choosing B-cell epitopes because they will generate both cell-mediated and humoral response. B-cell epitopes are the group of amino acid sequence present on the cell surface identified by specific antibodies or BCR that generate a humoral or cell-mediated response. The generated antibodies will neutralize the virus, and cell-mediated immunity will kill infected cells. The idea of choosing T cell epitopes was because it was presented by APC, bound by MHC molecules to generated immune response. Tcell identifies MHC class I and II in the cell surface. Class I MHC molecule generally present Peptide molecule between 8 and 10 amino acid length to be identified by Cytotoxic T cells having marker CD8+. Class II MHC molecule has a peptide length of 12to25 amino acid, identified by Helper T cell with marker CD4+. Ifa sufficient epitope is presented, the T cell may generate a significant adaptive immune response specific against the virus-specific pathogen. Because of this, all these epitopes from three proteins of SARS-CoV-2 were assembled to make three final constructs with the help of linkers and three different adjuvants. For vaccine construct- I non-allergenicscorewas-0.424, Non-toxin, and antigenicity score was 0.5292. The construct-II non-allergenic score was −1.0073, Non-toxic, the antigenic score was 0.59512. Construct-III Non-allergenic score was 1.029, Non-toxic, the antigenic score was 0.4861. All the three vaccines constructs show highly antigenic, non-toxic and non-allergen by passing the threshold cut-offs. Previous studies have shown that using immunoinformatic tools, epitopes were

Synthesized from the parasite proteome such as Leishmania braziliens is and assessed *in- vitro* found to be immunogenic and elicit an immune response (Silva et al., 2016). Another study was shown to generate T cell response using T cell epitopes derived from Mycobacterium tuberculosis proteome using an immunoinformatic approach (Khan et al., 2014). Based on these studies, the present study attempted to design and develop the vaccine construct against SARS-CoV-2 virus using computational and immunoinformatic approach. Furthermore, the molecular weight of construct I, II, III was identified as 44.033 kDa, 35.753 kDa, 32.768 kDa that shows molecular weight with in acceptable values for construction of vaccine. All three constructs show basic in nature and Instability index values of three shows all are thermally

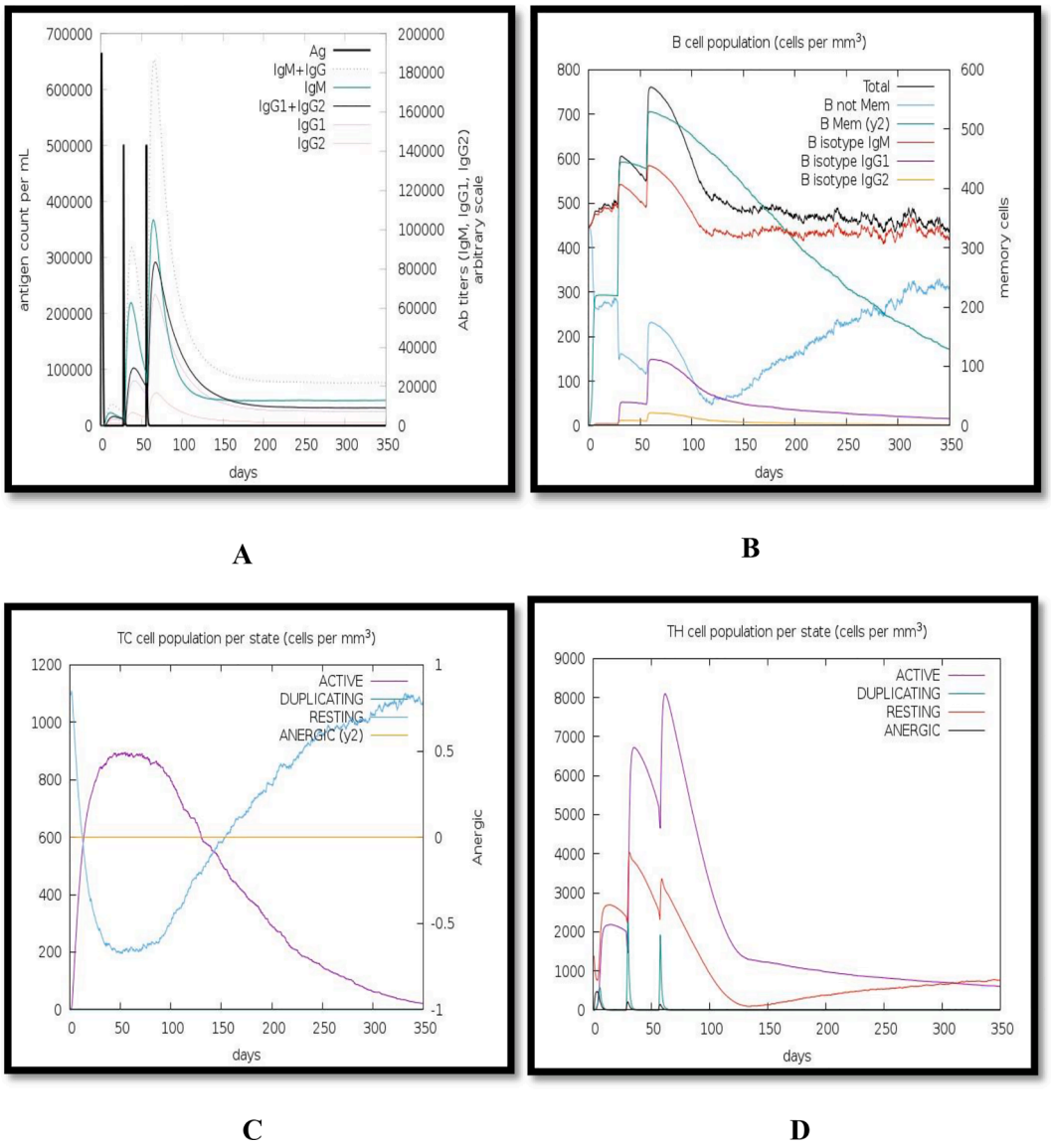

Fig. 9. C-ImmSim demonstration of antigenic peptide through immune simulation. (A) Feedback of antigenic injections w.r.t. production of immunoglobulin (black vertical lines); specific subclasses are indicated as coloured peaks. (B) Concentration of B-cell population after three exposures. (C) Indicating concentration of TC cells. (D) evolution of TH cells.

stable. Two and three-dimensional structures of three vaccines constructs were identified was using online servers, which were further preceded by structural validations.

Docking analysis was conducted to comprehend the immune response of TLR8 of three vaccine constructs. By comparing three vaccine constructs, construct III shows the lowest energy and is taken into further consideration for immune simulations. Immune simulation studies with repeated exposure of antigen show a significant increase in the immune response. Initially, there was a high level of IgM present; when exposed to Secondary and tertiary responses, there was also an increase in IgM and IgG. From three exposures, memory B cell remain high for several months in the case of T h and T c cells shows massive high response Interferon-gamma and IL-2 remains at a peak which shows there was an enormous production of Ig. Immune simulations of the final vaccine construct show better result in the modelled environment; hence by compilation of these new immunoinformatic approach in this study to the formation of potential, non-toxic, non-allergenic, antigenic with high immune response to control SARS-CoV-2.

5. Conclusion

In this study, an immunoinformatic approach employed to develop a multi-epitope vaccine construct against SARS-CoV-2. The different vaccine constructs are docked with TLR 8, showing good binding affinity. The predicted humoral and cell-mediated responses are also significant using simulation studies. Physiochemical structure analysis and immune simulation of vaccine construct are performed to check how vaccine constructs behave in cell environment and how much immune response is generated. Preferred vaccine construct will need *in-vivo* validations further to ensure its activity, stability and enhancement of immune response. This research will be assisted infection control by enhancing the immune response against SARS-CoV-2.

Author contributions

A.K. conceived the idea. A.S. & S.P. performed data retrieval, analysis and characterization. A.S., S.P. & A.P. Performed methodology, software, data curation and characterization under the guidance of A.K. The manuscript was written by A.K., A.S., S.P., A.P. & S.K. S.K. and A.K.

contributed to the analysis and discussion of the results. Review & editing done by A.K. & S.K.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.imbio.2021.152134.

References

Astuti, I., Ysrafil, 2020. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): an overview of viral structure and host response. Diabetes Metab. Syndr. 14 (4), 407–412.

Bhattacharya, M., et al., 2020. A SARS-CoV-2 vaccine candidate: in-silico cloning and validation. Inf. Med. Unlocked 20, 100394.

Bhattacharya, M., Sharma, A.R., Patra, P., Ghosh, P., Sharma, G., Patra, B.C., Lee, S., Chakraborty, C., 2020. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. J. Med. Virol. 92 (6), 618–631.

Braun, J., et al., 2020. SARS-CoV-2-reactive T cells in healthy donors and patients with. COVID-19. 587 (7833), 270–274.

Chiarella, P., et al., 2009. Recent advances in epitope design for immunotherapy of cancer. Recent Pat. Anticancer Drug Discov. 4 (3), 227–240.

Creed, M.A., et al., 2020. Mild COVID-19 infection despite chronic B cell depletion in a patient with aquaporin-4-positive neuromyelitis optica spectrum disorder. Mult. Scler. Relat. Disord. 44, 102199.

Dai, Y., Chen, H., Zhuang, S., Feng, X., Fang, Y., Tang, H., Dai, R., Tang, L., Liu, J., Ma, T., Zhong, G., 2020. Immunodominant regions prediction of nucleocapsid protein for SARS-CoV-2 early diagnosis: a bioinformatics and immunoinformatics study. Pathogens Global Health 114 (8), 463–470.

Dermime, S., Gilham, D.E., Shaw, D.M., Davidson, E.J., Meziane, E.-K., Armstrong, A., Hawkins, R.E., Stern, P.L., 2004. Vaccine and antibody-directed T cell tumour immunotherapy. Biochim. Biophys. Acta 1704 (1), 11–35.

Doytchinova, I.A., Flower, D.R., 2007. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. Vaccine 25 (5), 856–866.

EL-Manzalawy, Y., Dobbs, D., Honavar, V., 2008. Predicting linear B-cell epitopes using string kernels. J. Mol. Recognit. 21 (4), 243–255.

Groot, A.S.D., Rappuoli, R., 2004. Genome-derived vaccines. Expert Rev. Vaccines 3 (1), 59–76.

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava, G.P.S., Patterson, R.L., 2013. In silico approach for predicting toxicity of peptides and proteins. PLoS ONE 8 (9), e73957.

Heo, L., H. Park, Seok, C., GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic Acids Res, 2013. 41(Web Server issue): p. W384-8.

Janice Oh, H.-L., Ken-En Gan, S., Bertoletti, A., Tan, Y.-J., 2012. Understanding the T cell immune response in SARS coronavirus infection. Emerg. Microbes Infect. 1 (1), 1–6.

Ju, B., et al., Potent human neutralizing antibodies elicited by SARS-CoV-2 infection. bioRxiv, 2020: p. 2020.03.21.990770.

Khailany, R.A., Safdar, M., Ozaslan, M., 2020. Genomic characterization of a novel SARS-CoV-2. Gene Rep 19, 100682.

Khan, M.K., Zaman, S., Chakraborty, S., Chakravorty, R., Alam, M.M., Bhuiyan, T.R., Rahman, M.J., Fernández, C., Qadri, F., Seraj, Z.I., 2014. In silico predicted mycobacterial epitope elicits in vitro T-cell responses. Mol. Immunol. 61 (1), 16–22.

Korber, B., LaBute, M., Yusim, K., McIntyre, J., 2006. Immunoinformatics comes of age. PLoS Comput. Biol. 2 (6), e71.

Kumar, S., et al., 2021. Identification of phytochemicals as potential therapeutic agents that binds to Nsp15 protein target of coronavirus (SARS-CoV-2) that are capable of inhibiting virus replication. Phytomedicine 85, 153317.

Kumar, A., Panwar, A., Batra, K., Dey, S., Maan, S., 2020. In-silico analysis of angiotensin converting enzyme 2 (ACE2) of livestock, pet and poultry animals to determine its susceptibility to SARS-CoV- 2 infection. Comb. Chem. High Throughput Screen 23.

Larsen, M.V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., Nielsen, M., 2007. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinf. 8 (1).

Lee, S.J., Shin, S.J., Lee, M.H., Lee, M.-G., Kang, T.H., Park, W.S., Soh, B.Y., Park, J.H., Shin, Y.K., Kim, H.W., Yun, C.-H., Jung, I.D., Park, Y.-M., Gao, J.-X., 2014. A potential protein adjuvant derived from mycobacterium tuberculosis rv0652 enhances dendritic cells-based tumor immunotherapy. PLoS ONE 9 (8), e104351.

Mei, H.F., et al., 2012. β-defensin 2 as an adjuvant promotes anti-melanoma immune responses and inhibits the growth of implanted murine melanoma in vivo. PLoS ONE 7 (2), e31328.

Meloen, R.H., Langeveld, J.P.M., Schaaper, W.M.M., Slootstra, J.W., 2001. Synthetic peptide vaccines: unexpected fulfillment of discarded hope? Biologicals 29 (3-4), 233–236.

Patronov, A., Doytchinova, I., 2013. T-cell epitope vaccine design by immunoinformatics. Open Biol. 3 (1), 120139. https://doi.org/10.1098/rsob.120139.

Purcell, A.W., McCluskey, J., Rossjohn, J., 2007. More than one reason to rethink the use of peptides in vaccine design. Nat. Rev. Drug Discov. 6 (5), 404–414.

Rahman, N., Ali, F., Basharat, Z., Shehroz, M., Khan, M.K., Jeandet, P., Nepovimova, E., Kuca, K., Khan, H., 2020. Vaccine design from the ensemble of surface glycoprotein epitopes of SARS-CoV-2: an immunoinformatics approach. Vaccines 8 (3), 423. https://doi.org/10.3390/vaccines8030423.

Rana, A., Akhter, Y., 2016. A multi-subunit based, thermodynamically stable model vaccine using combined immunoinformatics and protein structure based approach. Immunobiology 221 (4), 544–557.

Saha, S., Raghava, G.P., AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res, 2006. 34(Web Server issue): p. W202-9.

Salajegheh Tazerji, S., et al., 2020. Transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) to animals: an updated review. J. Transl. Med. 18 (1), 358.

Sanche, S., Lin, Y.T., Xu, C., Romero-Severson, E., Hengartner, N., Ke, R., 2020. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. Emerg. Infect. Dis. 26 (7), 1470–1477.

Sarkar, B., Ullah, M.A., Johora, F.T., Taniya, M.A., Araf, Y., 2020. Immunoinformatics-guided designing of epitope-based subunit vaccines against the SARS Coronavirus-2 (SARS-CoV-2). Immunobiology 225 (3), 151955. https://doi.org/10.1016/j.imbio.2020.151955.

Scarselli, M., Giuliani, M.M., Adu-Bobie, J., Pizza, M., Rappuoli, R., 2005. The impact of genomics on vaccine design. Trends Biotechnol. 23 (2), 84–91.

Silva, F.E., et al., 2016. Combination of in silico methods in the search for potential CD4(+) and CD8(+) T cell epitopes in the proteome of Leishmania braziliensis. Front. Immunol. 7, 327.

Struyf, T., et al., Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease. Cochrane Database Syst. Rev., 2020. 7(7): p. Cd013665.

Tilocca, B., Soggiu, A., Greco, V., Piras, C., Arrigoni, N., Ricchi, M., Britti, D., Urbani, A., Roncada, P., 2020. Immunoinformatic-based prediction of candidate epitopes for the diagnosis and control of paratuberculosis (Johne's Disease). Pathogens 9 (9), 705. https://doi.org/10.3390/pathogens9090705.

Tilocca, B., et al., Computational Immune Proteomics Approach to Target COVID-19. 2020. 19(11): p. 4233-4241.

Wang, C., et al., 2020. Alveolar macrophage dysfunction and cytokine storm in the pathogenesis of two severe COVID-19 patients. EBioMedicine 57, 102833.

Wiederstein, M., Sippl, M.J., ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res., 2007. 35(Web Server issue): p. W407-10.

Wilkins, M.R., et al., 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. 112, 531–552.

Xu, B., et al., 2020. Suppressed T cell-mediated immunity in patients with COVID-19: A clinical retrospective study in Wuhan. *China.* J Infect 81 (1), e51–e60.

Yang, J., Zhang, Y., 2015. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res. 43 (W1), W174–W181.

Zheng, J., Lin, X., Wang, X., Zheng, L., Lan, S., Jin, S., Ou, Z., Wu, J., 2017. In Silico analysis of epitope-based vaccine candidates against hepatitis B virus polymerase protein. Viruses 9 (5), 112. https://doi.org/10.3390/v9050112.

Zhu, Z., Chakraborti, S., He, Y., Roberts, A., Sheahan, T., Xiao, X., Hensley, L.E., Prabakaran, P., Rockx, B., Sidorov, I.A., Corti, D., Vogel, L., Feng, Y., Kim, J.-O., Wang, L.-F., Baric, R., Lanzavecchia, A., Curtis, K.M., Nabel, G.J., Subbarao, K., Jiang, S., Dimitrov, D.S., 2007. Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal antibodies. Proc. Natl. Acad. Sci. U.S.A. 104 (29), 12123–12128.