

Available online at www.sciencedirect.com

ScienceDirect

Biomedical Journal

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

Review Article

Development of multi-epitope peptide-based vaccines against SARS-CoV-2



Hui Xuan Lim ^a, Jianhua Lim ^a, Seyed Davoud Jazayeri ^a,
Sibrandes Poppema ^b, Chit Laa Poh ^{a,*}

^a Centre for Virus and Vaccine Research, School of Science and Technology, Sunway University, Selangor, Malaysia

^b Sunway University, Selangor, Malaysia



Dr. Chit Laa Poh

ARTICLE INFO

Article history:

Received 28 August 2020

Accepted 25 September 2020

Available online 1 October 2020

Keywords:

SARS-CoV-2

CD4+ T-cell

CD8+ T-cell

B-cell

Epitopes

ABSTRACT

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused a pandemic involving so far more than 22 million infections and 776,157 deaths. Effective vaccines are urgently needed to prevent SARS-CoV-2 infections. No vaccines have yet been approved for licensure by regulatory agencies. Even though host immune responses to SARS-CoV-2 infections are beginning to be unravelled, effective clearance of virus will depend on both humoral and cellular immunity. Additionally, the presence of Spike (S)-glycoprotein reactive CD4+ T-cells in the majority of convalescent patients is consistent with its significant role in stimulating B and CD8+ T-cells. The search for immunodominant epitopes relies on experimental evaluation of peptides representing the epitopes from overlapping peptide libraries which can be costly and labor-intensive. Recent advancements in B- and T-cell epitope predictions by bioinformatic analysis have led to epitope identifications. Assessing which peptide epitope can induce potent neutralizing antibodies and robust T-cell responses is a prerequisite for the selection of effective epitopes to be incorporated in peptide-based vaccines. This review discusses the roles of B- and T-cells in SARS-CoV-2 infections and experimental validations for the selection of B-, CD4+ and CD8+ T-cell epitopes which could lead to the construction of a multi-epitope peptide vaccine. Peptide-based vaccines are known for their low immunogenicity which could be overcome by incorporating immunostimulatory adjuvants and nanoparticles such as Poly Lactic-co-Glycolic Acid (PLGA) or chitosan.

* Corresponding author. Centre for Virus and Vaccine Research, School of Science and Technology, Sunway University, 5, Jalan Universiti, Bandar Sunway, Kuala Lumpur, Selangor, 47500, Malaysia.

E-mail address: pohcl@sunway.edu.my (C.L. Poh).

Peer review under responsibility of Chang Gung University.

<https://doi.org/10.1016/j.bj.2020.09.005>

2319-4170/© 2020 Chang Gung University. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

SARS-CoV-2 first emerged in December 2019 in Wuhan (Hubei Province), causing a pandemic with over 22 million infections and 776,157 deaths globally (World Health Organization). Clinical manifestations of COVID-19 vary broadly, ranging from fever, cough and fatigue to headache, hemoptysis, diarrhea, dyspnea, and lymphopenia. Severe infections were reported to lead to pneumonia, acute respiratory distress syndrome, acute cardiac injury and death [1]. The genome size of the SARS-CoV-2 is approximately 29.9 kb and its genome structure is homologous to SARS-CoV-1. The 5' region consists of two-thirds of the genome, comprising the *orf1ab* that encodes the *orf1ab* polyprotein, while the 3' region is composed of one-third of the genome encoding structural proteins which include surface (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. Additionally, the SARS-CoV-2 virus contains 6 accessory proteins: ORF1ab, NS3, NS6, NS7a, NS7b, and NS8 [2].

Currently, there are over 125 vaccine candidates against SARS-CoV-2 in development worldwide, with at least 16 of these being registered for clinical trials [3]. mRNA-1273 is a novel lipid nanoparticle (LNP)-encapsulating mRNA-based vaccine that encodes for a full-length, prefusion stabilized spike (S) protein of SARS-CoV-2. It is co-developed by Moderna and the Vaccine Research Center at the National Institute of Allergy and Infectious Diseases (NIAID) of NIH. It is currently in phase III clinical trial. A vectored-vaccine, AZD1222, which used a chimpanzee viral vector (ChAdOx1) encoding the spike (S) protein of SARS-CoV-2 was developed by Oxford University and AstraZeneca, and it is currently in phase II clinical trial. CoronaVac, formerly known as PiCoVacc, is an inactivated vaccine developed by Sinovac Biotech. It constitutes the whole SARS-CoV-2 virus which carries the receptor binding domain (RBD) within the S protein as the main immunogen. It is inactivated with β -propiolactone and mixed with an alum adjuvant. Sinopharm has just begun testing their inactivated vaccine candidate in UAE. Both Sinovac's and Sinopharm's inactivated vaccine candidates are currently in phase III clinical trials. INO-4800 constitutes a DNA plasmid encoding the S protein that will be delivered by electroporation via the CELLCTRA® 2000 electroporator. It is developed by Inovio Pharmaceuticals and is currently in phase I/II clinical trials. Furthermore, Novavax has developed a recombinant subunit-based vaccine, NVX-CoV2373, that is currently in phase I/II clinical trials. They have genetically modified Sf9 insect cells to express the recombinant S protein in prefusion state and incorporated it into a nanoparticle formulation containing Matrix-M adjuvant. In summary, 19 vaccine candidates with their associated characteristics that are in clinical trials I, II, and III against SARS-CoV-2 are presented in [Table 1].

Synthetic peptides are capable of eliciting immune responses to specific antigenic epitopes while avoiding adverse effects of immunization. Peptide vaccines can induce broad immunity against multiple serotypes of the virus by carrying immunodominant epitopes that are highly conserved between the serotypes of a virus. Peptide vaccines can be synthesized to contain multiple conserved epitopes that can target multiple serological variants or strains [4]. Theoretically, there are several advantages of peptide-based vaccines over conventional vaccines such as live-attenuated and inactivated vaccines. Synthetic peptide vaccines can be

constructed to contain only the antigenic region, thus reducing the reactogenic or allergenic effects of the vaccine. Synthesis of peptides can be achieved with relatively low cost, in conjunction with its higher stability and safety when compared to live attenuated vaccines. If FDA approved LAV and IV vaccines against SARS-CoV-2 become available for immunization, the demand for industrial production to cater for the entire world's population can pose limitations on its availability. Since SARS-CoV-2 is highly infectious, growth of large amounts of viruses poses safety issues in production facilities. Unlike conventional LAV or IV vaccines, the composition of the peptide vaccine can be altered with self- and non-self-antigens such as universal CD4+ T-cell epitopes to enhance immunogenicity [5]. Moreover, several peptide epitopes which are capable of activating a number of HLA restricted T-cell specificities can be incorporated with B-cell epitopes to form a universal vaccine formulation. Epitope-based vaccines were found to induce more potent responses than whole protein vaccines [6]. Effective adjuvants, nanoparticle delivery systems and immunogenic carrier proteins could be incorporated in the formulation to improve the immunogenicity of epitope vaccines [7,8].

The elicitation of protective antibodies requires the binding of peptides representing B-cell epitopes to B-cell receptors (BCRs) which lead to B-cell activation. B-cell epitopes are variable in length between 5 and 20 amino acids [9]. Linear B-cell epitopes are composed of contiguous residues in the primary structure while discontinuous B-cell epitopes comprise residues remotely located in the primary structure that are brought into proximity due to the folding of the protein [10]. For peptide-based vaccines, B-cell epitopes that are linear and immunogenic are preferentially selected for further investigations [11]. Cytokines secreted by activated CD4+ T-cells are required for further activation of B-cells to antibody-secreting plasma cells.

T-cell epitopes are presented on the surface of an antigen-presenting cell (APC) where they are bound to MHC molecules to induce cellular immunity. One of the key issues in T-cell epitope prediction is MHC binding, as it is considered a prerequisite for T-cell recognition. MHC class I molecules usually present peptides between 8 and 11 amino acids in length, whereas the peptides binding to MHC class II may have a length from 11 to 30 amino acids [12]. MHC class I molecules are encoded by the HLA-A, -B and -C genes which are widely expressed in most cell types in humans. Internalized antigens are degraded by cytosolic and nuclear proteasomes and the resulting peptides are transported into the endoplasmic reticulum by the transporter associated with antigen processing (TAP) protein complex. There, they can bind to human leukocyte antigen class I (HLA-I) molecules in complex with beta2-microglobulin (β 2m). MHC class II molecules are encoded by the HLA-DP, -DQ, -DR genes which are expressed on specialized cell types, including professional APCs such as B-cells, macrophages and dendritic cells. Peptides derived from degraded extracellular proteins are presented on HLA-II molecules for recognition by CD4+ T-cells [13]. The recognition of peptide epitopes by the T-cell receptors (TCR) of CD4+ and CD8+ T-lymphocytes leads to the induction of cellular immune response that plays a key role in the adaptive immune system. If sufficient quantities of specific T-cell epitopes are

Table 1 Several vaccine candidates for SARS-CoV-2 in clinical phase.

Candidate	Vaccine Characteristics	Lead Developer	Status	Clinical Trial Identifier
mRNA-1273	Lipid nanoparticle-encapsulated mRNA-based vaccine that encodes for the S protein.	Moderna	Phase III	NCT04470427
Inactivated vaccine	Inactivated novel Coronavirus pneumonia vaccine (Vero cells).	Sinopharm	Phase III	ChiCTR2000034780
CoronaVac	β -propiolactone-inactivated and alum-adjuvanted candidate vaccine.	Sinovac	Phase III	NCT04456595
AZD1222	Genetically altered and weakened chimpanzee adenovirus vaccine vector carrying the S protein.	The University of Oxford	Phase II/III	NCT04400838
Ad5-nCoV	Adenovirus type 5 vector that expresses S protein.	CanSino Biologicals	Phase II	NCT04341389
Adjuvanted Recombinant Vaccine	Recombinant new Coronavirus vaccine produced in CHO Cells.	Anhui Zhifei Longcom Biologic Pharmacy	Phase II	NCT04466085
Ad26COVS1	Non-replicating Adenovirus type 26 vector.	Johnson & Johnson - Janssen	Phase I/II	NCT04436276
BNT162	4 mRNA vaccine candidates including uridine containing mRNA, nucleoside-modified mRNA, and self-amplifying mRNA that encodes trimerized S protein receptor binding domain antigen formulated with lipid nanoparticles.	BioNTech RNA Pharmaceuticals GmbH	Phase I/II	NCT04380701
INO-4800	DNA plasmid encoding S protein delivered by electroporation.	Inovio Pharmaceuticals	Phase I/II	NCT04447781
NVX-CoV2373	Recombinant S protein in prefusion state expressed in genetically engineered Sf9 insect cells and incorporated into a nanoparticle formulation and saponin-based Matrix-M adjuvant.	Novavax	Phase I/II	NCT04368988
GX-19	DNA vaccine candidate.	Genexine	Phase I/II	NCT04445389
Gam-COVID-Vac	Non-replicating viral vector.	Gamaleya Research Institute	Phase I/II	NCT04437875
LV-SMENP-DC	Dendritic cells modified with lentiviral vector expressing synthetic minigene based on domains of selected viral proteins; administered with antigen-selected viral proteins.	Shenzhen Geno-Immune Medical Institute	Phase I/II	NCT04276896
bacTRL-Spike	DNA plasmid expressing trimeric S protein and a hybrid transporter protein within <i>Bifidobacterium longum</i> .	Symvivo Corporation	Phase I	NCT04334980
CVnCOV	mRNA-based vaccine mixed with lipid nanoparticles.	CureVac	Phase I	NCT04449276
SCB-2019	Recombinant SARS-CoV-2 trimeric S protein subunit vaccine.	Clover Biopharmaceuticals	Phase I	NCT04405908
COVAX-19	Adjuvanted monovalent recombinant COVID19 vaccine.	Vaxine Pty Ltd.	Phase I	NCT04453852
Pathogen-specific aAPC	Artificial antigen-presenting cells modified with lentiviral vector expressing synthetic minigene based on domains of selected viral proteins.	Shenzhen Geno-Immune Medical Institute	Phase I	NCT04299724
Adjuvanted SARS-CoV-2 Sclamp Protein Subunit Vaccine	Molecular clamp stabilized Spike protein with MF59 adjuvant	University of Queensland	Phase I	ACTRN12620000674932

presented by MHC I and II molecules, the activated T-cells will trigger potent immune responses specific for the pathogen. Identification of peptides presenting T-cell epitopes on MHC I and MHC II is integral in the design and construction of epitope-based vaccines. HLA molecules exhibit high polymorphisms and hundreds of different alleles exist in the human population. Significant overlaps of peptide-binding to different HLAs have been demonstrated. Specifically, HLA-A02, A03 and B07 were reported to cover approximately 90% of the world's population [14]. HLA alleles have been shown to confer different degrees of viral susceptibility and severity of diseases. For example, HLA-B*46:01 was identified as the HLA allele with the fewest predicted binding peptides derived from SARS-CoV-2 and interestingly, HLA-B*15:03 showed the greatest binding capacity to highly conserved SARS-CoV-2 peptide epitopes. When analyzing binding affinity predictions across 145 different HLA-A, B and C alleles, alleles HLA-A*02:02, HLA-B*15:03 and HLA-C*12:03 were the top HLA alleles presenting conserved peptides of SARS-CoV-2 [15].

The roles of B- and T-cells in SARS-CoV-2 infection

The production of epitope-specific antibodies is a primary mechanism of protection against viruses. Neutralizing antibodies can inhibit infection by blocking host–cell interactions and entry of viruses. Antibodies also mediate the destruction of infected cells by inducing opsonization for phagocytosis, complement activation, and antibody-dependent cell-mediated cytotoxicity. Coronavirus-induced antibody responses are quite variable and rather short-lived [16,17]. Patients of SARS-CoV-2 displayed an antibody response within the first 10 days after the onset of symptoms [18]. Specifically, SARS-CoV-2 S-specific IgM antibodies were detectable from day 4 after the onset of symptoms while SARS-CoV-2 S-specific IgG antibodies were measurable from day 7 onwards [19]. SARS-CoV-2 IgM reached the highest point within 3 weeks of infection and gradually declined for the following 4 weeks. IgG was persistently detectable for seven weeks [20]. Further studies are required to elucidate the persistence of antibodies and their roles in the prevention of secondary infections. Previous studies reported the crucial role of T-cells in the elimination of SARS-CoV-1 [21]. SARS-CoV-1-specific memory T-cells were found to persist in the peripheral blood of SARS patients for up to 6 years post-infection despite a lack of specific memory B-cell response [22]. Ng et al. (2016) demonstrated the presence of memory T cell responses in SARS-recovered individuals up to 11 years post-infection [23]. These findings suggested that SARS-CoV-1-specific T-cell response is important for protection against re-infection of SARS-CoV-1. For SARS-CoV-1 (2002/03), it was shown that CD4+ T-cell responses correlated well with positive outcomes [16,24]. Helper T-cells stimulate B-cells and other immune defenders into action, whereas CD8+ cytotoxic T-cells target and eliminate virus-infected cells to prevent virus replication and facilitate virus clearance. Inducing memory T-cells with high proliferation potential should further enhance the development of efficacious vaccines. The severity of the viral disease may depend on the strength of these T-cell responses [25].

The total number of CD4+ and CD8+ T-cells was found to be dramatically reduced in COVID19 patients, especially those requiring intensive care. Expression of immune-inhibitory surface markers such as PD-1 and Tim-3 are known to be associated with the functional exhaustion of T-cells. ICU patients with COVID-19 disease showed markedly higher percentages of PD-1+CD8+ and CD4+ T-cells as well as Tim-3+CD4+ T-cells, indicating that SARS-CoV-2 could drive T-cell exhaustion in COVID-19 patients [26]. Studies also showed that the levels of IL-2, IL-7, IL-10, TNF- α , G-CSF, IP-10, MCP-1 and MIP-1A were significantly higher in COVID-19 patients [26,27]. Increased cytokine levels (IL-6, IL-10, and TNF- α), lymphopenia (low levels of CD4+ and CD8+ T cells), and decreased IFN- γ expression in CD4+ T-cells were reported in studies that characterized the cytokine storm in patients with severe COVID-19 [28]. CD8+ T- and NK-cells from COVID-19 patients had increased expression of the inhibitory receptor NKG2A with diminished productions of CD107a, IFN- γ , IL-2 and granzyme B which indicated the functional exhaustion of NK- and CD8+ T-cells [29].

Bioinformatics approach for B- and T-cell epitope prediction

Peptide-based vaccines comprising short immunogenic peptides have the ability to elicit potent and targeted immune responses. Bioinformatics approach offers a new *in silico* strategy for mapping and identification of potential B- and T-cell epitopes in the protein sequences of viruses which are useful for the development of effective multi-epitope vaccines. Peptide antigens bound to MHC molecules are the target of T-cells. Various bioinformatics tools that allow prediction of the binding affinity of peptides to MHC class I and II molecules have been developed. Here, we present our *in silico* predictions of potential T-cell epitopes. Epitopes were predicted from the spike (S), membrane (M), and nucleocapsid (N) proteins. First, the protein sequences were retrieved from the NCBI database (<https://www.ncbi.nlm.nih.gov/nuccore/MN996531.1/>). Each sequence is checked on its average antigenic propensity using the antigenic peptide prediction tool (<http://imed.med.ucm.es/Tools/antigenic.pl>). The protein sequences were considered for CD4+ and CD8+ T-cell epitope selections, respectively. CD4+ T-cell epitopes were predicted using the MHC-II epitope prediction tool from the Immune Epitope Database (IEDB, <http://tools.iedb.org/mhcii/>). This server uses a combination of different methods including Artificial Neural Network (ANN) and Quantitative Affinity Matrix (QAM) for finding the binding affinity of peptides for class II MHC. A total of 4047 possible epitopes from S, M, and N proteins were determined using this server. These CD4+ T-cell epitopes were filtered based on their lowest percentile rank and IC₅₀ values. Only epitopes with percentile ranks less than 5.0 and IC₅₀ values of below 50 nM were considered. Potential CD4+ T-cell epitopes were further refined based on their immunogenicity score using the IEDB CD4+ T-cell immunogenicity prediction tool (<http://tools.iedb.org/CD4episcore/>). All predicted epitopes have immunogenicity scores above 80.0. These CD4+ T-cell epitopes were also predicted based on their capability to induce Th1 immune

Table 2 SARS-CoV-2 CD4+ T-cell epitopes predicted in S, M, and N proteins.

Spike	Membrane	Nucleocapsid
VVLSFELLHAPATVC	GLMWLSYFIASFRLF	AQFAPSASAFFGMSR
QLLIRAAEIRASANL	NLVIGFLFLTWCILL	AALALLLDRLNQL
GNYNLYRFLFRKSNL		
FVFLVLLPLVSSQCV		
SSGWTAGAAAYVGY		
ITSGWTFGAGAAALQI		
NFTISVTTEILPVS		

*Percentile rank of ≤ 5.0 was adjusted to be the cut-off indicator for IEDB MHC-II epitope prediction tool.

response accompanied by IFN- γ production using the IFN epitope server (<http://crdd.osdd.net/raghava/ifnepitope/>). Here, we used a hybrid approach that combines motif-based models from the pattern discovery software MERCI and Support Vector Machine (SVM) based models to identify IFN- γ producing peptides. Positive IFN- γ producing epitopes were favoured and used as metrics to select better candidates for downstream analysis. Simultaneously, CD8+ T-cell epitopes were predicted using the NetCTL1.2 server (<http://www.cbs.dtu.dk/services/NetCTL/>) and epitopes with prediction scores over 1.0 were selected. Epitopes were further filtered based on their percentile rank using the IEDB MHC-I epitope prediction tool (<http://tools.iedb.org/mhci/>) that also uses QAM and ANN methods to predict binding affinity. A total of 11 CD4+ T-cell epitopes [Table 2] and 13 CD8+ T-cell epitopes [Table 3] were predicted in this manner.

Global population coverage of both HLA class I and II alleles were investigated against the selected CD4+ and CD8+ epitopes respectively using the IEDB population coverage analysis tool (<http://tools.iedb.org/population/>). The 11 selected CD4+ T-cell epitopes were shown to cover 99.36% of global population alleles whereas the 13 CD8+ T-cell epitopes exhibited 91.21% global population coverage. All epitopes were checked to be non-toxic using the ToxinPred module (https://webs.iitd.edu.in/raghava/toxinpred/multi_submit.php). Protein BLAST analysis was conducted to determine homology of these epitopes with human proteins (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&BLAST_SPEC=&LINK_LOC=blasttab). None of the epitopes shared identity with human protein sequences as no alignment score above 67.0 was observed.

Table 3 SARS-CoV-2 CD8+ T-cell epitopes predicted in S, M, and N proteins.

Spike	Membrane	Nucleocapsid
YLQPRTFLL	ATSRTLSYY	LLLDRLNQL
KIADYNYKL	RIAGHHLGR	QFAPSASAF
SIHAYTMSL	YFIASFRLF	
VLNDILSRL		
FIAGLIAIV		
WTAGAAAYY		
TSNQVAVLY		
GVYFASTEKS		

*The cut-off score was adjusted to ≥ 1.0 for NetCTL1.2.

Therefore, this eliminates chances of the epitopes being recognized as self-epitopes.

In addition, we have also predicted linear B-cell epitopes in the S protein using *in silico* analysis. Linear B-cell epitopes are favored over conformational B-cell epitopes because linear B-cell epitopes contain peptides which can represent antigens for immunizations and antibody production. Using the S protein sequence from the NCBI database, we used the IEDB linear B-cell epitope prediction tool (<http://tools.iedb.org/bcell/>) to determine conserved sequences of the B-cell epitopes. The recognition of B-cell epitopes is dependent on the prediction of linear epitopes, its antigenicity, and its surface accessibility. Therefore, three different prediction methods that are available in the IEDB prediction software were utilized. BepiPred Linear Epitope Prediction 2.0 method was used to predict the linear B-cell epitopes, while the Kolasker & Tongaokar Antigenicity method [30] and Emini Surface Accessibility Prediction method were used to evaluate antigenicity and surface accessibility of the linear B-cell epitopes. BepiPred2.0 uses a Random Forest algorithm trained on epitopes and non-epitopes amino acids determined from antibody-antigen crystal structures and performs a sequential prediction smoothing in the end. Lastly, the predicted epitopes were confirmed by ABCPred (<https://webs.iitd.edu.in/raghava/abcpred/index.html>), which applied ANN using fixed-length patterns trained on 700 B-cell epitopes and 700 random non-epitopes in its prediction algorithm with a 65.93% accuracy.

A threshold score of 0.500 was assigned for the BepiPred method and 34 linear B-cell epitopes were predicted in this manner. Next, the antigenicity of these B-cell epitopes was evaluated using a window size of 7 amino acids and a threshold score of 1.041 based on the Kolasker method. The surface accessibility of the B-cell epitopes was examined using a window size of 6 amino acids and a threshold score of 1.000 on the Emini tool. All residues with a score of above 1.000 on the Emini tool were selected as it indicates an increased probability for the residue to be located on the surface. Finally, we confirmed our predictions in the ABCPred server, with a threshold of 0.51 and using an over-lapping filter. A total of seven linear B-cell epitopes were predicted after filtering based on their antigenicity, surface accessibility, and confirmation using ABCPred [Table 4]. These seven epitopes have an average ABCPred score of 0.75 and are exposed on the surface of the virus, in addition to their high antigenicity score. This indicates their potential to elicit humoral immune responses.

Various publications have also predicted a number of T-cell epitopes of SARS-Cov-2 with bioinformatic approaches depending on high immunogenicity scores, binding affinity (IC_{50}) to HLA molecules, HLA coverage and IFN- γ production. Here, we list some of the notable results and cross-reference them with our predicted epitopes. Kiyotani et al. reported potential epitopes derived from the SARS-CoV-2 for HLAs commonly present in the Japanese population through comprehensive bioinformatic screening. The peptides were predicted using NetMHCv4.0 and NetMHCpanv4.0 software based on the high binding affinity to HLA class I and II which led to induction of CD8+ and CD4+ T-cell responses, respectively. They identified 781 HLA-class I and 418 HLA-class II epitopes that have high homologies to SARS-CoV-1. Two

Table 4 SARS-CoV-2 B-cell epitopes predicted in S protein.

Sequence	Start-End (Amino Acid Length)	ABCPred Score
SQCVNLTTRTQLPPAYTNSFTRGVY	13–37 (25)	0.90
DPFLGVYYHKNNKSWME	138–154 (17)	0.68
NLDSKVGGNYNLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTN	440–501 (62)	0.95
NCTEVPVAIHADQLTPT	616–632 (17)	0.81
ASYQTQTNSPRRARSVASQ	672–690 (19)	0.65
KQIYKTPPIKDFGGF	786–800 (15)	0.66
SCCKFDEDDSEPVKLG	1252–1267 (16)	0.59

*The cut-off score was adjusted to ≥ 0.500 for BepiPred2.0.

SARS-CoV-2-derived HLA-class I epitopes in the ORF1ab protein were found with high coverage (83%) of the Japanese population-based on HLA-A frequency [31]. Baruah et al. identified five CD8+ T-cell epitopes and five discontinuous B-cell epitopes in the S glycoprotein. All five CD8+ T-cell epitopes were classified as strong binders to their corresponding MHC class I supertype representatives (A2, A3, B7, B44, and B62) which accounted for the major HLA-A and B distributions in China. Importantly, these epitopes were observed to bind with the peptide-binding groove of MHC class I molecules during molecular dynamic simulations. Therefore, these epitopes could serve as potent T-cell epitopes for the development of SARS-CoV-2 vaccines. Sequential B-cell epitopes in the S protein were identified with BepiPred 2.0 and overlapping filters with ABCPred. To further identify discontinuous B-cell epitopes, Ellipro was used to model the S protein structure in Raptor X [32]. Sahoo et al. identified 36 epitopes for class I MHC and 25 epitopes for class II MHC molecules from the ORF1ab polypeptide, surface glycoprotein, membrane glycoprotein and nucleocapsid phosphoprotein of SARS-CoV-2 based on the maximum binding score to HLA molecules. These epitopes were predicted by using Propred and Propred1 immunoinformatics tools which cover the maximum number of HLAs compared to other epitope prediction tools. Among the seven T-cell epitopes identified in the S protein, peptide KIADYNYKL was predicted to bind to the highest number of MHC class I alleles [33]. Parvez et al. applied NetMHCpan, NetCTL and NetMHCII servers to predict T-cell epitopes while BepiPred 2.0 and ABCPred were used for prediction of linear B-cell epitopes in the SARS-CoV-2 genome [34]. The prediction of class I MHC epitopes was made based on the binding score of NETMHCpan4.0, followed by the affinity to be processed by the cytosolic pathway provided by NetCTL 1.2. Epitopes having a high binding affinity to MHC class I molecules and a high ability to be transported by TAP will have a higher probability to function as cytotoxic T-lymphocyte (CTL) epitopes. A total of 54 MHC class I epitopes have been identified and four promiscuous epitopes which can bind to different alleles were found in the spike protein and they are WTAGAAAYY, CVA-DYSVLY, VRFPNITNL and TRFQTLAL. Twenty-one CD4+ T-cell epitopes were identified based on the binding affinity to

MHC class II by using NetMHCII 2.3. There were 24 out of 45 B-cell epitopes found in the S protein of SARS-CoV-2, making it the highest number of predicted B-cell epitopes across all the SARS-CoV-2 proteins [34]. The immunoinformatics approach has been used by Mishra et al. to identify potential T-cell epitopes from structural and non-structural proteins for immunotherapy against SARS-CoV-2. Two different prediction algorithms, NetCTLpan and PickPocket, were used to generate consensus epitopes against 12 HLA supertypes (HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*24:02, HLA-A*26:01, HLA-B*07:02, HLA-B*08:01, HLA-B*27:05, HLA-B*39:01, HLA-B*40:01, HLA-B*58:01, HLA-B*15:01) and 57 epitopes have been identified in the key proteins of SARS-CoV-2. Many of these epitopes formed a tight cluster around consensus sequences MGYINVFAPFTIYSLLLC and KVSIIWNLDYIINLI across the ORF6 and ORF8 proteins [35]. Grifoni et al. listed dominant SARS-CoV-1 T-cell epitopes based on the response factor score and HLA restrictions. Fourteen SARSCoV-2 T-cell epitopes with 100% identity to SARS-CoV-1 were identified when SARS-CoV-2 T-cell epitopes were mapped against SARS-CoV-1. Only one out of the ten identified SARS-CoV-2 B-cell epitope sequences, FGA-GAALQIPFAMQMAYRFNGI, shared a 100% identity with SARS-CoV-1 after sequence alignment was performed [36]. Fast et al. identified a list of high potential T-cell epitopes through an analysis of human antigen presentations and ranked them based on their likely coverage of presentations by MHC-I and MHC-II alleles. Importantly, the epitope SYGFQPTNGVGYQPY was highly ranked based on MHC presentations and contains predicted promising B-cell epitopes. This epitope is localized near the S protein receptor-binding domain (RBD), suggesting its potential to trigger B- and T-cell responses [37]. *In silico* analysis identified 17 peptides representing both B- and T-cell epitopes and 13 peptides representing only T-cell epitopes based on the highest HLA scores, as well as B- and T-cell epitope counts. These findings provided us with valuable information to design highly conserved B- and T-cell epitope-based peptide vaccines against SARS-CoV-2 [38]. However, these epitopes have been generated using peptide prediction algorithms and may not represent the natural peptides presented on HLA class I and class II molecules of the SARS-CoV-2

Table 5 Predicted SARS-CoV-2 T-cell epitopes in the S protein.

CD4+ T-cell			CD8+ T-cell		
Sequence	No of studies	References	Sequence	No of studies	References
AAAYVGYLQPRFL	3	[34,37,38]	YLQPRFLL	7	[31,32,34,35,37,38,44]
APHGVVFLHVTYVPA	2	[37,38]	KIADYNYKL	5	[33,34,38,44,45]
SYGFQPTNGVGYQPY	2	[37,38]	RLFRKSNLK	5	[34,36,38,44,45]
HWFVTQRNFYEPQII	2	[36,38]	NYNLYRLF	5	[31,34,36,38,41]
YLYRLFRKSNLKPFERDI	2	[36,38]	WTAGAAAYY	4	[32,34,37,38]
GVVFLHVTYVPAQEK	2	[34,38]	GVYFASTEK	4	[32,34,38,44]
IIAYTMSLGAENSV	2	[34,36]	QIYKTPPIK	4	[34,36,38,45]
EFVFNIDGYFKIYS	2	[38,44]	VLNDILSRL	3	[34,36,38]
VYYHKNNKS	2	[38,45]	GVYYHKNNK	3	[34,38,45]
VFNATRFAS	2	[34,45]	FIAGLIAIV	3	[34,36,38]
YQTQNSPR	2	[38,45]	SHIAYTMSL	3	[36–38]
FKNHTSPDV	2	[38,45]	SPRRARVA	3	[38,44,45]
			FERDISTEI	3	[36,38,45]
			RNFYEPQII	3	[36,38,45]
			VVFLHVTYV	3	[37,38,41]
			IAPGQTGKI	2	[33,38]
			APGQTGKIA	2	[33,38]
			NIDGYFKIY	2	[34,38]
			CVADYSVLY	2	[34,38]
			RLDKVEAEV	2	[34,38]
			VTYVPAQEK	2	[34,38]
			VYSTGNSVF	2	[34,38]
			QYIKWPWYI	2	[34,38]
			YFPLQSYGF	2	[34,38]
			RFDNPVLPF	2	[34,38]
			EYVSQPFLM	2	[34,36]
			PYRVVLSF	2	[34,38]
			VYDPLQPEL	2	[34,38]
			FVFNIDGY	2	[34,38]
			VRFPNITNL	2	[34,38]
			TRFQTLAL	2	[34,36]
			TRTQLPPAY	2	[34,38]
			HADQLTPTW	2	[34,38]
			RSFIEDLLF	2	[34,38]
			NLNEIDL	2	[36,38]
			YGFQPTNGV	2	[37,38]
			IPTNFTISV	2	[37,44]
			KVGGNYNYL	2	[36,45]
			EGFNCYFPL	2	[38,45]
			RVYSTGNSV	2	[38,45]
			LGAENSVAY	2	[36,44]
			ELDSFKEEL	2	[38,45]
			FKNHTSPDV	2	[38,45]
			DEDDSEPV	2	[38,45]

infected cells. Additionally, prediction algorithms sort potential peptides based on the predicted MHC binding scores. Most often, only the top scoring or dominant peptides are chosen for follow up studies, but it is likely that some of the epitopes that are naturally associated with class I molecules do not follow such sorting criteria. Potent peptides could have been missed with such stringent cut-offs, including those peptides that have lower binding affinities. Therefore, the functional activities of these predicted peptides representing CD4+ and CD8+ T-cell epitopes should be further evaluated by experimental validations such as the secretion of IFN- γ and the expression of a degranulation marker such as CD107a and granzyme B by flow cytometry.

Studies on the SARS-CoV-1 epidemic in 2003 have shown that adaptive immune responses towards the spike (S)

glycoprotein were protective [39,40]. Hence, characterization and detection of SARS-CoV-2 spike glycoprotein (S)-reactive CD4+ T-cells in peripheral blood are likely to be critical in inducing protective antibody responses. Braun et al. (2020) demonstrated the presence of S-reactive CD4+ T-cells in 83% of COVID-19 patients and 34% of SARS-CoV-2 seronegative healthy donors (HD). Two overlapping peptide pools spanning the entire S protein that comprised different counts of MHC II epitopes were identified. The peptide pools comprised 21 and 13 predicted MHC II epitopes covering the N- and C-terminal regions, respectively. S-reactive CD4+ T-cells were identified by co-expression of 4-1BB and CD40L on the PBMCs isolated from COVID-19 patients after peptide stimulations. The majority of S-reactive T-cells in COVID-19 patients co-expressed CD38 and HLA-DR which indicated effector T-cell responses

Table 6 Predicted and cross-referenced CD4+ and CD8+ T-Cell and linear B-cell epitopes in the S protein.

Sequence	Linear B Cell	Start-End (Amino Acid Length)	CD4+ T-Cell Epitopes			CD8+ T-Cell Epitopes		Global HLA Population Coverage
			Percentile Rank	IC ₅₀ (nM)	IFN-γ Production Result	Immunogenicity Score	Prediction Score	
WTAGAAAYVGYLQPRTELLKY	No	258–279 (22)	3.80	42	Positive	97.729	1.5152	99.63%
NYNLYRLFRKSNLKPFDISTEI	Yes	448–472 (25)	0.2	8	Positive	81.4673	1.9496	97.95%
SYGFQPTNGVGYQPYRVVLSFELLHAPAT	Yes	494–523 (30)	0.24	5	Positive	99.222	1.8786	98.57%
FPQSAPHGVVFLHVTYVPAQEK	No	1052–1073 (22)	7.9	67	Positive	98.6684	1.043	96.41%
VPAQEKNFTTAPAICHGDKAHFPREGVFEV	No	1068–1096 (29)	8.9	155	Positive	99.9474	N/A	28.63%
CASYQTQTNSPRRARSVASQSIIAYTMSL	Yes	671–699 (29)	5	27	Positive	99.0918	1.369	91.87%
GTKREDNPVLPFNDGVYFAST	No	75–95 (21)	9.3	367	Positive	95.4416	1.0633	43.07%

during viral infections [17]. Nine MHC I epitopes clustering in the receptor-binding domain and S2 domain were identified by stimulating the splenocytes with pools of 15-mer overlapping peptides spanning the SARS-CoV-2, SARS-CoV-1, and MERS-CoV S protein. T-cell responses were validated by using IFN-γ ELISPOT analysis [41]. Grifoni et al. reported the presence of SARS-CoV-2-specific CD8+ and CD4+ T-cells in ~70% and 100% of convalescent samples, respectively, from predominantly mild COVID-19 patients. CD8+ T-cell responses were detected against immunogenic peptides not only from the S protein but also from M, N, nsp6 and ORF3a. CD4+ T cell responses were also detected in 40–60% of unexposed individuals, suggesting pre-existing SARS-CoV-2-cross-reactive T cell responses in certain healthy individuals [42]. For B-cell epitopes, Poh et al. have identified two linear epitopes on the SARS-CoV-2 S protein that elicited neutralizing antibodies in COVID-19 patients. Potent B-cell epitopes in the sera of COVID-19 and SARS patients were assessed by using pools of overlapping linear B-cell peptides spanning the entire S protein of SARS-CoV-2 or SARS-CoV-1. Two distinct peptide pools from the SARS-CoV-2 S peptide library, pools S14 and S21, were strongly recognized by sera from COVID-19 patients [43].

In an effort to select promising T-cell epitopes that were predicted by bioinformatics, CD4+ and CD8+ T-cell epitopes that were predicted to be present in the spike (S) protein of SARS-CoV-2 from a total of 11 different publications were compared [Table 5]. Among the top seven highly potent CD8+ T-cell epitopes (reported in at least four studies), both YLQPRTEFL and WTAGAAAYVGYLQPRTEFLKYKKKKK reported in the study of Feng et al. Additionally, the top CD4+ T-cell epitope “AAAYVGYLQPRTEFL” predicted from bioinformatics analysis was also present in the P6 peptide sequence, suggesting that this peptide might contain the potent epitopes needed to trigger effective CD4+ and CD8+ T-cell responses. Short peptides such as the P10 peptide “NYNYLYRLFRKSNLKPFDISTEI” reported by Feng et al. contained 3 potent CD8+ T-cell epitopes which were reported in at least three studies. Comparative analysis of prediction from different servers will narrow down the number of epitopes that have to be verified through experimental approaches for the selection of immunodominant CD4+ and CD8+ T-cell epitopes.

Some of the epitopes presented in [Table 5] can be observed to overlap with epitopes from [Tables 2 and 3], further confirming that the predicted T-cell epitopes are promising epitopes that bind to MHC class I and II molecules. A few potential epitopes that represent B- and T-cell epitopes from the S protein were identified after amalgamating information gathered from [Tables 2–5] and are presented in [Table 6]. Epitope sequences “WTAGAAAYVGYLQPRTEFLKY”, “NYNYLYRLFRKSNLKPFDISTEI”, “SYGFQPTNGVGYQPYRVVLSFELLHAPAT”, “FPQSAPHGVVFLHVTYVPAQEK”, and “CASYQTQTNSPRRARSVASQSIIAYTMSL” are the top candidates as they not only exhibit positive IFN-γ production, but they also exhibit high global HLA population coverages. Furthermore, these five epitopes are also capable of inducing both CD4+ and CD8+ T-cell responses, in contrast to the epitope sequence “VPAQEKNFTTAPAICHGDKAHFPREGVFEV” which is implicated to only induce CD4+ T-cell responses.

Even though both epitope sequences “VPAQEKNFTTA-PAICHGKAHFPRE GVFV” and “GTKRFDNPVLPFDNGVYFAST” show high immunogenicity and binding affinity scores, they have relatively low global HLA population coverages. Moreover, epitope sequences “NYNYLYRLFRKSNLKPFFERDISTEI”, “SYGFQPTNGVGYQPYRVVLSFELLHAPAT”, and “CASYQTQTNSPRRARSVASQSIIAYTMSL” contain overlapping B-cell epitope sequences presented in [Table 4]. All of the epitopes listed in [Table 4] contain B-cell epitopes predicted by the ABCPred server with threshold value of 0.51. However, our stringent cut offs ensure that the B-cell epitopes are linear and not conformational, highly antigenic, and are easily accessible from the surface.

T-cell epitope prediction servers such as the IEDB T-cell prediction server, NetMHC, NetMHCpan, NetMHCIIpan, and NetCTL are commonly utilized. These servers leverage on data-driven analysis over structure-based methods and utilize machine learning to overcome problems with non-linearity of peptide-MHC binding data. This allows researchers to formulate predictions based on the binding affinity of peptides to MHC molecules in addition to addressing the issue of the discrimination of MHC binders to nonbinders. All the aforementioned servers employ ANN algorithms while IEDB uses QAM in addition to ANN which further enhances binding affinity predictions. Furthermore, pan-MHC methods such as NetMHCpan and NetMHCIIpan were used to further improve predictions of peptide binding affinities to uncharacterized HLA alleles. On the other hand, common servers such as BepiPred2.0 and ABCPred are used to predict linear B-cell epitopes. Similar to the common T-cell epitope predicting servers, BepiPred2.0 and ABCPred leverage on machine learning algorithms as well. BepiPred uses a Random Forest algorithm whereas ABCPred uses ANN. These methods have been known to outperform B-cell epitope prediction servers that are solely based on amino acid propensity scales.

Design of multi-epitope peptide-based vaccines

In silico prediction revealed peptides presented by HLA class I and class II molecules to CD4+ and CD8+ T-cells, respectively. Predicted immunodominant T-cell epitopes can be used for rational design of multi-epitope peptide-based vaccines but effectiveness as immunogenic epitopes when presented as short peptides need to be validated. Alternatively, synthetic genes can be constructed to encode these immunogenic peptides which can be cloned into vectors like pcDNA or pVAX to generate DNA vaccines which can continuously express the antigens [46,47]. A DNA-based vaccine, INO-4800, was designed to encode the SARS-CoV-2 S protein. The immunogenicity of INO-4800 was demonstrated in multiple animal models. The DNA vaccine was able to induce antibodies that blocked S protein binding to host receptors and elicited good INF- γ levels on Day 10 post immunization. The use of adjuvants has been reported to further enhance the immunogenicity of vaccines [48,49]. For example, Montanide ISA 51 in combination with a cocktail of HIV-derived peptide antigens (containing a mixture of peptide specifying B- and T-cell epitopes) enhanced the immunogenicity of the antigenic peptide [50]. The epitopes can also be fused with adjuvant proteins

[e.g., toll-like receptor (TLR) ligands] which can be encapsulated or displayed on the surface of nanoparticles [e.g., poly lactic-co-glycolic acid (PLGA)] to elicit sustained T-cell responses and increase long term protection [8,51,52].

Several successful studies have led to the rational designs of epitope-based peptide vaccines. For example, Flunisyn™, Flu-V and Multimeric-001 (M-001) peptide-based influenza vaccine candidates are currently in clinical Phase I, IIb and III evaluations, respectively. FP-01.1 (Flunisyn™) is a novel synthetic influenza A vaccine consisting of six fluorocarbon-modified 35-mer conserved epitopes of influenza A NP, M, PB1 and PB2 proteins. The peptide vaccine carried multiple CD4+ and CD8+ T-cell epitopes and was shown to generate robust CD4+ and CD8+ T-cell immunity in phase I clinical trial [53]. Flu-V is a multi-epitope peptide vaccine comprising 4 peptides representing CD8+ T-cell epitopes originating from conserved regions in internal proteins M1, NPA, NPB and M2 from influenza A and influenza B. A single dose of FLU-v with adjuvant Montanide ISA 51 led to a significant reduction in the number of influenza symptoms in 40% of participants when compared to those receiving placebo (64.3%) [54]. M-001 is a DNA vaccine carrying nine conserved genes encoding B- and T-cell epitopes from the HA (four B and one CD4+ T-cell epitopes), NP (Two CD8+ and one CD4+ T-cell epitope), and M1 (one peptide that contains both B and CTL epitope) proteins of both influenza type A and type B strains. The genes encoding nine conserved linear epitopes (repetitions of three epitopes) were cloned into a vector and a recombinant protein (M001) was expressed in *Escherichia coli*. Administration of M001 vaccine with Montanide ISA 51VG adjuvant successfully induced production of antibodies and IFN- γ in humans [55]. Vacc-Flu is another peptide-based influenza vaccine candidate which contained one peptide to induce B-cell responses and three peptides to induce T-cell responses to conserved regions in the M2 and NP influenza A virus proteins. Peptide selection was conducted using a proprietary peptide design platform technology focusing on responses to human leukocyte antigen (HLA)-restricted epitopes. Immunization of Vacc-Flu into C57BL/6 and HLA-A2 transgenic mice resulted in the induction of IFN- γ -producing T-cells and influenza-specific antibodies which protected mice from influenza infection [56].

Peptides often lack immunogenicity and are susceptible to enzymatic digestion. Therefore, effective vaccine delivery and adjuvant systems which can enhance the immunogenicity and protect the peptide vaccines will need to be incorporated in the vaccine formulation [57]. PLGA is a non-toxic, FDA and European Medicines Agency approved polymer and is widely used as a vehicle for drug and vaccine delivery [58]. The size of the nanoparticle can also affect the uptake efficiency and the induction of immune response by APCs. Smaller PLGA NPs (350 nm) enable better internalization by DCs and led to a sustained cellular immune response in mice [59]. PLGA-NP encapsulated peptide antigens are being protected from enzymatic degradation *in vivo* until they are taken up by antigen-presenting cells (APCs) and led to induction of robust B- and T-cell responses [60,61]. Intranasal delivery of PLGA NP-entrapped inactivated porcine reproductive and respiratory syndrome virus (PRRSV) vaccine adjuvanted with soluble *Mycobacterium tuberculosis* whole cell lysate was observed to enhance antigen-specific IFN- γ secreting T-cell responses and

also led to enhanced production of high-affinity neutralizing antibodies which induced a cross-protective immune response in pigs [62]. A multi-peptide vaccine containing five overlapping peptide sequences covering amino acids 693–777 (unit B/C) of the E2 protein of classical swine fever virus (CSFV) was shown to completely protect pigs from lethal CSFV challenge [63]. Entrapment of H1N1 influenza virus-derived peptides in PLGA nanoparticles was demonstrated to enhance T-cell responses and vaccine efficacy in pigs [64].

Epitope-based peptide vaccine was also designed for the Middle East Respiratory Syndrome Coronavirus (MERS-COV). Several conserved B- and T-cell epitopes were identified in the S protein of MERS-COV which might elicit protection against MERS-COV infection. In particular, MHC class-I peptide YKLQPLTFL and MHC class-II peptide YCILEPRSG were predicted as highly antigenic whilst QLQMGFGITVQYGT displayed the highest antigenicity-score among the B-cell epitopes [65]. However, validation of these peptides as T-cell epitopes is needed.

Two different adverse syndromes might be associated with vaccine-enhanced COVID-19 disease: antibody-dependent enhancement (ADE) and vaccine-associated enhanced respiratory disease (VAERD). ADE is a phenomenon in which binding of a virus to non-neutralizing antibodies enhances its entry into host cells through Fc receptors. VAERD occurs when there is increased production of Th2 cytokines (IL-4, IL-5 and IL-13) which can lead to airway hyper-responsiveness and lung histopathology [66]. Most SARS-CoV-2 vaccines under development are aiming to elicit an immune response against the spike protein but previous studies showed that vaccine platforms expressing the whole SARS-CoV-1 S glycoprotein or using the whole inactivated virus was inducing enhanced respiratory disease which resulted in adverse lung pathology in animal models [67,68]. Inactivated MERS-CoV vaccine has been shown to increase eosinophil infiltrations and Th2 cytokine productions, leading to lung immunopathology after challenge in mice [69]. Thus, the usage of truncated proteins such as the receptor-binding domain (RBD) or the S1 subunit of the S protein could lead to the induction of neutralizing antibodies and avoid ADE. Therefore, the precise selection of CD4+, CD8 T- and B-cell epitopes is important to develop a safe and efficacious vaccine against SARS-CoV-2.

Conclusions

DNA and RNA vaccines are new technologies which enable rapid production of unconventional vaccines. These vaccine platforms have never been applied to any previously licensed human vaccines. The specific delivery method to attain high immunogenicity is currently one of the frontier problems of DNA vaccines while mRNA vaccines are poor in stability and may lack immunogenicity due to low translational efficiency [70]. Other platforms such as inactivated vaccines (IVs) and live attenuated vaccines (LAVs) are established platforms approved by the US FDA for many licensed human vaccines. Most of the licensed vaccines approved by the US Food and Drug Administration against viral infections to date are live-attenuated or inactivated vaccines. Despite the ability of LAVs to elicit lifelong immunity, it is associated with the risk

of reversion to the wild type and requires extensive safety monitoring. Moreover, the large-scale productions of LAVs and IVs are laborious and require containment facilities for growing the live SARS-CoV-2 and subsequent purification processes. To meet the world's demand of millions of doses of vaccine, there will be significant safety risks involved in the growth of the pathogen and downstream processing.

China's leading vaccine manufacturers Sinovac and Sinopharm have provided reassuring information on their inactivated vaccine candidates. The phase II clinical trial results of Sinovac's vaccine candidate, CoronaVac, showed that it induced neutralizing antibodies 14 days after vaccination. The neutralizing seroconversion rate of CoronaVac was over 90% in 600 healthy volunteers in phase II of their phase I/II study conducted in Brazil. This infers that CoronaVac was capable of inducing a positive immune response. Moreover, Sinopharm reported that their inactivated vaccine candidate triggered a strong neutralizing antibody response in all 1120 participants in the 28-day dosing schedule which only employed a middle-strength dosage in their phase I/II study. Additionally, the seroconversion rate for the 14-day and 21-day schedules of the mid-dose was 97.6% and 100% at 28 days.

The selection of B-cell, CD4+ and CD8+ T-cell epitopes based on bioinformatic predictions will require further experimental validations such as *ex-vivo* stimulations by the respective peptides to measure functional T-cell responses. Immunogenicity study of peptides are required to confirm their ability to elicit long term protection in animal models. For B-cell epitopes, quantifications and comparisons of the B-cell epitope-specific IgG titres against the RBD domain, or to other conformational epitopes are required by isolating monoclonal antibodies (mAB)s targeting the selected epitopes. Furthermore, the persistence level of these antibodies against linear and other conformational epitopes have to be thoroughly investigated. Phylogenetic analysis of the SARS-CoV-2 S protein should also be performed to determine the evolutionary relationship of SARS-CoV-2 with other coronaviruses such as SARS-CoV and MERS-CoV. This would help to locate highly conserved epitopes that could be used for construction of multi-epitope vaccine that can confer broad protection against coronaviruses.

The success of both mRNA-1273 by Moderna and Ad5-nCoV by CanSino in their phase I/II clinical trials have provided valuable information. mRNA-1273 evoked neutralizing antibody titer levels generally observed in convalescent sera within their initial 45 participants. It is also generally safe and provided complete protection in a mouse challenge model. Ad5-nCoV was capable of significantly increasing neutralizing antibodies after 28 days post-vaccination, in addition to specific T-cell responses. Furthermore, no serious adverse effects were reported, and only mild to moderate adverse effects were observed in 75%–83% of the participants [71,72].

Several multi-epitope vaccines against SARS-CoV-2 were designed by connecting multiple predicted B-, CD4+ and CD8+ T-cell epitopes to an adjuvant using EAAAK linkers [44,73]. However, the ability of multi-epitope peptide-based vaccine in eliciting immune responses and conferring protection against SARS-CoV-2 requires a thorough investigation. *In silico* approaches of identifying specific immunogenic epitopes would help to rationally design a multi-epitope peptide vaccine that

could serve the urgent need for an effective SARS-CoV-2 vaccine.

In conclusion, we have analysed the SARS-CoV-2 genome for both B- and T-cell epitope candidates. We then further reinforced our predictions with pre-existing data from various bioinformatics publications. Finally, we have indicated seven multi-epitope sequences where five of the sequences we believe could be integral in the development of an effective peptide-based vaccine that is highly immunogenic and has a wide global population coverage. However, *in vitro* evaluation and *in vivo* trials are necessary to confirm the effectiveness of these peptide-based vaccines. Such experimental validations will be essential to help prioritize the selection of immunogenic epitopes and hasten vaccine designs during this unprecedented crisis.

Conflicts of interest

The authors declare that there is no conflicts of interest.

Acknowledgments

This study was funded by the Sunway University Research Centre Grant 2020 (STR-RCTR-CVVR-01-2020) to the Centre for Virus and Vaccine Research (CVVR).

REFERENCES

- [1] Sharma A, Tiwari S, Deb MK, Marty JL. Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2): a global pandemic and treatment strategies. *Int J Antimicrob Agents* 2020;56:106054.
- [2] Khailany RA, Safdar M, Ozaslan M. Genomic characterization of a novel SARS-CoV-2. *Gene Rep* 2020;19:100682.
- [3] Funk CD, Laferrrière C, Ardakani A. A snapshot of the global race for vaccines targeting SARS-CoV-2 and the COVID-19 pandemic. *Front Pharmacol* 2020;11:937.
- [4] Li W, Joshi MD, Singhania S, Ramsey KH, Murthy AK. Peptide vaccine: progress and challenges. *Vaccines* 2014;2:515–36.
- [5] Li S, Peng L, Zhao W, Zhong H, Zhang F, Yan Z, et al. Synthetic peptides containing B- and T-cell epitope of dengue virus-2 E domain III provoked B- and T-cell responses. *Vaccine* 2011;29:3695–702.
- [6] Kao DJ, Hodges RS. Advantages of a synthetic peptide immunogen over a protein immunogen in the development of an anti-pilus vaccine for *Pseudomonas aeruginosa*. *Chem Biol Drug Des* 2009;74:33–42.
- [7] Lei Y, Zhao F, Shao J, Li Y, Li S, Chang H, et al. Application of built-in adjuvants for epitope-based vaccines. *PeerJ* 2019;6:e6185.
- [8] Kelly HG, Kent SJ, Wheatley AK. Immunological basis for enhanced immunity of nanoparticle vaccines. *Expert Rev Vaccines* 2019;18:269–80.
- [9] El-Manzalawy Y, Dobbs D, Honavar V. Predicting flexible length linear B-cell epitopes. *Comput Syst Bioinformatics Conf* 2008;7:121–32.
- [10] Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res* 2006;2:2.
- [11] Chandra S, Singh TR. Linear B cell epitope prediction for epitope vaccine design against meningococcal disease and their computational validations through physicochemical properties. *Netw Model Anal Health Inform Bioinform* 2012;1:153–9.
- [12] Malonis RJ, Lai JR, Vergnolle O. Peptide-based vaccines: current progress and future challenges. *Chem Rev* 2020;120:3210–29.
- [13] Cruz-Tapias P, Castiblanco J, Anaya JM. Major histocompatibility complex: antigen processing and presentation. In: Anaya JM, Shoenfeld Y, Rojas-Villarraga A, Levy RA, Cervera R, editors. *Autoimmunity: From Bench to Bedside*. Bogota (Colombia): El Rosario University Press; 2013.
- [14] Zhao L, Zhang M, Cong H. Advances in the study of HLA-restricted epitope vaccines. *Hum Vaccines Immunother* 2013;9:2566–77.
- [15] Nguyen A, David JK, Maden SK, Wood MA, Weeder BR, Nellore A, et al. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *J Virol* 2020;94:e00510–20.
- [16] Li CKf, Wu H, Yan H, Ma S, Wang L, Zhang M, et al. T cell responses to whole SARS coronavirus in humans. *J Immunol* 2008;181:5490–500.
- [17] Braun J, Loyal L, Frensch M, Wendisch D, Georg P, Kurth F, et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* 2020;587:270–4.
- [18] Özcürümez MK, Ambrosch A, Frey O, Haselmann V, Holdenrieder S, Kiehntopf M, et al. SARS-CoV-2 antibody testing—questions to be asked. *J Allergy Clin Immunol* 2020;146:35–43.
- [19] Liu X, Wang J, Xu X, Liao G, Chen Y, Hu CH. Patterns of IgG and IgM antibody response in COVID-19 patients. *Emerg Microb Infect* 2020;9:1269–74.
- [20] Xiao AT, Gao C, Zhang S. Profile of specific antibodies to SARS-CoV-2: the first report. *J Infect* 2020;81:147–78.
- [21] Janice Oh HL, Ken-En Gan S, Bertoletti A, Tan YJ. Understanding the T cell immune response in SARS coronavirus infection. *Emerg Microb Infect* 2012;1:e23.
- [22] Tang F, Quan Y, Xin ZT, Wrammert J, Ma MJ, Lv H, et al. Lack of peripheral memory B cell responses in recovered patients with severe acute respiratory syndrome: a six-year follow-up study. *J Immunol* 2011;186:7264–8.
- [23] Ng OW, Chia A, Tan AT, Jodi RS, Leong HN, Bertoletti A, et al. Memory T cell responses targeting the SARS coronavirus persist up to 11 years post-infection. *Vaccine* 2016;34:2008–14.
- [24] Channappanavar R, Zhao J, Perlman S. T cell-mediated immune response to respiratory coronaviruses. *Immunol Res* 2014;59:118–28.
- [25] Leslie M. T cells found in coronavirus patients 'bode well' for long-term immunity. *Science* 2020;368:809–10.
- [26] Diao B, Wang C, Tan Y, Chen X, Liu Y, Ning L, et al. Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). *Front Immunol* 2020;11:827.
- [27] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506.
- [28] Pedersen SF, Ho YC. SARS-CoV-2: a storm is raging. *J Clin Invest* 2020;130:2202–5.
- [29] Zheng M, Gao Y, Wang G, Song G, Liu S, Sun D, et al. Functional exhaustion of antiviral lymphocytes in COVID-19 patients. *Cell Mol Immunol* 2020;17:533–5.
- [30] Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 1990;276:172–4.
- [31] Kiyotani K, Toyoshima Y, Nemoto K, Nakamura Y. Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. *J Hum Genet* 2020;65:569–75.

- [32] Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. *J Med Virol* 2020;92:495–500.
- [33] Biswajit Sahoo KK, Rai Neeraj K, Kumar Chaudhary Dharmendra. Identification of T-cell epitopes in proteins of novel human coronavirus, SARS-CoV-2 for vaccine development. *Int J Appl Biol Pharm* 2020;37–45.
- [34] Parvez S, Preeti S. Prediction of T and B cell epitopes in the proteome of SARS-CoV-2 for potential use in diagnostics and vaccine design. *bioRxiv* 2020.07.14.202887. [Preprint] 2020 [cited 2020 Sept 21]. Available from: <https://doi.org/10.26434/chemrxiv.12116943.v1>.
- [35] Seema M. T cell epitope-based vaccine design for pandemic novel coronavirus 2019-nCoV. 10.26434/chemrxiv.12029523. [Preprint] 2020 [cited 2020 Sept 21]. Available from: <https://doi.org/10.26434/chemrxiv.12029523.v2>.
- [36] Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. *Cell Host Microbe* 2020;27:671–80. e2.
- [37] Fast E, Altman RB, Chen B. Potential T-cell and B-cell epitopes of 2019-nCoV. *bioRxiv* 2020.02.19.955484 [Preprint] 2020 [cited 2021 Mar 21]. Available from: <https://doi.org/10.1101/2020.02.19.955484>.
- [38] Feng Y, Qiu M, Liu L, Zou S, Li Y, Luo K, et al. Multi-epitope vaccine design using an immunoinformatics approach for 2019 novel coronavirus (SARS-CoV-2). *bioRxiv* 2020:2020.03.03.962332 [Preprint] 2020 [cited 2020 Sept 21]. Available from: <https://doi.org/10.1101/2020.03.03.962332>.
- [39] Yang ZY, Kong WP, Huang Y, Roberts A, Murphy BR, Subbarao K, et al. A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. *Nature* 2004;428:561–4.
- [40] Zhu Z, Chakraborti S, He Y, Roberts A, Sheahan T, Xiao X, et al. Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal antibodies. *Proc Natl Acad Sci USA* 2007;104:12123–8.
- [41] Smith TRF, Patel A, Ramos S, Elwood D, Zhu X, Yan J, et al. Immunogenicity of a DNA vaccine candidate for COVID-19. *Nat Commun* 2020;11:2601.
- [42] Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* 2020;181:1489–501. e15.
- [43] Poh CM, Carissimo G, Wang B, Amrun SN, Lee CY, Chee RS, et al. Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients. *Nat Commun* 2020;11:2806.
- [44] Kalita P, Padhi AK, Zhang KY, Tripathi T. Design of a peptide-based subunit vaccine against novel coronavirus SARS-CoV-2. *Microb Pathog* 2020;145:104236.
- [45] Bhattacharya M, Sharma AR, Patra P, Ghosh P, Sharma G, Patra BC, et al. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immunoinformatics approach. *J Med Virol* 2020;92:618–31.
- [46] Xu Q, Cui N, Ma X, Wang F, Li H, Shen Z, et al. Evaluation of a chimeric multi-epitope-based DNA vaccine against subgroup J avian leukosis virus in chickens. *Vaccine* 2016;34:3751–6.
- [47] Zhao C, Sun Y, Zhao Y, Wang S, Yu T, Du F, et al. Immunogenicity of a multi-epitope DNA vaccine against hantavirus. *Hum Vaccines Immunother* 2012;8:208–15.
- [48] Rueckert C, Guzmán CA. Vaccines: from empirical development to rational design. *PLoS Pathog* 2012;8. e1003001-e.
- [49] Yano A, Onozuka A, Matin K, Imai S, Hanada N, Nisizawa T. RGD motif enhances immunogenicity and adjuvanicity of peptide antigens following intranasal immunization. *Vaccine* 2003;22:237–43.
- [50] Graham BS, McElrath MJ, Keefer MC, Rybczyk K, Berger D, Weinhold KJ, et al. Immunization with cocktail of HIV-derived peptides in montanide ISA-51 is immunogenic, but causes sterile abscesses and unacceptable reactogenicity. *PLoS One* 2010;5:e11995.
- [51] Karuturi BVK, Tallapaka SB, Yeapuri P, Curran SM, Sanderson SD, Vetro JA. Encapsulation of an EP67-conjugated CTL peptide vaccine in nanoscale biodegradable particles increases the efficacy of respiratory immunization and affects the magnitude and memory subsets of vaccine-generated mucosal and systemic CD8(+) T cells in a diameter-dependent manner. *Mol Pharm* 2017;14:1469–81.
- [52] Rueda F, Eich C, Cordobilla B, Domingo P, Acosta G, Albericio F, et al. Effect of TLR ligands co-encapsulated with multiepitopic antigen in nanoliposomes targeted to human DCs via Fc receptor for cancer vaccines. *Immunobiology* 2017;222:989–97.
- [53] Francis JN, Bunce CJ, Horlock C, Watson JM, Warrington SJ, Georges B, et al. A novel peptide-based pan-influenza A vaccine: a double blind, randomised clinical trial of immunogenicity and safety. *Vaccine* 2015;33:396–402.
- [54] Pleguezuelos O, James E, Fernandez A, Lopes V, Rosas LA, Cervantes-Medina A, et al. Efficacy of FLU-v, a broad-spectrum influenza vaccine, in a randomized phase IIb human influenza challenge study. *NPJ Vaccines* 2020;5:22.
- [55] Atsmon J, Kate-Ilovitz E, Shaikovich D, Singer Y, Volokhov I, Haim KY, et al. Safety and immunogenicity of multimeric-001-a novel universal influenza vaccine. *J Clin Immunol* 2012;32:595–603.
- [56] Herrera-Rodriguez J, Meijerhof T, Niesters HG, Stjernholm G, Hovden AO, Sørensen B, et al. A novel peptide-based vaccine candidate with protective efficacy against influenza A in a mouse model. *Virology* 2018;515:21–8.
- [57] Purcell AW, McCluskey J, Rossjohn J. More than one reason to rethink the use of peptides in vaccine design. *Nat Rev Drug Discov* 2007;6:404–14.
- [58] Lü JM, Wang X, Marin-Muller C, Wang H, Lin PH, Yao Q, et al. Current advances in research and clinical applications of PLGA-based nanotechnology. *Expert Rev Mol Diagn* 2009;9:325–41.
- [59] Silva AL, Rosalia RA, Varypataki E, Sibuea S, Ossendorp F, Jiskoot W. Poly-(lactic-co-glycolic-acid)-based particulate vaccines: particle uptake by dendritic cells is a key parameter for immune activation. *Vaccine* 2015;33:847–54.
- [60] Panyam J, Labhasetwar V. Biodegradable nanoparticles for drug and gene delivery to cells and tissue. *Adv Drug Deliv Rev* 2003;55:329–47.
- [61] Nixon DF, Hioe C, Chen PD, Bian Z, Kuebler P, Li ML, et al. Synthetic peptides entrapped in microparticles can elicit cytotoxic T cell activity. *Vaccine* 1996;14:1523–30.
- [62] Binjawadagi B, Dwivedi V, Manickam C, Ouyang K, Wu Y, Lee LJ, et al. Adjuvanted poly(lactic-co-glycolic) acid nanoparticle-entrapped inactivated porcine reproductive and respiratory syndrome virus vaccine elicits cross-protective immune response in pigs. *Int J Nanomed* 2014;9:679–94.
- [63] Dong XN, Wei K, Liu ZQ, Chen YH. Candidate peptide vaccine induced protection against classical swine fever virus. *Vaccine* 2002;21:167–73.
- [64] Hiremath J, Kang KI, Xia M, Elaish M, Binjawadagi B, Ouyang K, et al. Entrapment of H1N1 influenza virus derived conserved peptides in PLGA nanoparticles enhances T cell response and vaccine efficacy in pigs. *PLoS One* 2016;11:e0151922.
- [65] Tahir Ul Qamar M, Saleem S, Ashfaq UA, Bari A, Anwar F, Alqahtani S. Epitope-based peptide vaccine design and target site depiction against Middle East Respiratory Syndrome

- Coronavirus: an immune-informatics study. *J Transl Med* 2019;17:362.
- [66] Graham BS. Rapid COVID-19 vaccine development. *Science* 2020;368:945–6.
- [67] Bolles M, Deming D, Long K, Agnihothram S, Whitmore A, Ferris M, et al. A double-inactivated severe acute respiratory syndrome coronavirus vaccine provides incomplete protection in mice and induces increased eosinophilic proinflammatory pulmonary response upon challenge. *J Virol* 2011;85:12201–15.
- [68] Tseng CT, Sbrana E, Iwata-Yoshikawa N, Newman PC, Garron T, Atmar RL, et al. Immunization with SARS coronavirus vaccines leads to pulmonary immunopathology on challenge with the SARS virus. *PloS One* 2012;7:e35421.
- [69] Agrawal AS, Tao X, Algaissi A, Garron T, Narayanan K, Peng BH, et al. Immunization with inactivated Middle East Respiratory Syndrome coronavirus vaccine leads to lung immunopathology on challenge with live virus. *Hum Vaccines Immunother* 2016;12:2351–6.
- [70] Amanat F, Krammer F. SARS-CoV-2 vaccines: status report. *Immunity* 2020;52:583–9.
- [71] Zhu FC, Li YH, Guan XH, Hou LH, Wang WJ, Li JX, et al. Safety, tolerability, and immunogenicity of a recombinant adenovirus type-5 vectored COVID-19 vaccine: a dose-escalation, open-label, non-randomised, first-in-human trial. *Lancet* 2020;395:1845–54.
- [72] Jackson LA, Anderson EJ, Roupael NG, Roberts PC, Makhene M, Coler RN, et al. An mRNA vaccine against SARS-CoV-2 - preliminary report. *N Engl J Med* 2020;383:1920–31.
- [73] Kar T, Narsaria U, Basak S, Deb D, Castiglione F, Mueller DM, et al. A candidate multi-epitope vaccine against SARS-CoV-2. *Sci Rep* 2020;10:10895.