



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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# A vaccine built from potential immunogenic pieces derived from the SARS-CoV-2 spike glycoprotein: A computational approximation

Jose Marchan <sup>\*</sup>

Experimental Medicine Centre, Venezuelan Institute for Scientific Research, Caracas 1020-A, Venezuela

## ARTICLE INFO

### Keywords:

SARS-CoV-2  
Spike  
COVID-19  
Epitope  
Vaccine

## ABSTRACT

Coronavirus Disease 2019 (COVID-19) represents a new global threat demanding a multidisciplinary effort to fight its etiological agent—severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In this regard, immunoinformatics may aid to predict prominent immunogenic regions from critical SARS-CoV-2 structural proteins, such as the spike (S) glycoprotein, for their use in prophylactic or therapeutic interventions against this highly pathogenic betacoronavirus. Accordingly, in this study, an integrated immunoinformatics approach was applied to identify cytotoxic T cell (CTC), T helper cell (THC), and Linear B cell (BC) epitopes from the S glycoprotein in an attempt to design a high-quality multi-epitope vaccine. The best CTC, THC, and BC epitopes showed high viral antigenicity and lack of allergenic or toxic residues, as well as CTC and THC epitopes showed suitable interactions with HLA class I (HLA-I) and HLA class II (HLA-II) molecules, respectively. Remarkably, SARS-CoV-2 receptor-binding domain (RBD) and its receptor-binding motif (RBM) harbour several potential epitopes. The structure prediction, refinement, and validation data indicate that the multi-epitope vaccine has an appropriate conformation and stability. Four conformational epitopes and an efficient binding between Toll-like receptor 4 (TLR4) and the vaccine model were observed. Importantly, the population coverage analysis showed that the multi-epitope vaccine could be used globally. Notably, computer-based simulations suggest that the vaccine model has a robust potential to evoke and maximize both immune effector responses and immunological memory to SARS-CoV-2. Further research is needed to accomplish with the mandatory international guidelines for human vaccine formulations.

## 1. Introduction

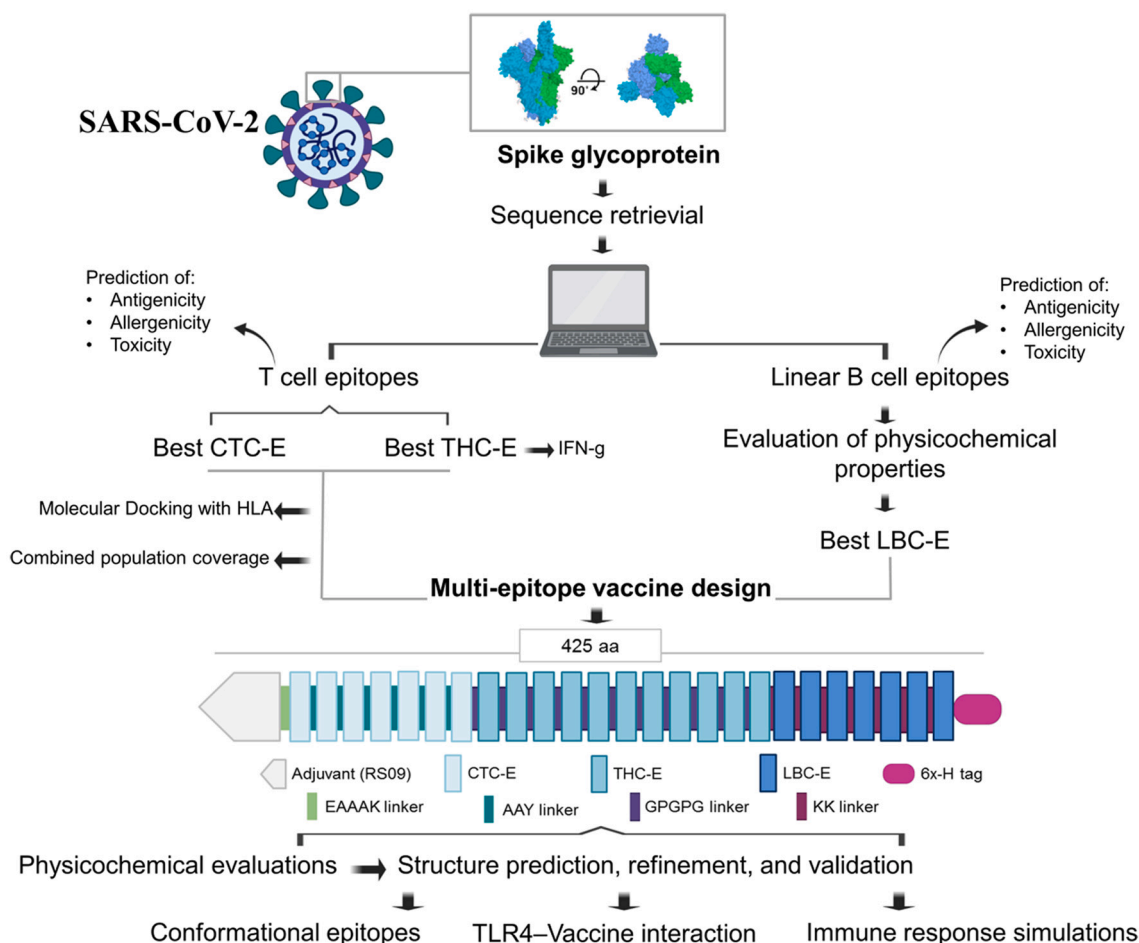
On 31st December 2019, a dramatic increase in the number of patients with a potentially fulminant respiratory disease was reported in Wuhan, China. The etiological agent was eventually identified as a novel highly pathogenic betacoronavirus—severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Lu et al., 2020). Thereafter, SARS-CoV-2 caused an overwhelming wave of coronavirus disease 2019 (COVID-19) cases across Asia, Europe, Oceania, the Americas, and Africa, which led to the World Health Organization declared COVID-19 as a pandemic on 11th March 2020 (World Health Organization, n.d.). Unfortunately, at the time of this research, no clinically approved treatment is available to fight SARS-CoV-2, whose rapid spread has generated explosives waves of COVID-19 all over the world (World Health Organization, n.d.).

The SARS-CoV-2 outer membrane is decorated with several structural proteins, including the S glycoprotein, the membrane protein, and

the envelope protein (Lu et al., 2020). The S glycoprotein forms homotrimers containing both a receptor-binding domain (RBD) and a receptor-binding motif (RBM) (Lan et al., 2020a). The latter mediates contacts with human angiotensin-converting enzyme 2 (hACE2), thereby allowing SARS-CoV-2 entry into host cell (Lan et al., 2020a). This critical role in viral pathogenesis turns the SARS-CoV-2 S glycoprotein into an attractive target for vaccine development (Amanat and Krammer, 2020).

Multi-epitope vaccines designed from immunoinformatics tools could aid to elicit a protective immune response against SARS-CoV-2, as reported previously for other infectious agents (Zhou et al., 2009). In this regard, recent data indicate that the SARS-CoV-2 S glycoprotein harbours prominent immunologically active regions, which may serve as candidates for multi-epitope vaccine models (Bhattacharya et al., 2020). Accordingly, the present study aimed to design a multiple-epitope vaccine construct against SARS-CoV-2 using for this purpose an integrated *in silico* approach.

<sup>\*</sup> Corresponding author at: Experimental Medicine Centre, Venezuelan Institute for Scientific Research (IVIC), Apartado postal 20632, Caracas 1020-A, Venezuela.  
E-mail address: [josemarchanalvarez@gmail.com](mailto:josemarchanalvarez@gmail.com).



**Fig. 1.** Overall experimental workflow (made with Biorender.com). Best epitopes predicted from the SARS-CoV-2 S glycoprotein were selected to design the multi-epitope vaccine construct, which was subjected to further *in silico* evaluations. CTC-E: cytotoxic T cell epitope. THC-E: T helper cell epitope. LBC-E: linear B cell epitopes. IFN-g: Interferon gamma. aa: amino acid. 6x-H: polyhistidine tag.

## 2. Materials and methods

### 2.1. Protein sequence retrieval

Taking into account that the SARS-CoV-2 S glycoprotein represents the major target for vaccine development (Amanat and Krammer, 2020), the present work focused only on such viral spike. The complete amino acid sequence of the SARS-CoV-2 S glycoprotein was retrieved from Uniprot (<http://www.uniprot.org/>) in FASTA format (accession number: P0DTC2). Fig. 1 summarises the *in silico* experimental work.

### 2.2. Prediction of allergenicity, toxicity, and viral antigenicity

Potential epitopes from the S glycoprotein were subjected to allergenic evaluation using AllergenFP (<http://ddg-pharmfac.net/AllergenFP/index.html>) (Dimitrov et al., 2014), whereas toxicity was predicted using the ToxinPred server (<http://crdd.osdd.net/raghava/toxinpred/>) (Gupta et al., 2013). Finally, viral antigenicity was calculated from the Vaxijen server (threshold: 0.5) ([www.ddg-pharmfac.net/vaxijen/](http://www.ddg-pharmfac.net/vaxijen/)) (Doytchinova and Flower, 2007).

### 2.3. Immunogenicity

CTC, THC, and BC epitopes were predicted and the best were selected for the final vaccine design (Fig. 1). To achieve this aim, multiple prediction tools were used to improve the rate of true positives. Furthermore, the algorithms parameters were chosen based on the

recommendations from the software developers/authors.

#### 2.3.1. Prediction of CTC and THC epitopes

Peptides that interact with HLA-I and HLA-II molecules commonly have 9 and 15 amino acids in length, respectively (Owen et al., 2013). In consequence, 9-mer and 15-mer peptides were considered in this work as CTC and THC epitopes, respectively (Fig. 1). These epitopes were identified using the Immune Epitope Database and Analysis Resource (IEDB-AR) (<http://tools.immuneepitope.org/main/>) (Kim et al., 2012).

To cross-validate binding peptides to HLA molecules, several methods were applied. In this regard, the 9-mer binding peptides to HLA-I were predicted using the artificial neural network (ANN) method (Tenzer et al., 2005), the Consensus method (Moutafsi et al., 2006), and the NetMHCpan method (Hoof et al., 2009). The prediction of the 15-mer binding peptides to HLA-II was performed using the Consensus method (Wang et al., 2008), the NetMHCIIpan method (Nielsen et al., 2008), and the SMM-align method (Nielsen et al., 2007). The above algorithms generated a prediction output based on a percentile rank—peptides with a small percentile rank have high affinity by HLA alleles. This percentile rank is produced on IEDB-AR by comparing the IC50 of each predicted peptide against random peptides from SWISS-PROT database. In this work, epitopes were selected by following this guideline as well as by using a percentile rank cut-off  $\leq 20$  as recommended previously (Paul et al., 2015), which has also been successfully applied in other *in silico* studies focused on SARS-CoV-2 (Grifoni et al., 2020a; Marchan, 2020).

In addition, binding peptides to HLA-II were also chosen by their

**Table 1**  
Evaluation of potential T cell epitopes.

ID	Epitope sequence	Position (star-end)	Percentile rank	Viral antigenicity	Allergenicity	Toxicity	IFN-g inducitor	Selected for vaccine design
<i>CTC-E (9-mer peptides):</i>								
E1	EPLVDLPIG	224–232	4.1	0.46	Negative	Negative	ne	No
E2	APGQTGKIA	411–419	1.8	1.20	Positive	Negative	ne	No
<b>E3</b>	<b>VVLSFELLH</b>	<b>511–519</b>	<b>1.5</b>	<b>1.41</b>	<b>Negative</b>	Negative	<b>ne</b>	<b>Yes</b>
E4	FPLQSYGFQ	490–498	0.4	0.46	Positive	Negative	ne	No
E5	ATRFASVYA	344–352	0.2	0.09	Negative	Negative	ne	No
<b>E6</b>	<b>VDLPIGINI</b>	<b>227–235</b>	<b>0.01</b>	<b>1.38</b>	<b>Negative</b>	Negative	<b>ne</b>	<b>Yes</b>
<b>E7</b>	<b>FTISVTTEI</b>	<b>718–726</b>	<b>0.01</b>	<b>0.85</b>	<b>Negative</b>	Negative	<b>ne</b>	<b>Yes</b>
E8	SVYAWNRRK	349–357	0.01	0.76	Positive	Negative	ne	No
E9	YLPRTFLL	269–277	0.02	0.45	Positive	Negative	ne	No
<b>E10</b>	<b>VRFPNITNL</b>	<b>327–335</b>	<b>0.03</b>	<b>1.11</b>	<b>Negative</b>	Negative	<b>ne</b>	<b>Yes</b>
E11	FERDISTEI	464–472	0.04	0.74	Positive	Negative	ne	No
E12	TLDKSTQSL	109–117	0.04	1.06	Positive	Negative	ne	No
E13	KIADYNYKL	417–425	0.04	1.66	Positive	Negative	ne	No
E14	YLPRTFLL	269–277	0.04	0.45	Positive	Negative	ne	No
<b>E15</b>	<b>YSKHTPINL</b>	<b>204–212</b>	<b>0.05</b>	<b>1.05</b>	<b>Negative</b>	Negative	<b>ne</b>	<b>Yes</b>
E16	VGYLQPRTF	267–275	0.06	1.22	Positive	Negative	ne	No
E17	TKSFTVEK	302–310	0.06	0.08	Positive	Negative	ne	No
<b>E18</b>	<b>FEYVSQPFL</b>	<b>168–176</b>	<b>0.07</b>	<b>0.63</b>	<b>Negative</b>	Negative	<b>ne</b>	<b>Yes</b>
<b>E19</b>	<b>GFQPTNGVG</b>	<b>496–504</b>	<b>5.9</b>	<b>0.64</b>	<b>Negative</b>	Negative	<b>ne</b>	<b>Yes</b>
E20	VRQIAPGQT	407–415	18.0	0.86	Positive	Negative	ne	No
<i>THC-E (15-mer peptides):</i>								
E21	ITRFQTLALHRSYL	235–270	0.02	0.11	Negative	Negative	Yes	No
E22	QSLIVNNATNVVIVK	115–129	0.02	0.43	Negative	Negative	No	No
E23	LSFELLHAPATVCGP	513–527	0.03	0.51	Negative	Negative	No	No
<b>E24</b>	<b>VVLSFELLHAPATVC</b>	<b>511–525</b>	<b>0.03</b>	<b>0.86</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
E25	SLIVNNATNVVIVK	116–130	0.03	0.47	Negative	Negative	Yes	No
<b>E26</b>	<b>KTQSLIVNNATNVV</b>	<b>113–127</b>	<b>0.17</b>	<b>0.63</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
<b>E27</b>	<b>AIPTNFTISVTTEIL</b>	<b>713–727</b>	<b>0.40</b>	<b>0.68</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
<b>E28</b>	<b>SFVIRGDEVRIAPG</b>	<b>399–413</b>	<b>0.51</b>	<b>0.58</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
E29	TPINLVRDLPPQGFSA	208–222	0.51	0.55	Negative	Negative	No	No
<b>E30</b>	<b>TRFASVYAWNRRKRIS</b>	<b>345–358</b>	<b>0.52</b>	<b>0.50</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
E31	IPNFTISVTTEILP	714–728	0.52	0.83	Negative	Negative	No	No
E32	PTESIVRFPNITNLC	322–336	0.64	0.25	Negative	Negative	No	No
<b>E33</b>	<b>ECSNLLQYGSFCTQ</b>	<b>748–762</b>	<b>0.72</b>	<b>0.76</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
E34	LQIPFAMQMAYRFNG	894–908	0.73	0.72	Negative	Negative	No	No
E35	HTPINLVRDLPPQGF	207–221	0.74	0.39	Negative	Negative	No	No
E36	ADYSVLYNSASFSTF	363–377	0.85	0.22	Negative	Negative	No	No
E37	SKTQSLIVNNATNV	112–126	0.99	0.62	Positive	Negative	No	No
E38	TDEMIAQYTSALLAG	866–880	1.10	0.16	Negative	Negative	Yes	No
E39	QMAYRFNGIGVTQNV	901–915	1.10	1.03	Negative	Negative	No	No
E40	QIPFAMQMAYRFNGI	895–909	0.44	0.96	Negative	Negative	No	No
E41	AALQIPFAMQMAYRF	892–906	0.44	0.91	Negative	Negative	No	No
<b>E42</b>	<b>LEPLVDLPIGINITR</b>	<b>223–237</b>	<b>28.00</b>	<b>1.01</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
<b>E43</b>	<b>DEVRIAPGQTGKIA</b>	<b>405–419</b>	<b>27.00</b>	<b>0.98</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
<b>E44</b>	<b>EVRQIAPGQTGKIAD</b>	<b>406–420</b>	<b>33.00</b>	<b>1.34</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
<b>E45</b>	<b>VRQIAPGQTGKIADY</b>	<b>407–421</b>	<b>41.00</b>	<b>1.30</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
<b>E46</b>	<b>RQIAPGQTGKIADYN</b>	<b>408–422</b>	<b>49.00</b>	<b>1.52</b>	<b>Negative</b>	Negative	<b>Yes</b>	<b>Yes</b>
E47	EGFNCYFPLQSYGFQ	484–498	7.20	0.57	Negative	Positive	Yes	No

E: Epitope; CTC-E: cytotoxic T cell epitope; THC-E: T helper cell epitope; ne: not evaluated; Epitopes highlighted in boldface were selected for final vaccine design.

potential to induce interferon-gamma (IFN-g) (Fig. 1), which is a cytokine necessary to fight viral infections (Owen et al., 2013). Epitopes with a high potential to induce the production of IFN-g were selected using the IFNepitope server (<http://crdd.osdd.net/raghava/ifnepitope/>) (Dhanda et al., 2013). This website harbours three models (motif based, SVM based and hybrid approach), which has been trained on 10,433 experimentally validated IFN-gamma inducing and non-inducing MHC class II peptides (Dhanda et al., 2013).

### 2.3.2. Prediction of linear BC epitopes

BCPRED (<http://ailab.ist.psu.edu/bcpred/>) (Saha and Raghava, 2006) was used to predict linear BC epitopes based on several physicochemical properties: hydrophilicity, flexibility, accessibility, and antigenicity propensity (threshold = 1 for each parameter). Simultaneously, the S glycoprotein amino acid sequence was also subjected to iBCE-EL (<http://thegleelab.org/iBCE-EL/>) (Manavalan et al., 2018) and BepiPred-2.0 (<http://www.cbs.dtu.dk/services/BepiPred/>) (Jespersen et al., 2017) for additional predictions of linear BC epitopes.

### 2.3.3. Three-dimensional (3D) interaction between HLA alleles and viral peptides: Molecular docking

To evaluate the presentation of the best epitopes in the context of HLA molecules, a molecular docking study was conducted (Fig. 1). Taking into account that HLA-A\*02:01 and HLA-DRB1\*01:01 were predicted as common interacting HLA alleles, they were selected for this purpose. The molecular docking simulation and the Gibbs free energy ( $\Delta G$ ) of the HLA-viral peptide complexes were evaluated as recently reported (Marchan, 2020).

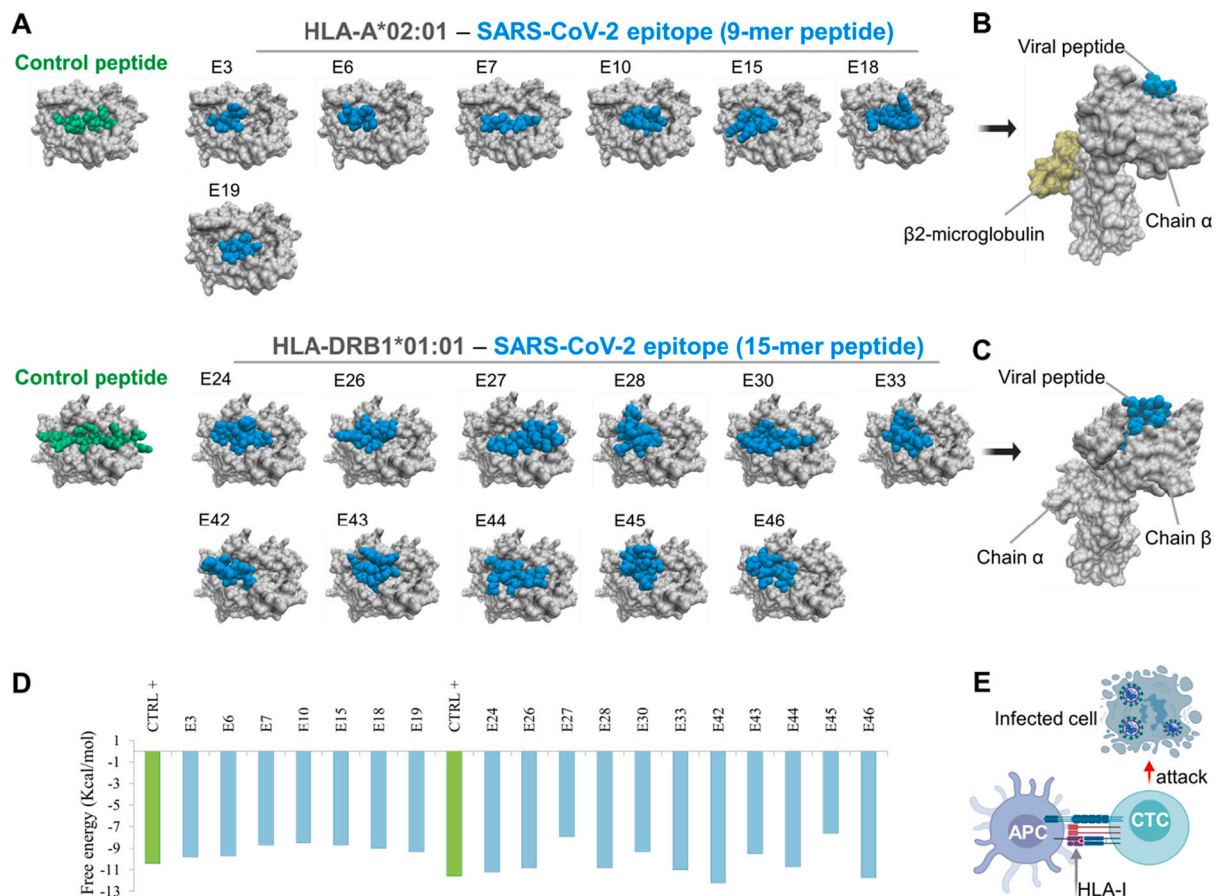
### 2.4. Design of the multi-epitope vaccine against SARS-CoV-2

High potential CTC, THC, and linear BC epitopes were selected to generate the amino acid sequence of the multi-epitope vaccine. The CTC and THC epitopes were linked together using AAY and GPGPG linkers, respectively, whereas linear BC epitopes were connected by KK linkers (Fig. 1). Moreover, a TLR4 agonist, known as RS09 (Sequence: APPHALS) (Shanmugam et al., 2012), was added as an adjuvant at the N-terminus by using an EAAAK linker (Fig. 1). For future validation

**Table 2**  
Evaluation of potential linear B cell epitopes.

ID	Epitope sequence	Position (start-end)	Predicted on BCPRED, iBCE-EL, and BepiPred-2.0	Viral antigenicity	Allergenicity	Toxicity	Selected for vaccine design
E1	<b>TTLDSKTQSL</b>	127–136	Yes	0.99	Negative	Non-Toxin	Yes
E2	<b>MDLEGKQGNFKNLREF</b>	196–211	Yes	0.83	Negative	Non-Toxin	Yes
E3	<b>PDPSPSKRS</b>	826–835	Yes	0.84	Negative	Non-Toxin	Yes
E4	<b>ILDITPCSFGGVSVITPG</b>	603–620	Yes	1.10	Negative	Non-Toxin	Yes
E5	YQPYRVVLSFELLH	524–538	Negative on iBCE-EL	0.97	Negative	Non-toxin	No
E6	<b>FSTFKCYGVSP</b>	270–281	Yes	0.81	Negative	Non-toxin	Yes
E7	VYYHKNNKSWMESEFRVYSS	162–181	Yes	0.30	Negative	Non-toxin	No
E8	GDEVRQIAPGQTGKI	423–437	Negative on iBCE-EL	0.97	Negative	Non-toxin	No
E9	<b>NLDSKVGGNYNY</b>	459–470	Yes	1.09	Negative	Non-toxin	Yes
E10	<b>GFQPTNGVGYQPYR</b>	496–509	Yes	0.72	Negative	Non-toxin	Yes

E: epitope; Epitopes highlighted in boldface were selected for final vaccine design.



**Fig. 2.** Screenshots of the HLA-viral peptide complexes. (A) Top view of HLA class I (HLA-A\*02:01) and HLA class II (HLA-DRB1\*01:01) presenting 9-mer and 15-mer viral peptides, respectively. HLA alleles, viral peptides, and control peptides are shown in grey, cyan, and green, respectively. Epitopes are named according to the nomenclature established in Table 1. (B–C) Representative lateral views of HLA alleles interacting with viral peptides. Importantly, peptides formed a bulge that project from the HLA groove, which could lead to a more direct interaction with immune cell receptors such as the T cell receptor. (D) Free energy for each HLA-viral peptide interaction. Note that the values of control peptides (CTRL+) are shown in green bars. (E) Diagram showing the presentation of viral peptides in the context of HLA-I alleles to Cytotoxic T Cell (CTC), which may target SARS-CoV-2 infected cells. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

studies, the vaccine molecule must be expressed *in vitro* and then purified. Therefore, a polyhistidine-tag (6×-H tag) was included at the C-terminus (Fig. 1), which would allow its purification (Loughran et al., 2017).

**Table 3**  
Physicochemical parameters of the multiple-epitope vaccine construct.

Parameter	Value	Comment
Number of aa	425	Suitable
Molecular weight	44 kDa	Average
Theoretical pI	9.59	Slightly basic
Negatively charged residues (N + Q)	25	
Positively charged residues (R + K)	43	Charged positive
Extinction coefficient (at 280 nm in water)	31,080 M <sup>-1</sup> cm <sup>-1</sup>	–
The instability index (II)	26.42	Stable
Aliphatic index	74.82	Thermostable
GRAVY	–0.306	Hydrophilic
Solubility	0.703269	Soluble
Estimated half-life:		Satisfactory
Mammalian reticulocytes ( <i>in vitro</i> )	4.4 h	
Yeast ( <i>in vivo</i> )	>20 h	
<i>Escherichia coli</i> ( <i>in vivo</i> )	>10 h	

aa: amino acid; GRAVY: Grand average of hydropathicity.

#### 2.4.1. Physicochemical evaluation and general studies

The ProtParam tool (<https://web.expasy.org/protparam/>) (Wilkins et al., 1999) was used to examine relevant physicochemical parameters of the multi-epitope vaccine. To reconfirm its viral antigenicity and lack of allergenicity and toxicity, the web tools described in Section 2.2 were applied. In addition, the vaccine solubility was predicted using the SOLpro server (<http://scratch.proteomics.ics.uci.edu/>) (Magnan et al., 2009).

#### 2.4.2. Structure prediction, refinement, and validation

PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) (Buchan and Jones, 2019) and GalaxyWEB (<http://galaxy.seoklab.org/>) (Ko et al., 2012) were utilized to predict the secondary and tertiary structure, respectively, of the multi-epitope vaccine construct. The best model was refined with GalaxyWEB (Ko et al., 2012). The vaccine structure was validated by comparing with experimentally validated 3D protein structure. In this regard, the vaccine structure was submitted to ProSA-web (<https://prosa.services.came.sbg.ac.at/prosa.php>), which provides a general quality score for a given structure (Wiederstein and Sippl, 2007). Furthermore, the Ramachandran plot was created on PROCHECK website (<https://servicesn.mbi.ucla.edu/PROCHECK/>) whereby the protein structure can be validated according to energetically allowed and disallowed dihedral angles psi and phi of amino acid residues (Laskowski et al., 1993).

#### 2.4.3. Prediction of conformational BC

B-cells are also considered professional antigen presenting cells and they can initiate this function by recognizing the antigen through B-cell receptors (Owen et al., 2013). Therefore, conformational epitopes of the multi-epitope vaccine construct were predicted from Ellipro (<http://tool.s.iedb.org/ellipro/>) (Ponomarenko et al., 2008), which represents the protein structure as an ellipsoid and calculates protrusion indexes for protein residues outside of such ellipsoid (Ponomarenko et al., 2008). Minimum levels of 0.8 and a distance of 6.0 Å were applied. The epitopes were visualized with the VMD software (Version 1.9.3) to illustrate their position and 3D structure as previously reported (Marchan, 2020).

#### 2.4.4. Innate immune receptor-Vaccine interaction: Molecular docking

Since TLR4 may serve as a sensor for the recognition of coronaviruses S glycoproteins (Lester and Li, 2014), this germline-encoded pattern recognition receptor was selected for the docking study. The 3D structure of TLR4 was obtained from PDB (accession number: 4G8A). The refined model of the multi-epitope vaccine was used as a ligand. The TLR4-Vaccine docking simulation and its 3D visualization were performed as recently reported (Marchan, 2020).

#### 2.4.5. *In silico* immune response simulations

To further characterize the potential immune response of the multi-

epitope vaccine, immune simulations were performed using the C-ImmSim server (<http://150.146.2.1/C-IMMSIM/index.php>) (Rapin et al., 2010). Three injections were applied four weeks apart as described previously (Nain et al., 2020). Furthermore, 12 injections were applied four weeks apart to simulate repeated exposure to the potential immunogen. The Simpson index D was used to interpret the diversity of the immune response.

#### 2.4.6. Population coverage

Global population coverage of the multi-epitope vaccine construct was calculated from IEDB-AR (<http://tools.iedb.org/population/>) (Bui et al., 2006). The HLA allele genotypic frequencies available on IEDB-AR were obtained from Allele Frequency Database (AFD) (<http://www.allelefrequencies.net/>). At present, AFD contains allele frequencies from 115 countries and 21 different ethnicities (<http://www.allelefrequencies.net/>). Those 115 countries were selected and the HLA-I and HLA-II interacting alleles predicted in this work were included to perform an HLA combined analysis. A highest result means that the vaccine harbours a broad HLA coverage and, therefore, it could be applied all over the world. The results were shown on a world map using Datawrapper (<https://www.datawrapper.de/>).

### 3. Results

#### 3.1. Prediction of T cell epitopes

The *in silico* approach (Fig. 1) allowed predict a total of 47 T cell epitopes; however, 7 cytotoxic T cell (CTC) and 11 T helper cell (THC) epitopes were identified as the best (Table 1). These epitopes showed a potent viral antigenicity—ranging from 0.63 to 1.52—and lack of allergenic or toxic residues in their sequences (Table 1). Moreover, THC epitopes were characterized by their potential capability to induce IFN- $\gamma$  (Table 1). Although “EGFNCFYFPLQSYGFQ” (E47 in Table 1) could be categorized as a strong potential THC epitope, it was identified as a probable inducer of toxicity. Therefore, this epitope was not included in the amino acid sequence of the multi-epitope vaccine.

##### 3.1.1. HLA-I and HLA-II interacting alleles

The selected CTC epitopes (Table 1) showed promiscuous affinity by several HLA-I alleles, including HLA-A\*01:01, HLA-A\*02:01, HLA-A\*03:01, HLA-A\*11:01, HLA-A\*23:01, HLA-A\*25:01, HLA-A\*30:01, HLA-A\*68:01, HLA-A\*74:01, HLA-B\*07:02, HLA-B\*08:01, HLA-B\*13:01, HLA-B\*13:02, HLA-B\*14:02, HLA-B\*15:01, HLA-B\*15:02, HLA-B\*18:01, HLA-B\*27:02, HLA-B\*35:03, HLA-B\*40:01, HLA-B\*58:01, HLA-C\*01:02, HLA-C\*02:02, HLA-C\*02:09, HLA-C\*03:02, HLA-C\*03:03, HLA-C\*03:04, HLA-C\*04:01, HLA-C\*05:01, HLA-C\*06:02, HLA-C\*07:01, HLA-C\*08:01, HLA-C\*12:02, HLA-C\*12:03, HLA-C\*14:02, HLA-C\*15:02, HLA-C\*16:01, and HLA-C\*17:01. Likewise, the selected THC epitopes (Table 1) showed common interaction with the following HLA-II alleles: HLA-DRB1\*01:01, HLA-DRB1\*01:03, HLA-DRB1\*07:01, HLA-DRB1\*15:01, HLA-DRB1\*01:20, HLA-DRB3\*02:02, HLA-DRB4\*01:01, and HLA-DRB5\*01:01.

#### 3.2. Prediction of linear B cell (BC) epitopes

A total of 10 linear BC epitopes of varying amino acid lengths were predicted (Table 2). Most of the epitopes showed robust viral antigenicity ( $\geq 0.5$ ), as well as, they were identified as non-allergenic and non-toxic (Table 2). However, only 7 epitopes were selected for the vaccine design due to they were predicted simultaneously by 3 different web tools (BCPred, iBCE-EL, and BepiPred-2.0) (Table 2). Interestingly, overlapping residues were observed between some linear BC and T cell epitopes.



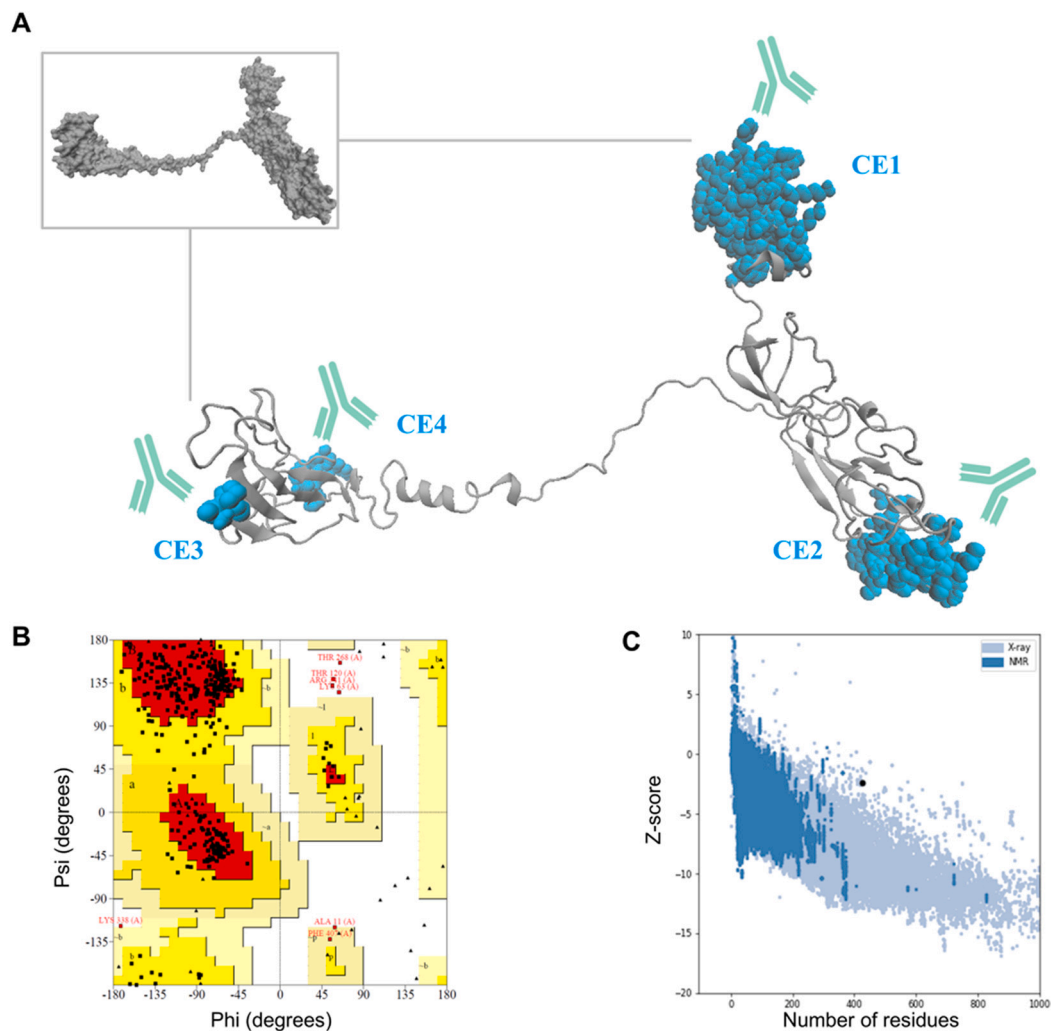
Fig. 3. Secondary structure prediction of the multi-epitope vaccine against SARS-CoV-2. This graphic was obtained from PSIPRED.

### 3.3. HLA allele-viral peptide interaction: Molecular docking

To evaluate the presentation of the best epitopes in the context of HLA, molecular docking simulations were conducted. For this purpose, HLA-A\*02:01 and HLA-DRB1\*01:01 were chosen as representative alleles.

HLA-I and HLA-II alleles were docked with CTC and THC epitopes, respectively, using the Cluspro server, which has been recently applied to successfully dock epitopes from SARS-CoV-2 non-structural proteins into HLA molecules (Marchan, 2020). The inspection on VDM software

allowed observing different binding patterns wherein viral peptides rightly interact with the active site residues of the HLA groove in a similar way to control peptides (Fig. 2A). Moreover, several viral peptides (e.g., E18 and E33) formed a bulge that projected from their respective HLA allele (Fig. 2B and C), which is relevant, for instance, to activate CTC against SARS-CoV-2 infected cells (Owen et al., 2013) (Fig. 2E). Importantly, each HLA-viral peptide complex showed robust potential interactions (free energy values  $-7 < \text{kcal/mol}^{-1}$ ) comparable to control peptides (Fig. 2D).



**Fig. 4.** Tertiary structure prediction and validation of the multi-epitope vaccine. (A) Tertiary structure after refinement showed in space-filling model (inset) and best conformational B cell epitopes (CE) identified in the vaccine construct. Of note, these epitopes (cyan atoms) showed high probability scores ( $>0.8$ ), thereby suggesting a considerable accessibility for antibodies (showed in turquoise). (B) Ramachandran plot of the 3D refined structure. (C) Z-score plot obtained from ProSA-web. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 3.4. Design of the multi-epitope vaccine against SARS-CoV-2: General evaluations

To design the amino acid sequence of the multi-epitope vaccine, epitopes were organized using several linkers (Fig. 1). This sequence is constituted by 425 amino acid residues (Fig. 1).

Of particular note, several epitopes selected for the vaccine design (E19, E42, E43, E44, and E45 in Table 1; E10 in Table 2) harbour residues that are usually involved in the interaction between the SARS-CoV-2S glycoprotein and hACE2 (Lan et al., 2020a; Ortega et al., 2020; Lan et al., 2020b). For instance, N501—which is present in the amino acid sequence of E19 (Table 1) and E10 (Table 2)—has been recently described as one of the critical hACE2-binding residues in SARS-CoV-2 (Lan et al., 2020a).

The vaccine showed a strong viral antigenicity (0.64), as well as neither allergenic nor toxic residues were observed in its amino acid sequence. Furthermore, the physicochemical properties examined with the ProtParam tool, including molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index, and GRAVY, were computed as conventional results (Table 3).

#### 3.4.1. Structure prediction, refinement, and validation

The vaccine construct was analysed using the PSIPRED server to predict its secondary structure, which identified 309, 65, and 51 amino acids forming coil, helix, and strand regions, respectively (Fig. 3).

The predicted tertiary structure was subjected to refinement using the GalaxyRefine server. The output showed four potential models. Model 1 (Fig. 4A) was classified as the best according to the web tool. Therefore, this model was selected for further analysis. In this regard, the Ramachandran plot (Fig. 4B) showed that 98.1 of residues were located in allowed regions, whereas the remaining residues were observed in disallowed regions (1.9%). In addition, the Z-score value ( $-2.4$ ) (Fig. 4C) suggests that the vaccine structure is similar to native proteins of comparable size.

#### 3.4.2. Prediction of conformational BC epitopes

Four conformational BC epitopes (CE) were predicted using Ellipro (Fig. 4A). These CE showed high probability scores—CE1: 0.911, CE2: 0.819, CE3: 0.803, and CE4: 0.803 suggesting a considerable accessibility for antibodies (Fig. 4A). Likewise, these results also confirm the immunogenic potential of the multi-epitope vaccine construct.



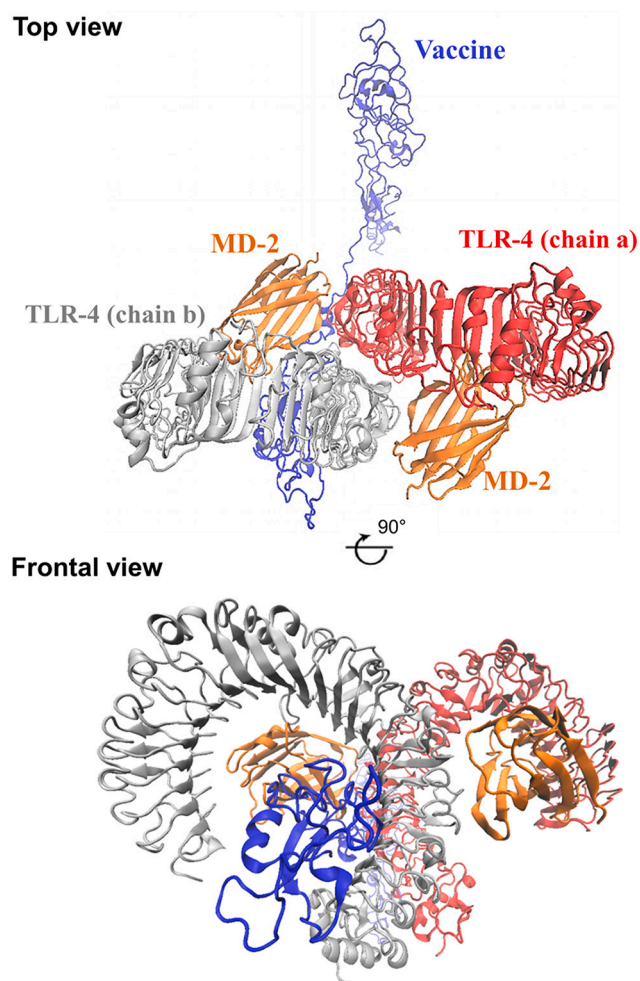


Fig. 5. Docked complex of TLR4, MD-2 (myeloid differentiation factor 2), and the multi-epitope vaccine.

#### 3.4.3. TLR4-vaccine model interaction: Molecular docking

The vaccine model showed a suitable interaction (Fig. 5) between the TLR-4 chain B and the myeloid differentiation factor 2 (MD-2) molecule, which are known to initiate the cascade signalling pathways *in vivo* (Owen et al., 2013), thereby suggesting that such a vaccine model could elicit an appropriate immune response against SARS-CoV-2. Importantly, the adjuvant inserted in the vaccine sequence was observed in the interaction zone with the TLR4 and MD-2, which indicates the relevance of such an adjuvant in the potential efficacy of the present vaccine model.

#### 3.4.4. Immune response simulations

The immune response simulations with the multi-epitope vaccine construct (3 doses given 4 weeks apart) showed cell-mediated and humoral responses. As expected, increased number and activity of Natural Killer (NK) cells—a relevant line of attack against viruses (Owen et al., 2013), and macrophages were observed (Fig. 6). Regarding the adaptive immune response, CTC and THC populations showed a proliferative burst, effector cell generation, and a dramatic cell number contraction (Fig. 6). Importantly, IL-2, which is necessary for T cell activation and optimal proliferation (Owen et al., 2013), was amplified after each dose (Fig. 6). Moreover, the vaccine model increased BC and plasma cell populations, particularly immunoglobulin M (IgM) and IgG1 isotypes (Fig. 6). In this regard, titres of IgM, IgG1, and IgG2 were higher in the secondary and tertiary response compared to primary response (Fig. 6). Of note, immunogen concentrations decreased after antibody response

(Fig. 6). Notably, repeated exposure with 12 injections (given 4 weeks apart) increased the IgG1 levels and stimulated CTC and THC populations (Fig. S1). Taken together, these results suggest that the multi-epitope vaccine could evoke and maximize both effector responses and immunological memory to SARS-CoV-2.

#### 3.4.5. Population coverage

To investigate whether the multi-epitope vaccine may be used in different ethnic groups or globally, a population coverage analysis was performed. Remarkably, the multi-epitope vaccine construct showed high global population coverage: 99.69%. For instance, several countries with positive reports of COVID-19 (>6000 cases) (World Health Organization, n.d.), obtained the highest values, including, Australia, Brazil, Ecuador, Chile, China, France, Germany, India, Iran, Israel, Italy, Japan, Mexico, Morocco, Peru, Philippines, Russia, Singapore, South Korea, Spain, Sweden, USA, UK, etc., (Fig. 7).

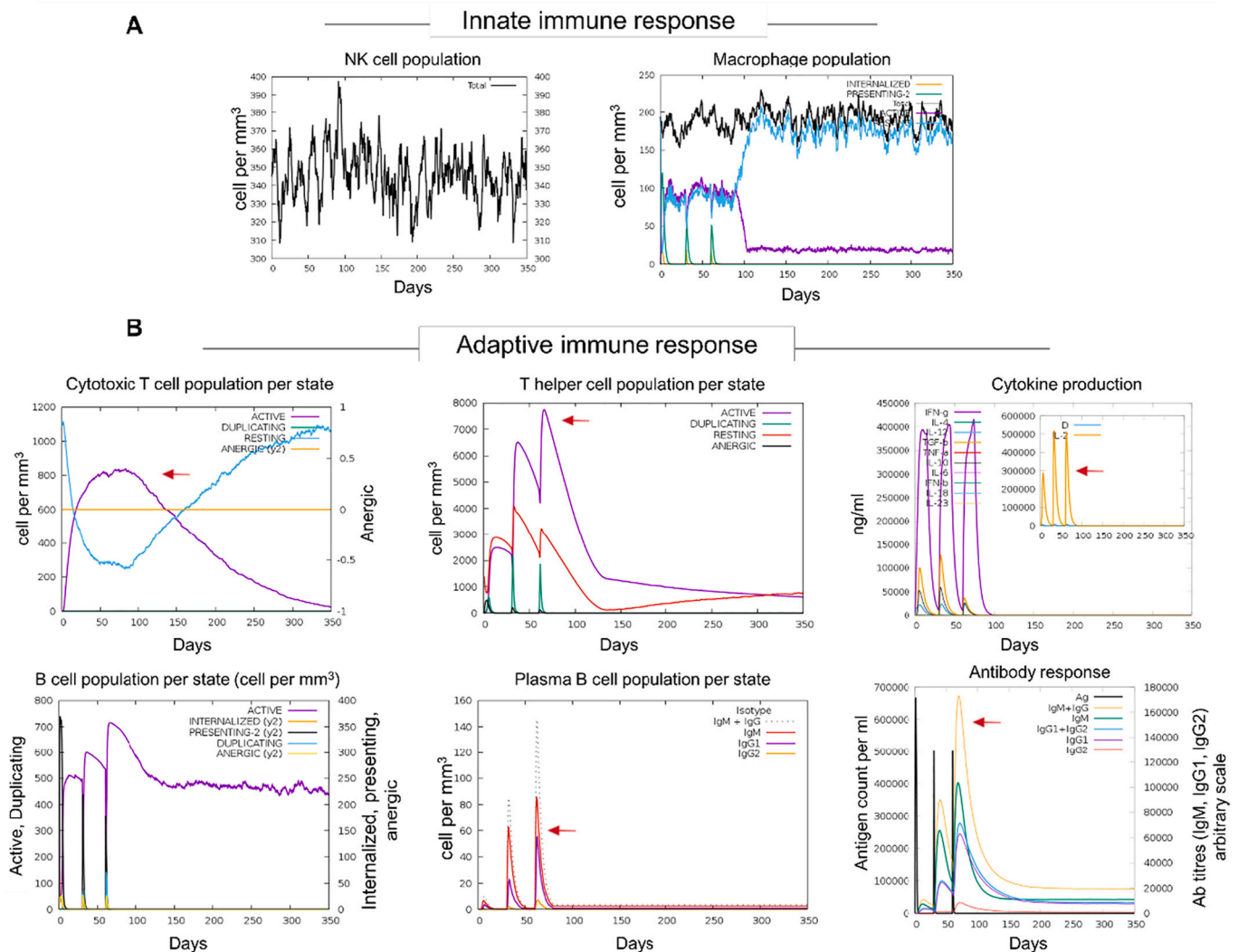
## 4. Discussion

Immunoinformatics represents a valuable tool whereby the limitations in the selection of appropriate antigens and immunodominant epitopes may be overcome (Sharma et al., 2019). Previous *in silico*-based reports have shown that the SARS-CoV-2S glycoprotein contains potential epitopes (Gupta et al., 2013). Therefore, researchers have recently attempted to design epitope-based vaccine candidates against SARS-CoV-2 (Gupta et al., 2013). In the present study, highly potential B and T cell epitopes from the SARS-CoV-2 glycoprotein were predicted and the best selected to design a high-quality multi-epitope vaccine candidate. Remarkably, this vaccine harbours 2 epitopes (E19 in Table 1 and E10 in Table 2) that could evoke immune responses against SARS-CoV-2 RBM—the main responsible for virus entry into human cells (Lan et al., 2020a) whereas 4 epitopes (E43, E44, E45, and E46 in Table 1) may direct the immune attack against other regions of SARS-CoV-2 RBD. These results are consistent with *in vitro* data that have demonstrated the antigenicity of the SARS-CoV-2 S glycoprotein (Walls et al., 2020).

The T cell epitopes included in the vaccine sequence accomplish with relevant requisites to design a suitable multi-epitope vaccine candidate. Firstly, they showed a marked antigenicity, immunogenicity, and lack of allergenic or toxic residues. Secondly, the THC epitopes were predicted as potent inducers of IFN- $\gamma$ —a crucial cytokine for CTC activation (Owen et al., 2013). Thirdly, both CTC and THC epitopes properly interacted with the groove of HLA-I and HLA-II alleles, respectively, which is in agreement with other computer-based reports (Marchan, 2020), thereby suggesting that the T cell epitopes identified and selected in the present study could be successfully presented in the context of HLA molecules. In addition, most of the peptides arched away from the HLA alleles and are, therefore, more exposed, which in turn suggests that they could interact more directly with the T-cell receptor, thereby possibly leading to a proper activation of T cells (Owen et al., 2013).

The purpose of an adjuvant is to make a vaccine “detectable” for antigen-presenting cells such as dendritic cells (Owen et al., 2013). Here, the TLR4 adjuvant known as RS09 (Shanmugam et al., 2012) was included in the multi-epitope vaccine sequence. The molecular docking simulation showed that the multi-epitope vaccine rightly interacts with this innate immune receptor in a similar way to previous works (Pandey et al., 2018).

Notably, this study shows, by immunoinformatics simulations, the induction of both innate and adaptive responses to SARS-CoV-2. In this regard, NK cell and macrophage activation were detected, as well as high production of typical antibodies (IgM and IgG), cytokines (IFN- $\gamma$  and IL-2), and a proliferative burst of CTC and THC were observed after three injections. The generation and increase of plasma cells were also documented. Furthermore, B and T cell populations decreased along with immunogen levels. These data is comparable to previous investigations that have been focused on vaccine development against



**Fig. 6.** Immune response simulations with the multi-epitope vaccine (3 doses given 4 weeks apart). (A) Innate immune response after *in silico* simulations. NK cells and Macrophages were properly activated. (B) The adaptive immune response was clearly elicited after vaccine shots on the C-ImmSim server, including the cytotoxic T cell, T helper, B cell, and antibody responses as well as the cytokine production (red arrows). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

*Mycobacterium ulcerans* (Nain et al., 2020) and filarial diseases (Shey et al., 2019), as well as are in agreement, at least partially, with a recent study that demonstrated a positive correlation between robust CD4+THC responses with anti-SARS-CoV-2 IgG and IgA titres of COVID-19 convalescent patients (Grifoni et al., 2020b). These immune responses were directed to the SARS-CoV-2 S glycoprotein (Grifoni et al., 2020b).

Recently, Kar et al. (2020) have reported a similar potential vaccine model, whose epitopes were also derived from the S glycoprotein. However, there are important differences to point out. For instance, it is not clear why they authors docked the entire vaccine model with HLA class I and class II alleles, which are known to interact only with peptides of a relatively small amino acid length (Owen et al., 2013) (9-mer for HLA class I molecules and 15-mer for HLA class II molecules) as has shown in the present and other similar works focused on SARS-CoV-2 and vaccine development (Marchan, 2020). In addition, the docking simulations between the vaccine model and the TLR-4 showed that the vaccine is not interacting in the region TLR-4 and MD-2 (Kar et al., 2020), which is pivotal for antigen recognition and initiation of the immune response (Park et al., 2009). Importantly, the population coverage of the vaccine molecule reported in the present study is higher compared to the vaccine model from Kar et al. (2020) (99.69% vs 95.78%). On the other hand, Tahir Ul Qamar et al. (2020) selected seven

SARS-CoV-2 proteins, including the S glycoprotein, to design a vaccine construct. Interestingly, they did not detect the S glycoprotein as antigenic on the Vaxijen website, which contrast with previous reports in the literature wherein the S glycoprotein has been demonstrated to induce immune responses (Grifoni et al., 2020a). This result may be due to the fact that complex proteins harbour both antigenic and non-antigenic regions and in order to identify and select the former is necessary to apply an integrated approach in which different algorithms are used to validate the predicted output (Grifoni et al., 2020a; Marchan, 2020).

This work was limited by A) the population coverage analysis did not include some countries, particularly from Africa, Central America, Eastern Europe, and Central Asia. This was mainly due to data not available concerning the HLA allele frequencies. Nevertheless, the highest population coverage was observed in several of the worst-hit countries by COVID-19 (e.g. Brazil, China, France, Italy, Iran, Peru, Spain, USA, etc.) (World Health Organization, n.d.). B) This study did not explore whether the epitopes used for vaccine design are conserved in other beta-coronaviruses. However, former reports have already demonstrated that SARS-CoV-2 shares 79.5% and 50% sequence identity to SARS-CoV and MERS-CoV, respectively (Lu et al., 2020).

In summary, this research provides a novel multi-epitope vaccine built from high potential epitopes derived from the SARS-CoV-2 S

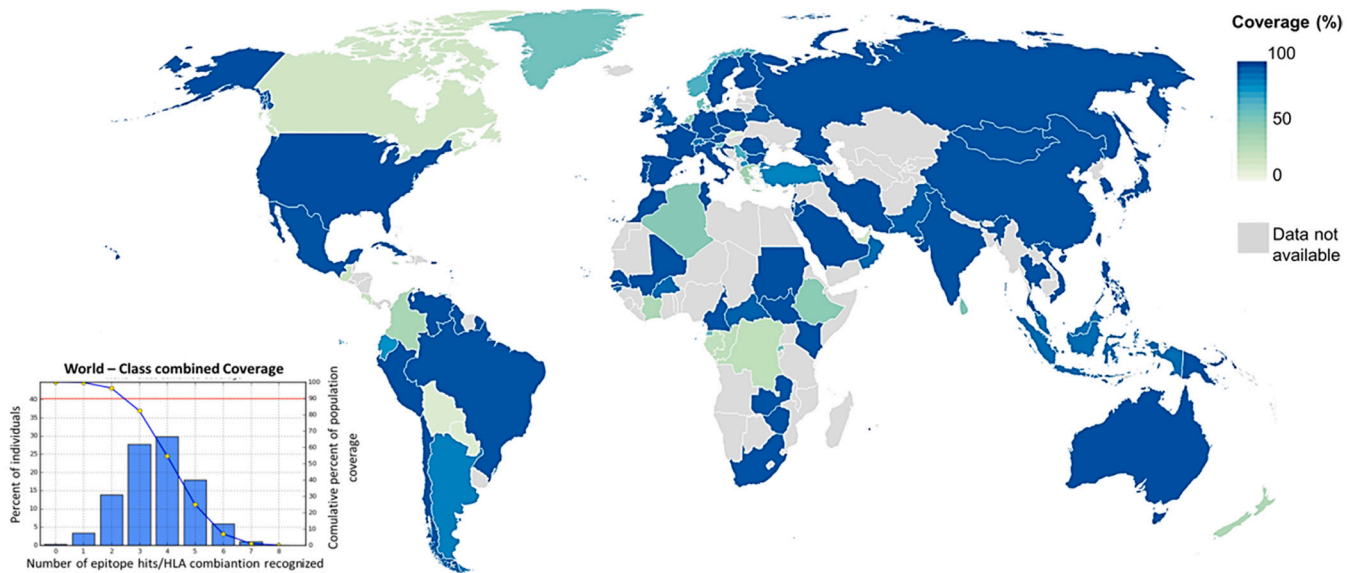


Fig. 7. Combined HLA population coverage analysis of the multi-epitope vaccine against SARS-CoV-2.

glycoprotein. This immunoinformatics study suggests that such multi-epitope vaccine could activate and generate robust humoral and cell-mediated responses in a simultaneous manner against SARS-CoV-2, as well as the population coverage analysis indicates that it could be used globally. However, further rigorous *in vitro* and *in vivo* studies are imperative to confirm its immunogenic properties, safety, and efficacy, which—of course—would imply months, even years.

### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Declaration of Competing Interest

I have no conflict of interest to declare.

### Acknowledgement

I wish to thank the support of the Venezuelan Institute for Scientific Research (IVIC).

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jim.2022.113216>.

### References

- Amanat, F., Krammer, F., 2020. SARS-CoV-2 vaccines: status report. *Immunity* 52, 583–589. <https://doi.org/10.1016/j.immuni.2020.03.007>.
- Bhattacharya, M., Sharma, A.R., Patra, P., Ghosh, P., Sharma, G., Patra, B.C., Lee, S.S., Chakraborty, C., 2020. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immunoinformatics approach. *J. Med. Virol.* <https://doi.org/10.1002/jmv.25736>. Advance online publication.
- Buchan, D., Jones, D.T., 2019. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* 47, W402–W407. <https://doi.org/10.1093/nar/gkz297>.
- Bui, H.H., Sidney, J., Dinh, K., Southwood, S., Newman, M.J., Sette, A., 2006. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinform.* 7, 153. <https://doi.org/10.1186/1471-2105-7-153>.
- Dhanda, S.K., Vir, P., Raghava, G.P., 2013. Designing of interferon-gamma inducing MHC class-II binders. *Biol. Direct* 8, 30. <https://doi.org/10.1186/1745-6150-8-30>.
- Dimitrov, I., Naneva, L., Doytchinova, I., Bangov, I., 2014. AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics* 30, 846–851. <https://doi.org/10.1093/bioinformatics/btt619>.

- Doytchinova, I.A., Flower, D.R., 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* 8, 4. <https://doi.org/10.1186/1471-2105-8-4>.
- Grifoni, A., Sidney, J., Zhang, Y., Scheuermann, R.H., Peters, B., Sette, A., 2020a. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. *Cell Host Microbe* 27, 671–680. <https://doi.org/10.1016/j.chom.2020.03.002>.
- Grifoni, A., Weiskopf, D., Ramirez, S.I., Mateus, J., Dan, J.M., Moderbacher, C.R., Rawlings, S.A., Sutherland, A., Premkumar, L., Jadi, R.S., Marrama, D., de Silva, A.M., Frazier, A., Carlin, A., Greenbaum, J.A., Peters, B., Krammer, F., Smith, D.M., Crotty, S., Sette, A., 2020b. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell*. <https://doi.org/10.1016/j.cell.2020.05.015>.
- Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Open Source Drug Discovery Consortium, Raghava, G.P., 2013. In silico approach for predicting toxicity of peptides and proteins. *PLoS One* 8, e73957. <https://doi.org/10.1371/journal.pone.0073957>.
- Hoof, I., Peters, B., Sidney, J., Pedersen, L.E., Sette, A., Lund, O., Buus, S., Nielsen, M., 2009. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* 61, 1–13. <https://doi.org/10.1007/s00251-008-0341-z>.
- Jespersen, M.C., Peters, B., Nielsen, M., Marcattili, P., 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* 45, W24–W29. <https://doi.org/10.1093/nar/gkx346>.
- Kar, T., Narsaria, U., Basak, S., et al., 2020. A candidate multi-epitope vaccine against SARS-CoV-2. *Sci. Rep.* 10, 10895. <https://doi.org/10.1038/s41598-020-67749-1>.
- Kim, Y., Ponomarenko, J., Zhu, Z., Tamang, D., Wang, P., Greenbaum, J., Lundegaard, C., Sette, A., Lund, O., Bourne, P.E., Nielsen, M., Peters, B., 2012. Immune epitope database analysis resource. *Nucleic Acids Res.* 40, W525–W530. <https://doi.org/10.1093/nar/gks438>.
- Ko, J., Park, H., Heo, L., Seok, C., 2012. GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res.* 40, W294–W297. <https://doi.org/10.1093/nar/gks493>.
- Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., Wang, X., 2020a. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581 (7807), 215–220. <https://doi.org/10.1038/s41586-020-2180-5>.
- Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., Wang, X., 2020b. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 581, 215–220. <https://doi.org/10.1038/s41586-020-2180-5>.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283–291.
- Lester, S.N., Li, K., 2014. Toll-like receptors in antiviral innate immunity. *J. Mol. Biol.* 426, 1246–1264. <https://doi.org/10.1016/j.jmb.2013.11.024>.
- Loughran, S.T., Bree, R.T., Walls, D., 2017. Purification of polyhistidine-tagged proteins. *Methods Mol. Biol. (Clifton, N.J.)* 1485, 275–303. [https://doi.org/10.1007/978-1-4939-6412-3\\_14](https://doi.org/10.1007/978-1-4939-6412-3_14).
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., Chen, J., Meng, Y., Wang, J., Lin, Y., Yuan, J., Xie, Z., Ma, J., Liu, W.J., Wang, D., Xu, W., Holmes, E.C., Gao, G.F., Wu, G., Chen, W., Shi, W., Tan, W., 2020. Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus

- origins and receptor binding. *Lancet*. 395 (10224), 565–574. [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8).
- Magnan, C.N., Randall, A., Baldi, P., 2009. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics*. 25, 2200–2207. <https://doi.org/10.1093/bioinformatics/btp386>.
- Manavalan, B., Govindaraj, R.G., Shin, T.H., Kim, M.O., Lee, G., 2018. iBCE-EL: a new ensemble learning framework for improved linear B-cell epitope prediction. *Front. Immunol.* 9, 1695. <https://doi.org/10.3389/fimmu.2018.01695>.
- Marchan, J., 2020. Conserved HLA binding peptides from five non-structural proteins of SARS-CoV-2—an in silico glance. *Hum. Immunol.* 81 (10–11), 588–595. <https://doi.org/10.1016/j.humimm.2020.08.001>.
- Moutaftsi, M., Peters, B., Pasquetto, V., Tschärke, D.C., Sidney, J., Bui, H.H., Grey, H., Sette, A., 2006. A consensus epitope prediction approach identifies the breadth of murine T(CD8+) cell responses to vaccinia virus. *Nat. Biotechnol.* 24, 817–819. <https://doi.org/10.1038/nbt1215>.
- Nain, Z., Karim, M.M., Sen, M.K., Adhikari, U.K., 2020. Structural basis and designing of peptide vaccine using PE-PGRS family protein of *Mycobacterium ulcerans*—an integrated vaccinomics approach. *Mol. Immunol.* 120, 146–163. <https://doi.org/10.1016/j.molimm.2020.02.009>.
- Nielsen, M., Lundegaard, C., Lund, O., 2007. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinform.* 8, 238. <https://doi.org/10.1186/1471-2105-8-238>.
- Nielsen, M., Lundegaard, C., Blicher, T., Peters, B., Sette, A., Justesen, S., Buus, S., Lund, O., 2008. Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. *PLoS Comput. Biol.* 4, e1000107. <https://doi.org/10.1371/journal.pcbi.1000107>.
- Ortega, J.T., Serrano, M.L., Pujol, F.H., Rangel, H.R., 2020. Role of changes in SARS-CoV-2 spike protein in the interaction with the human ACE2 receptor: an in silico analysis. *EXCLI J.* 19, 410–417. <https://doi.org/10.17179/excli2020-1167>.
- Owen, J., Punt, J., Stranford, S., Pat, J., 2013. *Kuby Immunology, Seventh ed.* W.H. Freeman and Company, New York.
- Pandey, R.K., Bhatt, T.K., Prajapati, V.K., 2018. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. *Sci. Rep.* 8, 1125. <https://doi.org/10.1038/s41598-018-19456-1>.
- Park, B.S., Song, D.H., Kim, H.M., Choi, B.S., Lee, H., Lee, J.O., 2009. The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* 458 (7242), 1191–1195. <https://doi.org/10.1038/nature07830>.
- Paul, S., Lindestam Arlehamn, C.S., Scriba, T.J., Dillon, M.B., Oseroff, C., Hinz, D., McKinney, D.M., Carrasco Pro, S., Sidney, J., Peters, B., Sette, A., 2015. Development and validation of a broad chemist for prediction of HLA class II restricted T cell epitopes. *J. Immunol. Methods* 422, 28–34. <https://doi.org/10.1016/j.jim.2015.03.022>.
- Ponomarenko, J., Bui, H.H., Li, W., Füsseder, N., Bourne, P.E., Sette, A., Peters, B., 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform.* 9, 514. <https://doi.org/10.1186/1471-2105-9-514>.
- Rapin, N., Lund, O., Bernaschi, M., Castiglione, F., 2010. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One* 5 (4), e9862. <https://doi.org/10.1371/journal.pone.0009862>.
- Saha, S., Raghava, G.P., 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins*. 65, 40–48. <https://doi.org/10.1002/prot.21078>.
- Shanmugam, A., Rajoria, S., George, A.L., Mittelman, A., Suriano, R., Tiwari, R.K., 2012. Synthetic Toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants. *PLoS One* 7, e30839. <https://doi.org/10.1371/journal.pone.0030839>.
- Sharma, M., Krammer, F., García-Sastre, A., Tripathi, S., 2019. Moving from empirical to rational vaccine design in the ‘omics’ era. *Vaccines*. 7, 89. <https://doi.org/10.3390/vaccines7030089>.
- Shey, R.A., Ghogomu, S.M., Esoh, K.K., Nebangwa, N.D., Shintouo, C.M., Nongley, N.F., Asa, B.F., Ngale, F.N., Vanhamme, L., Souopgui, J., 2019. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci. Rep.* 9, 4409. <https://doi.org/10.1038/s41598-019-40833-x>.
- Tahir Ul Qamar, M., Rehman, A., Tusleem, K., Ashfaq, U.A., Qasim, M., Zhu, X., Fatima, I., Shahid, F., Chen, L.L., 2020. Designing of a next generation multi-epitope based vaccine (MEV) against SARS-CoV-2: Immunoinformatics and in silico approaches. *PLoS One* 15 (12), e0244176. <https://doi.org/10.1371/journal.pone.0244176>.
- Tenzen, S., Peters, B., Bulik, S., Schöor, O., Lemmel, C., Schatz, M.M., Kloetzel, P.M., Rammensee, H.G., Schild, H., Holzhütter, H.G., 2005. Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. *Cell. Mol. Life Sci.* 62, 1025–1037. <https://doi.org/10.1007/s00018-005-4528-2>.
- Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., Veesler, D., 2020. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*. 181, 281–292. <https://doi.org/10.1016/j.cell.2020.02.058>.
- Wang, P., Sidney, J., Dow, C., Mothé, B., Sette, A., Peters, B., 2008. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput. Biol.* 4, e1000048. <https://doi.org/10.1371/journal.pcbi.1000048>.
- Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35, W407–W410. <https://doi.org/10.1093/nar/gkm290>.
- Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* 112, 531–552. <https://doi.org/10.1385/1-59259-584-7-531>.
- World Health Organization. Coronavirus disease (COVID-19) Pandemic. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (Accessed 19 May 2020).
- Zhou, W.Y., Shi, Y., Wu, C., Zhang, W.J., Mao, X.H., Guo, G., Li, H.X., Zou, Q.M., 2009. Therapeutic efficacy of a multi-epitope vaccine against *Helicobacter pylori* infection in BALB/c mice model. *Vaccine*. 27, 5013–5019. <https://doi.org/10.1016/j.vaccine.2009.05.009>.