*American Journal of Microbiological Research, 2018, Vol. 6, No. 3, 94-114*
Available online at http://pubs.sciepub.com/ajmr/6/3/5

DOI:10.12691/ajmr-6-3-5

# Multi Epitopes Vaccine Prediction against Severe Acute Respiratory Syndrome (SARS) Coronavirus Using Immunoinformatics Approaches

**Yassir A. Almofti[1,*], Khoubieb Ali Abd-elrahman[2], Sahar Abd Elgadir Gassmallah[3], Mohammed Ahmed Salih[4]**

[1]Department of Biochemistry and Molecular Biology, College of Veterinary Medicine, University of Bahri, Khartoum, Sudan
[2]Department of pharmaceutical technology, College of Pharmacy,
University of Medical Science and Technology (MUST) Khartoum, Sudan
[3]Department of Medical laboratory, Sudan University of Science and Technology, Khartoum, Sudan
[4]Department of Bioinformatics, Africa city of Technology, Khartoum, Sudan
*Corresponding author: yamofti99@gmail.com

**Abstract** Efforts for developing vaccine against Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) is crucial in prevention of SARS re-emergence. The global outbreak of SARS was contained since 2003. However concerns remain over the possibility of future recurrences, especially with recent reports of laboratory-acquired infections and the presence of sporadic cases, raising a serious concern. SARS-CoV spike S protein (1255aa) is an important target in developing safe and effective vaccines. In this study multiple bio-informatics and immuno-informatics implementation tools from NCBI and IEDB were used for epitopes prediction from spike S protein. The predicted epitopes were further assessed for population coverage against the whole world population. Our results demonstrated that the epitopes $_{38}$-**RGVYYPDEI**-$_{46}$, $_{200}$-**YQPIDVVRD**-$_{208}$ and $_{388}$-**VVKGDDVRQ**-$_{396}$ elicit and stimulate B cell since they got higher score in Emini and Kolaskar and tongaonker software. For T-cell: the epitopes $_{47}$-**FRSDTLYLT**-$_{55, 195}$-**YVYKGYQPI**-$_{203}$ and $_{880}$-**FAMQMAYRF**-$_{888}$ were found to interact with both MHC-1 and MHC-II alleles. Moreover $_{851}$-**MIAAYTAAL**-$_{859}$ showed higher affinity to MHC-1 alleles while $_{782}$-**FNFSQILPD**-$_{790}$ interacted only with MHC-II alleles. The population coverage epitope set for MHC-1 and MHC-II predicted epitopes was **82.16%** and **99.97%** respectively. All predicted epitopes against T cell (MHC-I/MHC-II) demonstrated strong potentiality as promising peptides vaccine with population coverage epitope set against the whole world of **100%**. Taken together eight epitopes were proposed to interact with B and T cells and act as peptide vaccine against SARS-CoV virus. In vitro and in vivo studies are recommended to prove the effectiveness of these epitopes as a peptide vaccine.

***Keywords:*** *SARS, NCBI, IEDB, Insilico prediction, Immunoinformatics*

**Cite This Article:** Yassir A. Almofti, Khoubieb Ali Abd-elrahman, Sahar Abd Elgadir Gassmallah, and Mohammed Ahmed Salih, "Multi Epitopes Vaccine Prediction against Severe Acute Respiratory Syndrome (SARS) Coronavirus Using Immunoinformatics Approaches." *American Journal of Microbiological Research*, vol. 6, no. 3 (2018): 94-114. doi: 10.12691/ajmr-6-3-5.

## 1. Introduction

Severe acute respiratory syndrome (SARS) is a viral pulmonary infection caused by SARS coronavirus (SARS-CoV) [1]. The disease is considered as an emerging infectious disease that spread to more than 30 countries in 2003. More than 8000 cases, including almost 800 deaths, were reported during the outbreak period with a mortality rate of ~10% [1,2,3]. The global outbreak of SARS was contained since the end of SARS outbreak in 2003. But concerns remain over the possibility of future recurrences, especially with recent reports of laboratory-acquired infections and the presence of a number of SARS sporadic cases, raising a serious concern [4,5,6,7]. Additionally, the presence of natural reservoir, such as bats, considered as risk factors for reoccurrence of SARS-CoV like virus from animals to humans [8]. The symptoms of SARS infection is characterized by acute onset of fever accompanied with cough, difficulty and shortness of breathing. These symptoms initiated as nonspecific symptoms of a flu-like illness one week after exposure to the virus [9,10]. A variety of therapeutic efforts were introduced for the remedy of the disease in Asia and Canada. However, no treatment demonstrated efficacy against the infection [11,12].

Like the other coronaviruses, SARS CoV is an enveloped virus containing a large, positive-stranded RNA genome that encodes viral replicase proteins and structural proteins including spike S, membrane, envelope, nucleocapsid, and several uncharacterized proteins [13,14,15]. The initiation of the infection started with the attachment of the spike S protein to specific host receptor accompanied by stimulation of conformational change in the S protein [13,16]. The S protein of SARS-CoV is a type I transmembrane glycoprotein with a predicted length of 1255 amino acids

[13,16]. The spike S glycoprotein consists of a leader (amino acids 1 to 14), an ectodomain (amino acids 15 to 1190), a membrane spanning domain (amino acids 1191 to 1227), and a short intracellular tail (amino acids 1227 to 1255) [17]. Moreover the S proteins of coronaviruses are also considered as major antigenic determinants that induce neutralizing antibodies [16,18,19]. It is noteworthy that a particular amino acids sequence region within the spike S proteins, termed receptor binding domain (RBD), is considered as a functional domain responsible for virus binding to the target cell receptor [17,20,21], may contain neutralizing epitopes [22,23] and acted as a potent inducer of neutralizing antibodies in the immunized rabbits [24]. Thus most existing vaccine candidates against SARS CoV were based on the spike S protein and RBD region [16,17,19,25].

Vaccinations for preventing SARS infection have been reported in multiple studies. The inactivated virus was used as the first-generation vaccine because it is easy to generate whole killed virus particles. SARS-CoV inactivated with formaldehyde, UV light, and β-propiolactone could induce virus-neutralizing antibodies in immunized animals [26,27,28,29]. In addition to that evaluations of an inactivated whole virus vaccine in ferrets and nonhuman primates and a virus-like-particle vaccine in mice induced protection against infection but the challenged animals exhibited an immunopathologic lung lesion [30]. Other vaccines including the spike S protein preparations, virus-like particles (VLPs), plasmid DNA and a number of vectors containing genes of SARS-CoV proteins were used [31,32,33]. Moreover strategies for using combination of different vaccines improved the humoral or cellular immune responses. For example prime with SARS-CoV S protein expressing DNA and boost with adenoviral vector encoding S protein induced optimal CD8+ T cell immunity. While boosting with inactivated SARS-CoV plus adjuvant stimulated the CD4+ T cell immunity and the antibody response [7,34]. Also a recombinant adeno-associated virus encoding RBD of SARS-CoV S protein (RBD-rAAV) induced SARS-CoV-specific IgG antibody with neutralizing activity [35]. Although these vaccines induced potent neutralizing and protective responses in the immunized animals, but safety of the inactivated vaccine is a serious concern. For instance workers are exposed to risk while dealing with concentrated live SARS-CoV, incomplete virus inactivation may cause SARS outbreaks among the vaccinated populations, and some viral proteins may induce harmful immune or inflammatory responses, even causing SARS-like diseases [12,36]. Thus the need for a safer and efficacious vaccine without future complications is highly recommended. In this study we aimed to use the immunoinformatic approaches found in the Immune Epitope Database (IEDB) to predict epitopes from spike S protein of SARS-CoV that elicit the human immune system and acted as safer efficacious vaccine.

# 2. Materials and Methods

## 2.1. Protein Sequences Retrieval and Alignment Tool

A total of 131 sequences of the spike S protein of SARS-CoV were retrieved from the NCBI database [37] on May 1st 2017. The majority of the samples were retrieved from China (58 strains), USA (48 strains), Taiwan (14 strains), Italy (3 strains), Canada (2 strains), Germany (2 strains), Japan (2 strains) and unnamed protein product (2 strains). The retrieved strains, their accession numbers, country and date of collection were shown in table (1). The spike S protein sequences of the retrieved strains were aligned to obtain the conserved regions using multiple sequence alignment (MSA). The sequences were aligned with the aid of Clustal W in the BioEdit program, version 7.0.9.0. [38]

## 2.2. Evolution Analysis

The retrieved sequences were subjected to evolutionary divergence analysis using the spike S protein of the SARS-CoA virus. The phylogenetic tree was constructed to determine the common ancestor of each strain using MEGA 7.0.26 (7170509-x86_64).

## 2.3. Determination of the Conserved Regions of Spike S protein

The conserved regions of the candidate epitopes were analyzed by different prediction software tools obtained by Immune Epitope Database (IEDB) analysis [39]. The glycoprotein precursor E2 [NP_828851.1] of the spike S protein was used as reference sequence for the IEDB software analysis.

## 2.4. B Lymphocytes Epitopes Prediction

Tools from IEDB were used to identify the B cell antigenicity, including Bepipred for linear epitope analysis [40], Emini for surface accessibility [41] and Kolaskar and Tongaonkar for antigenicity scale [42].

**Table 1. 131retrieved strains with accession numbers, countries and year of collection**

| Accession number | Country | Year |
|---|---|---|
| *NP_828851.1 | Canada Toronto | 2003 |
| AAP30030.1 | China | 2003 |
| AAR91586.1 | China | 2003 |
| AAP33697.1 | Germany-Frankfurt | 2003 |
| AAP50485.1 | Italy | 2003 |
| BAC81404.1 | China | 2003 |
| BAC81390.1 | China | 2003 |
| BAC81376.1 | China | 2003 |
| BAC81362.1 | China | 2003 |
| BAC81348.1 | China | 2003 |

| Accession number | Country | Year |
|---|---|---|
| AAP72986.1 | Italy | 2003 |
| AAQ94060.1 | Italy | 2003 |
| P59594.1 | USA Full-Spike glycoprotein | 2003 |
| AAP13441.1 | USA-Atlanta | 2003 |
| AAR33050.1 | China_Hangzhou | 2003 |
| AAP37017.1 | Taiwan | 2003 |
| AAQ01609.1 | Taiwan | 2003 |
| AAQ01597.1 | Taiwan | 2003 |
| AAP97882.1 | Taiwan | 2003 |
| AAR87600.1 | Taiwan | 2003 |
| AAR87589.1 | Taiwan | 2003 |
| AAR87578.1 | Taiwan | 2003 |
| AAR87567.1 | Taiwan | 2003 |
| AAR87556.1 | Taiwan | 2003 |
| AAR87545.1 | Taiwan | 2003 |
| AAR87534.1 | Taiwan | 2003 |
| AAR87523.1 | Taiwan | 2003 |
| AAR87512.1 | Taiwan | 2003 |
| AAR87501.1 | Taiwan | 2003 |
| AAS00003.1 | China-Guangzhou | 2003 |
| AAP51227.1 | China | 2003 |
| AAT52330.1 | China | 2004 |
| AAX16192.1 | China-Wuhan | 2004 |
| AAY60793.1 | China | 2004 |
| AAY60780.1 | China | 2004 |
| AAT74874.1 | China-Beijing | 2004 |
| AAT76147.1 | China | 2004 |
| AAS75868.1 | Germany-Frankfurt | 2004 |
| AAU81608.1 | USA-Atlanta | 2004 |
| CAI58832.1 | unnamed_protein_product_synthetic_construct | 2005 |
| ABD72982.1 | China_Hong_Kong | 2006 |
| CAL49859.1 | unnamed_protein_product_synthetic_construct | 2006 |
| BAE93401.1 | Japan-Frankfurt | 2006 |
| BAF42873.1 | Japan-Frankfurt1 | 2006 |
| ABD72972.1 | China_Hong_Kong | 2006 |
| ABD72984.1 | China_Hong_Kong | 2006 |
| ABD72969.1 | China_Hong_Kong | 2006 |
| ABF65836.1 | USA-Rockville | 2006 |
| ABD72968.1 | China-Hong_Kong | 2006 |
| ABD72998.1 | China_Hong_Kong | 2006 |
| ABD72996.1 | China_Hong_Kong | 2006 |
| ABD72994.1 | China_Hong_Kong | 2006 |
| ABD72991.1 | China_Hong_Kong | 2006 |
| ABD72989.1 | China_Hong_Kong | 2006 |
| ABD72987.1 | China_Hong_Kong | 2006 |
| ABD72983.1 | China_Hong_Kong | 2006 |
| ABD72981.1 | China_Hong_Kong | 2006 |
| ABD72980.1 | China_Hong_Kong | 2006 |
| ABD72978.1 | China_Hong_Kong | 2006 |
| ABD72975.1 | China_Hong_Kong | 2006 |
| ABI96958.1 | Canada-CV7 | 2006 |
| ABD73002.1 | China-Hongkong | 2006 |
| ABD73000.1 | China_Hong_Kong | 2006 |
| ABD72997.1 | China_Hong_Kong | 2006 |
| ABD72976.1 | China_Hong_Kong | 2006 |
| ABD72974.1 | China_Hong_Kong | 2006 |
| ABD72973.1 | China_Hong_Kong | 2006 |
| ABD72971.1 | China_Hong_Kong | 2006 |
| ABD72970.1 | China_Hong_Kong | 2006 |
| ABD72988.1 | China_Hong_Kong | 2006 |
| ABD72995.1 | China_Hong_Kong | 2006 |

| Accession number | Country | Year |
|---|---|---|
| ABD72979.1 | China_Hong_Kong | 2006 |
| ABD72992.1 | China_Hong_Kong | 2006 |
| ABD72977.1 | China_Hong_Kong | 2006 |
| ABD72993.1 | China_Hong_Kong | 2006 |
| ABD72990.1 | China_Hong_Kong | 2006 |
| ABD72986.1 | China_Hong_Kong | 2006 |
| ABD72985.1 | China_Hong_Kong | 2006 |
| ABD73001.1 | China_Hong_Kong | 2006 |
| ABY85258.1 | USA | 2007 |
| ABD72999.1 | China_Hong_Kong | 2007 |
| ACB69905.1 | China | 2008 |
| ACB69872.1 | China | 2008 |
| ACB69860.1 | China | 2008 |
| ACB69883.1 | China | 2008 |
| ACB69894.1 | China | 2008 |
| ACB69849.1 | China | 2008 |
| AFM43867.1 | USA_Tennessee | 2012 |
| AFR58728.1 | USA | 2012 |
| AFR58700.1 | USA | 2012 |
| AFR58672.1 | USA | 2012 |
| AFR58714.1 | USA | 2012 |
| AFR58742.1 | USA | 2012 |
| AFR58686.1 | USA | 2012 |
| AGT21078.1 | USA_Nashville | 2013 |
| AGT21318.1 | USA | 2013 |
| AGT21303.1 | USA | 2013 |
| AGT21288.1 | USA | 2013 |
| AGT21273.1 | USA | 2013 |
| AGT21258.1 | USA | 2013 |
| AGT21243.1 | USA | 2013 |
| AGT21228.1 | USA | 2013 |
| AGT21213.1 | USA_Nashville | 2013 |
| AGT21198.1 | USA_Nashville | 2013 |
| AGT21183.1 | USA_Nashville | 2013 |
| AGT21168.1 | USA_Nashville | 2013 |
| AGT21153.1 | USA_Nashville | 2013 |
| AGT21138.1 | USA_Nashville | 2013 |
| AGT21123.1 | USA_Nashville | 2013 |
| AGT21108.1 | USA_Nashville | 2013 |
| AGT21093.1 | USA_Nashville | 2013 |
| AGT21063.1 | USA_Nashville | 2013 |
| AGT21048.1 | USA_Nashville | 2013 |
| AGT21033.1 | USA_Nashville | 2013 |
| AGT21018.1 | USA_Nashville | 2013 |
| AGT21003.1 | USA_Nashville | 2013 |
| AGT20988.1 | USA_Nashville | 2013 |
| AGT20973.1 | USA_Nashville | 2013 |
| AGT20958.1 | USA_Nashville | 2013 |
| AGT20943.1 | USA_Nashville | 2013 |
| AGT20928.1 | USA_Nashville | 2013 |
| AGT20913.1 | USA_Nashville | 2013 |
| AGT20898.1 | USA_Nashville | 2013 |
| AGT20883.1 | USA_Nashville | 2013 |
| AGT20868.1 | USA_Nashville | 2013 |
| AGT20853.1 | USA_Nashville | 2013 |
| AGT20838.1 | USA_Nashville | 2013 |
| AGT20823.1 | USA_Nashville | 2013 |
| AGT20808.1 | USA_Nashville | 2013 |
| AGT20793.1 | USA_Nashville | 2013 |
| ALK02457.1 | China | 2015 |

*The reference sequence.

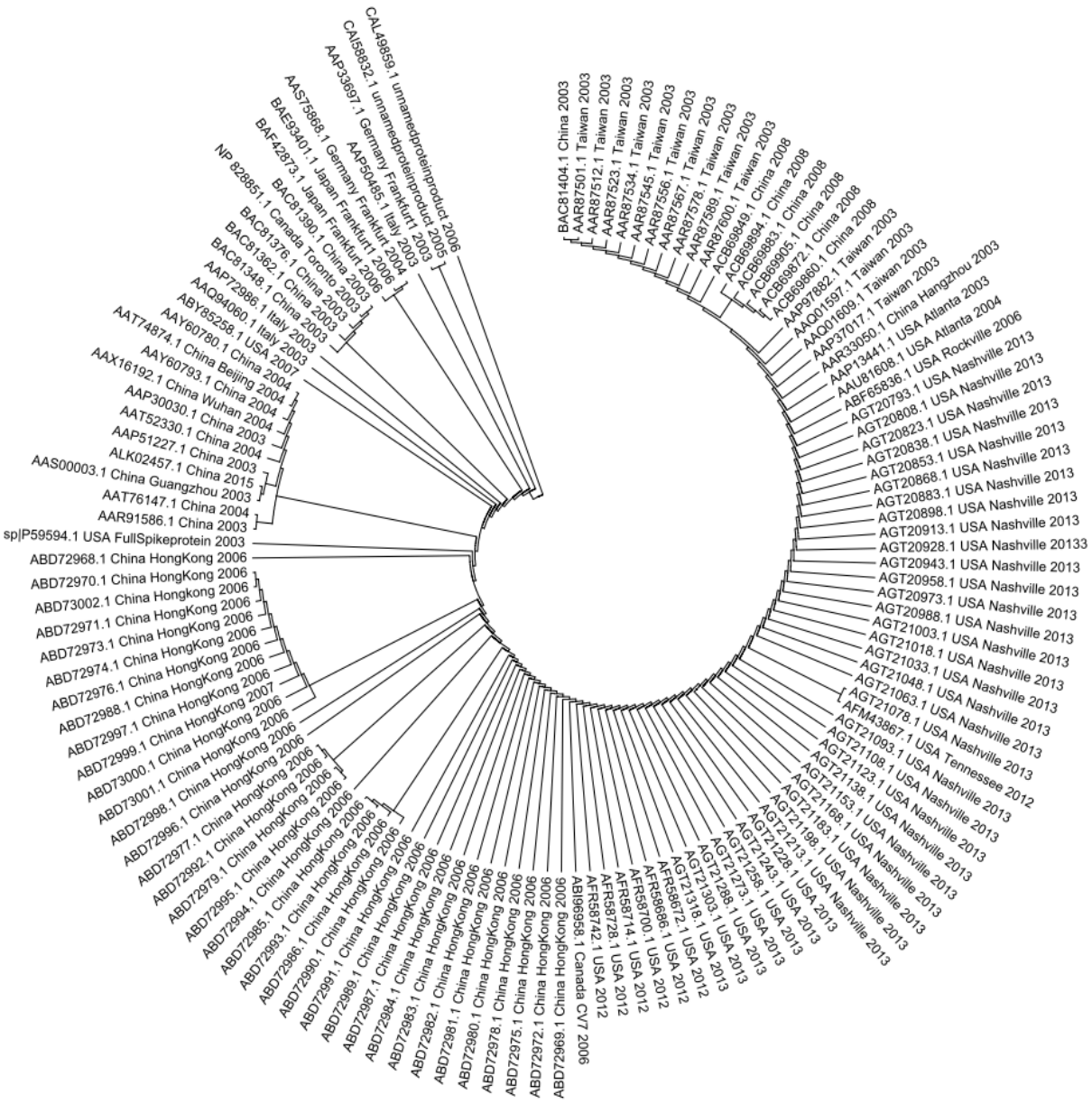


**Figure 1.** Phylogenetic tree of the 131 retrieved strains. The retrieved strains demonstrated divergence in their common ancestors

#### 2.4.1. B cell Linear Epitopes Prediction

BepiPred from IEDB [43] was used as linear B-cell epitopes prediction with a default threshold value of -0.075.

#### 2.4.2. B cell Surface Accessibility Prediction

Emini surface accessibility prediction tool from IEDB was used [43]. The surface accessible epitopes were predicted from the conserved regions with the default threshold value 1.000.

#### 2.4.3. B-cell Epitopes Antigenicity Prediction

Kolaskar and tongaonker antigenicity method [43] was used to determine the antigenic sites with a default threshold value of 1.039.

### 2.5. T-lymphocytes Epitopes Prediction

The IEDB was used for the identification of the T cell epitopes prediction. The prediction method included the major histocompatibility complex class I and Π (MHC-I, MHC-Π).

#### 2.5.1. MHC-I Binding Predictions

Analysis of epitopes binding to MHC-I molecules was assessed by the software of IEDB MHC-I prediction tool [44]. The prediction method was obtained by Artificial Neural Network (ANN), Stabilized Matrix Method (SMM) or Scoring Matrices derived from combinatorial peptide libraries [42-49]. Before the prediction step, epitopes length was set as 9mers. The conserved epitopes that bind to alleles at score equal to or less than 100 half-maximal inhibitory concentrations (IC50) were selected for further analysis [49,50].

#### 2.5.2. MHC-II Binding Predictions

Analysis of epitopes binding to MHC-II molecules was achieved by the IEDB MHC-II prediction tool [51]. The neural networks align (NN-align) that allow for simultaneous identification of the MHC-II binding core

epitopes and binding affinity was used. All conserved epitopes that bind to many alleles at score equal to or less than 1000 half-maximal inhibitory concentration (IC50) were selected for further analysis [52].

## 2.6. Population Coverage

For the calculation of the population coverage for all potential MHC-I and II epitopes bindings, the IEDB tools [53] was used. The S glycoprotein of SARS-CoA virus was assessed for population coverage against the whole world with selected MHC-I and MHC-II interacted alleles.

## 2.7. Homology Modeling

Raptor X protein structure prediction server was used for creation the 3D structure of the spike S protein of SARS-CoA virus [54]. The reference sequence [NP_828851.1] was used as an input and Chimera 1.8 was used as a tool to visualize the selected epitopes belonging to B cell and T cell (MHC-I and MHC-II) [55]. Homology modeling was used for visualization of the surface accessibility of the B lymphocytes predicted candidate epitopes as well as for visualization of all predicted T cell epitopes in the structural level [56].

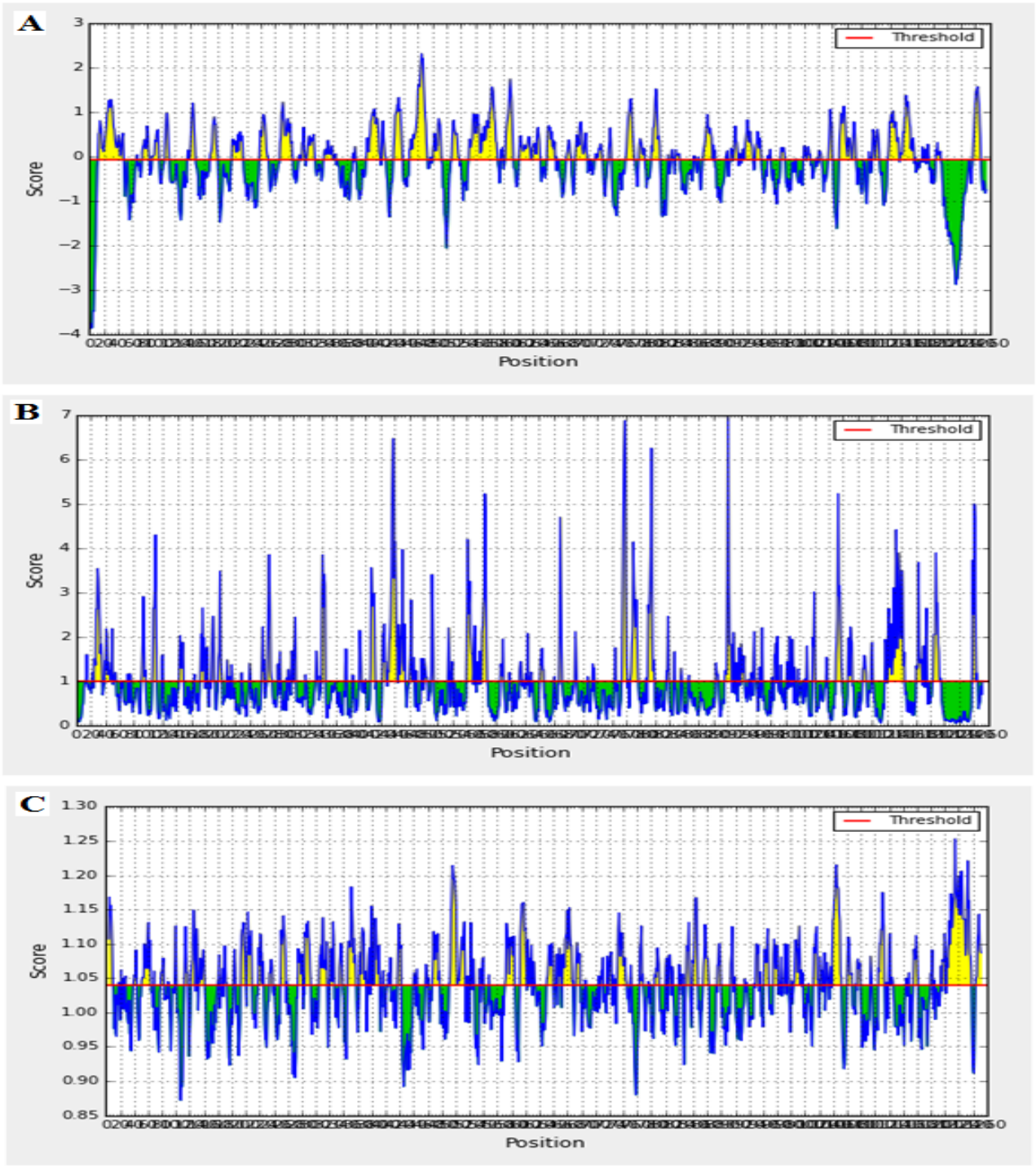


**Figure 2.** Prediction of B-cell epitopes by different IEDB scales (**A-** Bepipred linear epitope prediction, **B-** Emini surface accessibility, **C-** Kolaskar and Tongaonkar antigenicity prediction). Regions above threshold (red line) are proposed to be a part of B cell epitope while regions below the threshold (red line) are not

# 3. Results

## 3.1. Phylogenetic Analysis

Figure 1 provided the phylogenetic relationship of the 131 retrieved strains of the spike S protein of SARS-CoA viruses. The phylogeny demonstrated evolutionary divergence among the retrieved strains of SARS spike S protein.

## 3.2. B Cell Epitopes Prediction

The spike S glycoprotein was subjected to Bepipred linear epitope prediction, Emini surface accessibility and Kolaskar and Tongaonkar antigenicity prediction methods from IEDB. The thresholds of Bepipred linear epitope, Emini surface accessibility and Kolaskar and Tongaonkar antigenicity were shown in Figure 2. In Bepipred linear epitope prediction method; the average score of spike S protein to B lymphocytes was -0.075 (minimum: -3.871 and maximum: 2.320). Values equal to or greater than the default threshold -0.075 were predicted as conserved linear epitope. Table 2 showed that 66 epitopes were predicted by Bepipred method as a linear epitopes. In Emini surface accessibility prediction; the average score of SARS spike S protein was 1.000 (minimum: 0.036 and maximum: 7.000). Values equal to or greater than the default threshold 1.000 were regarded potentially on the surface. Emini surface accessibility method predicted 35 epitopes on the surface that have potential binding to B lymphocytes cells (Table 2). In Kolaskar and Tongaonkar antigenicity prediction method; the average score of SARS spike S protein was 1.039 (minimum: -3.921 and maximum: 1.183). Values equal to or greater than the default threshold 1.039 were considered as antigenic epitopes. This method predicted 29 antigenic epitopes with potential binding to B lymphocytes cells (Table 2). Accordingly three conserved epitopes were successfully predicted to elicit the B cell lymphocytes since they were conserved among all retrieved strains, got higher score values in Emini surface accessibility and Kolaskar and Tongaonkar antigenicity prediction methods. These three epitopes were $_{38}$-**RGVYYPDEI**-$_{46}$ and $_{200}$-**YQPIDVVRD**-$_{208}$ and $_{388}$-**VVKGDDVRQ**-$_{396}$. The three dimension structural (3D) level of these epitopes was shown in Figure 3.

**Table 2. B-cell epitopes prediction, the position of peptides is according to the position of amino acids in the spike S protein of SARS CoA virus.**

| Peptide | Start | End | Length | Emini[a] | Kolaskar [b] |
|---|---|---|---|---|---|
| GSDL | 13 | 16 | 4 | 0.753 | 1.001 |
| CTTFDDVQAPNY | 19 | 30 | 12 | 0.992 | 1.043 |
| RGVYYPDEIFR | 38 | 48 | 11 | 1.692 | 1.032 |
| *RGVYYPDEI | 38 | 46 | 9 | 1.535 | 1.043 |
| YFAATEKSNV | 88 | 97 | 10 | 1.234 | 1.024 |
| TMNNKS | 106 | 111 | 6 | 2.647 | 0.871 |
| GTQT | 145 | 148 | 4 | 1.471 | 0.927 |
| SLDV | 169 | 172 | 4 | 0.565 | 1.128 |
| EKSGNF | 174 | 179 | 6 | 1.71 | 0.922 |
| YQPIDVVRDLPSGFN | 200 | 214 | 15 | 0.767 | 1.055 |
| *YQPIDVVRD | 200 | 208 | 9 | 1.227 | 1.085 |
| PAQD | 240 | 243 | 4 | 1.862 | 1.002 |
| WGTSAA | 245 | 250 | 6 | 0.549 | 0.969 |
| DENGTITDAVDCSQNPLA | 267 | 284 | 18 | 0.648 | 1.006 |
| GIYQT | 298 | 302 | 5 | 0.897 | 1.022 |
| FNAT | 329 | 332 | 4 | 0.837 | 0.96 |
| FPSVYAW | 334 | 340 | 7 | 0.478 | 1.095 |
| VVKGDDVRQIAPGQTGVIAD | 388 | 407 | 20 | 0.199 | 1.05 |
| *VVKGDDVRQ | 388 | 396 | 9 | 1.062 | 1.064 |
| NYKLPD | 409 | 414 | 6 | 2.87 | 1.008 |
| RNIDAT | 426 | 431 | 6 | 1.438 | 0.94 |
| LRPFERDISNVPFSP | 448 | 462 | 15 | 1.403 | 1.028 |
| GKPCTPPA | 464 | 471 | 8 | 0.989 | 1.048 |
| *GKPCTPP | 464 | 470 | 7 | 1.221 | 1.045 |
| YGFY | 481 | 484 | 4 | 0.867 | 1.072 |
| GIGYQ | 488 | 492 | 5 | 0.615 | 1.015 |
| ATVCGPKLS | 508 | 516 | 9 | 0.272 | 1.1 |
| LTGTGVLTPSSKRFQPFQQFGRDV | 532 | 555 | 24 | 0.865 | 1.026 |
| DFTDSVRDPKTSEIL | 557 | 571 | 15 | 2.5 | 1.002 |
| VITPGTN | 583 | 589 | 7 | 0.575 | 1.01 |
| NCTDV | 602 | 606 | 5 | 0.509 | 1.069 |
| AIHADQLTP | 609 | 617 | 9 | 0.719 | 1.054 |
| YSTGNNVFQT | 622 | 631 | 10 | 1.389 | 0.991 |
| VDTSYECDIP | 642 | 651 | 10 | 0.704 | 1.068 |
| *VDTSY | 642 | 646 | 5 | 1.24 | 1.066 |

| Peptide | Start | End | Length | Emini[a] | Kolaskar [b] |
|---|---|---|---|---|---|
| TSQKSI | 669 | 674 | 6 | 1.683 | 1.005 |
| LGADSSIAYSNNTIAIP | 681 | 697 | 17 | 0.24 | 1.021 |
| MAKTSV | 713 | 718 | 6 | 0.768 | 1.021 |
| DSTE | 727 | 730 | 4 | 2.305 | 0.909 |
| EQDRNTRE | 755 | 762 | 8 | 13.355 | 0.877 |
| QMYKTPTLK | 769 | 777 | 9 | 3.956 | 0.999 |
| ILPDPLK | 787 | 793 | 7 | 0.821 | 1.082 |
| ECLGDI | 821 | 826 | 6 | 0.237 | 1.067 |
| LVSGT | 859 | 863 | 5 | 0.387 | 1.086 |
| QKQIANQ | 902 | 908 | 7 | 2.551 | 0.995 |
| QESLTTTSTALG | 917 | 928 | 12 | 1.155 | 0.997 |
| DVVNQNAQA | 932 | 940 | 9 | 1.01 | 1.038 |
| KVEAEV | 968 | 973 | 6 | 0.893 | 1.077 |
| *KVEA | 968 | 971 | 4 | 1.07 | 1.057 |
| NLAAT | 1005 | 1009 | 5 | 0.645 | 1.013 |
| RVDF | 1021 | 1024 | 4 | 0.866 | 1.053 |
| PQAAPH | 1035 | 1040 | 6 | 1.538 | 1.063 |
| *PQAAP | 1035 | 1039 | 5 | 1.395 | 1.054 |
| VPSQERNFTTAPAICHE | 1050 | 1066 | 17 | 0.833 | 1.035 |
| *VPSQE | 1050 | 1054 | 5 | 1.522 | 1.065 |
| KAYFP | 1068 | 1072 | 5 | 1.399 | 1.062 |
| *KAYF | 1068 | 1071 | 4 | 1.13 | 1.062 |
| IITT | 1096 | 1099 | 4 | 0.422 | 1.03 |
| NTFVS | 1101 | 1105 | 5 | 0.66 | 1.034 |
| TVYDPLQPELDSFKEE | 1118 | 1133 | 16 | 4.603 | 1.026 |
| *TVYDPLQPE | 1118 | 1126 | 9 | 2.299 | 1.063 |
| DKYFKNHTSPDVD | 1135 | 1147 | 13 | 7.114 | 0.997 |
| GDISGI | 1149 | 1154 | 6 | 0.288 | 0.988 |
| LGKYEQY | 1185 | 1191 | 7 | 2.592 | 1.035 |
| *LGKY | 1185 | 1188 | 4 | 1.054 | 1.054 |
| KFDEDDSEPVL | 1237 | 1247 | 11 | 2.765 | 1.003 |

*peptides revealed higher score if they were shorten in all tools.
a: default threshold value 1.000
b: default threshold value 1.039.

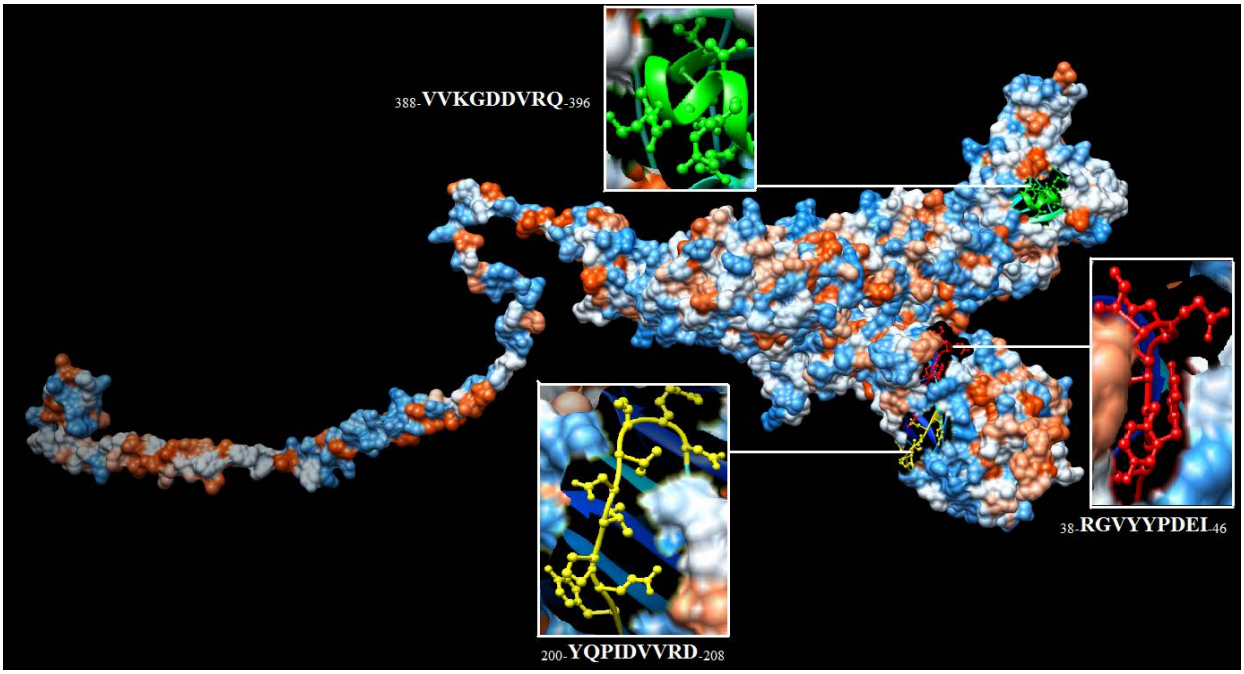


**Figure 3.** Position of proposed conserved B cell epitopes in structural level of spike S protein of SARS-CoV. Three epitopes were predicted to interact with B cell. The epitopes showed conservancy, surface accessibility and antigenicity using IEDB software

## 3.3. T Lymphocytes Epitopes Binding Prediction

### 3.3.1. MHC-I Binding Predictions

SARS spike S protein was analyzed using IEDB MHC-1 binding prediction tool to predict T lymphocytes epitopes that have binding affinity with MHC-I alleles based on Artificial Neural Network (ANN) with half-maximal inhibitory concentration (IC50) ≤ 100. As shown in Table 3 a total of 118 epitopes were found to interact with MHC-I alleles. The epitopes $_{47}$**FRSDTLYLT**$_{55}$, $_{195}$**YVYKGYQPI**$_{203}$, $_{880}$**FAMQMAYRF**$_{888}$ and $_{851}$**MIAAYTAAL**$_{859}$ interacted with the highest number of alleles and the best binding affinity to MHC-1 alleles (Table 5). The three dimensional structural level (3D) of these epitopes within spike S protein of SARS-CoA was shown in Figure 4.

**Table 3. List of epitopes that had binding affinity with MHC-I alleles. The position of peptides is according to position of amino acids in spike S protein of SARS CoA virus**

| Peptide | Start | End | Allele | ic50 | Percentile |
|---|---|---|---|---|---|
| VYYPDEIFR | 40 | 48 | HLA-A*31:01 | 33.19 | 0.4 |
| EIFRSDTLY | 45 | 53 | HLA-A*26:01 | 59.76 | 0.1 |
| | | | HLA-A*68:01 | 37.16 | 0.5 |
| FRSDTLYLT | 47 | 55 | HLA-C*07:01 | 85.79 | 0.1 |
| LTQDLFLPF | 54 | 62 | HLA-A*02:06 | 63.88 | 0.7 |
| | | | HLA-A*32:01 | 36.03 | 0.3 |
| | | | HLA-B*15:01 | 9.05 | 0.1 |
| LPFYSNVTG | 60 | 68 | HLA-B*35:01 | 43.15 | 0.3 |
| PFYSNVTGF | 61 | 69 | HLA-A*23:01 | 87.46 | 0.2 |
| FAATEKSNV | 89 | 97 | HLA-C*12:03 | 57.82 | 0.2 |
| AATEKSNVV | 90 | 98 | HLA-C*03:03 | 45.43 | 0.4 |
| IIINNSTNV | 115 | 123 | HLA-A*02:06 | 54.51 | 0.7 |
| GTQTHTMIF | 145 | 153 | HLA-A*32:01 | 58.32 | 0.3 |
| HTMIFDNAF | 149 | 157 | HLA-A*32:01 | 20.41 | 0.3 |
| | | | HLA-B*35:01 | 36.33 | 0.3 |
| MIFDNAFNC | 151 | 159 | HLA-A*02:06 | 71.85 | 0.7 |
| NAFNCTFEY | 155 | 163 | HLA-A*29:02 | 27.81 | 0.2 |
| | | | HLA-A*68:01 | 26.6 | 0.5 |
| | | | HLA-B*35:01 | 5.29 | 0.1 |
| | | | HLA-C*12:03 | 36.74 | 0.2 |
| KSGNFKHLR | 175 | 183 | HLA-A*31:01 | 5.24 | 0.1 |
| FVFKNKDGF | 185 | 193 | HLA-C*12:03 | 85.14 | 0.2 |
| KNKDGFLYV | 188 | 196 | HLA-A*30:01 | 29.63 | 0.4 |
| GFLYVYKGY | 192 | 200 | HLA-A*29:02 | 57.81 | 0.2 |
| YVYKGYQPI | 195 | 203 | HLA-A*02:01 | 47.04 | 0.5 |
| | | | HLA-A*02:06 | 6.14 | 0.1 |
| | | | HLA-A*68:02 | 79.07 | 0.7 |
| | | | HLA-C*03:03 | 5.33 | 0.1 |
| | | | HLA-C*12:03 | 6.97 | 0.1 |
| | | | HLA-C*14:02 | 12.14 | 0.2 |
| GYQPIDVVR | 199 | 207 | HLA-A*31:01 | 64.99 | 0.4 |
| LPLGINITN | 222 | 230 | HLA-B*35:01 | 70.7 | 0.4 |
| LGINITNFR | 224 | 232 | HLA-A*31:01 | 46.88 | 0.4 |
| | | | HLA-A*68:01 | 27.98 | 0.5 |
| GTSAAAYFV | 246 | 254 | HLA-A*02:06 | 41.09 | 0.7 |
| | | | HLA-A*68:02 | 10.33 | 0.2 |
| SAAAYFVGY | 248 | 256 | HLA-A*29:02 | 41.82 | 0.2 |
| | | | HLA-B*15:01 | 58.73 | 0.2 |
| | | | HLA-B*35:01 | 51.03 | 0.3 |
| AAAYFVGYL | 249 | 257 | HLA-A*68:02 | 37.12 | 0.6 |
| | | | HLA-C*03:03 | 37.17 | 0.4 |
| AAYFVGYLK | 250 | 258 | HLA-A*03:01 | 28.63 | 0.2 |
| | | | HLA-A*11:01 | 6.23 | 0.2 |
| | | | HLA-A*30:01 | 95.31 | 0.5 |
| | | | HLA-A*31:01 | 55.03 | 0.4 |
| | | | HLA-A*68:01 | 13.67 | 0.2 |
| SQNPLAELK | 279 | 287 | HLA-A*11:01 | 36.62 | 0.4 |
| SVKSFEIDK | 289 | 297 | HLA-A*11:01 | 42.8 | 0.4 |
| | | | HLA-A*30:01 | 30.96 | 0.4 |
| GIYQTSNFR | 298 | 306 | HLA-A*11:01 | 75.48 | 0.4 |

| Peptide | Start | End | Allele | ic50 | Percentile |
|---|---|---|---|---|---|
| | | | HLA-A*31:01 | 43.73 | 0.4 |
| | | | HLA-A*68:01 | 44.48 | 0.6 |
| VRFPNITNL | 314 | 322 | HLA-C*06:02 | 31.01 | 0.1 |
| | | | HLA-C*07:01 | 20.44 | 0.1 |
| | | | HLA-C*12:03 | 94.62 | 0.2 |
| FPNITNLCP | 316 | 324 | HLA-B*53:01 | 39.11 | 0.1 |
| NLCPFGEVF | 321 | 329 | HLA-B*15:01 | 83.41 | 0.2 |
| FPSVYAWER | 334 | 342 | HLA-A*68:01 | 89.6 | 0.6 |
| FSTFKCYGV | 361 | 369 | HLA-A*68:02 | 9.14 | 0.2 |
| KCYGVSATK | 365 | 373 | HLA-A*30:01 | 86.6 | 0.5 |
| FVVKGDDVR | 387 | 395 | HLA-A*68:01 | 18.49 | 0.3 |
| QIAPGQTGV | 396 | 404 | HLA-A*68:02 | 48.85 | 0.7 |
| GQTGVIADY | 400 | 408 | HLA-A*30:02 | 48.56 | 0.2 |
| GVIADYNYK | 403 | 411 | HLA-A*11:01 | 16.27 | 0.3 |
| | | | HLA-A*68:01 | 96.67 | 0.6 |
| VIADYNYKL | 404 | 412 | HLA-A*02:01 | 50.27 | 0.5 |
| HGKLRPFER | 445 | 453 | HLA-A*31:01 | 51.58 | 0.4 |
| YWPLNDYGF | 475 | 483 | HLA-A*23:01 | 64.95 | 0.2 |
| | | | HLA-A*24:02 | 62.83 | 0.5 |
| | | | HLA-C*14:02 | 88.11 | 0.3 |
| WPLNDYGFY | 476 | 484 | HLA-A*29:02 | 89.75 | 0.2 |
| | | | HLA-B*35:01 | 11.45 | 0.2 |
| YQPYRVVVL | 491 | 499 | HLA-A*02:06 | 99.74 | 0.8 |
| | | | HLA-B*39:01 | 79.44 | 0.3 |
| | | | HLA-C*12:03 | 42.2 | 0.2 |
| | | | HLA-C*14:02 | 55.72 | 0.2 |
| LIKNQCVNF | 519 | 527 | HLA-B*15:01 | 65.04 | 0.2 |
| CVNFNFNGL | 524 | 532 | HLA-A*68:02 | 23.74 | 0.6 |
| SSKRFQPFQ | 541 | 549 | HLA-A*30:01 | 34.51 | 0.4 |
| KRFQPFQQF | 543 | 551 | HLA-B*27:05 | 92.24 | 0.2 |
| FTDSVRDPK | 558 | 566 | HLA-A*68:01 | 38.38 | 0.5 |
| YQDVNCTDV | 598 | 606 | HLA-A*02:06 | 35.16 | 0.7 |
| | | | HLA-C*05:01 | 89.62 | 0.4 |
| FQTQAGCLI | 629 | 637 | HLA-A*02:06 | 17.15 | 0.6 |
| TQAGCLIGA | 631 | 639 | HLA-A*02:06 | 5.98 | 0.1 |
| TSQKSIVAY | 669 | 677 | HLA-B*15:01 | 84.15 | 0.2 |
| SIVAYTMSL | 673 | 681 | HLA-A*02:01 | 37.41 | 0.5 |
| | | | HLA-A*02:06 | 37.01 | 0.7 |
| | | | HLA-A*68:02 | 76.94 | 0.7 |
| | | | HLA-C*14:02 | 24.85 | 0.2 |
| MSLGADSSI | 679 | 687 | HLA-B*58:01 | 70.21 | 0.3 |
| LGADSSIAY | 681 | 689 | HLA-B*15:01 | 81.62 | 0.2 |
| | | | HLA-B*35:01 | 5.64 | 0.1 |
| SIAYSNNTI | 686 | 694 | HLA-A*32:01 | 64.24 | 0.3 |
| | | | HLA-A*68:02 | 61.13 | 0.7 |
| AYSNNTIAI | 688 | 696 | HLA-C*14:02 | 17.12 | 0.2 |
| IAIPTNFSI | 694 | 702 | HLA-A*02:06 | 42.71 | 0.7 |
| | | | HLA-B*58:01 | 17.41 | 0.3 |
| | | | HLA-C*03:03 | 12.91 | 0.2 |
| | | | HLA-C*12:03 | 68.85 | 0.2 |
| EVMPVSMAK | 707 | 715 | HLA-A*11:01 | 13.87 | 0.3 |
| | | | HLA-A*68:01 | 6.85 | 0.1 |
| MPVSMAKTS | 709 | 717 | HLA-B*35:01 | 45.74 | 0.3 |
| KTSVDCNMY | 715 | 723 | HLA-A*30:02 | 13.12 | 0.1 |
| TSVDCNMYI | 716 | 724 | HLA-A*68:02 | 6.06 | 0.2 |
| GSFCTQLNR | 739 | 747 | HLA-A*11:01 | 51.4 | 0.4 |
| | | | HLA-A*31:01 | 50.93 | 0.4 |
| | | | HLA-A*68:01 | 73.21 | 0.6 |
| FCTQLNRAL | 741 | 749 | HLA-C*03:03 | 46.82 | 0.4 |
| NTREVFAQV | 759 | 767 | HLA-A*02:06 | 74.34 | 0.7 |
| | | | HLA-A*68:02 | 9.64 | 0.2 |
| EVFAQVKQM | 762 | 770 | HLA-A*26:01 | 45.94 | 0.1 |
| VFAQVKQMY | 763 | 771 | HLA-A*29:02 | 42.2 | 0.2 |

| Peptide | Start | End | Allele | ic50 | Percentile |
|---|---|---|---|---|---|
| FAQVKQMYK | 764 | 772 | HLA-A*11:01 | 75.27 | 0.4 |
| | | | HLA-A*68:01 | 30.49 | 0.5 |
| KQMYKTPTL | 768 | 776 | HLA-A*02:01 | 47.7 | 0.5 |
| | | | HLA-A*02:06 | 27.15 | 0.6 |
| | | | HLA-A*32:01 | 11.86 | 0.3 |
| | | | HLA-C*14:02 | 46.48 | 0.2 |
| QMYKTPTLK | 769 | 777 | HLA-A*03:01 | 8.58 | 0.1 |
| | | | HLA-A*11:01 | 62.58 | 0.4 |
| SQILPDPLK | 785 | 793 | HLA-A*11:01 | 66.07 | 0.4 |
| RSFIEDLLF | 797 | 805 | HLA-B*57:01 | 17.84 | 0.3 |
| | | | HLA-B*58:01 | 11.93 | 0.2 |
| FIEDLLFNK | 799 | 807 | HLA-A*11:01 | 46.69 | 0.4 |
| | | | HLA-A*68:01 | 42.86 | 0.6 |
| FMKQYGECL | 815 | 823 | HLA-B*08:01 | 15.86 | 0.2 |
| | | | HLA-C*03:03 | 44.61 | 0.4 |
| | | | HLA-C*12:03 | 52.69 | 0.2 |
| | | | HLA-C*14:02 | 84.19 | 0.3 |
| LPPLLTDDM | 843 | 851 | HLA-B*35:01 | 78.12 | 0.4 |
| LLTDDMIAA | 846 | 854 | HLA-A*02:01 | 32.91 | 0.5 |
| | | | HLA-A*02:06 | 11.59 | 0.2 |
| LTDDMIAAY | 847 | 855 | HLA-A*01:01 | 2.53 | 0.1 |
| | | | HLA-A*29:02 | 85.6 | 0.2 |
| | | | HLA-B*35:01 | 25.26 | 0.3 |
| | | | HLA-C*05:01 | 17.95 | 0.4 |
| MIAAYTAAL | 851 | 859 | HLA-A*02:01 | 14.39 | 0.3 |
| | | | HLA-A*02:06 | 12.92 | 0.4 |
| | | | HLA-A*32:01 | 70.26 | 0.3 |
| | | | HLA-A*68:02 | 4.15 | 0.2 |
| | | | HLA-B*07:02 | 56.9 | 0.2 |
| | | | HLA-B*08:01 | 45.43 | 0.2 |
| | | | HLA-B*15:01 | 55.03 | 0.2 |
| | | | HLA-B*35:01 | 32.34 | 0.3 |
| | | | HLA-B*39:01 | 17.67 | 0.3 |
| | | | HLA-C*03:03 | 24.7 | 0.4 |
| | | | HLA-C*14:02 | 15.88 | 0.2 |
| IAAYTAALV | 852 | 860 | HLA-A*68:02 | 32.33 | 0.6 |
| YTAALVSGT | 855 | 863 | HLA-A*68:02 | 14.77 | 0.3 |
| AALQIPFAM | 874 | 882 | HLA-B*35:01 | 26.99 | 0.3 |
| | | | HLA-C*03:03 | 16.12 | 0.2 |
| LQIPFAMQM | 876 | 884 | HLA-A*02:06 | 16.5 | 0.6 |
| | | | HLA-B*15:01 | 36.09 | 0.1 |
| IPFAMQMAY | 878 | 886 | HLA-A*29:02 | 26.63 | 0.2 |
| | | | HLA-B*18:01 | 55.68 | 0.1 |
| | | | HLA-B*35:01 | 2.99 | 0.1 |
| | | | HLA-B*53:01 | 41.77 | 0.1 |
| FAMQMAYRF | 880 | 888 | HLA-A*29:02 | 82.71 | 0.2 |
| | | | HLA-B*08:01 | 74.55 | 0.2 |
| | | | HLA-B*35:01 | 5.03 | 0.1 |
| | | | HLA-B*53:01 | 13.26 | 0.1 |
| | | | HLA-B*58:01 | 32.13 | 0.3 |
| | | | HLA-C*03:03 | 39.73 | 0.4 |
| | | | HLA-C*12:03 | 50.95 | 0.2 |
| QMAYRFNGI | 883 | 891 | HLA-A*02:06 | 45.76 | 0.7 |
| AYRFNGIGV | 885 | 893 | HLA-A*30:01 | 22.18 | 0.4 |
| YRFNGIGVT | 886 | 894 | HLA-C*06:02 | 90.89 | 0.1 |
| | | | HLA-C*12:03 | 57.74 | 0.2 |
| KQIANQFNK | 903 | 911 | HLA-A*11:01 | 25.44 | 0.3 |
| TTTSTALGK | 921 | 929 | HLA-A*11:01 | 71.54 | 0.4 |
| | | | HLA-A*68:01 | 28.77 | 0.5 |
| TTSTALGKL | 922 | 930 | HLA-A*68:02 | 54.53 | 0.7 |
| VVNQNAQAL | 933 | 941 | HLA-C*03:03 | 49.2 | 0.4 |
| | | | HLA-C*14:02 | 79.63 | 0.3 |
| AQALNTLVK | 938 | 946 | HLA-A*11:01 | 77.12 | 0.4 |

| Peptide | Start | End | Allele | ic50 | Percentile |
|---|---|---|---|---|---|
| RLDKVEAEV | 965 | 973 | HLA-A*02:01 | 38.95 | 0.5 |
| | | | HLA-C*05:01 | 59.54 | 0.4 |
| AEVQIDRLI | 971 | 979 | HLA-B*44:03 | 63.1 | 0.1 |
| LITGRLQSL | 978 | 986 | HLA-A*02:06 | 95.7 | 0.7 |
| RLQSLQTYV | 982 | 990 | HLA-A*02:01 | 16.66 | 0.4 |
| QTYVTQQLI | 987 | 995 | HLA-B*58:01 | 92.25 | 0.3 |
| ASANLAATK | 1002 | 1010 | HLA-A*11:01 | 19.32 | 0.3 |
| | | | HLA-A*68:01 | 84.38 | 0.6 |
| ATKMSECVL | 1008 | 1016 | HLA-A*30:01 | 93.02 | 0.5 |
| FCGKGYHLM | 1024 | 1032 | HLA-C*03:03 | 95.84 | 0.4 |
| HLMSFPQAA | 1030 | 1038 | HLA-A*02:01 | 30.89 | 0.4 |
| QAAPHGVVF | 1036 | 1044 | HLA-B*15:01 | 40.36 | 0.1 |
| | | | HLA-B*35:01 | 22.61 | 0.3 |
| | | | HLA-C*03:03 | 17.46 | 0.2 |
| | | | HLA-C*12:03 | 19.58 | 0.2 |
| GVVFLHVTY | 1041 | 1049 | HLA-A*29:02 | 19.76 | 0.2 |
| VVFLHVTYV | 1042 | 1050 | HLA-A*02:01 | 36.56 | 0.5 |
| | | | HLA-A*02:06 | 21.97 | 0.6 |
| | | | HLA-A*68:02 | 51.51 | 0.7 |
| | | | HLA-C*12:03 | 77.09 | 0.2 |
| VTYVPSQER | 1047 | 1055 | HLA-A*31:01 | 78.84 | 0.5 |
| | | | HLA-A*68:01 | 56.35 | 0.6 |
| KAYFPREGV | 1068 | 1076 | HLA-C*12:03 | 52.24 | 0.2 |
| AYFPREGVF | 1069 | 1077 | HLA-C*07:02 | 59.02 | 0.2 |
| | | | HLA-C*14:02 | 39.35 | 0.2 |
| FPREGVFVF | 1071 | 1079 | HLA-B*07:02 | 19.14 | 0.2 |
| | | | HLA-B*35:01 | 3.52 | 0.1 |
| | | | HLA-B*53:01 | 28.8 | 0.1 |
| | | | HLA-C*12:03 | 48.55 | 0.2 |
| FVFNGTSWF | 1077 | 1085 | HLA-A*02:06 | 28.53 | 0.6 |
| | | | HLA-A*26:01 | 12.13 | 0.1 |
| | | | HLA-B*15:02 | 35.34 | 0.1 |
| | | | HLA-B*35:01 | 10.81 | 0.2 |
| | | | HLA-C*03:03 | 82.72 | 0.4 |
| | | | HLA-C*05:01 | 56.8 | 0.4 |
| | | | HLA-C*12:03 | 10.47 | 0.2 |
| | | | HLA-C*14:02 | 54.24 | 0.2 |
| GTSWFITQR | 1081 | 1089 | HLA-A*11:01 | 39.88 | 0.4 |
| | | | HLA-A*31:01 | 11.83 | 0.3 |
| | | | HLA-A*68:01 | 10.23 | 0.1 |
| QIITTDNTF | 1095 | 1103 | HLA-B*15:01 | 66.32 | 0.2 |
| VYDPLQPEL | 1119 | 1127 | HLA-C*14:02 | 66.93 | 0.2 |
| YEQYIKWPW | 1188 | 1196 | HLA-B*18:01 | 4.19 | 0.1 |
| | | | HLA-B*44:02 | 12.91 | 0.1 |
| | | | HLA-B*44:03 | 38.77 | 0.1 |
| QYIKWPWYV | 1190 | 1198 | HLA-A*23:01 | 15.67 | 0.2 |
| | | | HLA-A*24:02 | 66.28 | 0.5 |
| YIKWPWYVW | 1191 | 1199 | HLA-C*12:03 | 33.46 | 0.2 |
| WPWYVWLGF | 1194 | 1202 | HLA-B*35:01 | 9.09 | 0.1 |
| | | | HLA-B*53:01 | 55.52 | 0.2 |
| MTSCCSCLK | 1219 | 1227 | HLA-A*03:01 | 87.4 | 0.2 |
| | | | HLA-A*11:01 | 17.44 | 0.3 |
| | | | HLA-A*30:01 | 15.66 | 0.4 |
| | | | HLA-A*31:01 | 98.56 | 0.5 |
| | | | HLA-A*68:01 | 7.86 | 0.1 |
| VLKGVKLHY | 1246 | 1254 | HLA-B*15:01 | 96.09 | 0.2 |

### 3.3.2. MHC-II binding predictions

SARS spike S protein was analyzed using IEDB MHC- II binding prediction tool to predict T lymphocytes epitopes that have binding affinity with MHC-II alleles based on NN-align with half-maximal inhibitory concentration (IC50) ≤ 1000. Total of 559 core epitopes were found to interact with MHC-II alleles. Table 4 demonstrated the best four core epitopes that interacted with MHC-II alleles. The core epitopes $_{47}$-**FRSDTLYLT**-$_{55}$, $_{195}$-**YVYKGYQPI**-$_{203}$, $_{880}$-**FAMQMAYRF**-$_{888}$ and

$_{782}$-**FNFSQILPD**-$_{790}$ interacted with higher number of MHC-II alleles (Table 5). The three dimensional structural level (3D) of these epitopes within spike S protein of the SARS-CoA was shown in Figure 4. The other core epitopes and their corresponding alleles that interacted with MHC-ΠI were supplemented in an extra sheet.

**Table 4. List of top four epitopes that had binding affinity with MHC-II alleles. The position of peptides is according to position of amino acids in spike S protein of SARS-CoA virus**

| Core Sequence | Peptide Sequence | Start | End | Allele | IC50 | Rank |
|---|---|---|---|---|---|---|
| **FAMQMAYRF** | AALQIPFAMQMAYRF | 874 | 888 | HLA-DPA1*01:03/DPB1*02:01 | 335.3 | 17.32 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 138.2 | 2.73 |
| | | | | HLA-DRB1*01:01 | 8.9 | 4.3 |
| | | | | HLA-DRB1*04:01 | 351.9 | 22.86 |
| | | | | HLA-DRB1*04:05 | 281.6 | 20.88 |
| | | | | HLA-DRB1*07:01 | 24.8 | 4.79 |
| | | | | HLA-DRB1*11:01 | 390.8 | 27.93 |
| | | | | HLA-DRB1*15:01 | 53.6 | 5.48 |
| | | | | HLA-DRB3*01:01 | 73.2 | 3.66 |
| | | | | HLA-DRB5*01:01 | 21 | 5.1 |
| | ALQIPFAMQMAYRFN | 875 | 889 | HLA-DPA1*01/DPB1*04:01 | 962.4 | 21.37 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 358.9 | 17.99 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 173.7 | 3.65 |
| | | | | HLA-DRB1*01:01 | 7.8 | 3.36 |
| | | | | HLA-DRB1*04:01 | 371.9 | 23.72 |
| | | | | HLA-DRB1*04:05 | 244.2 | 18.99 |
| | | | | HLA-DRB1*07:01 | 34.3 | 6.4 |
| | | | | HLA-DRB1*11:01 | 275.5 | 23.78 |
| | | | | HLA-DRB1*15:01 | 48.8 | 4.95 |
| | | | | HLA-DRB3*01:01 | 83.6 | 3.99 |
| | | | | HLA-DRB5*01:01 | 17.3 | 4.26 |
| | LQIPFAMQMAYRFNG | 876 | 890 | HLA-DPA1*01:03/DPB1*02:01 | 396.4 | 18.98 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 168.5 | 15.71 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 236.4 | 5.22 |
| | | | | HLA-DRB1*01:01 | 6.9 | 2.55 |
| | | | | HLA-DRB1*04:01 | 460.3 | 27.34 |
| | | | | HLA-DRB1*04:05 | 259.1 | 19.77 |
| | | | | HLA-DRB1*07:01 | 59.8 | 9.94 |
| | | | | HLA-DRB1*09:01 | 75.8 | 5.21 |
| | | | | HLA-DRB1*11:01 | 115.4 | 14.97 |
| | | | | HLA-DRB1*15:01 | 55.1 | 5.64 |
| | | | | HLA-DRB3*01:01 | 112.6 | 4.82 |
| | | | | HLA-DRB5*01:01 | 17.2 | 4.23 |
| | QIPFAMQMAYRFNGI | 877 | 891 | HLA-DPA1*01/DPB1*04:01 | 987.8 | 21.66 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 299.6 | 16.26 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 181.6 | 16.58 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 280.5 | 6.29 |
| | | | | HLA-DRB1*01:01 | 6.1 | 1.81 |
| | | | | HLA-DRB1*04:01 | 428.9 | 26.1 |
| | | | | HLA-DRB1*04:05 | 245.6 | 19.05 |
| | | | | HLA-DRB1*07:01 | 81.9 | 12.4 |
| | | | | HLA-DRB1*09:01 | 164.3 | 10.95 |
| | | | | HLA-DRB1*11:01 | 67.6 | 10.62 |
| | | | | HLA-DRB1*15:01 | 52.2 | 5.32 |
| | | | | HLA-DRB3*01:01 | 124.1 | 5.14 |
| | | | | HLA-DRB5*01:01 | 14.9 | 3.69 |
| | IPFAMQMAYRFNGIG | 878 | 892 | HLA-DPA1*01:03/DPB1*02:01 | 383.7 | 18.66 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 247.4 | 20.44 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 358.6 | 8.11 |
| | | | | HLA-DRB1*01:01 | 8 | 3.54 |
| | | | | HLA-DRB1*04:01 | 634.1 | 33.3 |
| | | | | HLA-DRB1*04:05 | 372.5 | 24.79 |
| | | | | HLA-DRB1*07:01 | 127.8 | 16.54 |
| | | | | HLA-DRB1*09:01 | 225.6 | 14.24 |
| | | | | HLA-DRB1*15:01 | 65.1 | 6.68 |
| | | | | HLA-DRB3*01:01 | 293.8 | 8.6 |
| | | | | HLA-DRB5*01:01 | 22.8 | 5.49 |

| Core Sequence | Peptide Sequence | Start | End | Allele | IC50 | Rank |
|---|---|---|---|---|---|---|
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 519.6 | 21.93 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 327.8 | 24.31 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 466.9 | 10.48 |
| | | | | HLA-DRB1*01:01 | 11.1 | 6.03 |
| | | | | HLA-DRB1*04:01 | 635.7 | 33.35 |
| | PFAMQMAYRFNGIGV | 879 | 893 | HLA-DRB1*04:04 | 297.1 | 26.12 |
| | | | | HLA-DRB1*04:05 | 636.3 | 33.23 |
| | | | | HLA-DRB1*07:01 | 109.5 | 15.06 |
| | | | | HLA-DRB1*09:01 | 166.4 | 11.08 |
| | | | | HLA-DRB1*15:01 | 62.8 | 6.46 |
| | | | | HLA-DRB3*01:01 | 639.9 | 13.6 |
| | | | | HLA-DRB5*01:01 | 34 | 7.58 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 959.1 | 21.32 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 563.5 | 22.89 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 389.4 | 26.82 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 520.3 | 11.6 |
| | | | | HLA-DRB1*01:01 | 14.6 | 8.36 |
| | FAMQMAYRFNGIGVT | 880 | 894 | HLA-DRB1*04:01 | 832.7 | 38.92 |
| | | | | HLA-DRB1*04:04 | 394.8 | 30.75 |
| | | | | HLA-DRB1*04:05 | 700.5 | 34.91 |
| | | | | HLA-DRB1*07:01 | 156.8 | 18.73 |
| | | | | HLA-DRB1*09:01 | 219.8 | 13.96 |
| | | | | HLA-DRB1*15:01 | 119.9 | 11.57 |
| | | | | HLA-DRB5*01:01 | 45.4 | 9.35 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 7.4 | 0.77 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 37.9 | 4.49 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 927.2 | 38.68 |
| | | | | HLA-DQA1*03:01/DQB1*03:02 | 997.9 | 17.66 |
| | | | | HLA-DQA1*04:01/DQB1*04:02 | 467.2 | 7.24 |
| | LKYFGGFNFSQILPD | 776 | 790 | HLA-DQA1*05:01/DQB1*02:01 | 502 | 11.21 |
| | | | | HLA-DQA1*05:01/DQB1*03:01 | 243.5 | 25.12 |
| | | | | HLA-DRB1*04:04 | 52.1 | 5.97 |
| | | | | HLA-DRB1*04:05 | 14.2 | 0.65 |
| | | | | HLA-DRB1*07:01 | 32.4 | 6.11 |
| | | | | HLA-DRB1*11:01 | 198.3 | 20.2 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 8.3 | 0.9 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 41.9 | 4.93 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 400 | 23.64 |
| | | | | HLA-DQA1*03:01/DQB1*03:02 | 865.2 | 15.42 |
| | | | | HLA-DQA1*04:01/DQB1*04:02 | 382.9 | 5.76 |
| | KYFGGFNFSQILPDP | 777 | 791 | HLA-DQA1*05:01/DQB1*02:01 | 417.1 | 9.41 |
| | | | | HLA-DQA1*05:01/DQB1*03:01 | 253.8 | 25.66 |
| | | | | HLA-DRB1*04:04 | 45.8 | 5.09 |
| **FNFSQILPD** | | | | HLA-DRB1*04:05 | 14.8 | 0.71 |
| | | | | HLA-DRB1*07:01 | 74.1 | 11.6 |
| | | | | HLA-DRB1*11:01 | 176.1 | 18.98 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 10.8 | 1.25 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 301.5 | 19.38 |
| | | | | HLA-DQA1*03:01/DQB1*03:02 | 848.8 | 15.14 |
| | YFGGFNFSQILPDPL | 778 | 792 | HLA-DQA1*04:01/DQB1*04:02 | 463.6 | 7.18 |
| | | | | HLA-DQA1*05:01/DQB1*03:01 | 272.4 | 26.64 |
| | | | | HLA-DRB1*04:04 | 38.8 | 4.13 |
| | | | | HLA-DRB1*11:01 | 129.1 | 15.98 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 19.6 | 2.34 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 277 | 18.19 |
| | FGGFNFSQILPDPLK | 779 | 793 | HLA-DQA1*04:01/DQB1*04:02 | 662 | 10.54 |
| | | | | HLA-DQA1*05:01/DQB1*03:01 | 420.6 | 33.05 |
| | | | | HLA-DRB1*04:04 | 39 | 4.17 |
| | | | | HLA-DRB1*11:01 | 79.6 | 11.88 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 27.6 | 3.19 |
| | GGFNFSQILPDPLKP | 780 | 794 | HLA-DPA1*02:01/DPB1*01:01 | 56.5 | 6.03 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 398.4 | 23.57 |
| | | | | HLA-DQA1*05:01/DQB1*03:01 | 717 | 41.92 |

| Core Sequence | Peptide Sequence | Start | End | Allele | IC50 | Rank |
|---|---|---|---|---|---|---|
| | | | | HLA-DRB1*04:04 | 38.5 | 4.07 |
| | | | | HLA-DRB1*11:01 | 87.6 | 12.65 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 81.1 | 7.27 |
| | GFNFSQILPDPLKPT | 781 | 795 | HLA-DPA1*02:01/DPB1*01:01 | 85.2 | 9.04 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 610.9 | 30.83 |
| | | | | HLA-DRB1*04:04 | 46.7 | 5.22 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 364.2 | 18.14 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 146.1 | 14.11 |
| | FNFSQILPDPLKPTK | 782 | 796 | HLA-DPA1*03:01/DPB1*04:02 | 613.5 | 25.91 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 832.5 | 36.57 |
| | | | | HLA-DRB1*04:04 | 54 | 6.23 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 99 | 5.26 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 45.2 | 4.75 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 33.5 | 3.29 |
| | | | | HLA-DPA1*02:01/DPB1*05:01 | 195.2 | 4.31 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 54.9 | 6.23 |
| | | | | HLA-DQA1*05:01/DQB1*02:01 | 457.6 | 10.28 |
| | YYPDEIFRSDTLYLT | 41 | 55 | HLA-DQA1*05:01/DQB1*03:01 | 941.4 | 46.84 |
| | | | | HLA-DRB1*01:01 | 327.9 | 49.68 |
| | | | | HLA-DRB1*03:01 | 65.5 | 3.58 |
| | | | | HLA-DRB1*04:01 | 17 | 0.71 |
| | | | | HLA-DRB1*04:05 | 41 | 3.83 |
| | | | | HLA-DRB1*11:01 | 428.4 | 29.07 |
| | | | | HLA-DRB3*01:01 | 3.4 | 0.02 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 73.2 | 4.18 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 38.7 | 4.22 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 27.8 | 2.57 |
| | | | | HLA-DPA1*02:01/DPB1*05:01 | 187.2 | 4.12 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 31.1 | 3.71 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 728.6 | 34.06 |
| | YPDEIFRSDTLYLTQ | 42 | 56 | HLA-DQA1*05:01/DQB1*02:01 | 670.5 | 14.63 |
| | | | | HLA-DRB1*01:01 | 173.8 | 38.54 |
| | | | | HLA-DRB1*03:01 | 44.6 | 2.62 |
| | | | | HLA-DRB1*04:01 | 12.9 | 0.39 |
| | | | | HLA-DRB1*04:05 | 38.7 | 3.56 |
| | | | | HLA-DRB1*11:01 | 254.8 | 22.91 |
| **FRSDTLYLT** | | | | HLA-DRB3*01:01 | 3.2 | 0.01 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 70.5 | 4.05 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 39.4 | 4.28 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 25.2 | 2.23 |
| | | | | HLA-DPA1*02:01/DPB1*05:01 | 159.2 | 3.46 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 22.1 | 2.58 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 343.6 | 21.29 |
| | PDEIFRSDTLYLTQD | 43 | 57 | HLA-DQA1*05:01/DQB1*02:01 | 820.4 | 17.4 |
| | | | | HLA-DRB1*01:01 | 176.5 | 38.79 |
| | | | | HLA-DRB1*03:01 | 37.9 | 2.24 |
| | | | | HLA-DRB1*04:01 | 10 | 0.2 |
| | | | | HLA-DRB1*04:05 | 38 | 3.48 |
| | | | | HLA-DRB1*11:01 | 165.4 | 18.35 |
| | | | | HLA-DRB3*01:01 | 3.1 | 0.01 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 54.2 | 3.3 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 33.2 | 3.73 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 20.7 | 1.67 |
| | | | | HLA-DPA1*02:01/DPB1*05:01 | 127.6 | 2.69 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 17.6 | 1.95 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 321.4 | 20.3 |
| | DEIFRSDTLYLTQDL | 44 | 58 | HLA-DRB1*01:01 | 109 | 31.42 |
| | | | | HLA-DRB1*03:01 | 33.4 | 1.98 |
| | | | | HLA-DRB1*04:01 | 7.6 | 0.09 |
| | | | | HLA-DRB1*04:05 | 32.7 | 2.83 |
| | | | | HLA-DRB1*11:01 | 102.8 | 13.95 |
| | | | | HLA-DRB3*01:01 | 3 | 0.01 |

| Core Sequence | Peptide Sequence | Start | End | Allele | IC50 | Rank |
|---|---|---|---|---|---|---|
| | | | | HLA-DPA1*01/DPB1*04:01 | 53.2 | 3.24 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 35.8 | 3.96 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 23.9 | 2.07 |
| | | | | HLA-DPA1*02:01/DPB1*05:01 | 94.8 | 1.87 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 21.7 | 2.53 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 462.8 | 26.02 |
| | EIFRSDTLYLTQDLF | 45 | 59 | HLA-DRB1*01:01 | 126.3 | 33.55 |
| | | | | HLA-DRB1*03:01 | 72.8 | 3.9 |
| | | | | HLA-DRB1*04:01 | 8.9 | 0.15 |
| | | | | HLA-DRB1*04:05 | 37.2 | 3.38 |
| | | | | HLA-DRB1*11:01 | 159.5 | 17.99 |
| | | | | HLA-DRB3*01:01 | 3.4 | 0.02 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 37 | 2.38 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 23.2 | 2.74 |
| | | | | HLA-DPA1*02:01/DPB1*05:01 | 81.8 | 1.54 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 26.6 | 3.18 |
| | | | | HLA-DQA1*01:02/DQB1*06:02 | 644 | 31.78 |
| | IFRSDTLYLTQDLFL | 46 | 60 | HLA-DRB1*01:01 | 161.2 | 37.32 |
| | | | | HLA-DRB1*03:01 | 200.9 | 8.07 |
| | | | | HLA-DRB1*04:01 | 11.9 | 0.32 |
| | | | | HLA-DRB1*04:05 | 33.7 | 2.96 |
| | | | | HLA-DRB1*11:01 | 290.2 | 24.37 |
| | | | | HLA-DRB3*01:01 | 4.2 | 0.06 |
| | | | | HLA-DRB1*03:01 | 528.6 | 14.53 |
| | | | | HLA-DRB1*04:01 | 26.4 | 1.56 |
| | FRSDTLYLTQDLFLP | 47 | 61 | HLA-DRB1*04:05 | 43.7 | 4.14 |
| | | | | HLA-DRB1*11:01 | 630.6 | 34.12 |
| | | | | HLA-DRB3*01:01 | 5.6 | 0.16 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 209.2 | 8.79 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 51.6 | 5.26 |
| | NKDGFLYVYKGYQPI | 189 | 203 | HLA-DRB1*01:01 | 105.9 | 31 |
| | | | | HLA-DRB1*07:01 | 73 | 11.47 |
| | | | | HLA-DRB3*01:01 | 314.7 | 8.94 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 194 | 8.36 |
| | | | | HLA-DQA1*05:01/DQB1*03:01 | 673.1 | 40.83 |
| | KDGFLYVYKGYQPID | 190 | 204 | HLA-DRB1*01:01 | 109 | 31.42 |
| | | | | HLA-DRB1*07:01 | 106.7 | 14.79 |
| | | | | HLA-DRB3*01:01 | 286.2 | 8.47 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 159.5 | 7.34 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 40.4 | 4.36 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 240.8 | 20.08 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 590.3 | 25.45 |
| | DGFLYVYKGYQPIDV | 191 | 205 | HLA-DQA1*05:01/DQB1*03:01 | 397.4 | 32.19 |
| | | | | HLA-DRB1*01:01 | 48.8 | 20.96 |
| | | | | HLA-DRB1*13:02 | 731.7 | 21.74 |
| **YVYKGYQPI** | | | | HLA-DRB3*01:01 | 279 | 8.35 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 142.5 | 6.8 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 38.5 | 4.2 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 217.6 | 18.78 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 485.8 | 23.3 |
| | GFLYVYKGYQPIDVV | 192 | 206 | HLA-DQA1*05:01/DQB1*03:01 | 361.8 | 30.75 |
| | | | | HLA-DRB1*01:01 | 28.1 | 14.76 |
| | | | | HLA-DRB1*13:02 | 581.6 | 19.02 |
| | | | | HLA-DRB3*01:01 | 287.1 | 8.48 |
| | | | | HLA-DRB4*01:01 | 874 | 37.84 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 143.2 | 6.82 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 39.4 | 4.28 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 207.3 | 18.17 |
| | FLYVYKGYQPIDVVR | 193 | 207 | HLA-DPA1*03:01/DPB1*04:02 | 400.6 | 21.24 |
| | | | | HLA-DQA1*05:01/DQB1*03:01 | 373.2 | 31.22 |
| | | | | HLA-DRB1*13:02 | 563.8 | 18.69 |
| | | | | HLA-DRB3*01:01 | 700 | 14.34 |

| Core Sequence | Peptide Sequence | Start | End | Allele | IC50 | Rank |
|---|---|---|---|---|---|---|
| | | | | HLA-DRB4*01:01 | 599.8 | 30.63 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 375.1 | 12.57 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 85.1 | 7.51 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 307.9 | 23.42 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 526.4 | 24.19 |
| | LYVYKGYQPIDVVRD | 194 | 208 | HLA-DQA1*05:01/DQB1*03:01 | 434.6 | 33.56 |
| | | | | HLA-DRB1*13:02 | 616.8 | 19.68 |
| | | | | HLA-DRB1*15:01 | 137.7 | 12.94 |
| | | | | HLA-DRB4*01:01 | 568 | 29.65 |
| | | | | HLA-DRB5*01:01 | 222.4 | 23.16 |
| | | | | HLA-DPA1*01/DPB1*04:01 | 660.4 | 17.4 |
| | | | | HLA-DPA1*01:03/DPB1*02:01 | 295.9 | 16.14 |
| | | | | HLA-DPA1*02:01/DPB1*01:01 | 518.2 | 31.31 |
| | | | | HLA-DPA1*03:01/DPB1*04:02 | 910 | 30.66 |
| | YVYKGYQPIDVVRDL | 195 | 209 | HLA-DQA1*05:01/DQB1*03:01 | 643 | 40.04 |
| | | | | HLA-DRB1*04:04 | 158.2 | 17.13 |
| | | | | HLA-DRB1*13:02 | 771.6 | 22.4 |
| | | | | HLA-DRB4*01:01 | 360.2 | 22.11 |
| | | | | HLA-DRB5*01:01 | 271.6 | 25.44 |

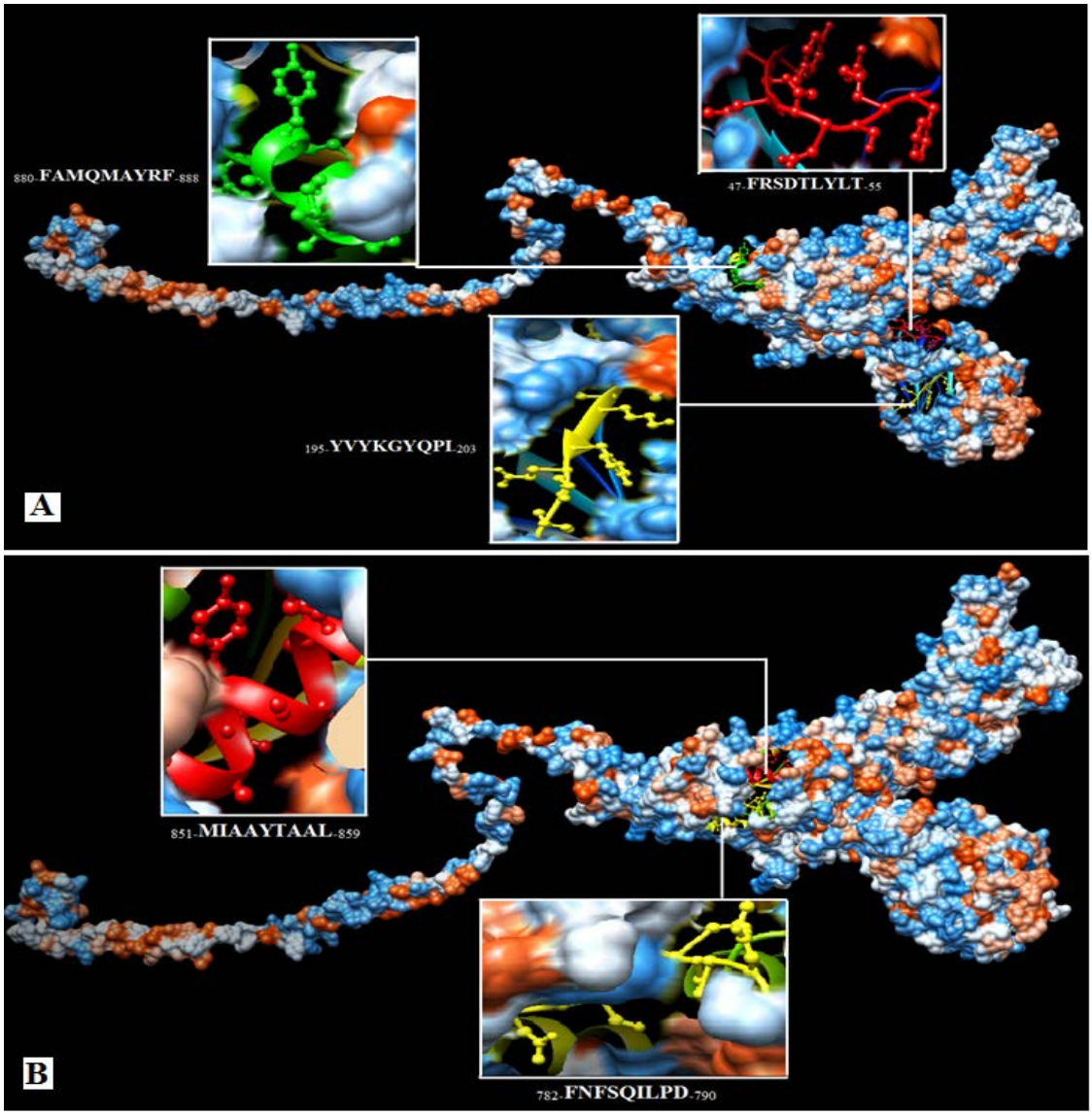


**Figure 4.** T cell proposed epitopes that interact with MHC-I and/ or MHC-II. (A): The epitopes $_{47}$-**FRSDTLYLT**-$_{55}$, $_{195}$-**YVYKGYQPI**-$_{203}$ and $_{880}$-**FAMQMAYRF**-$_{888}$ were found to interact with both MHC-1 and MHC-II alleles. (B): The epitope $_{851}$-MIAAYTAAL-$_{859}$ interacted with MHC-I alleles while the epitope $_{782}$-FNFSQILPD-$_{790}$ interacted only with MHC-II alleles. The positions of proposed epitopes were shown in the 3D structural level of spike S protein of SARS-CoV

**Table 5. The predicted epitopes and the number of the MHC-1 and/or MHC-II interacted alleles**

| Epitope | MHC-1 interacted alleles | Epitope | MHC-11 interacted alleles |
|---|---|---|---|
| FRSDTLYLT | 1 | FRSDTLYLT | 117 |
| YVYKGYQPI | 6 | YVYKGYQPI | 82 |
| FAMQMAYRF | 7 | FAMQMAYRF | 103 |
| MIAAYTAAL | 11 | FNFSQILPD | 84 |

**Table 6. The population coverage against the whole world for the predicted epitopes. The first three epitopes were found to be interacted with both MHC-1 and MHC-11 with good population coverage, while the epitopes MIAAYTAAL and FNFSQILPD demonstrated higher population coverage only in MHC-1 and MHC-11 respectively. The overall population coverage epitope set for all predicted epitopes in MHC-1 and MHC-11 was 100%**

| MHC-1 | | MHC-11 | | MHC-1/ MHC-11 | |
|---|---|---|---|---|---|
| Epitope | Population coverage | Epitope | Population coverage | Epitope | Population coverage |
| FRSDTLYLT | 19.44% | FRSDTLYLT | 99.77% | FRSDTLYLT | 99.81% |
| YVYKGYQPI | 54.45% | YVYKGYQPI | 99.21% | YVYKGYQPI | 99.64% |
| FAMQMAYRF | 39.96% | FAMQMAYRF | 98.88% | FAMQMAYRF | 99.33% |
| MIAAYTAAL | 70.80% | FNFSQILPD | 99.77% | MIAAYTAAL | 70.80% |
| | | | | FNFSQILPD | 99.77% |
| Epitope set | 82.16% | Epitope set | 99.97% | Epitope set | 100% |

## 3.4. Population Coverage

The suggested epitopes that demonstrated higher affinity to interact with MHC-I and MHC-II alleles and that bound to different sets of alleles were selected for population coverage analysis. Table 6 demonstrated the population coverage percentages for each epitope and their epitopes sets. The epitopes $_{47}$-**FRSDTLYLT**-$_{55}$, $_{195}$-**YVYKGYQPI**-$_{203}$, $_{880}$-**FAMQMAYRF**-$_{888}$ and $_{851}$-**MIAAYTAAL**-$_{859}$ interacted with most frequent MHC-I alleles and they demonstrated high percentage against the whole world population coverage with epitope set 82.16%**.** Strikingly the first three epitopes that interacted with MHC-1 in addition to the epitope $_{782}$-**FNFSQILPD**-$_{790}$ were interacted with most frequent MHC-II alleles and they demonstrated high percentage against the whole world population coverage with epitopes set of **99.97%.** The overall MHC-1 and MHC-II whole world population coverage for all suggested epitopes was **100%.**

# 4. Discussion

The main goal of vaccine design is to elicit immunity of an individual against particular pathogens by selectively stimulating antigen specific for B and T cells [57]. Vaccine mainly contains two classes of epitopes: a B-cell epitopes and a T-cell epitopes. The combination of these epitopes, vaccine is able to either induce specific humoral or cellular immune against specific pathogens [57]. Therefore in the present study the B and T cells were analyzed systemically against 131 retrieved spike S protein of SARS-CoV to obtain epitope vaccine candidates. Moreover for better determination of best vaccine candidate epitopes; the whole spike S protein was analyzed including the receptor binding domain (RBD). Noticeably the amino acids sequence from 1 to 510 of the SARS-CoV spike S glycoprotein represents a unique domain during SARS infection [16,17,19-25]. Also it contains the receptor binding domain, RBD (amino acids from 270 to 510), analogous to the S1subunit of other coronavirus S glycoproteins [17]. Thus RBD sequence was used to develop safe and effective SARS vaccines in multiple studies [17,18,19,58,59].

In this study eight epitopes were successfully predicted to interact with B and T cells. In our results three epitopes demonstrated high affinity to interact with B cell and were found within the region from 1 to 510 of spike S glycoprotein. These epitopes were $_{38}$-**RGVYYPDEI**-$_{46}$, $_{200}$-**YQPIDVVRD**-$_{208}$ and $_{388}$-**VVKGDDVRQ**-$_{396}$. Most importantly the latter epitope ($_{388}$-**VVKGDDVRQ**-$_{396}$) was positioned within the RBD region. Furthermore the three epitopes demonstrated conservancy in Bepipred as linear epitopes and got high scores (over the thresholds) in Emini surface accessibility and Kolaskar and Tongaonkar antigenicity method. Therefore they were recommended as promising peptides vaccine candidates against B cells. The epitopes illustrated in Table 2, were the only conserved epitopes from all retrieved strains of spike S glycoprotein that were available in NCBI database until May 2017 and have high probability of activating humoral immune response.

The immune response of T cell is considered as a long lasting response compared to B cell, where the antigen can easily escape the antibody memory response [60]. Vaccines that effectively generate cell-mediated response are needed to provide protection against the invading pathogen. Moreover the CD8+ T and CD4+ T cell responses play a major role in antiviral immunity [61]. Thus designing vaccine against T cell is much more important. In this study 118 and 559 conserved T cell epitopes were predicted to interact with both MHC-1 and MHC-II alleles respectively. Among them, five epitopes $_{47}$-**FRSDTLYLT**-$_{55}$, $_{195}$-**YVYKGYQPI**-$_{203}$, $_{880}$-**FAMQMAYRF**-$_{888}$, $_{782}$-**FNFSQILPD**-$_{790}$ and $_{851}$-**MIAAYTAAL**-$_{859}$ demonstrated good binding affinity against MHC-1 or MHC-II alleles.

The epitope $_{47}$-**FRSDTLYLT**-$_{55}$ was found positioned in the region from 1 to 510 that considered as a unique domain during SARS infection. This epitope successfully interacted with both MHC-1 and MHC-II alleles. Moreover it interacted with the highest number of alleles (Table 5) and provided the highest population coverage epitope set (Table 6). Also it was found tandemly

positioned with $_{38}$-**RGVYYPDEI**-$_{46}$ (epitope that predicted for B-cell). This result may reflect the importance of this region in the spike S protein during SARS infection since it contains these tandem epitopes. Also the epitope $_{195}$-**YVYKGYQPI**-$_{203}$ demonstrated unique features. Firstly it was found to be located within the unique domain (residues 1 to 510) of spike S protein [62]. Secondly it provided high binding affinity with both MHC-1 and MHC-II alleles with favorable epitope set (Table 6). Thirdly it partially overlapped with $_{200}$-**YQPIDVVRD**-$_{208}$ (epitope that predicted for B-cell) in its last four amino acids. This overlapping between these two epitopes may indicate the importance of their region during SARS infection.

In a study by Babcock et al. [17] described the construction and expression of a codon-optimized gene encoding the soluble ectodomain (amino acids 1 to 1190) in the S spike glycoprotein of SARS-CoV. They proposed that amino acids 1 to 510 termed S1and amino acids 511 to 1190 be called S2. In their study S1 and S2 exhibited very similar profiles for binding to the Vero E6 cell surface; that is; S1 binds to Vero E6 cells at least as well as S2 binds. Moreover they showed that antibodies specific for S2 were predicted to interfere with fusion of the viral envelope and host cell. They demonstrated this region (S2) as appropriate targets for monoclonal antibody development or as vaccine candidates [17]. In the present study the epitopes $_{782}$-**FNFSQILPD**-$_{790}$, $_{851}$-**MIAAYTAAL**-$_{859}$ and $_{880}$-**FAMQMAYRF**-$_{888}$ were located within S2 region and may act or interfere with fusion of the virus envelope to host cell. Moreover the epitope $_{880}$-**FAMQMAYRF**-$_{888}$ showed high binding affinity for both MHC-1 and MHC-II alleles with good population coverage against whole world. While, the epitope $_{851}$-**MIAAYTAAL**-$_{859}$ provided excellent results concerning binding affinity to MHC-1 alleles. It is noteworthy that this epitope interacted with eleven alleles in MHC-1 with population coverage 70.80%. However it only bound to 31 alleles in MHC-II with population coverage 86.76%, a percentage that is less compared to other predicted epitopes percentages (Table 6). Therefore it was only chosen as MHC-1 epitope. In the same manner the epitope $_{782}$-**FNFSQILPD**-$_{790}$ was successfully interacted with MHC-II alleles (84 alleles) with population coverage 99.77% but it did not interact with MHC-1 alleles. Thus it was only chosen as MHC-II epitope. Accordingly the five predicted epitopes demonstrated very promising features against T lymphocytes as a vaccine candidate against SARS-CoV. Moreover the population coverage against the whole world for the epitopes that interacted with MHC-I and MHC-II alleles provided coverage of 82.16% and 99.97% respectively. The overall epitope set for the MHC-I and MHC-II predicted epitopes showed excellent population coverage against whole world population (100%). Accordingly these epitopes were strongly recommended as promising epitopes vaccine candidate against T cell.

## 5. Conclusion

Developing effective and safe vaccine is strongly needed to prevent the infection and re-emergence of Severe Acute Respiratory Syndrome associated coronavirus (SARS-CoV). Vaccine design using an insilico prediction method is highly appreciated as it selects specific epitopes in protein than conventional peptide vaccine development methods. In this study three epitopes were successfully predicted to interact against B cells. Five epitopes were successfully predicted to interact against T cell with population coverage epitope set of 100%. These epitopes provided excellent results as promising vaccine against SARS-CoV. However in vitro and in vivo trials are required to achieve the effectiveness of these epitopes as vaccine candidates.

## Acknowledgments

Authors would like to thank the staff members of College of Veterinary Medicine, University of Bahri, Sudan for their cooperation and support.

## Competing Interest

The authors declare that they have no competing interests.

## Funding

No funding was received.

## References

[1] Gu J, Gong E, Zhang B, Zheng J, Gao Z, Zhong Y, Zou W, Zhan J, Wang S, Xie Z, Zhuang H, Wu B, Zhong H, Shao H, Fang W, Gao D, Pei F, Li X, He Z, Xu D, Shi X, Anderson VM, Leong AS. Multiple organ infection and the pathogenesis of SARS. *J Exp Med* 2005 Aug 1; 202(3):415-24. Epub 2005 Jul 25.

[2] Fouchier RA, Kuiken T, Schutten M, van Amerongen G, van Doornum GJ, van den Hoogen BG, Peiris M, Lim W, Stöhr K, Osterhaus AD. Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* 2003 May 15; 423(6937):240.

[3] WHO website: Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003 http://www.who.int/csr/sars/country/table2004_04_21/en/ Accessed 2004 April 21.

[4] Orellana C. Laboratory-acquired SARS raises worries on biosafety. *Lancet Infect Dis* 2004; 4(2):64.

[5] Normile D. Infectious diseases. Mounting lab accidents raise SARS fears. *Science* 2004; 304 (5671): 659-61.

[6] Normile D. Infectious diseases. SARS experts want labs to improve safety practices. *Science* 2003; 302 (5642):31.

[7] Lanying Du, Guangyu Zhao, Yongping Lin, Chris CS Chan, Yuxian He, Shibo Jiang, Changyou Wu, Dong-Yan Jin, Kwok-Yung Yuen, Yusen Zhou and Bo-Jian Zheng. Priming with rAAV encoding RBD of SARS-CoV S protein and boosting with RBD-specific peptides for T cell epitopes increase T cell responses and provide protection against SRAS-CoV infection. *Vaccine* 2008 Mar 20; 26(13): 1644-1651.

[8] Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science* 2005; 310 (5748): 676-9.

[9] Tong TR. SARS infection control. *Lancet* 2003 Jul 5; 362(9377):76; author reply 76-7.

[10] Ho PL, Tang XP, Seto WH. SARS: hospital infection control and admission strategies. *Respirology* 2003 Nov; 8 Suppl:S41-5.

[11] Fiona Fleck. SARS virus returns to China as scientists race to find effective vaccine. *Bull World Health Organ* 2004 Feb; 82(2): 152-153.

[12] Marshall E and M Enserink. Caution urged on SARS vaccines. *Science* 2004; 303: 944.

[13] Rota PA, MS Oberste, SS Monroe, WA Nix, R Campagnoli, JP Icenogle, S Penaranda, B Bankamp, K Maher, MH Chen, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 2003; 300:1394.

[14] Marra MA, SJ Jones, CR Astell, RA Holt, A Brooks-Wilson, YS Butterfield, J Khattra, JK Asano, SA Barber, SY Chan, et al. The genome sequence of the SARS-associated coronavirus. *Science* 2003; 300:1399.

[15] Qin E, Q Zhu, M Yu, B Fan, G Chang, B Si, B Yang, W Peng, T Jiang, B Liu, et al. A complete sequence and comparative analysis of a SARS-associated virus (isolate BJ01). *Chin. Sci. Bull* 2003; 48: 941.

[16] He Y, Lu H, Siddiqui P, Zhou Y, Jiang S. Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies. *J Immunol* 2005 Apr 15; 174(8):4908-15.

[17] Babcock GJ, DJ Esshaki, WD Thomas Jr, DM Ambrosino. Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor. *J. Virol* 2004; 78: 4552.

[18] Yee-Joo Tan, Eileen Teng, Shuo Shen, Timothy H. P. Tan, Phuay-Yee Goh, Burtram C. Fielding, Eng-Eong Ooi, Hwee-Cheng Tan, Seng Gee Lim, and Wanjin Hong. A novel Severe Acute Respiratory Syndrome Coronavirus Protein, U274, is transported to the cell surface and undergoes endocytosis. *J Virol* 2004 Jul; 78 (13): 6723-6734.

[19] Saif L J. Coronavirus immunogens. *Vet. Microbiol* 1993; 37: 285.

[20] Wong SK, W Li, MJ Moore, H Choe, M Farzan. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin converting enzyme 2. *J. Biol. Chem* 2004; 279: 3197.

[21] Xiao X, S Chakraborti, AS Dimitrov, K Gramatikoff, DS Dimitrov. The SARS-CoV S glycoprotein: expression and functional characterization. *Biochem. Biophys. Res. Commun* 2003; 312: 1159.

[22] Sui J, W Li, A Murakami, A Tamin, LJ Matthews, SK Wong, MJ Moore, AS Tallarico, M Olurinde, H Choe, et al. Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association. *Proc. Natl. Acad. Sci.* USA 2004; 101: 2536.

[23] Godet M, Grosclaude J, Delmas B, Laude H. Major receptor-binding and neutralization determinants are located within the same domain of the transmissible gastroenteritis virus (coronavirus) spike protein. *J Virol* 1994 Dec;68(12):8008-16.

[24] He Y, Y Zhou, S Liu, Z Kou, W Li, M Farzan, S Jiang. Receptor-binding domain of SARS-CoV spike protein induces highly potent neutralizing antibodies: implication for developing subunit vaccine. *Biochem. Biophys. Res. Commun* 2004; 324: 773

[25] Buchholz UJ, Bukreyev A, Yang L, Lamirande EW, Murphy BR, Subbarao K, et al. Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity. *Proc Natl Acad Sci* USA 2004; 101 (26): 9804-9.

[26] Xiong S, Wang YF, Zhang MY, Liu XJ, Zhang CH, Liu SS, et al. Immunogenicity of SARS inactivated vaccine in BALB/c mice. *Immunol Lett* 2004; 95: 139-43.

[27] He Y, Zhou Y, Siddiqui P, Jiang S. Inactivated SARS-CoV vaccine elicits high titers of spike protein-specific antibodies that block receptor binding and virus entry. *Biochem Biophys Res Commun* 2004; 325: 445-52.

[28] Chou TH, Wang S, Sakhatskyy PV, Mboudoudjeck I, Lawrence JM, Huang S, et al. Epitope mapping and biological function analysis of antibodies produced by immunization of mice with an inactivated Chinese isolate of severe acute respiratory syndrome–associated coronavirus (SARS-CoV). *Virology* 2005; 334: 134-43.

[29] Qu D, Zheng B, Yao X, Guan Y, Yuan ZH, Zhong NS, et al. Intranasal immunization with inactivated SARS-CoV (SARS-associated coronavirus) induced local and serum antibodies in mice. *Vaccine* 2005; 23: 924-31.

[30] Tseng C-T, Sbrana E, Iwata-Yoshikawa N, Newman PC, Garron T, et al. Immunization with SARS coronavirus vaccines leads to pulmonary immunopathology on challenge with the SARS virus. *PLoS ONE* 2012; 7 (4): e35421.

[31] Zhou Z, Post P, Chubet R, Holtz K, McPherson C, et al. A recombinant baculovirus-expressed S glycoprotein vaccine elicits high titers of SARS-associated coronavirus (SARS-CoV) neutralizing antibodies in mice. *Vaccine* 2006; 24: 3624-3631.

[32] Deming D, Sheahan T, Heise M, Yount B, Davis N, et al. Vaccine efficacy in senescent mice challenged with recombinant SARS-CoV bearing epidemic and zoonotic spike variants. *PLoS Medicine* 2006; 3: 2359-2375.

[33] Yasui F, Kai C, Kitabatake M, Inoue S, Yoneda M, et al. Prior immunization with severe acute respiratory syndrome (SARS)-associated coronavirus (SARS-CoV) nucleocapsid protein causes severe pneumonia in mice infected with SARS-CoV. *J Immunol* 2008; 181: 6337-6348.

[34] Kong WP, Xu L, Stadler K, Ulmer JB, Abrignani S, Rappuoli R, et al. Modulation of the immune response to the severe acute respiratory syndrome spike glycoprotein by gene–based and inactivated virus immunization. *J Virol* 2005; 79 (22): 13915-23.

[35] Du L, He Y, Wang Y, Zhang H, Ma S, Wong CK, et al. Recombinant adeno-associated virus expressing the receptor-binding domain of severe acute respiratory syndrome coronavirus S protein elicits neutralizing antibodies: Implication for developing SARS vaccines. *Virology* 2006; 353 (1): 6-16.

[36] Wang D and Lu J. Glycan arrays lead to the discovery of autoimmunogenic activity of SARS-CoV. *Physiol Genomics* 2004; 18: 245-8.

[37] National Center for Biotechnology Information, NCBI. https://www.ncbi.nlm.nih.gov/protein/?term=SARS+spike+protein+S.

[38] Hall T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, Vol. 41 (1999), pp. 95-98. Key: citeulike: 691774.

[39] Immune Epitope Database (IEDB) analysis. http://www.iedb.org/.

[40] Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res* 2006; 2 (1): 2.

[41] Emini E A, J V Hughes, D S Perlow, and J Boger. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol* 1985 Sep; 55(3): 836-839.

[42] Kolaskar AS and Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 1990; 276 (1-2): 172-174.

[43] Immune Epitope DataBase, IEDB. http://tools.iedb.org/bcell/.

[44] Immune Epitope Database (IEDB) MHC-I prediction tool. http://tools.iedb.org/mhci/n.

[45] Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P; et al. Immune epitope database analysis resource. *Nucleic Acids Res* 2012; 40: W525-530

[46] Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 2003; 12: 1007-1017.

[47] Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, et al. NetMHC-3.0: Accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res* 2008; 36: 509-512.

[48] Peters B and Sette A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics* 2005; 6:132.

[49] Abu-haraz AH, Abd-elrahman KA, Ibrahim MS, Hussien WH, Mohammed MS, et al. Multi epitope peptide vaccine prediction against Sudan Ebola virus using immuno-informatics approaches. *Adv Tech Biol Med* 2017; 5: 203.

[50] Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, et al. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome Res* 2008; 4: 2.

[51] Immune Epitope Database (IEDB) MHC-II prediction tool. http://tools.immuneepitope.org/mhcii/

[52] Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol* 2008; 4 (4): e1000048.

[53] Immune Epitope Database (IEDB) population coverage prediction tool. http://tools.iedb.org/population/

[54] Raptor X protein structure prediction server. http://raptorx.uchicago.edu/StructurePrediction/predict/.

[55] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. UCSF Chimera a visualization system for exploratory research and analysis. *J Comput Chem* 2004; 25: 1605-1612.

[56] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015 Jun; 10(6):845-58. Epub 2015 May 7

[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 (5): 404-14. PMID: 17473845.

[58] Yang ZY, WP Kong, Y Huang, A Roberts, BR Murphy, K Subbarao, GJ Nabel. A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. *Nature* 2004; 428: 561.

[59] Bisht H, A Roberts, L Vogel, A Bukreyev, PL Collins, BR Murphy, K Subbarao, B Moss. Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice. *Proc. Natl. Acad. Sci*. USA. 2004; 101: 6641.

[60] Black M, Trent A, Tirrell M, Olive C. Advances in the design and delivery of peptide subunit vaccines with a focus on toll-like receptor agonists. *Expert Rev Vaccines* 2010; 9: 157-173.

[61] Sesardic D. Synthetic peptide vaccines. *J Med Microbiol* 1993; 39: 241-242.

[62] Pang H, Y Liu, X Han, Y Xu, F Jiang, D Wu, X Kong, M Bartlam, Z Rao. Protective humoral responses to severe acute respiratory syndrome-associated coronavirus: implications for the design of an effective protein-based vaccine. *J. Gen. Virol* 2004; 85: 3109-31.