QadirVax-19: A multi epitope-based vaccine against COVID-19

Misha Mukhtar , Muhammad I. Qadir*

Institute of Molecular Biology & Biotechnology, Bahauddin Zakariya University, Multan, Pakistan

**Corresponding author: mrimranqadir@hotmail.com*

Abstract

An outbreak of COVID-19, caused by a novel virus named SARS-corona virus 2 (SARS-CoV-2), has become a global challenge which needs to be addressed immediately. It has high rate of transmission and severity of the disease varying from person to person. Researchers are trying to find effective vaccines and therapeutic targets to control this novel type of coronavirus. In the present study, surface glycoprotein was used to identify B cell and T cell epitopes that have strong immunogenic potential. Highly conserved region of the surface glycoprotein was identified from seven related human coronaviruses. Epitopes and their prominent features were predicted by using immunoinformatics tools. Epitopes which have high binding energy were joined together through linker to form an epitope-based subunit vaccine construct. Molecular docking was performed by MOE to predict the binding energies of construct with B cell receptor, MHC Class I and MHC Class II receptors. The sequence of the multi-epitope construct finally came out as LQYGSFCTQLNRGPGPGTFGAGAALQGPGPGNFTTAPAICGPGPGHWFVTQRNFAA YQYIKWPWYI. It was named as QadirVax-19. Multi-epitope construct was highly antigenic along with proteasomal cleavage sites and full population coverage. The highest binding energies were obtained which shows that the construct has the ability to produce stronger humoral and immune response against structural glycoproteins of SARS-CoV-2. *In-silico* cloning of the construct revealed its stable expression in *E. coli*. Our study suggests that reverse vaccinology approaches give the immunogenic profile of the epitopes which helped us in designing the subunit vaccine against the SARS-corona virus 2.

Keywords: B cell receptors; molecular docking; SARS-CoV-2; surface glycoprotein; T cell receptors.

1. Introduction

The first outbreak of COVID-19 started in Wuhan, Hubei, China in November 2019. On 13 January 2020, it was reported outside of China and then up to March 2, 2020, the virus spread to 67 territories. COVID-19 was declared as "global pandemic" when the virus significantly affected more than 118 countries by March 11, 2020 (Di Gennaro *et al.,* 2020). The actual impact of this outbreak is difficult to measure accurately because it is important to calculate both mild and serious cases (Lipsitch *et al.,* 2020). Age seems to be an important risk factor for the progression of disease. Different geographical regions show variation in the mortality rate. Immune profile of people plays an essential role in susceptibility to the virus (Albitar *et al.,* 2020).

Coronaviruses was first described by Tyrell and Bynoe in 1966. The formal name of the SARS-CoV-2 is given by the International Committee on Taxonomy of Viruses. Corona is a Latin word meaning "Crown". Coronavirus is called so due to the crown like appearance on its surface. Coronaviruses are of four types, alpha, beta, delta, and gamma coronaviruses. Human coronaviruses include two alpha coronaviruses (hCoV-NL63 and hCoV-229E) and five beta coronaviruses (hCoV-OC43, HKU1, MERS-CoV, SARS-CoV and SARS-CoV-2). SARS-CoV-2 is the human beta coronavirus and has sequence similarity with other human coronaviruses. Researchers are trying to find the main host and carriers of coronavirus-2. COVID-19 is likely to have a zoonotic origin (Rothan & Byrareddy, 2020). Homology studies show that genome of SARS-CoV and Bat coronaviruses are 79.5% and 96% similar to SARS-CoV-2 respectively (Wang *et al.,* 2020).

The modes of transmission of coronavirus are respiratory droplets, aerosols, and contaminated areas (Bassetti *et al.,* 2020). Mostly asymptomatic individuals and patients with mild symptoms are unnoticeable due to low sickness of these patients. Mildly symptomatic patients also contain the viral load and spread the same amount of virus. Computed Tomography scan's detail and virologic order of mildly symptomatic patients and asymptomatic patients is not completely described (Li *et al.,* 2020). Symptoms include breathing difficulty, dry coughing, fever, fatigue, nausea and myalgia. Severe complications are pneumonia, severe acute respiratory syndrome, acute liver, cardiac and kidney injury. The onset of disease from time of exposure to symptom is 2-7 days (Wiersinga *et al.,* 2020).

SARS-CoV-2 consists of single positive strand RNA with genome length of 29,881 nucleotides encoding 9860 amino acids. They are spherical enveloped viruses with a diameter of 60-140 nm. Surface (S protein), Enveloped (E protein), Matrix (M protein) and Nucleocapsid (N protein) are the structural proteins (Sarma *et al.* 2020). Envelope glycoprotein of SARS-CoV-2 is made up of 75 amino acids. Virus particles which lack the E protein are not able to infect the host and produce very low titers (Tilocca *et al.,* 2020). Membrane glycoprotein is 230 amino acids long. It maintains the shape and size of virions. All the other structural proteins of virus are assembled through M protein (Mahtarin *et al.,* 2020). The length of nucleocapsid protein is 419 amino acids. N protein is essential for replication and it moves into the host cell along with the RNA (Zeng *et al.,* 2020).

Surface glycoproteins bind with the angiotensin converting enzyme 2 (ACE-2) receptors of the host cells. Serine protease enzyme of host cell facilitates the entry of the virus by activating the S protein. After entry into cell, RNA genome of the virus is replicated and transcribed by the replicase-transcriptase complex. Structural proteins package the virus and release the mature viral particles (Ling *et al.* 2020). Surface proteins are responsible for virus attachment and recognition of receptors. They are the most highly conserved proteins among all other proteins and are considered as the main therapeutic target for producing vaccine. S proteins are 1273 aa long with the size of 180-200 kDa. Structure of S protein consists of N terminus domain, S1 subunit and S2 subunit. S1 subunit has receptor binding domain which cleaves off by furin protein. S2 subunit consists of Fusion peptide (FP), Heptad repeat 1 (HR1), Heptad repeat 2 (HR2), Transmembrane domain (TD) and cytoplasmic domain (CD). RBD region of S1 subunit is highly mutable and is not ideally used. Epitopes derived from the S2 subunit and fusion peptide are used to elicit a strong immune response (Huang *et al.,* 2020). Coronaviruses have the high potential of gene

modifications and possess extra plasticity (Woo *et al.,* 2009). Complete understanding of the SARS-CoV-2 is currently in progress.

Neutralizing antibodies and immune cells combat the viruses. T cells do not recognize the antigen directly. Major histocompatibility complex (MHC) molecules are responsible for presenting the antigens and protein fragments to T cells. T cells also activate B cells for producing antibodies which will target the Surface glycoprotein (Fast & Chen, 2020).

For the control of SARS-CoV-2 infection, vaccination is the most appropriate strategy. Studies reported that recombinant vectors, mRNA containing lipid nanoparticles and subunits of antigenic proteins were main strategies used in the formation of vaccines against COVID-19 (Hu *et al.,* 2020). Some vaccines are based on using the whole attenuated or killed form of the viruses. Due to the contagious nature of the virus and high transmission rate, such types of vaccines demand high biosafety levels. Therefore, it is not feasible for the researchers to develop it for the large-scale population. Virus-based vaccines are extremely immunogenic and have the probability of reversion to more virulent pathogen (Rahman MS *et al.,* 2020). Vaccines were produced using the replicating or non-replicating viral vectors such as adenoviruses. These viral vectors contain the protein part of SARS-CoV-2, mainly spike protein, for expressing into the cell. However, it can cause adverse reactions and mixed immunogenicity. In other cases, pre-existing immunity may present in the body against the carrier viral vectors. In nucleic acid (DNA or RNA) based vaccines, the major concerning issues are the genetic integration of nucleic acid and the stability of RNA. All the mentioned type of vaccines are more time consuming and require high cost (Rahman N *et al.,* 2020). Epitope based vaccines are relatively safe, more stable as compared to these vaccines. They allow to deliver the desired antigenic peptides of specified length (Purcell *et al.,* 2007).

With the revolution of Whole Genome Sequencing, it has become possible to find out the potential antigens as vaccine candidates. Classical vaccinology has certain limitations in term of selecting the protective antigens. The main advantage of reverse vaccinology is that it does not require the maintenance of growth conditions. Reverse vaccinology relies on *In-silico* analysis of genome and proteome. Moreover, the detection of gene expression is not needed. Molecular simulations and dockings are the main *In-silico* tools which give promising results (Mustafa *et al.,* 2020). Many web-based tools are designed on the basis of MHC binding affinity for the prediction of B and T cell epitopes and their population coverage. *In-silico* methods are used as an essential tool in modern era for drug discovery (Qadir *et al.,* 2018). Programs of molecular docking evaluate the binding of ligand with the receptor through the quantification of interactions (Pagadala *et al.,* 2017).

In terms of significance, our novel multi-epitope construct was designed only from the highly conserved and non-mutagenic residues of Surface glycoprotein. Moreover, each epitope in the construct has the ability to interact with both B cells and T cells. The construct is shorter in length which makes it more stable and easier to synthesize. The construct gave high binding energies equivalent to that of long length peptides.

The study focused on the molecular docking of the B cell and T cell epitopes which targeted the structural glycoproteins of SARS-CoV-2 by producing antibodies against it. The present study was based on the hypothesis that conserved region of the surface glycoprotein of SARS-CoV-2 was used to make a novel multi-epitope vaccine construct which will also be effective against new emerging strains of virus.

2. Materials and Methods

2.1 Identification of conserved region

The conserved region of the surface glycoprotein present among all the seven human coronaviruses was obtained through multiple sequence alignment on Clustal Omega. Surface glycoprotein sequences of seven coronaviruses 229E, NL63, OC43, HKU1, MERS-CoV, SARS-CoV and SARS-CoV-2 was retrieved in FASTA format from NCBI database with the accession numbers >AWH62679.1, >APF29063.1, >AWW13575.1, >BBA20983.1, >AMO03401.1, >ARO76382.1, >YP_009724390.1 respectively. Sequence similarity of surface glycoprotein and homology of all the seven coronaviruses was found by BLASTp and phylogenetic tree construction.

2.2 Prediction of B cell and T cell epitopes

B cell and T cell epitopes were predicted from the conserved region of the surface glycoprotein by using online epitope designing servers. Linear epitopes of B cell were determined through ABCpred on the basis of their percentile rank. All the epitopes of length ranging from 10 to 16 were predicted.
T cell epitopes of length 9mer for all the HLA alleles were obtained by using the NetMHC server (Jensen *et al.,* 2018). NetMHC 4.0 designed the cytotoxic T cell epitopes while NetMHC 2.3 was used for predicting helper T cell epitopes. Threshold value of percentile rank for the strong binders was set as less than 0.5.

2.3 Analysis of predicted epitopes

Antigenicity and Immunogenicity of the predicted epitopes was determined through VaxiJen 2.0 and Immune Epitope Database (IEDB) respectively (Doytchinova & Flower 2007) (Dhanda *et al.* 2018). Toxicity of each epitope was obtained by ToxinPred server (Gupta *et al.,* 2013). Epitopes having high antigenic value and immunogenicity score were further selected for molecular docking. Binding energies of the best predicted epitopes were calculated by docking on MOE. PDB structures of the epitopes were obtained from PEPFOLD server. Population coverage analysis of the epitopes of construct was performed for both MHC Class I and MHC Class II alleles by Population Coverage IEDB tool (Bui *et al.,* 2006).

2.4 Designing and Analysis of Multi-epitope construct

Best B cell and T cell epitopes having high binding energies were joined together through linkers to form a Multi-epitope construct. The purpose of the linker is to provide the flexibility and stability in the overall conformation of construct. Linkers are used to join the epitopes, thereby, distinguish their functional domains (Dong *et al.,* 2020). All the physicochemical properties, molecular weight, isoelectric point, instability index and extinction coefficient of the construct were measured by Expasy Protparam tool. Antigenicity and allergenicity was obtained by VaxiJen and AllerTop v. 2.0 respectively. NetChop IEDB tool was used for determining the proteasomal cleavage sites present within the construct (Nielsen *et al.,* 2005). Expasy Peptide cutter found out the positions of enzymatic cleavage sites of the construct as well as the enzymes which do not cut (Maillet 2020).

2.5 Structure predictions for Molecular docking

Secondary structure of the construct was predicted by SOPMA server. RaptorX structure prediction server was used for designing the PDB structure of the construct. Any error present in the 3D structure of the construct was further detected and validated by ProSA-web tool. Molecular docking of the construct with the B cell receptors and MHC molecules was performed by MOE (Vilar *et al.,* 2008). Interaction of the construct with the receptors was visualized on Pymol. Java Codon Adaption tool (JCat) carried out the reverse translation and codon optimization of the construct.

2.6 Physicochemical Properties of construct

All the physicochemical properties of multi-epitope construct were computed by ProtParam tool-Expasy and PepCalc tool. Amino acid sequence of the peptide was entered in both tools. These tools provided the results for different properties which included pI, molecular weight, amino acid and atomic composition, half-life, extinction coefficient, instability index, aliphatic index, GRAVY, and water solubility. ProtParam tool performed all the calculations on the basis of amino acid composition

2.7 *In-Silico* Cloning

Reverse translation of the multi-epitope based vaccine was carried out by Java Codon Adaption tool (JCat). Protein sequence of the construct was entered for codon optimization. SnapGene was used for restriction mapping of the construct and *In-silico* cloning.

3. Results

3.1 Conserved region of surface glycoprotein

From the position of amino acid 714 to 1240 in the surface glycoprotein, the region was shown as conserved among the human coronaviruses. This region encodes for the HR1, HR2 and fusion peptide of the S protein. Phylogenetic relationship analysis shpwed that S2 subunit of SARS-CoV-2 is 99% similar to bat SARS-like coronaviruses (ZXC21 and ZC45) and human SARS-CoV. pBLAST of S glycoproteins of human coronaviruses revealed that SARS-CoV-2 is 74.28% and 35.19% similar to SARS-CoV and MERS-CoV respectively. Human coronaviruses OC43 and HKU1 have percent identity of 37.63% and 35.29% with SARS-CoV-2. While the percent identity of S protein of human coronaviruses 229E and NL63 with SARS-CoV-2 is 31.32% and 30.57% respectively.

3.2 B cell and T cell epitopes

Epitopes having antigenic value greater than 0.5 and toxicity value less than 0.0 were selected. Best predicted epitopes were shown as antigenic, highly immunogenic and non-toxin. Among many predicted epitopes ten B cell epitopes, five helper T cell epitopes and five cytotoxic T cell epitopes were subjected to molecular docking. All the epitopes showed high binding energies with the receptors ranging from 8 to 10.

3.3 Multi-epitope Construct

One B cell epitope (LQYGSFCTQLNR) having the highest binding affinity among all the epitopes joined with two helper T cell epitopes (TFGAGAALQ, NFTTAPAIC) and two cytotoxic T cell epitopes (HWFVTQRNF, QYIKWPWYI) through linkers. The multi-epitope construct is the pentavalent molecules with the molecular weight of 7061.98 Daltons. GPGPG linker used to join the B cell epitope with the helper T cell epitopes. It also combined two helper T cell epitopes with each other and with cytotoxic T cell epitopes. AAY linker was present among the two cytotoxic T cell epitopes for joining them (Tahir ul Qamar *et al.,* 2020). The sequence of the multi-epitope construct finally came out as LQYGSFCTQLNR**GPGPG**TFGAGAALQ**GPGPG**NFTTAPAIC**GPGPG**HWFVTQRNF**A AY**QYIKWPWYI. It was named as QadirVax-19.

Each epitope in the construct has the cleavage sites for proteasome. Amino acid which has the value greater than threshold value of 0.5 represented the cleavage site. All the epitopes of multi-epitope construct showed the world population coverage of 100% in case of both MHC class I and MHC class II receptors. Protein structure validation tool calculated the ProSA-Z score as -4.8. It showed that the structure of multi-epitope vaccine was similar to the native protein structure.

3.4 Molecular docking

Molecular docking structures of the construct with their targets are given in Figure 1. The ligand was effectively in bonded form with the receptor. Predicted binding energies values given by MOE upon docking were shown to be greater with all the three receptors (Table 1). Multi-epitope construct gave higher binding energy with MHC class II receptor as compared to others. The binding energy with MHC class I receptor is more than that of B cell receptor. Overall, all the binding energies were considered high with the receptors. The higher binding energy and the interaction with all the receptor made the multi-epitope construct a good subunit vaccine candidate for eliciting a stronger immune response.

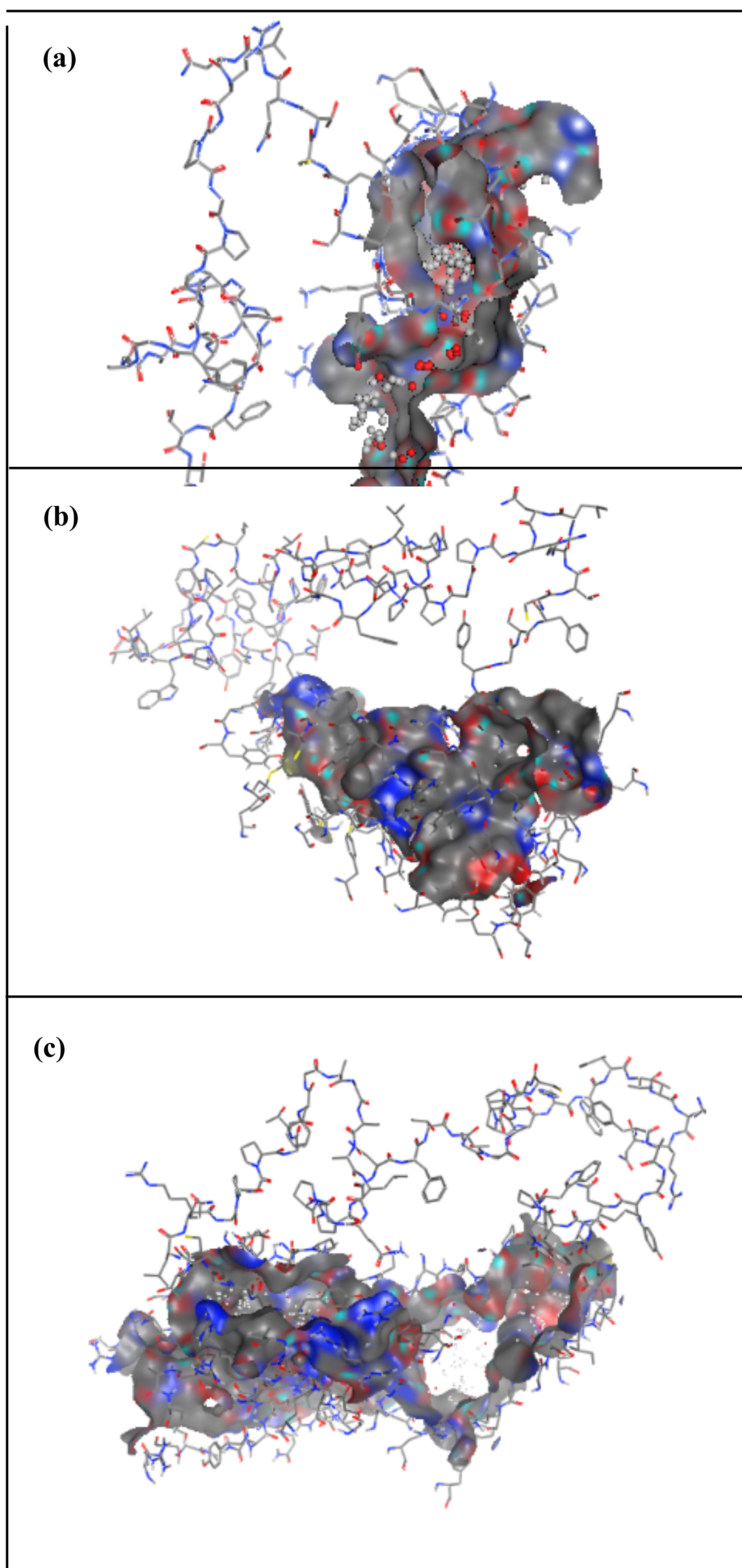

Fig. 1. Docking of the multi-epitope construct with **(a)** B cell receptor **(b)** MHC Class I receptor **(c)** MHC Class II receptor

Table 1. Binding energies of the construct by MOE

	B cell receptor	MHC Class I	MHC Class II
RSEQ	1	1	1
MSEQ	1	1	1
S	-11.8771	-15.2431	-22.275
RMSD_REFINE	5.553171	4.686814	5.042731
E_CONF	-406.839	-367.689	-366.265
E_PLACE	647.1857	467.0003	426.3775
E_SCORE1	49.96589	54.59181	55.68112
E_REFINE	-50.9681	-64.9148	-95.1125
E_SCORE2	-11.8771	-15.2431	-22.275

3.5 Physicochemical Properties

Physicochemical properties and other parameters for the construct were analysed. Multi-epitope construct in the present study was predicted as the probable antigen and non-allergen. Molecular weight, pI, net charge, solubility, and other properties predicted through bioinformatics tools are shown in Table 2. The extinction coefficient of the construct is the light absorbed by it at a specific wavelength. The half-life is the time taken by the construct after its synthesis to degrade in the cell. The prediction for half-life estimation is not applicable to such proteins which have modified N-terminal ends. The lower value of instability index indicated the construct as stable. Aliphatic index is the measure of the relative volume taken by the aliphatic amino acids present in the construct. This characteristic is useful for globular proteins in terms of increasing their thermostability. Hydropathicity of all the amino acids is added and then divided by the total number of residues in the construct. The value that obtained is referred as GRAVY. It represented the hydrophilic and hydrophobic properties of construct. Isoleucine and valine have the high hydrophobicity of 4.5 and 4.2 respectively. The antigenicity score of the construct was 0.5045. The theoretical isoelectric point of the construct was 9.24 with the net charge of +3 at pH 7. The extinction coefficient was predicted as 22585 M^{-1} cm^{-1}. The lower value of instability index 13.95 showed the construct as stable. The construct has the poor water solubility as predicted by the Protparam tool.

Table 2. Physicochemical properties of construct

Sr. No.	Properties of construct	Predicted Values
1	Antigenicity	0.5045
2	Allergenicity	Non-Allergen
3	Total number of amino acids	66
4	Molecular Weight (Da)	7061.98
5	Theoretical isoelectric point	9.24
6	Net Charge at pH 7	3
7	Total Atoms	969
8	Total number of negatively charged residues (Asp + Glu)	0
9	Total number of positively charged residues (Arg + Lys)	3
10	Extinction Coefficient (M^{-1} cm^{-1})	22585
11	Abs 0.1% (280 nm)	3.198
12	Half-life (mammalian reticulocytes, in vitro)	5.5 hours
13	Half-life (yeast)	3 min
14	Half-life (*E. coli*)	2 min
15	Instability Index	13.95
16	Aliphatic Index	50.45
17	Grand average of hydropathicity (GRAVY)	-0.200
18	Water Solubility	Poor

Different types of enzymes and proteases cleaved the construct at the predicted positions of cleavage sites (Table 3). These enzymes which do not cut the sequence are given in table also. The table shows that the construct will be destroyed by given orally and it will be stable when given directly to the blood circulation.

Table 3. Cleavage of the construct by enzymes and their cleavage sites by Expasy Peptide Cutter

Name of enzymes	Number of cleavages	Positions of cleavage sites in construct
Chymotrypsin -high specificity	11	3 6 19 33 47 48 54 57 59 64 65
Chymotrypsin-low specificity	14	1 3 6 10 19 25 33 47 48 54 57 59 64 65
Clostripain	2	12 52
Iodosobenzoic acid	3	47 62 64
LysC	1	61
LysN	1	60
NTCB (2-nitro-thiocyanobenzoic acid)	2	6 39
Pepsin (pH 1.3)	13	1 5 6 9 10 18 19 24 25 32 33 47 53
Pepsin (pH>2)	22	1 2 3 5 6 9 10 18 1924 25 32 33 46 47 53 56 57 58 59 62 65
Proteinase K	31	1 3 6 8 10 18 19 21 23 24 25 33 34 35 36 38 39 47 48 49 50 54 55 56 57 59 60 62 64 65 66
Thermolysin	17	5 9 18 20 22 23 24 32 37 38 47 48 53 54 55 59 65
Trypsin	3	12 52 61
Enzymes which do not cut		
Asp-N endopeptidase, Asp-N endopeptidase + N-terminal Glu, CNBr, Caspases 1-10, Enterokinase, Factor Xa, Formic acid, Glutamyl endopeptidase, Granzyme B, Hydroxylamine, Prolineendopeptidase, Staphylococcal peptidase I, Thrombin		

All the epitopes of multi-epitope construct showed the world population coverage of 100% in case of both MHC class I and MHC class II receptors (Figure 2).

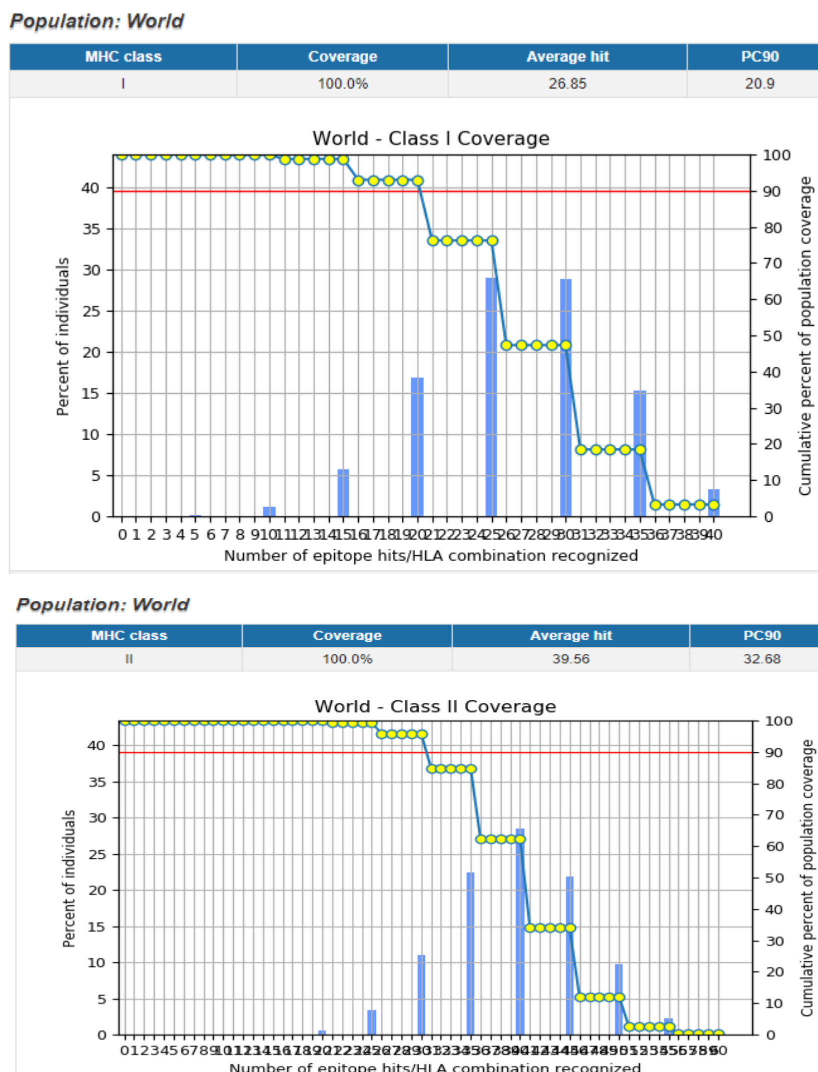

Fig. 2. Population coverage analysis of construct epitopes for MHC Class I and MHC Class II in the World population

3.6 *In-Silico* Cloning

In-Silico Cloning results are given in Figure 3. The length of the multi-epitope construct is 198 nucleotides after codon optimization. The optimum value of Codon Adaption Index (CAI) is 0.8 to 1.0 while optimum GC content value must be 30-70%. The CAI value and GC content for our multi-epitope construct was 1.0 and 61.1% respectively. The values in this range show the good expression of the construct in the *E. coli*.

The aim of the *In-Silico* cloning is to determine that whether the construct is suitable for expression into prokaryotes. pUC19 vector was selected from the SnapGene database because of its easy blue-white screening method of selection for the recombinants. Restriction site for Xba1 and BamH1 were placed at the first and last position at the construct by keeping in view the position and direction of these sites on vector. The orientation of our construct was in accordance with the direction of vector. Our construct was inserted into the lacZ gene of

the vector so that the recombinants will be selected due to the disruption of lacZ gene. Recombinants will produce white colonies on media plate containing X-Gal and IPTG inducer. Our construct will express under the lacZ promoter and terminator. The level of expression will be checked by designing the primers on SnapGene for PCR.

The length of the multi-epitope construct was 198 nucleotides after codon optimization. The optimum value of Codon Adaption Index (CAI) should be 0.8 to 1.0 while optimum GC content value must be 30-70%. The CAI value and GC content for our multi-epitope construct was 1.0 and 61.1% respectively. The values in this range show the good expression of the construct in the *E. coli* by *in-silico* cloning. The higher binding energies values predicted by MOE made the multi-epitope construct a good subunit epitope-based vaccine candidate.

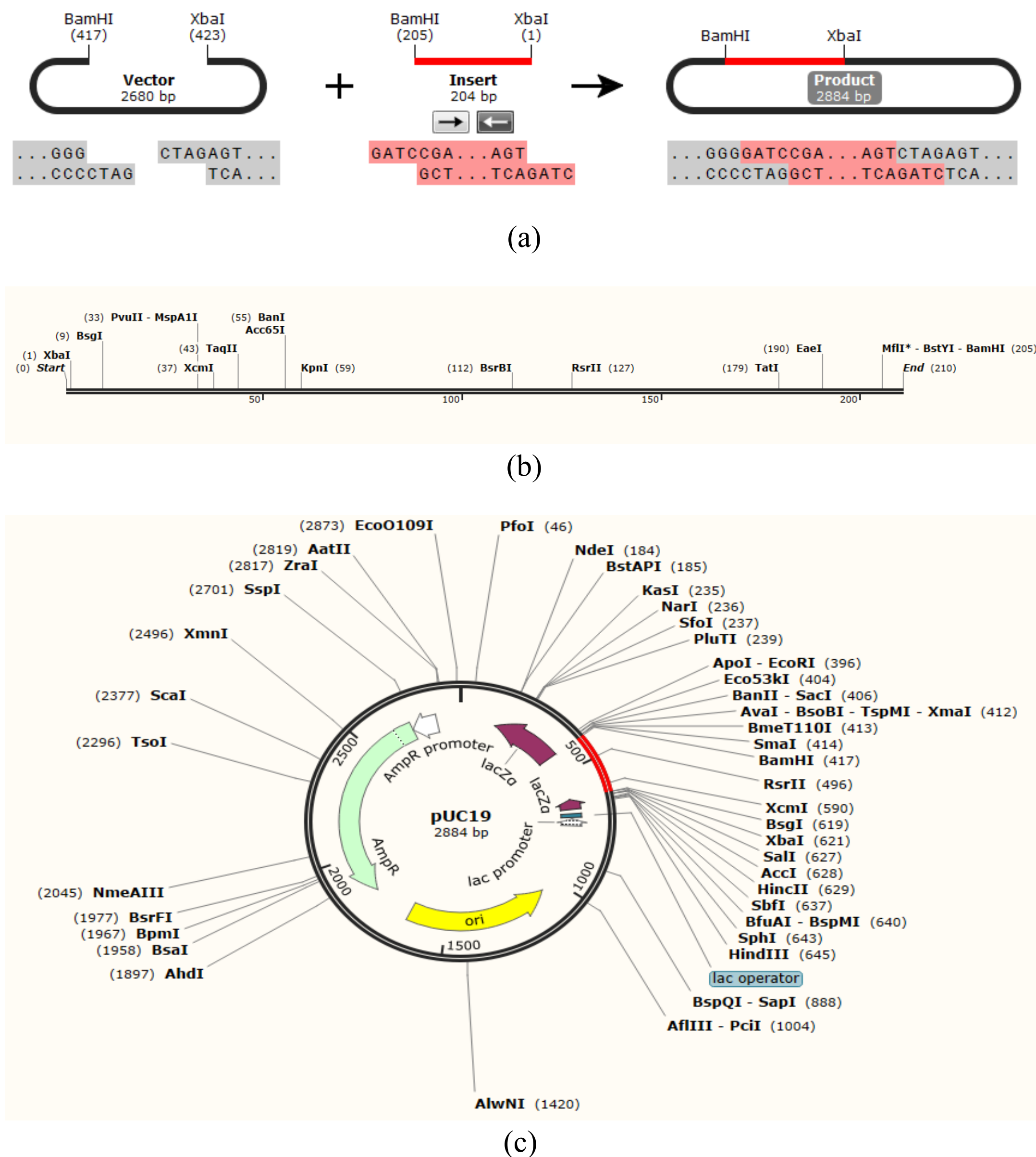

Fig. 3. *In-Silico* Cloning of multi-epitope vaccine construct into *E.coli* vector by SnapGene (a) Restriction map of multi-epitope construct (b) Insertion of vaccine construct into vector (c) Cloned product of pUC19 vector and multi-epitope insert

4. Discussion

Researchers are trying at their full potential to combat the deadly infectious virus, the Novel Coronavirus-2019 (Qadir, 2020). In order to understand the pathophysiology and transmission of virus, data is gathering by the researchers. The whole genome of the virus was sequenced and reported on January 29, 2020. Different types of vaccines with various mechanisms are currently in their trial stages. The work is being done for the construction of peptide-based subunit type of vaccines by many research groups (Sarkar *et al.,* 2021). Epitope-based subunit vaccines are preferred over the traditional vaccines in terms of better immunogenic control, experimental handling, and reproducibility. Peptides are used because of their chemical stability and ease of production (Patronov & Doytchinova, 2013). Different groups of workers are struggling to produce SARS-CoV-2 vaccines on the basis of epitope-based vaccine strategy taking the benefit from the genome sequence of the virus. However, none of these epitope-based vaccines are available on commercial level due to its low immunogenicity. Another reason is the immediate need of vaccine by massive population on large scale.

One of the major impediments of the COVID-19 vaccine is the emergence of new variants, not fully susceptible to the antibodies. Our study designed a highly antigenic epitope-based subunit vaccine against SARS-CoV-2 and its strains by using immunoinformatic tools and software. The conserved amino acids that present among all the seven human coronaviruses were retrieved for finding the best antigenic epitopes. S2 subunit of surface glycoprotein is less likely to undergo mutations as compared to S1 subunit. The present study identified and selected the epitopes from S2 subunit encoding HR1 and HR2 domains. Epitopes were sorted on the basis of their percentile rank comparing to threshold value. Antigenic epitopes have the value greater than 0.5. Toxicity and immunogenicity of all the epitopes were also predicted. Binding energies of ten B cell epitopes and ten T cell epitopes were *in-silico* predicted by using MOE 2014. Multi-epitope vaccine construct was designed by combining the five epitopes using GPGPG and AAY linkers. Physicochemical properties, proteasomal and enzymatic cleavage, allergenicity, antigenicity and population coverage of the vaccine construct were interpreted. The current study clarified the binding capability of the construct with the receptors by predicting its binding energies. The binding energy of the construct come out as -11.8771, -15.2431 and -22.275 with B cell receptor, MHC class I receptor and MHC class II receptor respectively.

Many different epitope-based vaccines were forecasted and determined by taking other structural proteins (nucleocapsid, membrane and envelope) of SARS-CoV-2. Chen at al identified the T cell epitopes from the nucleocapsid protein of SARS-CoV-2 (Chen *et al.,* 2020). Different studies were also performed on immunodominant epitopes and discontinuous B cell epitopes. TLR receptors are also utilizing for epitope-based vaccines in molecular simulation and docking in the literature (Banerjee *et al.,* 2020).

5. Conclusion

Different computational methods have paved the way and made it easier to design epitope-based subunit vaccines which are safe, cheap and less time consuming. These vaccines are

helpful in eliciting the stronger humoral and cellular immune responses which in turn fight with the deadly virus through producing antibodies against it. To conclude, Multi Epitope-Based Subunit vaccine against COVID-19 was designed through reverse vaccinology approaches and bioinformatics tools.

List of abbreviations

COVID-19: Coronavirus disease-2019; SARS: Severe Acute Respiratory Syndrome; CoV-2: Coronavirus-2; MERS: Middle East Respiratory Syndrome; ACE2: Angiotensin Converting Enzyme 2; S: Surface Glycoprotein; MHC: Major Histocompatibility Complex; HLA: Human Leukocyte Antigen; MOE: Molecular Operating Environment; IEDB: Immune Epitope Database; PDB: Protein Data Bank

ACKNOWLEDGEMENTS

Authors are thankful to Dr. Syed Aun Muhammad for his contribution in completion of this work.

References

Albitar O, Ballouze R, Ooi JP, Ghadzi SMS (2020). Risk factors for mortality among COVID-19 patients. Diabetes Res Clin Pract 166:108293. doi: 10.1016/j.diabres.2020.108293

Banerjee S, Majumder K, Gutierrez GJ, Gupta D, Mittal B (2020). Immuno-informatics approach for multi-epitope vaccine designing against SARS-CoV-2. BioRxiv 39:1087-1095. doi: 10.1016/j.vaccine.2021.01.011

Bassetti M, Vena A, Giacobbe DR (2020). The novel Chinese coronavirus (2019-nCoV) infections: Challenges for fighting the storm. Eur J Clin Invest 50:e13209. doi: 10.1111/eci.13209

Bui H-H, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A (2006). Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinform 7:1-5. doi : 10.1186/1471-2105-7-153

Chen H-Z, Tang L-L, Yu X-L, Zhou J, Chang Y-F, Wu X (2020). Bioinformatics analysis of epitope-based vaccine design against the novel SARS-CoV-2. Infect Dis Poverty 9:1-10. doi: 10.1186/s40249-020-00713-3

Di Gennaro F, Pizzol D, Marotta C, Antunes M, Racalbuto V, Veronese N, Smith L (2020). Coronavirus diseases (COVID-19) current status and future perspectives: a narrative review. Int J Environ Res Public Health 17:2690. doi: 10.3390/ijerph17082690

Dong R, Chu Z, Yu F, Zha Y (2020). Contriving multi-epitope subunit of vaccine for COVID-19: immunoinformatics approaches. Front Immunol 1784. doi: 10.3389/fimmu.2020.01784

Doytchinova IA, Flower DR (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinform 8:1-7. doi : 10.1186/1471-2105-8-4

Fast E, Chen B (2020) Potential T-cell and B-cell Epitopes of 2019-nCoV. J BioRxiv 10.1101/2020.02.19.955484

Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP, Consortium OSDD (2013). In silico approach for predicting toxicity of peptides and proteins. PloS One 8:e73957. doi : 10.1371/journal.pone.0073957

Hu B, Guo H, Zhou P, Shi Z-L (2020). Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 1-14. doi: 10.1038/s41579-020-00459-7

Huang Y, Yang C, Xu X-f, Xu W, Liu S-w (2020). Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. J Acta Pharmacologica Sinica 41:1-9. 10.1038/s41401-020-0485-4

Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, Sette A, Peters B, Nielsen M (2018). Improved methods for predicting peptide binding affinity to MHC class II molecules. J Immunol 154:394-406. doi : 10.1111/imm.12889

Li Y, Shi J, Xia J, Duan J, Chen L, Yu X, Lan W, Ma Q, Wu X, Yuan Y (2020). Asymptomatic and symptomatic patients with non-severe coronavirus disease (COVID-19) have similar clinical features and virological courses: a retrospective single center study. Front Microbiol 11:1570. doi: 10.3389/fmicb.2020.01570

Ling Y, Xu SB, Lin YX, Tian D, Zhu ZQ, Dai FH, Wu F, Song ZG, Huang W, Chen J, Hu BJ (2020). Persistence and clearance of viral RNA in 2019 novel coronavirus disease rehabilitation patients. Chinese M J. 133(09):1039-43.

Lipsitch M, Swerdlow DL, Finelli L (2020). Defining the epidemiology of Covid-19—studies needed. N Engl J Med 382:1194-1196. doi: 10.1056/NEJMp2002125

Mahtarin R, Islam S, Islam MJ, Ullah MO, Ali MA, Halim MA (2020). Structure and dynamics of membrane protein in SARS-CoV-2. J Biomol Struct Dyn 1-14. doi : 10.1080/07391102.2020.1861983

Maillet N (2020). Rapid Peptides Generator: fast and efficient in silico protein digestion. NAR genom bioinform 2:lqz004. doi : 10.1093/nargab/lqz004

Mustafa MI, Abdelmoneim AH, Makhawi AM (2020). Immunoinformatics Patterns and Characteristic of Epitope-Based Peptide Vaccine candidates against COVID-19. AIJR Preprints doi: 10.21467/preprints.164

Nielsen M, Lundegaard C, Lund O, Keşmir C (2005). The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. Immunogenetics 57:33-41. doi : 10.1007/s00251-005-0781-7

Pagadala NS, Syed K, Tuszynski J (2017). Software for molecular docking: a review. Biophys Rev 9:91-102. doi : 10.1007/s12551-016-0247-1

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

Purcell AW, McCluskey J, Rossjohn J (2007). More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov 6:404-414. doi: 10.1038/nrd2224

Qadir MI, Mushtaq H, Mobeen T (2018). *In-silico* study of potential carboxylic acid derivatives as D-glutamate ligase inhibitors in *Salmonella typhi*. Kuwait Journal of Science 45:100-107.

Qadir MI (2020). QadirVID-19: A fusion inhibitor against Novel Coronavirus-2019 for specific management of COVID-19. Pak J Pharm Sci 33(5): 2187-2191.

Rahman MS, Hoque MN, Islam MR, Akter S, Alam ARU, Siddique MA, Saha O, Rahaman MM, Sultana M, Crandall KA (2020). Epitope-based chimeric peptide vaccine design against S, M and E proteins of SARS-CoV-2, the etiologic agent of COVID-19 pandemic: an in silico approach. PeerJ 8:e9572. doi: 10.1101/2020.03.30.015164

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

Rothan HA, Byrareddy SN (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun 109:102433. doi: 10.1016/j.jaut.2020.102433

Sarkar B, Ullah A, Araf Y, Islam NN, Zohora US (2021). Immunoinformatics-guided Designing and In Silico Analysis of Epitope-based Polyvalent Vaccines against Multiple Strains of Human Coronavirus (HCoVs). Expert Rev Vaccines doi: 10.1080/14760584.2021.1874925

Sarma P, Shekhar N, Prajapat M, Avti P, Kaur H, Kumar S, Singh S, Kumar H, Prakash A, Dhibar DP, Medhi B (2021). *In-silico* homology assisted identification of

inhibitor of RNA binding against 2019-nCoV N-protein (N terminal domain). J Biomolecular Structure & Dynamics. 20;39(8):2724-32.

Tahir ul Qamar M, Rehman A, Tusleem K, Ashfaq UA, Qasim M, Zhu X, Fatima I, Shahid F, Chen L-L (2020). Designing of a next generation multiepitope based vaccine (MEV) against SARS-COV-2: Immunoinformatics and in silico approaches. PloS One 15:e0244176. doi: 10.1101/2020.02.28.970343

Tilocca B, Soggiu A, Sanguinetti M, Babini G, De Maio F, Britti D, Zecconi A, Bonizzi L, Urbani A, Roncada P (2020). Immunoinformatic analysis of the SARS-CoV-2 envelope protein as a strategy to assess cross-protection against COVID-19. Microbes Infect 22:182-187. doi : 10.1016/j.micinf.2020.05.013

Vilar S, Cozza G, Moro S (2008). Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery. Curr Top Med Chem 8:1555-1572. doi : 10.2174/156802608786786624

Wang L-s, Wang Y-r, Ye D-w, Liu Q-q (2020). A review of the 2019 Novel Coronavirus (COVID-19) based on current evidence. J International Journal of Antimicrobial Agents 55:105948. 10.1016/j.ijantimicag.2020.105948

Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC (2020). Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. JAMA 324:782-793. 10.1001/jama.2020.12839

Woo PC, Lau SK, Huang Y, Yuen K-Y (2009). Coronavirus diversity, phylogeny and interspecies jumping. Exp Biol Med 234:1117-1127. doi : 10.3181/0903-MR-94

Zeng W, Liu G, Ma H, Zhao D, Yang Y, Liu M, Mohammed A, Zhao C, Yang Y, Xie J (2020). Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochem Biophys Res Commun 527:618-623. doi : 10.1016%2Fj.bbrc.2020.04.136

Submitted: 27/12/2021
Revised: 22/04/2022
Accepted: 25/05/2022
DOI: 10.48129/kjs.17575