

Subunit Vaccine against Sars-Cov-2 using CTL and HTL epitopes.

Laila Zahra, f2018108014@umt.edu.pk

Yaser Daanial Khan, yaser.khan@umt.edu.pk

University of Management and Technology, CII Johar Town, Lahore, Pakistan.

Abstract:

Severe Acute Respiratory Syndrome coronavirus 2 which is widely known as Sars-Cov-2 is a deadly virus that is the main cause of Coronavirus Disease (COVID-19). This plague affects the human immune system badly and adversely affects the human body. This disease emerged from Wuhan, China, and spread all over the world in a very short period. The World Health Organization (WHO) has warned the world about its dire consequences and directed all the countries to take strict precautionary measures and still, there are lots of things that need to be done. This study focuses on the in silico methods which use immunoinformatic approaches to build epitope-based subunit vaccine for SARS-COV-2 that is used to produce several positive immune responses within the host cell. Various B-cells, T_c cells, and T_h cells containing different epitopes are considered for the inhibition of spike of SARS-COV-2. By following different approaches, eventually, the structure of the proposed vaccine consists of T_c , T_h cells, and B-cells joined by different linkers was designed. Currently having B-cell as well as IFN- γ made epitopes confirm the humoral and cell-mediated immune response developed by the proposed vaccine. An online server, PSIPRED is used to develop the model of vaccine. 15 antigenic epitopes were chosen from Spike protein to develop an effective vaccine. This vaccine was antiviral, non-allergic, and less toxic. The sequence of vaccine structure was then validated by different computational methods like Molecular Docking, RMSD, RMSF, and Molecular Dynamic Simulation. Java Codon adaption tool also known as JCat is used for maximal optimization of vaccine expression with vector.

Keywords: Spike glycoprotein; peptide vaccine; Reverse Vaccinology approach; novel coronavirus (2019-nCov).

1. Introduction:

SARS-COV-2 has become a constant challenge to be dealt with for the humans all over the world. COVID-19 which is a plague, comes from a virus named SARS-COV-2 [1]. The word “corona” has various transformed types. The term corona means crown because of the shape of this virus is similar to a crown. Many virologists associated “the specific, edge of estimation” on the outer surface of the virus with the solar corona. SARS-CoV-2 structure consist of single-stranded RNA, with 120 nanometers in diameter, and the genome size of this virus is about 29.9 KBS. Moreover structural proteins and nonstructural proteins are encrypted by orf1ab. Among these designed structures there are some additional proteins which include orf3a, orf7a, orf7b, orf8, orf10.

Structural proteins also comprises of Spike protein (S-Protein), Membrane protein (M-Protein), Envelope protein (E-Protein), and NUCLEOCAPSID protein (N-Protein) [2]. These 4 mentioned proteins play vital role in the mechanism of virus to gain entrance into the human cell and affecting ribosome. There are about 40 different species of Sars (see Appendix 1) and they are generally found in both humans and animals including birds [3]. This virus affects birds especially it primarily targets bats in the recent times, and from there it spreads into human's body. The SARS-COV-2 is believed to have started in bats, then spread into snakes and pangolins and later on traced in humans, due to large amount of meat consumptions in Chinese markets [4]. Five variants of coronavirus are already known since 2003. 27 cases of pneumonia have been reported in Wuhan city, China. On December 31, 2019 [5]. The virus spreads through human to human by different means. Infected human shows symptoms of illness, flu, breathing problems and sometimes lack of taste and smell is reported by virus effected person. [6]. Individuals who have previous medical history and humans with age of more than 60 years are primary target of this disease and they have more chance to contract the virus resulting in sickness and death [7]. Major signs of corona affected patient includes tiredness, dry cough, and fever. Some Other indications are windedness, a pain in a body, sore throat, and a runny nose. Many studies indicate that minor spread of this disease was not considered serious which later resulted huge broke out and thousands of deaths in Iran, Italy, USA, South Korea, Japan, and Pakistan [8].

According to WHO some steps for prevention of Sars-Cov-2 include regular hands washing with soap, water, or sanitizer, keeping distance of 1 meter between one person to other, avoid touching mouth and nose, while sneezing and coughing cover your mouth, in case of feeling unwell stay at home, abstaining yourself from smoking and different exercises that hinder the lungs, avoiding unnecessary travel and refraining from gathering. The increasing rate of this virus is caused due to various health and financial problems. Extra care as well as proven vaccination is required in order to overcome this disease and control spread of this virus [9].

In current situation, the only way to stop the spread of this disease is preparation of its vaccine. The vaccine can play a significant role to help the immune system of body to fight against viral pathogens like bacteria and viruses, and then protects the body from the disease [10]. More than twenty-five (25) dangerous and life threatening diseases are controlled through the vaccine including polio, tetanus, measles, diphtheria, influenza, meningitis, and typhoid. Currently, children receive their vaccine on time but still nearly twenty (20) million worldwide population miss-out, suffer ill-health and death on account of attacks by different severe diseases. Edward Jenner developed a technique for the cure of smallpox in 1796[11]. The other suggested vaccines were established early in the 20th century. These vaccines are used to fight against pertussis (1914), diphtheria (1926), and tetanus (1938). In Pakistan different companies that produce vaccines include Anson vaccine & PHAMA Islamabad, which develop a vaccine against Tetanus, Hepatitis B, and typhoid. BF bioscience Ltd Lahore and Get pharms, Karachi, PHERMEDIC laboratories Lahore and WERRICK pharmaceuticals, Islamabad also develop vaccines against Hepatitis B & C[12].

Nowadays due to the critical situation of COVID-19, our main focus was on developing under reverse volcanology by using different bioinformatics tools to understand the structure of Virus,

bacteria, cancer cell, Parasites, or allergens that could induce an immune response against a specific disease [13]. In reverse volcanology we identified a protein sequence from a pathogen and then these proteins pathogen have been evaluated using different parameters to select the antigenic portion and as a result this protein produces an immune response. Further, this segment is used to design vaccine for Sars-Cov-2[14].

In this study, structural protein (spike) S protein was used for the experimentation. Sequence of spike protein or spike trimmers of SARS-COV-2, which ranged from a class I virus containing mixture of HIV glycoprotein 160 (ENV), haemagglutinin (HA), Paramyxovirus F, and Ebola virus glycoprotein.[15]. The Spike protein creates monoclonal antibodies (mAbs) with the solid defusing to help in the development of cure of COVID-19. Spike glycoprotein contains seventy-six% (76%) amino acid sequence individually with SARS-COV-2 and 80% identity with the bat SAR-COV-2 ZXC21'S and ZC45'S glycoprotein. This study is carried out according to the secondary structure and we have predicted an epitope against SARS-COV-2 which can further be used as possible vaccine candidate[16]. To complete this, physically tested sequence data was taken from the UNIPROT database. An IEDB analysis tool was used for the prediction of B-cell and T-cell epitopes in order to find the best CTL and HTL epitope from a sequence of the spike protein of SARS-COV-2. In this research, online server CTL epitope NETCTL 1.2 s has been used in order to check the antigenicity of epitope VAXIJEN v2.0 server has been used, on the other hand in order to find the Non-allergen epitope, ALGPRED server has been used. Further, in order to check toxicity, TOXINPRED server was used, to check the immunogenicity, IEDB tool was used and the positive epitope was selected. By using the IEDB MHC-II module HTL, epitope has been selected in order to check the antigenicity of that epitope and VAXIJEN v2.0 server was used. During our study we used ALGPRED online server to check the Non-allergenicity, Overlapping was removed in these epitopes manually by using Notepad++. Online server TOXINPRED was used to check the toxicity of that epitope, to predict interferon-gamma, an online server IFNEPITOPE was used and positive epitope has been selected. AAY and GPGPG linker is used to construct the vaccine sequence. In this study after linking the positive epitope, secondary structure was produced using PSIPRED online server. The predicted epitope consistency is found to be a suitable vaccine candidate [17]. 15 antigenic epitopes were chosen from Spike protein to develop an effective vaccine. This vaccine was antiviral, non-allergic, and less toxic. The sequence of vaccine structure was then validated by different computational methods like Molecular Docking, RMSD, RMSF, and Molecular Dynamic Simulation. Java Codon adaption tool also known as JCat is used for maximal optimization of vaccine expression with vector.

2. Methods

The approaches used for peptide vaccine growth for Spike glycoprotein are discussed below in diagram 1. In the first step of the below diagram, database of protein has been illustrated. The second step indicates the sequence of spike glycoprotein. In the third step, by using the online servers B-cell and T-cell have been predicted. In addition to that, CTL and HTL were joined together by using AAY and GPGPG linker in the next step, while physicochemical properties have been checked by using EXPASY Tool PROTPARMA. At the end secondary structure has been built by using PSIPRED server.

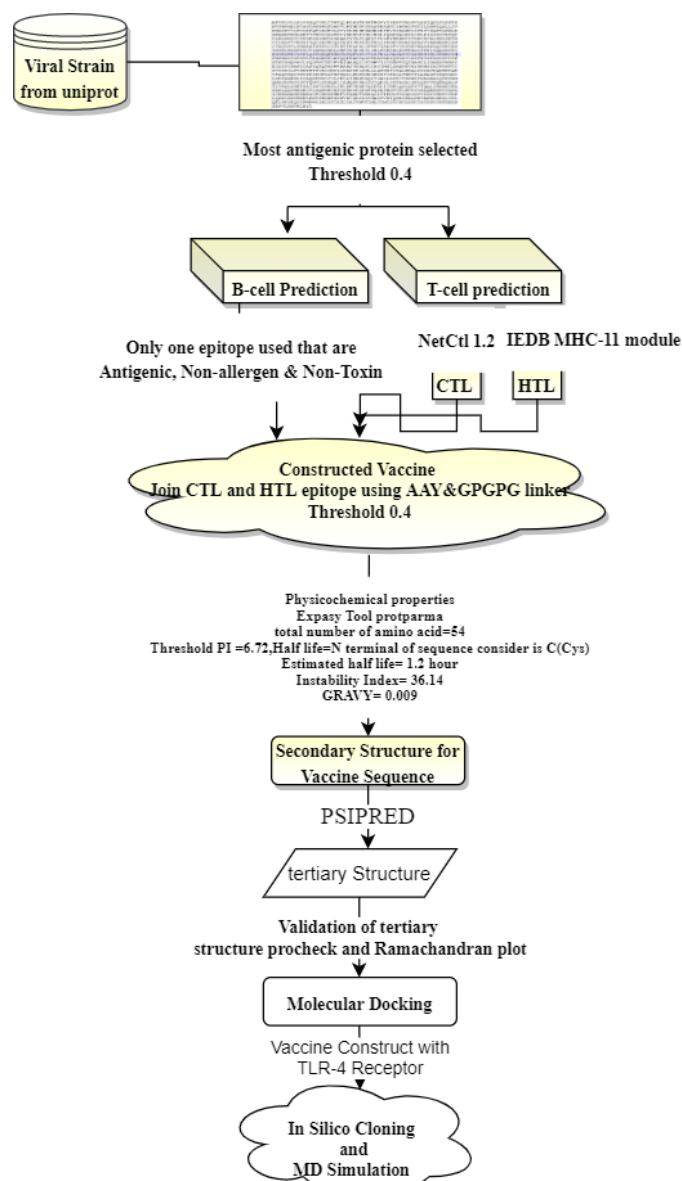


Figure: 1 Methodological Frame work

2.1 Retrieval of SARS_COV_2 Virus Protein Sequence

To find out the details about the Spike glycoprotein P0DTC2, record was obtained from the UNIPROT database [18]. S2 protein contains 1273 residue. The method of sequence analysis is generally based on the topographies, construction, and prediction of the sequence, which focuses on RNA, DNA, or large variety of peptide sequences by analytical methods. The organism of this protein is SARS-COV-2[19].

2.2 Epitope Prediction from IEDB

To find out the calculation of B-cell and T-cell the spike glycoprotein sequence has been used. Immune Epitope Database (IEDB) has been used to predict these epitopes. EXPASY (Bioinformatics Resource Portal) has been used for Molecular weight, the half-life of dissociation, and Isoelectric (PI) point[20].

2.3 Antigenicity prediction

VAXIJEN identifies the Viral, Bacterial, and Tumor antigens. VAXIJEN server has been designed in a language called Perl, with an interface designed in Hypertext Markup Language (HTML). Single sequence uploaded in Plain format, and multiple sequence is uploaded in FASTA format. VAXIJEN predicted the antigen possibility of every protein and antigen status (“probable Non-Antigen versus Probable Antigen”). In order to check the protein sequence in antigenic, online server named VAXIJEN v 2.0 has been used, with a default threshold value [21]. By using Notepad++, all the antigenic spike protein are organized and individual scores have been obtained.

2.3.1 Pre-processing of Auto Cross Covariance (ACC)

Auto-Cross Covariance transformation is followed to convert the sequence of the protein into fixed-length Vectors, and also used to measure the Correlation of two matrices. The protein used in this current study has different measurement, and an ACC is used to convert it to the same measurement. AJJ (lag) calculated using the equation mentioned below[22].

$$AJJ(l) = \sum_i^{n-l} \frac{Z_{j,i} * Z_{j,i+1}}{N-l}$$

In this current equation Index J was is used here to represent Z scale J=(1,2,3), and amount of amino acid in this sequence is represented by n, I is the position of the amino acid I=(1,2,...n) and l is the lag (1,2,3,4,5). In order to check the antigenicity of a sequence, small range of lag were used..

2.4 T-cell epitopes predication

T-cell epitopes are one of the important type of leucocyte also called white blood cells (WBCs) of the immune system, and play significant part to strengthen the immune system. In this research work, we observed two major subtypes: Cytotoxic T-lymphocytes (CTL) also called CD8+ “Killer” and Helper CD4+ helper cells.

2.4.1. Cytotoxic T-lymphocytes (CTL) identification

The protein that is antigenic will be used to predict the CTL epitope. To find out the best CTL-cell epitope NETCTL 1.2 online server was used, with a 0.75 threshold value. The tool expands all the left parameters by default by combining score[23]. Established on the overall score the 37 epitopes are selected for further evaluation.

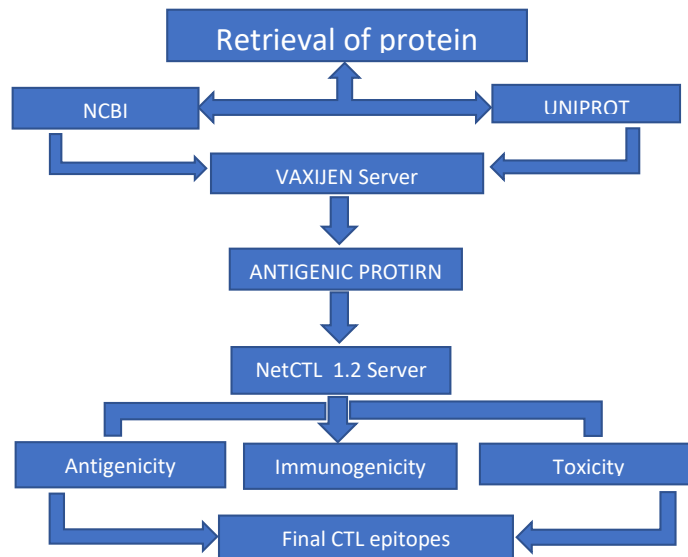


FIGURE: 2 Selection of CTL epitope

2.4.2 Epitope Toxicity and immunogenicity prediction

In order to find out the toxicity of those epitope, an online server TOXINPRED have been used and IEDB tool has been used to analyze the immunogenicity of those epitopes[24]. CTL epitope must be antigenic, immunogenic, and Non-toxin. For selecting the best CTL epitope, it checked three parameters which are: antigenicity, Toxicity, and immunogenicity.

2.5 Helper T- lymphocyte (HTL) epitope prediction

In this research IEDB MHC-11 module has used to predict a Helper T lymphocyte (HTL) epitopes. There are some parameters for the selection of HTLs epitopes similar to what we have already discussed earlier for predicting a CTL epitope. Special epitopes which are selected for this study must be Non-allergenic, Non-overlapping (overlapping done manually), Non-toxic, antigenic and also includes Interferon-gamma[25]. To check these parameters we used a different online server. For allergen city calculation, ALGPRED server is used, Non-overlapping HTL epitopes are selected by using Notepad++, to analyze the toxicity, a server named as TOXINPRED is used. Finally to observe interferon-gamma inducing capability, we use the IFNEPITOPE server.

2.6 Prediction of Linear B-cell:

In order to categorize potential antigens that would enhance the humeral protection, we find out B-cell epitope. To find out the antigenicity of B-cell, several tools of IEDB were followed. To classify B-cell antigenicity along with EMINI External prediction and the flexibility of epitope, KARPLUS and SCHULZ are used. In order to check the antigenicity scale, KOLASKAR and TONGAONKAR are followed, and for the prediction of linear epitope, BEPIRED is used [26]. For the antigenic segments, listed protein is converted into beta-turns fragments, the CHOU and FASMAN beta-turn prediction tool is used for this purpose.

2.7 Construct a vaccine sequence

The vaccine sequence is constructed manually by using Notepad++. We brought together the best CTL and HTL epitope. The CTL epitope is joined together using AAY linker and the HTL epitope

is combined using GPGPG linker. After that, allergen city is checked by the ALGPRED server[27]. To find out the physicochemical properties calculation, PROTPARAM which is a tool of EXPASY is used .By using PROTPARAM tool, different parameters like theoretical PI, Molecular weight, half-life, GRAVY average, and instability index were analyzed.

2.8 Structure prediction, Validation and Molecular docking with receptor

To obtain the model of proposed vaccine for the current study, PSIPRED server is used, while the tertiary model is developed by the help of trRosetta modeling suit. TrRosetta is fast algorithm which is used in residual deep neural networks to predict the inner-residue space and location distribution by using a given input [28]. Then conversion is carried out for the predicted distance and orientation distribution to flat restraints to design a 3D model based on very commonly used technique known as direct energy minimization. Proposed vaccine model with best TM score is validated with PROCHECK v 6.0 web server. Docking of vaccine with receptor is carried out by using PATCHDOCK web server with TLR4 receptor.

2.9 Molecular Dynamic Simulation

Molecular dynamic (MD) simulation is a potent way to study the molecular interaction and dynamic of the vaccine TLR-4 receptor. In this research work, Nanoscale Molecular Dynamic (NAMD) tool is used to perform simulation on TLR-4 receptor and vaccine construct sequence [29]. Parameter files were obtained using the CHARMM General Force Field (CGenFF) tool. Water molecules were added in the protein complex.

2.10 Codon optimization and in silico cloning:

In order to check the high level expression of vaccine codon optimization approach was used. Java codon adaptation tool (JCAT) was used in the codon system E. homo spinous to get the CAI (Codon optimization index) values and Gc contents of the improved sequence to define the level of protein expression. Pet28a (+) vector was used and for modelling the silico clone of vaccine SnapGene 5.2.4 software was used[30].

3. RESULT

3.1 Retrieval of SARS_COV_2 Virus Protein Sequence

To find out the positive epitopes of spike glycoprotein we retrieved the sequence of spike protein from UNIPROT database in FASTA format and then passed through CDHIT with threshold 0.9. Sp|P0DTCS|SPIKE_SARS2 Spike glycoprotein organism OX=2798050 GN=S PE=1 SV=1. From the numerous recognized homologs, we selected protein with length of 1273 amino acids. The antigenic score of the spike protein is greater than 0.4 which is shown in figure (3).

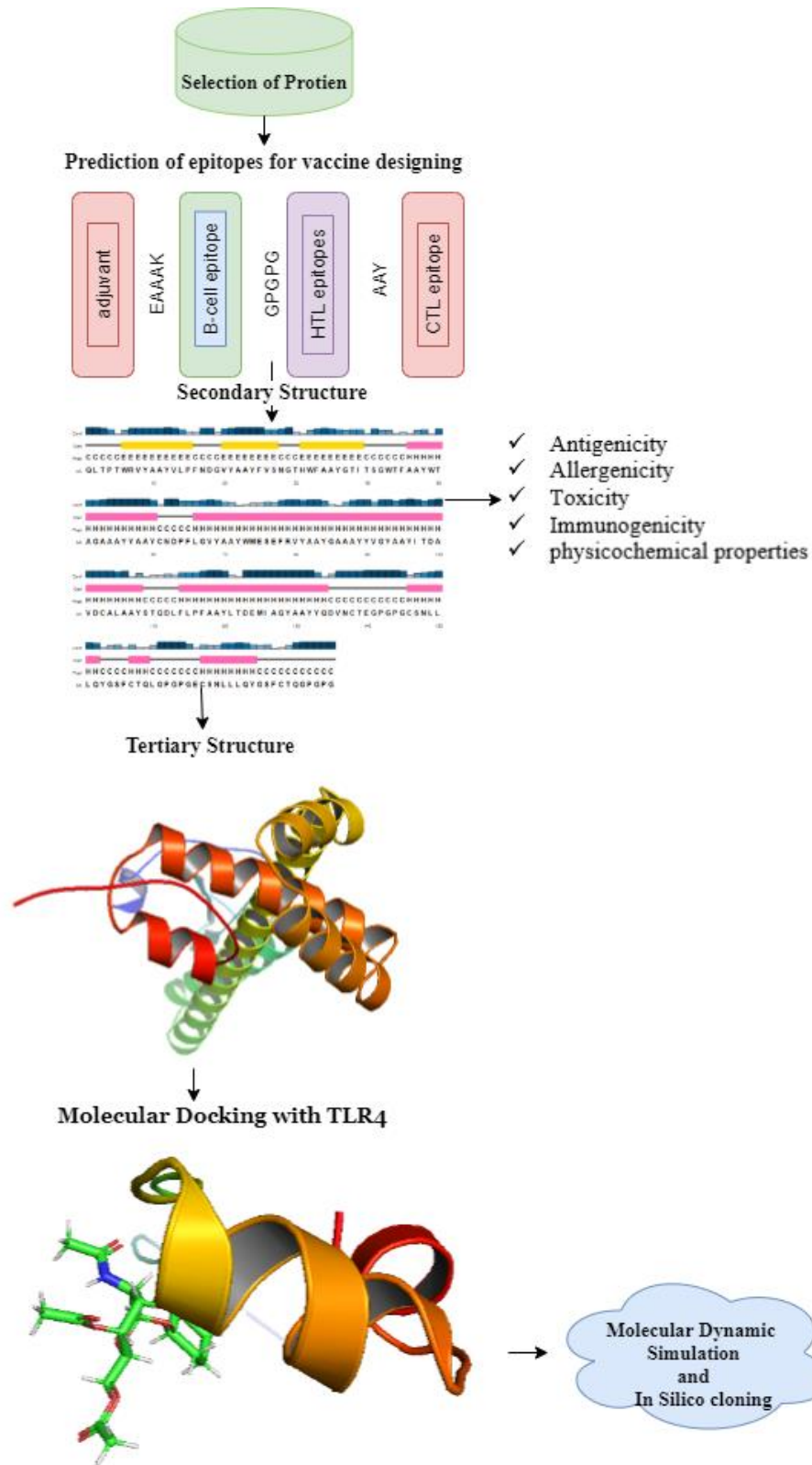


Figure3: Description of the epitope based vaccine using B-cell epitope, HTL epitope and CTL epitope

3.2 Antigenicity prediction

The online VAXIJEN v2.0 server is used to predict the most potent antigenic protein based on the ACC transformation of protein groupings into standard pathways of basic amino corrosive properties. The full antigen prediction value is 0.4683 (probable antigen) with the default threshold value. A current analysis of the proteomic and genomic examination of SARS-COV-2 originates the formation of its proteins, which includes spike glycoprotein, M, N, and E. These proteins have the capability to create SARS-COV-2 immune response [31]. Furthermore, current research suggested that epitope on the Spike protein of virus caused a strong immune response for T-cell in recent times; consequently, these might be potential vaccine candidates [32]. In the light of recent studies on other coronavirus, this study applied different immune informatics tools and Reverse volcanology techniques to identify an epitope with a secondary structure-based vaccine applicant exploiting the spike glycoprotein [33]. In order to find out the antigenicity of spike glycoprotein, we used the online server VAXIJEN v2.0. To find the suitable immunogenic vaccine design, the recent research focused on the prediction of spike protein epitopes and secondary structure of SARS-COV-2; while another research of VAXIJEN-ML website suggested spike protein as the suitable applicant [34]. The VAXIJEN-ML server combines biological data, but the VAXIJEN tool mostly integrates the physicochemical properties of protein sequence for its organization for the calculation of vaccine candidates [35].

3.3 Cytotoxic T-lymphocytes (CTL) identification

Cytotoxic T-lymphocytes are white blood cells that are important for MHC-1 also known as CD8+, T killer cell and T_c . In order to classify CTL cells, the protein sequence 37 epitope was predicted using the online NETCTL1.2 server with score and the MHC class 1 binding [36]. The antigenicity of each nominated by above mentioned tool, twenty-seven (27) epitopes were selected as antigenic. Over all prediction of each epitope are mentioned, which includes WTAGAAAYY= 0.6306 (Probable ANTIGEN), TSNQVAVLY= 0.4387 (Probable ANTIGEN), LTDEMIAQY= 0.1043 (Probable ANTIGEN), KTSVDCTMY= 1.1824 (Probable ANTIGEN), STECSNLLL= 0.4871 (Probable ANTIGEN), GAEHVNSY= 0.9347 (Probable ANTIGEN), WMESEFRVY= 0.2698 (Probable ANTIGEN), VASQSIIAY= 0.1366 (Probable ANTIGEN), CNDPFLGVY = 0.4295 (Probable ANTIGEN), GAAAYYVGY= 0.6604 (Probable ANTIGEN), ITDAVDCAL = 0.5260 (Probable ANTIGEN), RVDFCGKGY= 0.3717 (Probable ANTIGEN), MTSCCSCLK= 0.4270 (Probable ANTIGEN), STQDLFLPF= 0.6619 (Probable ANTIGEN), GTITSGWTF= 0.3272 (Probable ANTIGEN), VLKGVKLHY= 1.2378 (Probable ANTIGEN), ECSNLLLQY= 0.3331 (Probable ANTIGEN), NSASFSTFK= 0.1232 (Probable ANTIGEN), ASFSTFKCY= 0.2795 (Probable ANTIGEN), QLTPTWRVY= 1.2119 (Probable ANTIGEN), LSETKCTLK= 0.6883 (Probable ANTIGEN), VGGNYNYLY= 0.7432 (Probable ANTIGEN), VLPFNDGVY = 0.4642 (Probable ANTIGEN), FVSNGTHWF= 0.0807 (Probable ANTIGEN), YQDVNCTE= 1.5946 (Probable ANTIGEN). For further analysis these 27 epitopes were used and after that the researcher checked toxicity, which revealed that 23 epitopes were Non-toxin WTAGAAAYY, TSNQVAVLY, LTDEMIAQY, STECSNLLL, GAEHVNSY, WMESEFRVY, VASQSIIAY, CNDPFLGVY, GAAAYYVGY, ITDAVDCAL, RVDFCGKGY, STQDLFLPF, GTITSGWTF, VLKGVKLHY, ECSNLLLQY, NSASFSTFK, ASFSTFKCY, QLTPTWRVY, LSETKCTLK,

VGGNYNYLY, VLPFNDGVY, FVSNGTHWF, YQDVNCTE. In order to use the online server TOXINPRED, twenty three (23) epitopes were predicted and these epitopes were used in the immunogenicity analysis which also highlighted MHC class I collaboration. Subsequently, all peptides except GAEHVNSY interacted with the MHC class I alleles. And then by using IEDB immunogenicity tool, positive and negative CTL epitopes have been predicted for developed vaccine design. The high combined score epitopes can be used for vaccine design show in (Table 1).

serial no.	epitopes	epitopes length	NetCtI combined score	antigenicity score	immunogenicity score
1	QLTPTWRVY	9	0.7887	1.2119	0.31555
2	VLPFNDGVY	9	0.7675	0.4642	0.1815
3	FVSNGTHWF	9	0.7622	0.0807	0.16605
4	GTITSGWTF	9	0.9433	0.3272	0.16268
5	WTAGAAAYY	9	3.1128	0.6306	0.15259
6	CNDPFLGVY	9	1.3355	0.4295	0.15232
7	WMESEFRVY	9	1.9232	0.2698	0.14153
8	GAAAYYVGY	9	1.2194	0.6604	0.09963
9	ITDAVDCAL	9	1.168	0.526	0.08501
10	STQDLFLPF	9	1.0468	0.1043	0.06828
11	LTDEMIAQY	9	3.6616	0.6883	0.02757
12	YQDVNCTE	8	0.7502	1.5946	0.02445

Table 1: CTL epitopes with high combined score

3.4 HTL epitope prediction

Helper T-lymphocyte is very important cell in adaptive immunity, as they are play role for adaptive immune response. HTL epitopes are used in initiation of B-cell and CTL for killing stimulated target cells and antibody production. In this research work, IEDB MHC-11 module was used for prediction of HTL epitopes. Same sequence is used of spike glycoprotein set as 7-allele reference with 15 length rest of the parameter is left by default[34]. The result of MHC-11 predicted 8,812 epitopes with allele, SEQ_NUM, Start of epitope, End of epitope, Length of epitope, Method, SEQ_NUM start of epitope, end of epitope, length of epitope, method,peptide,percentile,rank,adjustedrank,COMBLIB_CORE,COMBLIB_SCORECOMBLIB_RANK,COMBLIB_ADJUSTED_RANK,smm_align,SMM_ALIGN_ADJUSTED_RANK,NN_ALIGN_CORE,nn_align,NN_ALIGN_ADJUSTED_RANK.

3.4.2 Allergenicity prediction

To find out the allergenicity of all predicted 8,812 epitopes one by one, we used ALGPRED online server with an SVM segment based on the amino acid configuration method. The result of these epitopes produced potential epitope, allergen epitope, and non-allergen epitope. In this research work, we used only Non- allergen epitopes for further analysis[37]. Total non-allergen epitopes were 957 but we used only peptides for analysis. The percentile rank started with zero because

lower the percentile rank means higher binding affinity. As shown in Table 2, only those (22) peptides were presented which have lower percentile rank.

Table:2 Non-Allergen peptide with allele					
Allele	Peptide	Percentile rank	Start	End	Length
HILA-DRB5*01:01	ITQRFRQTLLALHSY	0.26	235	249	15
HLIA-DRB5*01:01	INITRFQTLLALHRS	0.32	233	247	15
HLA-DRB5*01:01	NITRFQLHTLLARSY	0.32	234	248	15
HLA-DRB5*01:01	TRFQTLLALHRSYLT	0.35	236	250	15
HLA-DRB3*01:01	PINLVRDLPQGFSAL	0.49	209	223	15
HLA-DRB3*01:01	INLVRDLPQGFSALE	0.51	210	224	15
HLA-DRB3*01:01	KHTPINLVRDLPQGF	0.51	206	220	15
HLA-DRB3*01:01	TPINLVRDLPQGFS	0.51	208	222	15
HLA-DRB5*01:01	GINITRFQTLLALHR	0.52	232	246	15
HLA-DRB1*15:01	CSNLLLQYGSFCT	0.58	749	763	15
HLA-DRB5*01:01	RFQTLLALHRSYLT	0.58	237	251	15
HLA-DRB5*01:01	RFQTLLALHRSYLT	0.58	237	251	15
HLA-DRB1*03:01	TPINLVRDLPQGFS	0.59	208	222	15
HLA-DRB1*03:01	PINLVRDLPQGFSAL	0.6	209	223	15
HLA-DRB1*03:01	INLVRDLPQGFSALE	0.62	210	224	15
HLA-DRB1*15:01	ECSNLLLQYGSFCTQ	0.72	748	762	15
HLA-DRB4*01:01	LQIPFAMQMAYRFNG	0.73	894	908	15
HLA-DRB1*03:01	HTPINLVRDLPQGF	0.74	207	221	15
HLA-DRB1*03:01	KHTPINLVRDLPQGF	0.74	206	220	15
HLA-DRB4*01:01	ALQIPFAMQMAYRFN	0.81	893	907	15
HLA-DRB4*01:01	QIPFAMQMAYRFNGI	0.82	895	909	15
HLA-DRB4*01:01	AALQIPFAMQMAYRF	0.93	892	906	15

3.4.3 Homologous Determination

IN this research work for determination of homology modeling chose this non-overlap peptide sequence founded on the numeral of residues checked in the HLA class II peptide binding split [38]. Current research utilized the base bunch grouping in a challenge to lessen difficulties on the charging and furthest reaches of the noticed peptides that probably don't impersonate the probably restricting center contained in the bunch. Non-overlapping was done manually using the Notepad++ and seventeen (17) peptides were selected for further analysis ITRFQTLLALHRSYL, INITRFQTLLALHRS, NITRFQTLLALHRSY, TRFQTLLALHRSYLT, PINLVRDLPQGFSAL, INLVRDLPQGFSALE, KHTPINLVRDLPQGF, TPINLVRDLPQGFS, GINITRFQTLLALHR, CSNLLLQYGSFCTQL, RFQTLLALHRSYLT, ECSNLLLQYGSFCTQ, LQIPFAMQMAYRFNG, ALQIPFAMQMAYRFN, QIPFAMQMAYRFNGI, AALQIPFAMQMAYRF.

3.4.5 Prediction of Antigenicity

Antigenicity is the main part of an organic structure to fix a cluster of secure peptides with potential immunity, T cell receptor and antibodies. In the past, antigenicity was broadly used to refer to immunogenicity, and both are still used interchangeably[39]. In order to predict the antigenicity

with threshold value 4.0 online server VAXIIEN2.0 is used. By using this server we predicted antigenicity of some epitopes which is used to design SARS-COV-2 vaccine, some epitopes were antigenic and some were non-antigenic. The researcher had used only antigenic epitope for further investigation; results are shown in the table 3.

Table:3 antigenicity prediction with score			
SEQ no.	Peptide	probable result	Score
seq1	ITRFQTLLALHRSYL	NON-ANTIGEN	0.1131
seq2	INITRFQTLLALHRS	ANTIGEN	0.4118
seq3	NITRFQTLLALHRSY	NON-ANTIGEN	0.3117
seq4	TRFQTLLALHRSYLT	NON-ANTIGEN	0.3117
seq5	PINLVRDLPQGFSAL	NON-ANTIGEN	0.3117
seq6	INLVRDLPQGFSALE	ANTIGEN	0.5842
seq7	KHTPINLVRDLPQGF	ANTIGEN	0.5644
seq8	TPINLVRDLPQGFS	ANTIGEN	0.5531
seq9	GINITRFQTLLALHR	ANTIGEN	0.5582
seq10	CSNLLLQYGSFCTQL	ANTIGEN	0.6336
seq11	RFQTLLALHRSYLT	ANTIGEN	0.5470
seq12	ECSNLLLQYGSFCTQ	ANTIGEN	0.7635
seq13	LQIPFAMQMAYRFNG	ANTIGEN	0.7205
seq14	HTPINLVRDLPQGFS	NON-ANTIGEN	0.397
seq15	ALQIPFAMQMAYRFN	ANTIGEN	1.0112
seq16	QIPFAMQMAYRFNGI	ANTIGEN	0.9573
seq17	AALQIPFAMQMAYR	ANTIGEN	1.0216

3.4.6 Prediction of Toxicity

Toxicity is referred to as peptides, molecules, or protein that are capable of producing illness on interaction with body muscles coupling with genetic macromolecular such as cellular receptor or enzymes. For designing a potential vaccine structure it's mandatory to check toxicity [40]. As shown in table 4, in order to check the toxicity all antigenic epitopes were screened.

Peptide Sequence	SVM score	Predication	Hydrophobicity	Hydrophobicity	Hydrophobicity	Charge	Mol wt
INITRFQTLLALHRS	-0.87	Non-toxin	-0.13	0.24	-0.44	2.5	1783.34
INLVRDLPQGFSALE	-1.23	Non-toxin	-0.06	0.23	-0.13	-1	1672.14
KHTPINLVRDLPQGF	-0.94	Non-toxin	-0.17	-0.49	-0.06	1.5	1735.25
TPINLVRDLPQGFS	-1.4	Non-toxin	-0.07	0.05	-0.24	0	1628.08
GINITRFQTLLALHR	-1.1	Non-toxin	-0.1	0.27	-0.46	2.5	1753.32
CSNLLLQYGSFCTQL	-0.8	Non-toxin	0.02	0.57	-0.88	0	1690.19
RFQTLLALHRSYLT	-1.18	Non-toxin	-0.15	-0.07	-0.49	2.5	1816.36
ECSNLLLQYGSFCTQ	-0.22	Non-toxin	-0.06	0.08	-0.56	-1	1706.14
LQIPFAMQMAYRFNG	-1.19	Non-toxin	-0.01	0.2	-0.73	1	1787.35

ALQIPFAMQMAYRFN	-1.16	Non-toxin	-0.01	0.35	-0.76	1	1801.37
QIPFAMQMAYRFNGI	-1.37	Non-toxin	0	0.25	-0.73	1	1787.35
AALQIPFAMQMAYR	-1.23	Non-toxin	0.01	0.55	-0.69	1	1611.15

Table 4: Toxicity Prediction

3.4.7 Prediction of positive epitope using Interferon gamma

In order to find out the positive epitope, we used IFNEPITOPES tool. This tool has predicted 12 epitopes, among these, 10 epitopes are negative and two epitopes are positive. In order to predict a model, we used only two (2) positive epitopes, which include CSNLLLQYGSFCTQL and ECSNLLLQYGSFCTQ sequence. These epitopes are used for structure design of vaccine.

Table 5: prediction of positive peptide using Interferon gamma					
Serial No.	Epitope Name	Sequence	Method	Result	Score
1	peptide_5	GINITRFQTLALHR	SVM	POSITIVE	1
2	peptide_1	INITRFQTLALHRS	SVM	POSITIVE	-0.00715
3	peptide_2	INLVRDLPQGFSALE	SVM	NEGATIVE	-0.8355
4	peptide_3	KHTPINLVRDLPQGF	SVM	NEGATIVE	-0.90211
5	peptide_4	TPINLVRDLPQGFS	SVM	NEGATIVE	-0.65774
6	peptide_6	CSNLLLQYGSFCTQL	SVM	POSITIVE	0.286946
7	peptide_7	RFQTLALHRSYLTP	SVM	NEGATIVE	-0.02227
8	peptide_8	ECSNLLLQYGSFCTQ	SVM	POSITIVE	0.582941
9	peptide_9	LQIPFAMQMAYRFNG	SVM	NEGATIVE	-0.21948
10	peptide_10	ALQIPFAMQMAYRFN	SVM	NEGATIVE	-0.46894
11	peptide_11	QIPFAMQMAYRFNGI	SVM	NEGATIVE	-0.17853
12	peptide_12	AALQIPFAMQMAYR	SVM	NEGATIVE	-0.78174

3.5 Linear B cell predication

In order to understand the linear B-cell epitope, the method based on amino acid is followed. The commonly worldwide followed tool that has been used to calculate the linear B-cell epitope is IEDB B-cell Predication tool and this tool is implemented on a hidden Markov model[39]. For linear B-cell epitopes, total thirty-four (34) epitopes were predicted as shown in (Table 6). Thirty (30) epitopes were chosen based on amino acid lengths. The threshold length was set ≥ 7 and only epitopes fulfilling this criteria were further analyzed. In early studies researchers show that the CHAUS and FAMAN method was used for the prediction of β -turns and mostly the antigenic part of the spike protein remained as β -sheet[41]. To check the antigenicity, allergenicity, and toxicity only 1 B cell epitope was designed which contains KHTPINLVRDLPQGFS sequence in its structure.

No.	Start	End	Peptide	
1	13	37	SQCVNLTTRTQLPPAYTNSFTRGVY	25
2	59	81	FSNVTWFHAIHVSGTNGTKRFDN	23
3	97	98	KS	2
4	138	154	DPFLNKSXGVYYHKNME	17

Table:6 Liner B-cell predictions using IEDB B-cell prediction tool

5	177	189	MDLFKNLEGKQGN	13
6	206	221	KHTPINGFSLVRDLPQ	16
7	250	260	TSSSGPGDWT A	11
8	293	296	LDPL	4
9	304	322	KSQTSNFRFTVEKGIYVQP	19
10	329	363	FPNTRFASVYAWN RKITNLCPFGEVFNARISNCVA	35
11	369	393	YNSAKLNDLCSFSTFKCYGVSP TFT	25
12	404	426	GDEV RQNYKLIAPGQTGKIADYP	23
13	440	501	NVGGNYNYLYRLFRKSNLDSKLPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTN	62
14	516	536	ELLHGPKKSTNAPATVCLVKN	21
15	555	562	SFKKNLPF	8
16	580	583	QLTE	4
17	602	606	TNNTS	5
18	616	632	NCTADQLTPTEVPVAIH	17
19	634	644	RVYVFQSTGSN	11
20	656	666	VNCDIPINSYE	11
21	672	690	ASPRRARSVAYQTQTNSSQ	19
22	695	710	YTMVAYSNSLGAENSN	16
23	748	748	E	1
24	773	779	EQDQKNT	7
25	786	800	KQIDFGGFYKTPPIK	15
26	807	814	PKPSDPSK	8
27	828	842	LADAGFIKQYGDCLG	15
28	988	992	EAEVQ	5
29	1035	1043	GQSKRVDFC	9
30	1107	1118	RNFYEPQIITTD	12
31	1133	1172	VTVYDPLQPENNLDSFKEELDKYFKNHTSPDVDLGDISGI	40
32	1203	1206	LYGK	4
33	1252	1267	SCCDSEPV LKFDEDKG	16
34	1269	1269	K	1

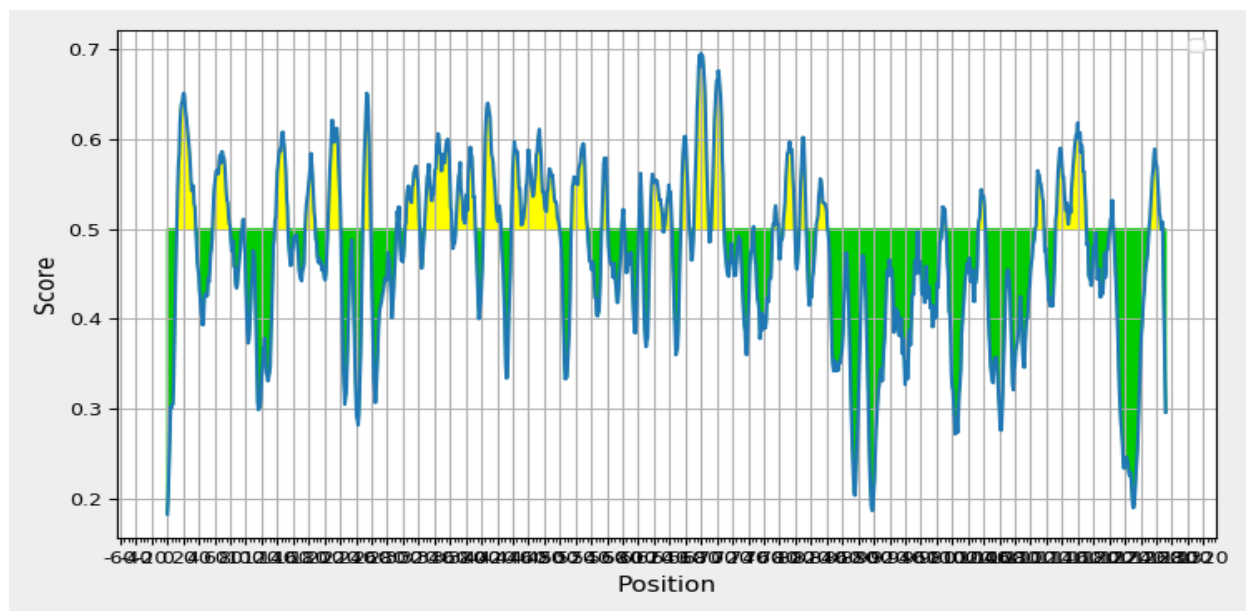


Figure: 4 B-cell epitope prediction using antigenicity

3.6: Designing a subunit vaccine

In this research work B-cell epitopes were considered as a template for CTL and HTL epitopes. A total of 12 CTL epitopes, 2 HTL epitopes, and 1 B-cell epitopes were selected and combined by the help of specific linker[42]. After joining these epitopes, the overall length of the linked sequence was found to be 211 amino acids. For constructing a vaccine design, antigenicity is checked by using online server VAXIJEN v2.0, with threshold 0.4. The protective Antigen=0.4496 (Probable ANTIGEN). Prediction of allergenicity is done by an online server with threshold 0.4. This method yielded a score of -0.19470702 [Threshold= -0.4] with positive and negative predictive value of 64.55% and 86.61% respectively.

3.6.1: physicochemical properties prediction

For prediction of the physicochemical properties calculation we used online tool named as PROTPARAM [43]. To use the PROTPARAM different parameters were set. Chain of 414 Amino acids, with Molecular weight of 45198.52, theoretical PI of 4.23. The probable half-life is 1.2 hours,. The instability index (II) is calculated 29.71 with aliphatic index of 70.53 and Grand average of hydropathicity (GRAVY) of 0.156. The II classifies that the protein is stable.

3.6.1.1 Amino acid composition:

Table: 7 Amino acid composition

Total number of negatively charged residues (Asp + Glu): 28

Total number of positively charged residues (Arg + Lys): 10

3.6.1.2 Atomic composition:

Carbon	C	2100
Hydrogen	H	2952

Ala (A)	78	18.8%
Arg (R)	6	1.4%
Asn (N)	14	3.4%
Asp (D)	16	3.9%
Cys (C)	14	3.4%
Gln (Q)	18	4.3%
Glu (E)	12	2.9%
Gly (G)	34	8.2%
His (H)	4	1.0%
Ile (I)	8	1.9%
Leu (L)	32	7.7%
Lys (K)	4	1.0%
Met (M)	4	1.0%
Phe (F)	22	5.3%
Pro (P)	20	4.8%
Ser (S)	18	4.3%
Thr (T)	28	6.8%
Trp (W)	10	2.4%
Tyr (Y)	52	12.6%
Val (V)	20	4.8%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Nitrogen	N	486
Oxygen	O	601
Sulfur	S	18

Table: 8 Atomic compositions

Formula: C₂₁₀₀H₂₉₅₂N₄₈₆O₆₀₁S₁₈

Total number of atoms: 6157

3.6.1.3 Extinction coefficients:

This protein has no residues of Trp. Study suggests that this could result in more than 10% error in the calculated extinction coefficient. Extinction coefficients are in units of M⁻¹ cm⁻¹, calculated at 280 nm in water. Ext. coefficient 133355 Abs 0.1 percent (=1 g/l) 2.950, assuming that all Cys residues form cystines. Ext. coefficient 132480 Abs 0.1 percent (=1 g/l) 2.931, assuming all Cys residues are reduced.

3.6.1.4 Estimated half-life:

The N-terminal of the sequence is considered is E (Glu). The calculated half-life is 1.2 hours (mammalian reticulocytes, in vitro).

>30 min (yeast, in vivo).

>10 hrs (Escherichia coli, in vivo).

3.7 Prediction of vaccine structure and validation

The structure of the vaccine was predicted by using online server PSIPRED. Prediction of the structure was predicted on the bases of the amino acid sequence of the protein. 211 amino acids long chain was evaluated in which 61 amino acids involved in coil formation and 97 amino acids took part in Alpha helix formation while only 27 amino acids were formatted in Beta standard. Overall secondary structure result indicates 32.97% are coils, 52.43% are Helix while 14.59% are standard as shown in (Fig 5). While the tertiary structure was predicted by the trRosetta modeling suit. The 3D model was developed by PROCHECK

server in order to check the validity, whereas Ramachandran plot were also generated. The output of Ramachandran plot showed 91.6% residues in suitable regions, 6.7% residues were in additional allowed regions, 1.1% residues were in allowed regions and 0.6% residues were in disallowed regions. There were 100% of non-glycine and non-proline residues, number of end-residues is 1, number of glycine residues is 17 and number of proline residues is 10. Based on analysis of 118 structures of at least 2.0 Angstroms and R-factor which is not greater than 20%, a good quality model would be considered if it is over 90% in favored region as it is shown in figure 7.

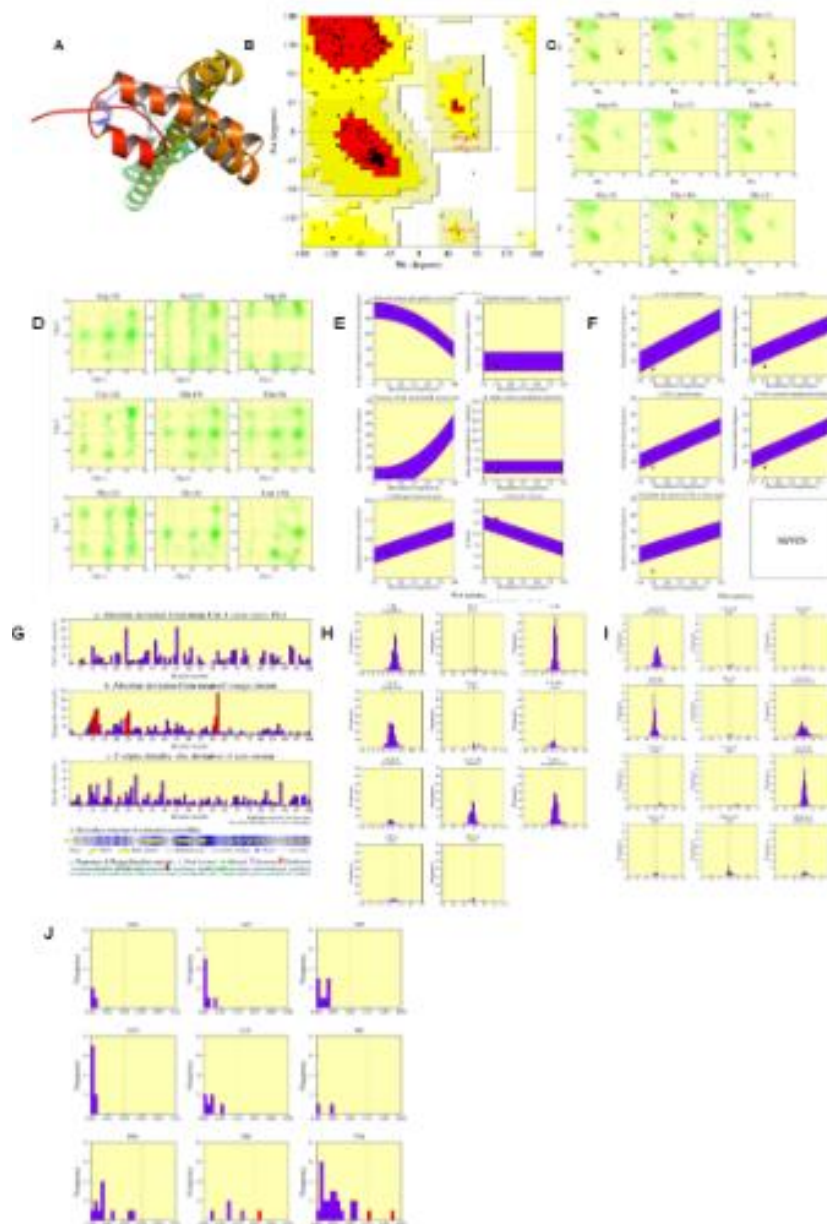


Figure: 5 (A) 3D model prevail from subunit vaccine protein, (B) Ramachandran plot shows favored regions, generously allowed regions, additional allowed region and Disallowed region (C)

Ramachandran plot for all residue type number of residue are shown in brackets, Score < - 3.00 out of 163 structure are labelled for favorable and unfavorable confirmation 2.0 structure **(D)** Chi 1- Chi 2 plots, **(E)** Main Chain parameters, **(F)** Side chain parameters stereo chemical parameters, number of data points, parameters values, Comparison values, number of band width from mean, **(G)** residue properties absolute deviation from mean Chi value, alpha chirality: abs. division of Zeta torsion and Model & estimation, **(H)** Main-chain bond length black bars > 2.0 deviation from mean, standard deviation and small molecules **(I)** main-chain bond angles and **(J)** RMS distance from planarity.

3.8 Vaccine Molecular Docking with Receptor

Vaccine molecular docking is carried out to assess the binding affinity which is -4.23 of protein with TLR4 receptor. PATCHDOCK analysis has constructed 6 receptors on the bases of energy score, but in this research work only 1 receptor was used in MD simulation with energy score of 3436, area 491.60, ACE -275.71 and transformation of -1.58 0.21 -2.64 127.32 278.33 152.46 as shown in figure 7.

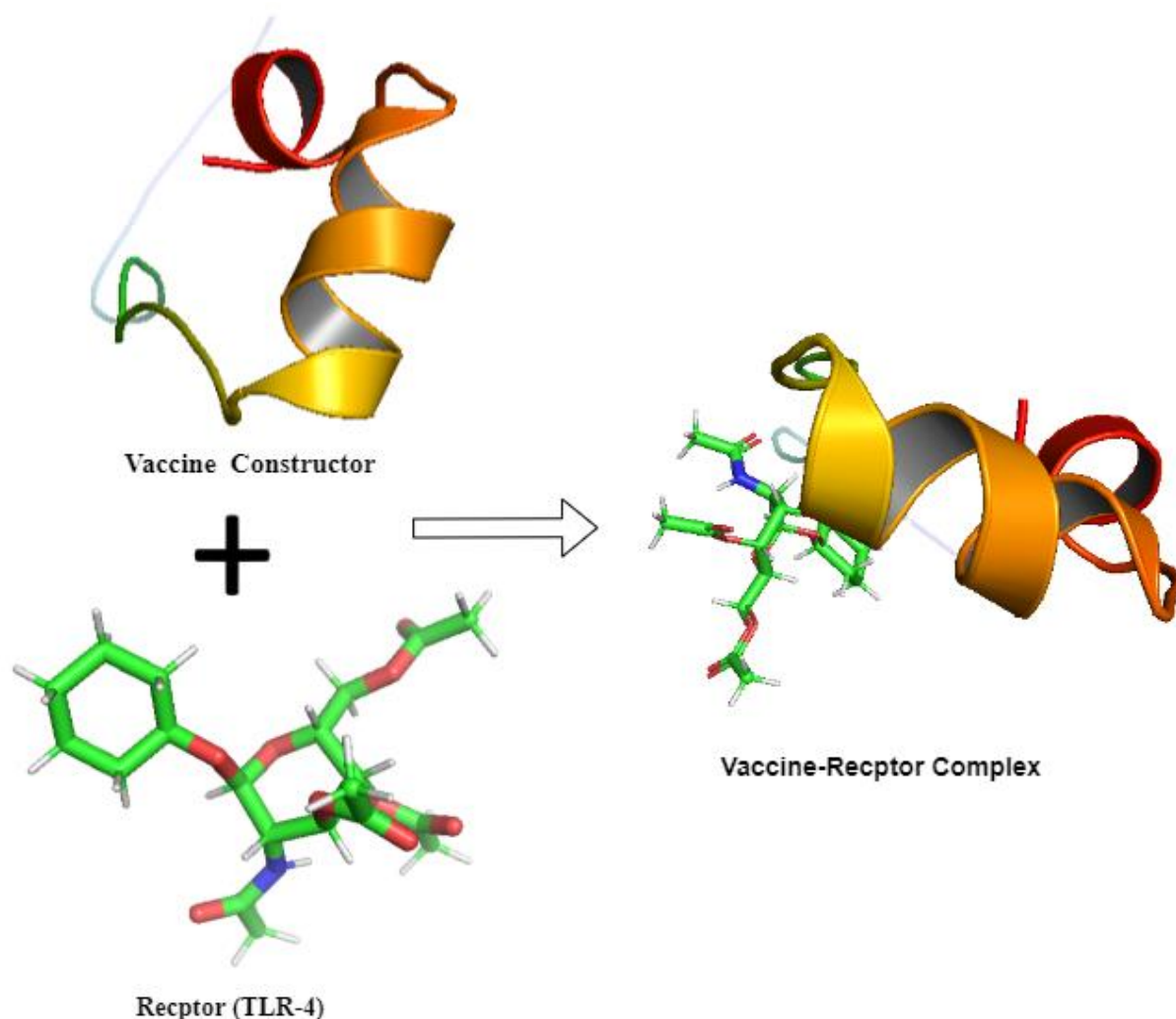


Figure: 6 Interaction between vaccine construct and TLR-4.

3.9 Molecular Dynamic Simulation

A Simulation method known as root mean square deviation (RMSD) is followed to estimate the range between vaccine construct and TLR-4 receptor. MD simulation is used to analyze the stability between TLR -4 and vaccine sequence. As shown in figure 7, interaction between TLR-4 and vaccine construct is verified. The RMSD data indicate that the TLR-4 receptor-vaccine construct were remain stable after 28 ns to the end of the experiment. The RMSD of vaccine construct and TLR-4 has remained stable from 28ns at 1.55 Å and after slight rise system was consistent at 28ns on 1.77 Å. Then the TLR-4 and vaccine construct were fluctuated downward on 18ns to 23ns with timescale 1.3 Å. The total number of interaction and stabilized structure indicates the stability of binding of vaccine with receptor.

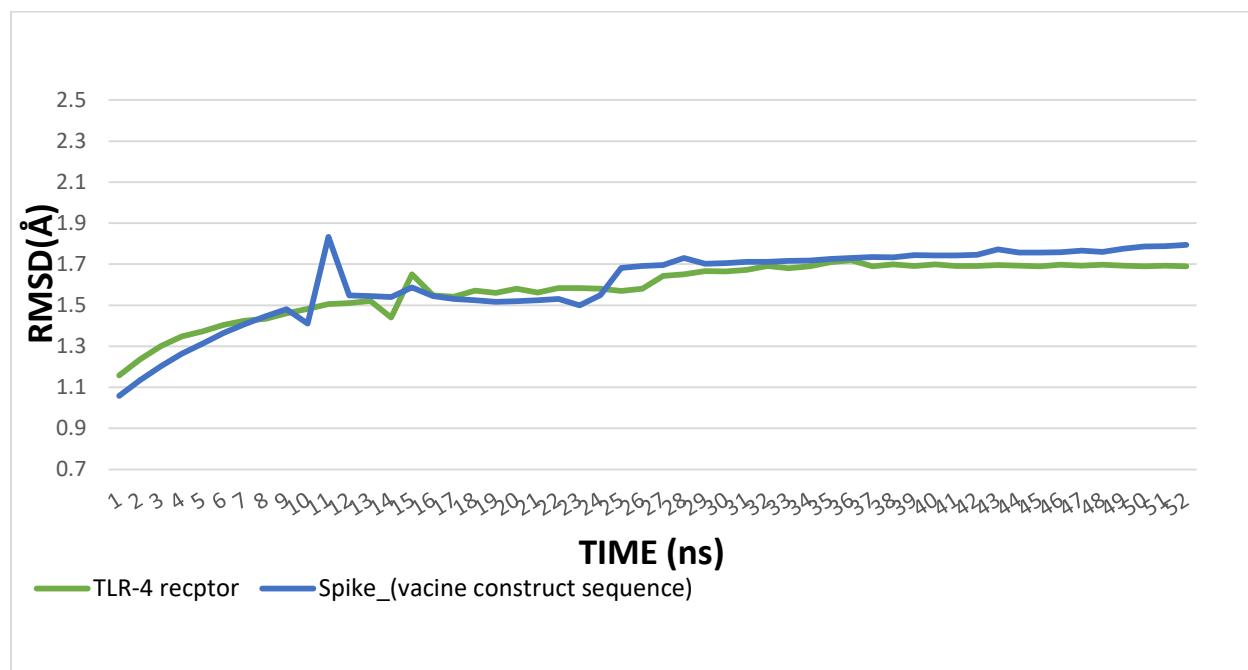


Figure 7. RMSD plot of TLR-4 and vaccine sequence with stability 29ns till 52ns. TLR-4 (green) and Vaccine construct sequence (Blue) is shown in graph.

A short evaluation has shown RMSF of spike protein residues with its tlr4. In Figure 8, spike and tlr4 have shown the variations between 1.2 Å to 1.7 Å. This illustrates that suggested compounds have retained good interaction with the key sites during Molecular Dynamic Simulation.

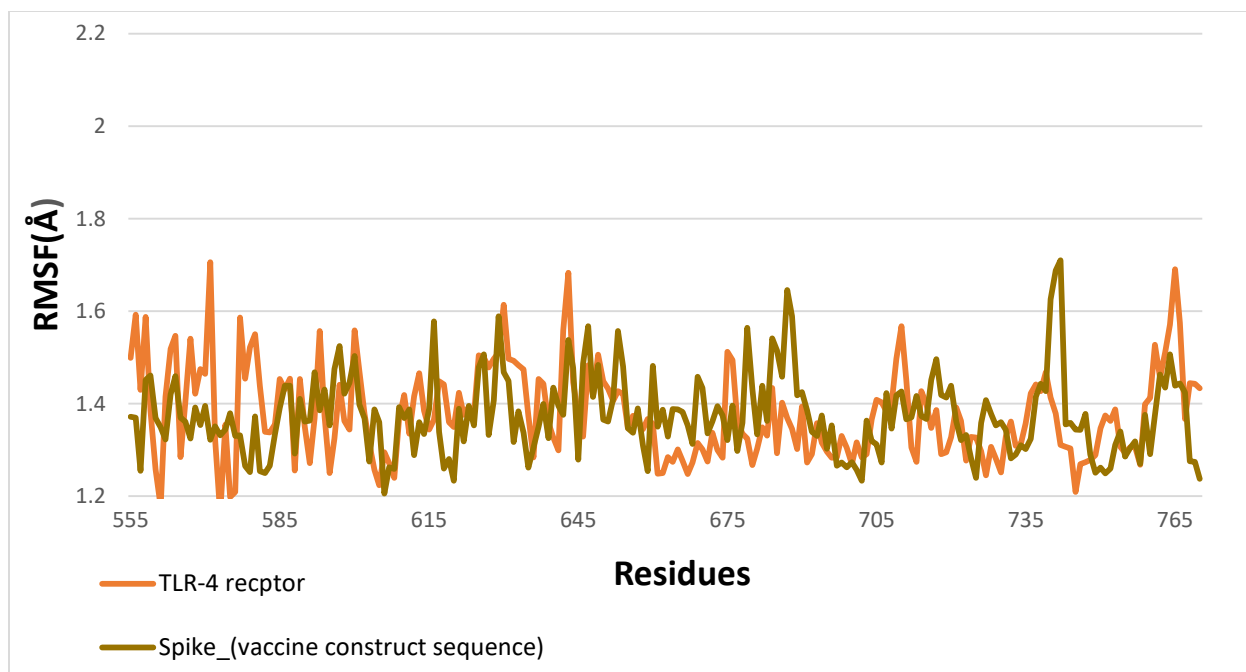


Figure 8: RMSF plot of TLR-4 (orange) and Vaccine construct (brown) sequence during the binding residues.

3.10 Codon optimization and in silico cloning:

Java Codon adaption tool (JCat) was used for maximal optimization of vaccine expression with vector. The codon optimization index (CAI) guarantees the relationship between codon usage and gene expression. The best GC-content value is in 40 to 60 and the best CAI value is less than 0.5. DNA sequence of protein used with CAI value is set to 0.95 and the GC-content is 69.40. The output of JCAT was further analyzed by using Snapgene software. For making a new clone of vaccine sequence vector Pet28a(+) and DNA sequence were inserted. The total value of Pet28a (+) is 5369bp and DNA sequence is 621 bp. Two common restriction sites are PpuM1 and ACu1, value of common restriction sites for vector is PpuM1 (2230), Acu1 (3772), and the value of common restriction site for DNA sequence is PpuM1 (405). After putting these common restriction sites the total value of vector is 1542bp and DNA sequence is 363bp. When both regions are added new cloned was generated using sNapGene 2.4.5 restriction tool as shown in figure 9.

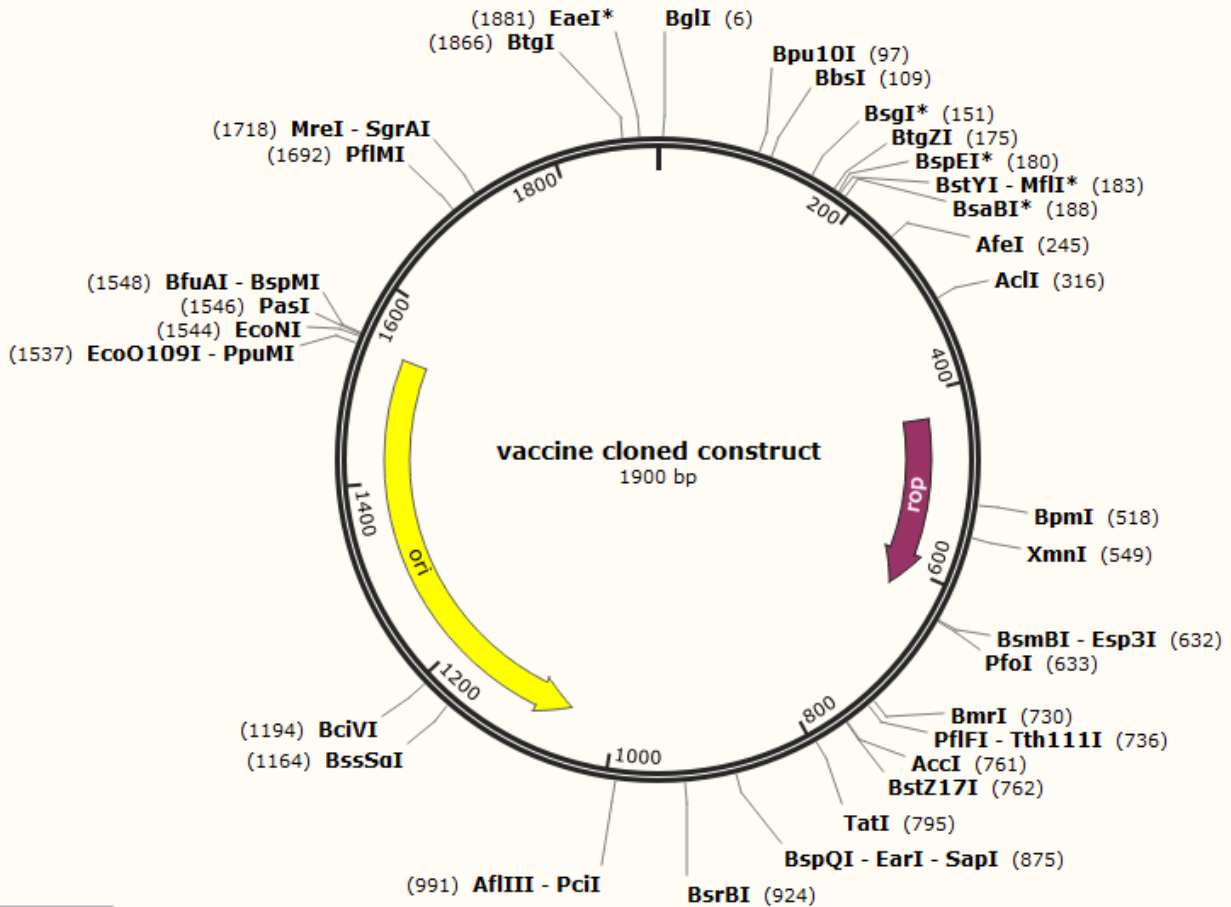


Figure 9: in silico map shows the insertion of vaccine protein optimization codon with Pet28a (+) vector.

4. Mutation against Non-variance

Mutation is transformation of the genetic material (genome) of a cell in a living organism or in a virus that is going to last and may be spread to the cell. The genomes of humans are made up of DNA, while the viral genomes may be DNA or RNA. DNA replication can transfer mutations in the DNA of a body cell in a multicellular organism (somatic mutation) to next cell, resulting in an irregularity in a sector or a patch of cells, such as cancer. Mutations are caused either by collisions during normal chemical DNA transactions, often during reproduction, or by exposure to high-energy electromagnetic radiation (e.g. ultraviolet or X-ray light) in the atmosphere, or by particulate or highly reactive chemicals. Since mutations are abrupt modifications, they are deadly in most cases, but in some conditions, it can be helpful. The primary cause of genetic diversity, is mutation in general [44][45]. UK variant of SARS-COV-2 carries a mutation at position N501Y, South African variant carries a mutation at position N501Y, E484K, K417N and Natural Variant carries a mutation at position D614G. Modified mutations of SARS-COV-2 by UNIPROT are at following locations of mutagenesis Q493N, Q493Y, N501T, D614G, NSPRRR679IL, NSPRRR680IL, NSPRRR681IL, NSPRRR682IL, NSPRRR683IL, NSPRRR684IL, PRRA681RRRK, PRRA682RRRK, PRRA683RRRK, PRRA684RRRK, and Natural Variant at

position D614G. The variation in the above sequences were not discovered in the epitopes under consideration after careful analysis. Epitopes under consideration with their starting and end points are mentioned as follows: CSNLLLQYGSFCTQL Start: 749, End: 763, ECSNLLLQYGSFCTQ Start: 748, End: 762, KHTPINLVRDLPQGFS Start: 206, End: 221.

5. Discussion

From March 18-22, there are no new cases reported in Chinese city Wuhan but few reports indicate that 5-10% retrieved patients in Wuhan test positive again; this strengthens a chance of a comeback of COVID-19 in areas where there is still strict lockdown[46,47]. The spread of virus can also be caused by asymptomatic carriers. A positive re-test, although may be because the actual test was false-negative, and the patient was not actually COVID negative. In this case, maybe a vaccine is a better idea than a drug [47]. The struggles to generate a vaccine against SARS-COV-2 is rapidly increasing. Two candidate vaccines are in phase I in clinical trials: first is an Adenovirus type-5 vector based vaccine, and second is LNP-encapsulated mRNA vaccine. Different studies analyze the immunogenicity and protection of these vaccines. Several candidates of vaccines are under clinical trial. However, these trials are still in progress, even in recent trials that vaccines have failed to make an impact. Recently, few groups have tried to design the subunit vaccine against the SARS-COV-2; however, their workflow engaged either using only HTL epitope or CTL epitope without contemplating the significance of B-cell epitope. In this research work we considered CTL epitopes, HTL epitopes and B-cell epitope for prediction of subunit vaccine. Based on immunoinformatic approaches, in this research main target was spike glycoprotein. To human respiratory epithelial cells by interacting with cell surface receptor Human ACE2, S protein mediates the entrance of the virus. There are two regions for S protein which are known as S1 and S2. For host cell receptor binding S1 is used and for membrane fusion S2 is used. The S protein is a key target for developing vaccine antibodies, therapeutic and diagnosis of coronavirus. However, for defense S protein is promising immunogenic, to ensure an optimal immune response the optimizing antigen design is critical. In this study, vaccine includes CTL, B-cell and HTL epitopes that are combined by specific linkers. Furthermore, the epitopes were tested for their antigenicity, allergenicity and toxicity. The development of an effective vaccine investigates the immunogenic correlation with SARS-COV-2. Molecular dynamic simulation and molecular docking shows the stability effectiveness, strength, and molecular structure of the vaccine receptor depicts the molecular interaction between the Vaccine model and TLR-4 receptor. Moreover, the epitope used to make the vaccine in this research work does not include epitopes that have undergone mutation changed. Though, such methodologies would not serve the importance due to the crisis and safety in the outbreak of disease. In controlling the disease and for designing a vaccine and computational prediction is helpful. The development of vaccines is an expensive and lengthy process, with a high rate of failure. Development of a vaccine takes multiple candidates and several years. The vaccine can be rapidly tested and released in the market. Studies suggest that the epitopes based vaccine can be good ploy against the SARS-COV-2 infection in humans. This research work suggests that experimental and synthesis evaluation of this vaccine determines its immunogenic protein.

6. Conclusion

Developments in the field of immune informatics are critically important in the development of the peptide-based vaccine calculations. In B-cells, CTL and HTL, viruses can survive so in addition to SARS-CoV-2, the epitopes that are predicted in this study can regain immunity. Efforts are made to follow the vital methods of immunity; thus the role of the host immune cells is to intensify the immune response in which sufficient information is passed to B-cells, CTL and HTL. In this study, using the immunoinformatic approach, the expected epitopes are able to interact with the vaccine. On the other hand, current study is just an initial strategy for the production of an epitope-based vaccine against the virus. The aim of this research work is to propose epitopes that may prove good candidate for an effective vaccine with the more extensive analysis of vaccine structure against SARS-CoV-2.

7. Bibliography

- [1] J. Esparza, “Three different paths to introduce the smallpox vaccine in early 19th century United States,” *Vaccine*, vol. 38, no. 12, pp. 2741–2745, 2020, doi: 10.1016/j.vaccine.2020.01.077.
- [2] S. K. Patel *et al.*, “A vaccine is not too far for COVID-19,” *J. Infect. Dev. Ctries.*, vol. 14, no. 5, pp. 450–453, 2020, doi: 10.3855/jidc.12744.
- [3] A. Rakib *et al.*, “Immunoinformatics-guided design of an epitope-based vaccine against severe acute respiratory syndrome coronavirus 2 spike glycoprotein,” *Comput. Biol. Med.*, vol. 124, p. 103967, 2020, doi: 10.1016/j.compbiomed.2020.103967.
- [4] W. H. Chen, P. J. Hotez, and M. E. Bottazzi, “Potential for developing a SARS-CoV receptor-binding domain (RBD) recombinant protein as a heterologous human vaccine against coronavirus infectious disease (COVID)-19,” *Hum. Vaccines Immunother.*, vol. 16, no. 6, pp. 1239–1242, 2020, doi: 10.1080/21645515.2020.1740560.
- [5] J. Esparza, “Early vaccine advocacy: Medals honoring Edward Jenner issued during the 19th century,” *Vaccine*, vol. 38, no. 6, pp. 1450–1456, 2020, doi: 10.1016/j.vaccine.2019.11.077.
- [6] G. Mujtaba, F. Ashraf, F. Waheed, and B. Saeed, “Effect of routing protocols and layer 2 mediums on bandwidth utilization and latency,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 2, pp. 517–530, 2019, doi: 10.14569/ijacsa.2019.0100266.
- [7] C. R. Damaso, “Revisiting Jenner’s mysteries, the role of the Beaugency lymph in the evolutionary path of ancient smallpox vaccines,” *Lancet Infect. Dis.*, vol. 18, no. 2, pp. e55–e63, 2018, doi: 10.1016/S1473-3099(17)30445-0.
- [8] D. Chaudhuri, J. Datta, S. Majumder, and K. Giri, “In silico designing of peptide based vaccine for Hepatitis viruses using reverse vaccinology approach,” *Infect. Genet. Evol.*, vol. 84, no. May, p. 104388, 2020, doi: 10.1016/j.meegid.2020.104388.
- [9] Y. Dong, T. Dai, Y. Wei, L. Zhang, M. Zheng, and F. Zhou, “A systematic review of SARS-CoV-2 vaccine candidates,” *Signal Transduct. Target. Ther.*, vol. 5, no. 1, 2020, doi: 10.1038/s41392-020-00352-y.

- [10] M. Czub, H. Weingartl, S. Czub, R. He, and J. Cao, "Evaluation of modified vaccinia virus Ankara based recombinant SARS vaccine in ferrets," *Vaccine*, vol. 23, no. 17–18, pp. 2273–2279, 2005, doi: 10.1016/j.vaccine.2005.01.033.
- [11] J. Zhang, H. Zeng, J. Gu, H. Li, L. Zheng, and Q. Zou, "Progress and prospects on vaccine development against sars-cov-2," *Vaccines*, vol. 8, no. 2, pp. 1–12, 2020, doi: 10.3390/vaccines8020153.
- [12] K. G. Andersen, A. Rambaut, W. I. Lipkin, E. C. Holmes, and R. F. Garry, "The proximal origin of SARS-CoV-2," *Nat. Med.*, vol. 26, no. 4, pp. 450–452, 2020, doi: 10.1038/s41591-020-0820-9.
- [13] L. van Dorp *et al.*, "Emergence of genomic diversity and recurrent mutations in SARS-CoV-2," *Infect. Genet. Evol.*, vol. 83, no. May, p. 104351, 2020, doi: 10.1016/j.meegid.2020.104351.
- [14] M. J. Mulligan *et al.*, "Phase 1/2 study of COVID-19 RNA vaccine BNT162b1 in adults," *Nature*, 2020, doi: 10.1038/s41586-020-2639-4.
- [15] S. J. Zost *et al.*, "Potently neutralizing and protective human antibodies against SARS-CoV-2," *Nature*, vol. 584, no. 7821, pp. 443–449, 2020, doi: 10.1038/s41586-020-2548-6.
- [16] R. H. See *et al.*, "Comparative evaluation of two severe acute respiratory syndrome (SARS) vaccine candidates in mice challenged with SARS coronavirus," *J. Gen. Virol.*, vol. 87, no. 3, pp. 641–650, 2006, doi: 10.1099/vir.0.81579-0.
- [17] A. Ramaiah and V. Arumugaswami, "Insights into Cross-species Evolution of Novel Human Coronavirus 2019-nCoV and Defining Immune Determinants for Vaccine Development," *bioRxiv*, pp. 1–15, 2020, doi: 10.1101/2020.01.29.925867.
- [18] J. Yu *et al.*, "DNA vaccine protection against SARS-CoV-2 in rhesus macaques," *Science* (80-.), vol. 369, no. 6505, pp. 806–811, 2020, doi: 10.1126/science.abc6284.
- [19] P. Pajer *et al.*, "Characterization of two historic smallpox specimens from a Czech museum," *Viruses*, vol. 9, no. 8, pp. 1–14, 2017, doi: 10.3390/v9080200.
- [20] P. Banerjee, A. O. Eckert, A. K. Schrey, and R. Preissner, "ProTox-II: A webserver for the prediction of toxicity of chemicals," *Nucleic Acids Res.*, vol. 46, no. W1, pp. W257–W263, 2018, doi: 10.1093/nar/gky318.
- [21] G. Zhou and Q. Zhao, "Perspectives on therapeutic neutralizing antibodies against the novel coronavirus sars-cov-2," *Int. J. Biol. Sci.*, vol. 16, no. 10, pp. 1718–1723, 2020, doi: 10.7150/ijbs.45123.
- [22] X. Ma *et al.*, "Ceftazidime/avibactam Improves the Antibacterial Efficacy of Polymyxin B Against Polymyxin B Heteroresistant KPC-2-Producing *Klebsiella pneumoniae* and Hinders Emergence of Resistant Subpopulation in vitro," *Front. Microbiol.*, vol. 10, 2019, doi: 10.3389/fmicb.2019.02029.
- [23] M. Prachar *et al.*, "COVID-19 Vaccine Candidates: Prediction and Validation of 174 SARS-CoV-2 Epitopes," 2020, doi: 10.1101/2020.03.20.000794.
- [24] A. Q. Gao, L. Bao, H. Mao, L. Wang, K. Xu, and M. Yang, "Title : Rapid development of an inactivated vaccine for SARS-CoV-2 Affiliation : Abstract ;," 2020.

- [25] B. Dearlove *et al.*, “A SARS-CoV-2 vaccine candidate would likely match all currently circulating variants,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 38, pp. 23652–23662, 2020, doi: 10.1073/pnas.2008281117.
- [26] N. M. A. Okba *et al.*, “SARS-CoV-2 Antibody Responses in COVID-19 Patients,” vol. 26, no. 7, 2020, [Online]. Available: <https://doi.org/10.1101/2020.03.11.987958>.
- [27] A. Saghazadeh and N. Rezaei, “Immune-epidemiological parameters of the novel coronavirus—a perspective,” *Expert Rev. Clin. Immunol.*, vol. 16, no. 5, pp. 465–470, 2020, doi: 10.1080/1744666X.2020.1750954.
- [28] J. Zhao *et al.*, “Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019,” *Clin. Infect. Dis.*, no. Xx, pp. 1–8, 2020, doi: 10.1093/cid/ciaa344.
- [29] S. F. de O. Tosta *et al.*, “Multi-epitope based vaccine against yellow fever virus applying immunoinformatics approaches,” *J. Biomol. Struct. Dyn.*, vol. 0, no. 0, p. 000, 2020, doi: 10.1080/07391102.2019.1707120.
- [30] A. Atapour *et al.*, “In Silico Designing a Candidate Vaccine Against Breast Cancer,” *Int. J. Pept. Res. Ther.*, vol. 26, no. 1, pp. 369–380, 2020, doi: 10.1007/s10989-019-09843-1.
- [31] M. I. Abdelmageed *et al.*, “Design of a Multiepitope-Based Peptide Vaccine against the e Protein of Human COVID-19: An Immunoinformatics Approach,” *Biomed Res. Int.*, vol. 2020, 2020, doi: 10.1155/2020/2683286.
- [32] C. Sun *et al.*, “SARS-CoV-2 and SARS-CoV Spike-RBD Structure and Receptor Binding Comparison and Potential Implications on Neutralizing Antibody and Vaccine Development,” no. December 2019, pp. 1–18, 2020, doi: 10.1101/2020.02.16.951723.
- [33] L. A. Jackson *et al.*, “An mRNA Vaccine against SARS-CoV-2 - Preliminary Report.,” *N. Engl. J. Med.*, 2020, doi: 10.1056/NEJMoa2022483.
- [34] M. Enayatkhani *et al.*, “Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study,” *J. Biomol. Struct. Dyn.*, vol. 0, no. 0, p. 000, 2020, doi: 10.1080/07391102.2020.1756411.
- [35] S. F. Ahmed, A. A. Quadeer, and M. R. McKay, “Preliminary identification of potential vaccine targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies,” *Viruses*, vol. 12, no. 3, 2020, doi: 10.3390/v12030254.
- [36] X. Chi *et al.*, “A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2,” *Science (80-.)*, vol. 369, no. 6504, pp. 650–655, 2020, doi: 10.1126/science.abc6952.
- [37] M. Bhattacharya *et al.*, “Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach,” *J. Med. Virol.*, vol. 92, no. 6, pp. 618–631, 2020, doi: 10.1002/jmv.25736.
- [38] A. Atapour *et al.*, “Designing a Fusion Protein Vaccine Against HCV: An In Silico Approach,” *Int. J. Pept. Res. Ther.*, vol. 25, no. 3, pp. 861–872, 2019, doi: 10.1007/s10989-018-9735-4.
- [39] D. Murphy, P. Reche, and D. R. Flower, “Selection-based design of in silico dengue epitope ensemble vaccines,” *Chem. Biol. Drug Des.*, vol. 93, no. 1, pp. 21–28, 2019, doi:

10.1111/cbdd.13357.

- [40] W. T. Choga *et al.*, “In Silico Prediction of Human Leukocytes Antigen (HLA) Class II Binding Hepatitis B Virus (HBV) Peptides in Botswana,” *Viruses*, vol. 12, no. 7, 2020, doi: 10.3390/v12070731.
- [41] P. Sharma, R. Kaur, A. K. Upadhyay, and V. Kaushik, “In-Silico Prediction of Peptide Based Vaccine Against Zika Virus,” *Int. J. Pept. Res. Ther.*, vol. 26, no. 1, pp. 85–91, 2020, doi: 10.1007/s10989-019-09818-2.
- [42] B. Ortega-Berlanga, L. Hernández-Adame, C. del Angel-Olarte, F. Aguilar, S. Rosales-Mendoza, and G. Palestino, “Optical and biological evaluation of upconverting Gd₂O₃:Tb³⁺/Er³⁺ particles as microcarriers of a Zika virus antigenic peptide,” *Chem. Eng. J.*, vol. 385, p. 123414, 2020, doi: 10.1016/j.cej.2019.123414.
- [43] The D614G mutaton in the SARS-COV-2 spike protien reduce S1 shedding and increases infectivity.
- [44] Geographic and genomic distribution of SARS-COV-2 mutations.
- [45] E. Prompetchara, C. Ketloy, T. Palaga, Immune responses in COVID-19 and potential vaccines: lessons learned from SARS and MERS epidemic, *Asian Pac. J.Allergy Immunol.* 38 (2020) 1-9.
- [46] L.Du, Y. He, Y. Zhou, S. Liu, B.j. Zheng, S. Jiang, The spike protien of SARS-CoV-atarget forvaccine and therapeutic development, *Nat. Rev. Microbiol.* 7 (2009)226-236.
- [47] C.K chang, M-h chen cm Fau- Chen Cm FAU - Chiang, Y-1 Chiang Mh Fau - Hsu, T-h Hsu Y1 Fau - Huang, T.H. Huang, Transient oligomerization of the SARS-Cov-2 N protein-implicationfor virus ribnucleoprotein packaging, *Plos One* 8 (5) (2013) e65045.