ISSN print: 1224-5984
ISSN online: 2248-3942

Rom Biotechnol Lett. 2021; 26(5): 2901-2915
doi: 10.25083/rbl/26.5/2901.2915

Received for publication, May, 12, 2021
Accepted, June, 9, 2021

Original paper

Using Immunoinformatics to Design a mRNA Vaccine against the Spike Glycoprotein of SARS-CoV-2

SHEREEN F. ELKHOLY [1,2]

[1] Biology Department, College of Science and Humanities, Prince Sattam bin Abdulaziz University, 11942 Alkharj, KSA.
[2] Department of Plant Molecular Biology, Plant Transformation and Biopharmaceuticals Lab, Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Centre (ARC), Giza, EGYPT

Abstract The rapid outbreak of the new coronavirus SARS-COV-2 has created a major public health challenge. Immunoinformatics tools had a clear effect in tracking the genetic sequence of the virus and monitoring mutations and design vaccines that are effective enough to produce antibodies. In our study, we resorted to the emerging discipline of immunoinformatics in order to design a multi-epitope mRNA vaccine against the spike glycoprotein of SARS-CoV-2. We screened the B cell and T cell epitopes of the Spike glycoprotein. we used ABC pred server to predict B cell epitope in the spike glycoprotein sequence and we used NetMHC-4.1 server to predict the T-cell epitope. Then we selected the B cell and T cell epitopes that fulfilled the antigenicity, non-toxicity, non-allergenicity, induction of both IL4 and IFN gamma. Finally, we designed multi-epitope mRNA Vaccine construct by linking 6 B lymphocytes epitopes (BL) with 6 cytotoxic T lymphocytes epitopes (CTL) together with helper T lymphocyte (HTL) epitope up-streamed by 5' cap and down-streamed by poly A tail. The vaccine was found to be antigenic, non-toxic, non-allergenic, capable of generating a robust immune response. Based on these parameters, this design can be considered a promising choice for a vaccine against SARS-CoV-2.

Keywords Bioinformatics, COVID -19, SARS-COV - 2, immunoinformatics, Types of vaccines

To cite this article: ELKHOLY SF. Using Immunoinformatics to Design a mRNA Vaccine against the Spike Glycoprotein of SARS-CoV-2. Rom Biotechnol Lett. 2021; 26(5): 2901.2915. DOI: 10.25083/rbl/26.5/2901.2915

*Corresponding author: SHEREEN F. ELKHOLY, Biology Department, College of Science and Humanities, Prince Sattam bin Abdulaziz University, 11942 Alkharj, KSA, email: s.elkholy@psau.edu.sa.

Introduction

The present world has witnessed the outbreak of many life-threatening human pathogens. (RAKIB et al, 2020). More recently in late December 2019, a cluster of pneumonia cases was reported in the city of Wuhan, Hubei province, China, which was of unknown cause. (ZHU et al, 2019) Similarly to patients with SARS and MERS, these patients showed symptoms of viral pneumonia, including fever, cough and chest discomfort, and in severe cases dyspnea and bilateral lung infiltration. (GRALINSKI et al .,2020). The current COVID-19 pandemic is a global concern and is spreading at an alarming rate and as more than 131 million cases along with over 2.8 million deaths have been reported globally On 5/4/2021. (World Health Organization). On 11 March, 2020, the World Health Organization (WHO) assessed that COVID-19 can be characterized as a pandemic. (RAKIB et al., 2020). The current COVID-19 pandemic is a global concern and is spreading at an alarming rate.

The epidemic of novel coronavirus disease 2019 (COVID-19) was caused by a new coronavirus occurred in December 2019, and now has spread worldwide and turned into a global pandemic (ZHU et al, 2019). It was first reported in Wuhan, Hubei province, China in December 2019 and gradually spread into all over the world (ZHU et al, 2020 , WU et al, 2019). Epidemiological studies showed that the elderly and patients with co-morbidities are becoming severely ill and ended up with poor prognosis and higher mortality rates (GORBALENYA et al, 2020).

SARS-CoV-2 is a single-stranded RNA-enveloped virus (LU et al, 2019). An RNA-based metagenomic next-generation sequencing approach has been applied to characterize its entire genome, which is 29,881 bp in length (GenBank no. NC_045512), encoding 9860 amino acids (Chen et al, 2019). Gene fragments express structural and non-structural proteins. The S, E, M, and N genes encode structural proteins, whereas non-structural proteins, such as 3-chymotrypsin-like protease, papain-like protease, and RNA-dependent RNA polymerase, are encoded by the ORF region (CHEN et al, 2019). A large number of glycosylated S proteins cover the surface of SARS-CoV-2 and bind to the host cell receptor angiotensin-converting enzyme 2 (ACE2), mediating viral cell entry, (LETKO et al, 2020). When the S protein binds to the receptor, TM protease serine 2 (TMPRSS2), a type 2 TM serine protease located on the host cell membrane, promotes virus entry into the cell by activating the S protein. Once the virus enters the cell, the viral RNA is released, polyproteins are translated from the RNA genome, and replication and transcription of the viral RNA genome occur via protein cleavage and assembly of the replicase–transcriptase complex. Viral RNA is replicated, and structural proteins are synthesized, assembled, and packaged in the host cell, after which viral particles are released (FEHR AND PERLMAN, 2015).

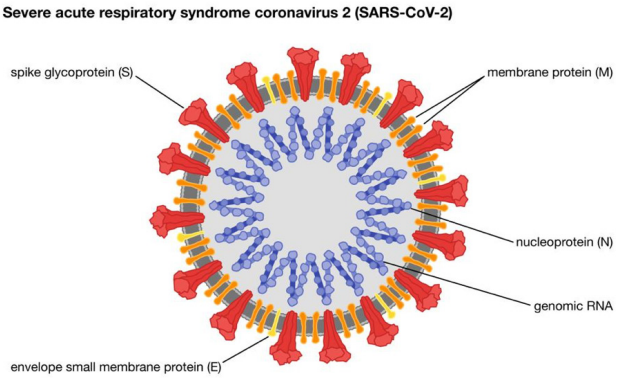

Figure 1. Structure of SARS-COV-2 with the four structural Proteins: S, M, N and E

Bioinformatics studies the storage, manipulation, and interpretation of biological data, especially data of nucleic acids and amino acids, and studies molecular rules and systems that govern or affect the structure, function, and evolution of various forms of life from computational approaches (RUI et al, 2007). In the context of diseases transmitted by microorganisms, such as SARS-CoV-2, the mapping of the genome of microorganisms collected from infected patients in different regions of the world also allows tracing a transmission profile, including its dissemination in different regions and countries, contributing to the search for strategies to combat the disease and monitor mutations (LU et al, 2019, Xingguang et al, 2020). The knowledge of the genome of a species based on the genetic sequencing technique is the starting point for the structure and function of its genes to be understood. Data from the NCBI GenBank® gene sequence database (https:// www.ncbi.nlm.nih.gov/ genbank) accessed on 08/04/2020 indicated more than 14,000 nucleotide sequences inserted since December 2019 for SARS-CoV-2, most of them coming from cities in China and USA (GUSTAVO et al, 2020). In one of the first studies on the new virus that caused respiratory infections in China, the researchers were able to determine from the sequencing of RNA obtained from broncho-alveolar fluid samples that the etiologic agent was RNA virus of the Coronaviridae family. In addition, by using bioinformatics tools, it was possible to perform a phylogenetic analysis, revealing 89.1% similarity of the nucleotide sequence of this virus with a group of coronaviruses of the genus Betacoronavirus already identified in bats in China, which gave evidence of the virus origin (CANRONG et al, 2020).

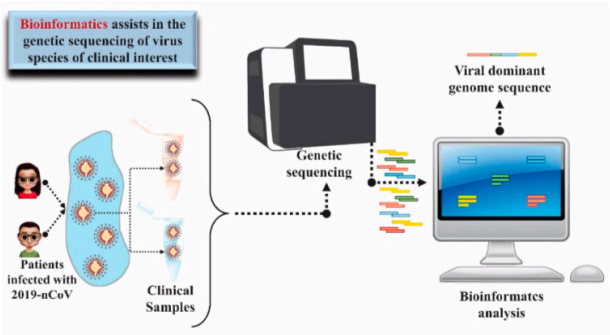

Figure 2. Bioinformatics analysis of patients infected with COVID-19

Immunoinformatics has paved the way for a better understanding of some infectious disease pathogenesis, diagnosis, immune system response and computational vaccinology. infectious agents with complex life cycle and antigenic variability and the need for personalized vaccination present additional challenges in vaccine development (ANGUS et al, 2020). T cell epitopes are relevant to vaccine safety, efficacy, antigen characterization, antigen engineering, and vaccine design. T cell epitopes deserve greater focus in vaccine development, because even in the absence of effective antibody response, T cell epitopes are important drivers of immune defense against pathogens and may also facilitate their escape from immune defense (DE GROOT et al,2020). Immunological studies produce data in colossal quantities. Also, with proteomics and genomics projects, extensive screening of pathogens and/or pathogen–host interaction, it has become increasingly necessary to store, manage and analyze these data. Immunoinformatics deals with computational techniques and resources used to study the immune functions. Statistical, computational, mathematical and biological knowledge and tools are applied in immunoinformatics in order to accurately and specifically store, and analyze data concerning the immune system and its functions. To handle evidence diversity, immunoinformatics uses tools that cut across several aspects of bioinformatics such as creation and management of databases, use and definition of both structural and functional signatures and the formation and application of predictive tools (ANGUS et al ,2020).

Immunoinformatics is a branch of bioinformatics that deals with the computational analysis of immunological data. It is also a powerful computational tool for designing vaccines (KHALILI et al., 2014). By predicting appropriate antigens, epitopes, carriers, and adjuvants for a vaccine, immunoinformatics is able to reduce the timeframe and cost of vaccine development (MARIA et al., 2017). Immunoinformatics approach has been followed for designing vaccines against many infectious agents such as Ebola virus, Human Immunodeficiency Virus (HIV-1), Herpes Simplex Virus (HSV) (BOUNDS et al., 2017, ABDULLAH et al., 2019 and HASAN et al., 2019).

Vaccination has been central in diminishing or eradicating multiple infectious diseases, such as smallpox or polio. However, producing vaccines is a long and complex process, and it has been difficult to implement vaccines against certain pathogens. Thus, designing new vaccines remains a major challenge for public health. To answer this challenge, there have been many improvements to design vaccines, such as using computational prediction. Development of nucleotide vaccines based on DNA, and the related molecule RNA, is another promising area of progress in the field (SCHLAKE et al., 2012). Conventional vaccines usually contain inactivated disease-causing organisms or proteins made by the pathogen (antigens), which work by mimicking the infectious agent. They stimulate the body's immune response, so it is primed to respond more rapidly and effectively if exposed to the infectious agent in the future. RNA vaccines use a different approach that takes advantage of the process that cells use to make proteins: cells use DNA as the template to make messenger RNA (mRNA) molecules, which are then translated to build proteins. An RNA vaccine consists of a mRNA strand that codes for a disease-specific antigen. Once the mRNA strand in the vaccine is inside the body's cells, the cells use the genetic information to produce the antigen. This antigen is then displayed on the cell surface, where it is recognized by the immune system. RNA vaccines are not made with pathogen particles or inactivated pathogen, so are non-infectious (PASCOLO 2006 and PROBST et al, 2007). Unlike DNA, RNA does not integrate itself into the host genome and the RNA strand in the vaccine is degraded once the protein is made (PROBST et al, 2007). As mRNA vaccines do not have to pass through nuclear envelope for translation, it possesses higher efficacy over DNA vaccines (PASCOLO 2006, WOLFF et al, 2000). A major advantage of RNA vaccines is that RNA can be produced in the laboratory from a DNA template using readily available materials, less expensively and faster than conventional vaccine production. (BLACKBURN et a., 2018).

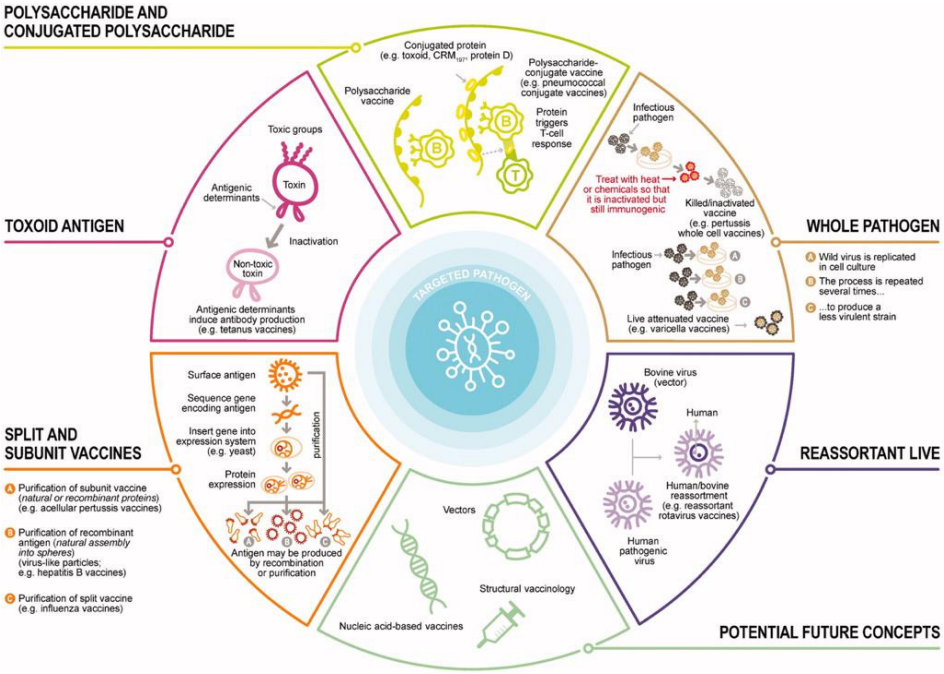

Figure 3. Types of vaccination

In our study, we detected the genetic sequence of the virus in gene bank with accession number: NC_045512. Then we focused on a genetic sequence of the surface glycoprotein with accession number YP_009724390.1. Since it is the immunogenic part of the virus that binds to the host cell receptor (ACE2), mediating viral cellentry and we determined the 3D configuration of the SARS-CoV-2 spike S- glycoprotein. Then we selected the B cell and T cell epitopes of the S Glycoprotein together with helper T lymphocyte epitope that fulfilled the following immunogenic features: Antigenicity, non-toxicity, non-allergenicity and induction of both IL4 and IFN gamma.

Finally, we designed multi-epitope RNA Vaccine construct by linking 6 B lymphocytes epitopes (BL) with 6 cytotoxic T lymphocytes epitopes (CTL) together with helper T lymphocyte (HTL) epitope up-streamed by 5' cap and down-streamed by poly A tail.

Materials and methods

Materials

National Center for Biotechnology Information (NCBI)

The National Center for Biotechnology Information (NCBI) formed in 1988 as a division of the National Library of Medicine (NLM) at the National Institutes of Health (NIH). Among other responsibilities, the NCBI facilitates the use of databases and software and performs research on advanced methods of computer-based information processing for analyzing the structure and function of biologically important molecules including proteins (A.G et al, 2009)

Virus Pathogen Resource (ViPR)

The Virus Pathogen Resource (ViPR; www.viprbrc.org) is a US National Institute of Allergy and Infectious Diseases (NIAID)-sponsored Bioinformatics Resource Center providing bioinformatics support for major human viral pathogens. The hepatitis C virus (HCV) portal of ViPR facilitates basic research and development of diagnostics and therapeutics for HCV, by providing a comprehensive collection of HCV-related data integrated from various sources, a growing suite of analysis and visualization tools for data mining and hypothesis generation, and personal workbench spaces for data storage and sharing (ZHANG et al, 2018).

Rankpep

This server predicts peptide binders to MHCI and MHCII molecules from protein sequence/s or sequence alignments using Position Specific Scoring Matrices (PSSMs). In addition, it predicts those MHCI ligands whose C-terminal end is likely to be the result of proteasomal cleavage.

ABC pred

The aim of ABCpred server is to predict B cell epitope(s) in an antigen sequence, using artificial neural network. This is the first server developed based on recurrent neural network (machine based technique) using fixed length patterns Algorithm: The machine-learning technique need fixed length patterns for training or testing whereas B-cell epitopes vary 5 to 30 residues as reported in literature (Bcipep database).

RCSB PDB

RCSB PDB: Enabling Breakthroughs in Scientific and Biomedical Research and Education. The Protein Data Bank (PDB) was established as the 1st open access digital data resource in all of Biology and Medicine (Historical Timeline). It is today a leading global resource for experimental data central to scientific discovery. Through an internet information portal and downloadable data archive, the PDB provides access to 3D structure data for large biological molecules (proteins, DNA, and RNA). These are the molecules of life, found in all organisms on the planet.

AllerTop V 2.0

AllerTOP is a server for in silico prediction of allergens based on the main physicochemical properties of proteins (IVAN et al, 2014)

Vaxijen 2.0

VaxiJen is the first server for alignment-independent prediction of protective antigens. It was developed to allow antigen classification solely based on the physicochemical properties of proteins without recourse to sequence alignment.

ToxinPred

It is an in silico method, which is developed to predict and design toxic/non-toxic peptides. The main dataset used in this method consists of 1805 toxic peptides

IFNepitope

IFNepitope is a web server to predict and design the epitope, which could induce the release of interferon gamma. The webserver has been developed on the basis of a dataset, which comprises of IFN-gamma inducing and non-inducing MHC class II binders.

NetMHC 4.1 server

The NetMHC-4.1 server predicts binding of peptides to any MHC molecule of known sequence using artificial neural networks (ANNs). Furthermore, the user can obtain predictions to any custom MHC class I molecule by uploading a full length MHC protein sequence. Predictions can be made for peptides of any length.

IL-4pred

Aim of this module is to map IL4 inducing peptides on a protein submitted by a user. Server first generate all possible overlapping peptides of desired length (selected by user) from a protein sequence submitted by user. Secondly, it computes IL4 potential of each overlapping, this way user can identify IL4 inducers in a given protein. Shorting facility allows to identify best regions in a protein that have ability to induce IL4.

Method

We used NCBI for screening of the whole genome of SARS-Cov-2 that compose about 29.9 Kbps. Then we identified the spike glycoprotein that emerges from the envelope and composes 1273 amino acids.

In this study, we focused on Spike glycoprotein since it is the immunogenic part of the virus that binds to the host cell receptor (ACE2) and facilitate the virus entry to the host cell. We used the PDB to determine the 3D configuration structure of the S glycoprotein.

We Screened the S glycoprotein for the LBL, CTL and HTL epitopes using immunoinformatics programs. thereafter we selected 6 best B cell epitopes (LBL) and 6 best T cell epitopes (CTL) and helper T lymphocyte epitope that fulfill certain criteria to activate both MHC class I and class II and induce both humoral and cell mediated immune responses.

Finally, we designed multi-epitope mRNA Vaccine construct by linking 6 B cell epitopes with 6 T cell epitopes together with HTL epitope up-streamed by 5' cap and down-streamed by poly A tail for protection of mRNA construct, initiate and facilitate its translation.

Results

The sequence of the whole genome of SARS-COV-2 and its gene bank accession No.:

We used the National Center for Biotechnology Information (NCBI) https://www.ncbi.nlm.nih.gov/nuccore/1798174254 to determine the sequence of the whole genome of SARS-COV-2 and its accession number.

We found that SARS-COV-2 composes around 29.9 Kbps with accession number NC_045512 and it encodes around 9960 amino acids (data not shown).

The amino acid sequence of the Spike glycoprotein and its accession number:

In this study we concentrated on the spike S glycoprotein as it is surface-exposed and mediates entry into host cells by binding to the host cell receptor (ACE2), it is the main target of neutralizing antibodies (Abs) upon infection and the focus of therapeutic and vaccine design. We used the National Center for Biotechnology Information (NCBI) https://www.ncbi.nlm.nih.gov/protein/YP_009724390.1 to determine the amino acid sequence of the surface glycoprotein and its accession number. We found that the S glycoprotein is composed of 1273 amino acids and its accession number is: YP_009724390.1 (F.g.4)

```
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPF
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYN
ENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV
YAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIAD
YNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYF
PLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL
PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT
PTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLG
AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGI
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDC
LGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG
VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDI
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM
SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVA
KNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD
SEPVLKGVKLHYT
```

Figure 4. The amino acid sequence of the S glycoprotein

The 3D structure of the S glycoprotein and its symmetry

We used the protein data bank RCSB PDB to determine the 3D configuration structure of the SARS-CoV-2 spike glycoprotein with ID: 6VXX.

It is tertiary complex structure glycoprotein with multiple alpha helices and beta plated sheets as shown in fig. 5a,b and linked to carbohydrates as indicated by the blue boxes (Fig. 5 a,b). We used the PDB website to determine the protein symmetry that highlights global, local, and helical symmetry among subunits. The view displays the symmetry axes, a polyhedron that reflects the symmetry, and a color scheme that emphasizes the symmetry of the S glycoprotein with and without the carbohydrates (Fig. 6 a,b).

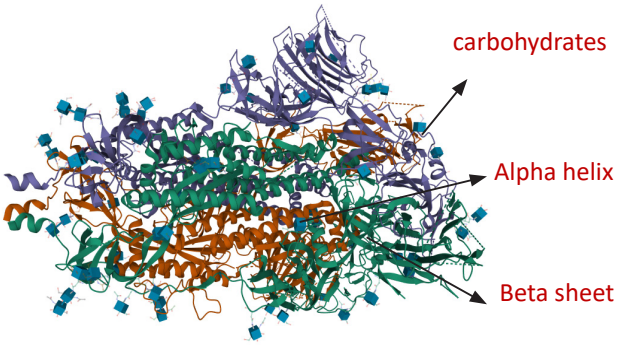

Figure 5a. The 3D structure of the spike glycoprotein with carbohydrates (blue boxes)

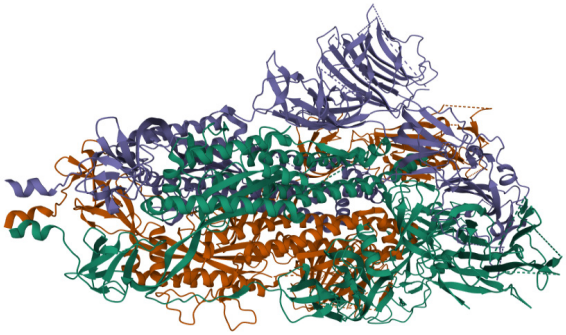

Figure 5b. The 3D structure of the spike glycoprotein without carbohydrates (blue boxes)

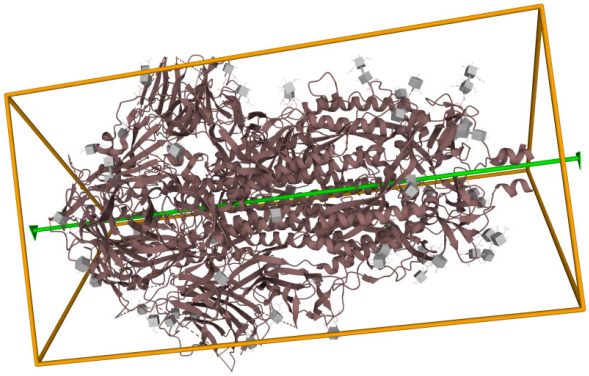

Figure 6a. The symmetry of the S glycoprotein with carbohydrates

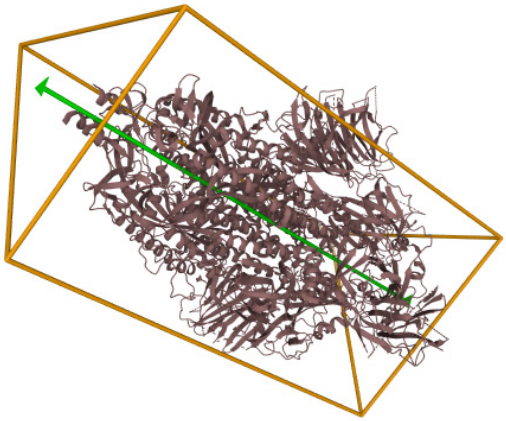

Figure 6b. The symmetry of the S glycoprotein without carbohydrates

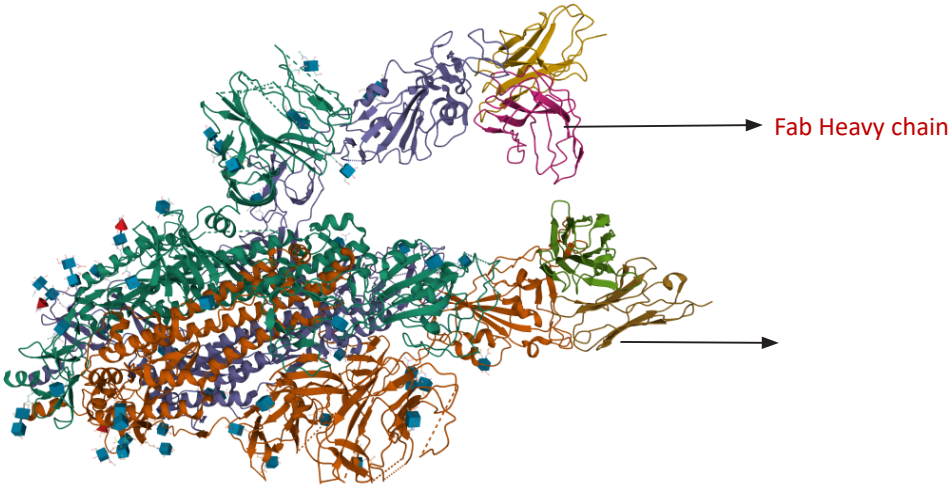

Figure 7. The interaction of the S glycoprotein with Fab fragment of the neutralizing antibody

The interaction of the S glycoprotein with Fab fragment of the neutralizing antibody

We used the protein data bank RCSB PDB to determine the 3D configuration structure of the interaction between the receptor binding domain (RBD) of the S glycoprotein and the Fab fragment of the neutralizing antibody with PDB ID: 6XCM. (Fig.7)

We found the Fab fragment of the neutralizing antibody is less complex in structure than the S glycoprotein since it is composed of β plated sheets without α helices. (Fig.7).

Prediction and assessment of CTL epitopes

CTLs also known as CD8+ T cells play a cardinal role in the battle against viral infections. The mechanism of CTL response against viruses has been elucidated in great

detail by various studies (Asquith *et al*., 2007). Class I MHC-bound epitopes generated from degraded fragments of viral proteins are recognized on the surface of infected cells by CTLs (Alberts *et al*.,2020). Prediction of potential CTL epitopes is an essential and widely used step in in silico vaccine design (Ali *et al*., 2019). There are various immunoinformatics tools that are used for CTL epitope prediction. In this study, we used NetCTL v1.2 server for the prediction of 9-mer CTL epitopes (Table 1) and used Rankpep MHC Class I for prediction of 9-mer CTL epitopes (Table 2). Only epitopes that showed a positive value for immunogenicity were kept for the next stage of evaluation. Toxicity and allergenicity of the immunogenic epitopes were checked using ToxinPred and AllerTOP 2.0 servers respectively. The non-toxic and non-allergenic epitopes were subjected to VaxiJen server for checking antigenicity. Inducibility of interferon γ (IFN γ) and interleukin-4 (IL-4) was checked using the IFNepitope and IL4pred servers respectively. Based on these criteria, only the antigenic, non-toxic, non-allergenic and all two cytokine inducing epitopes were selected for vaccine construction. Five epitopes selected from Table 1 are highlighted and marked by astrisks while one epitope is selected from Table 2.

The six selected CTL epitopes that are antigenic, non-toxic, non allergenic, IL-4 and IFN gamma inducer are presented in table 3.

Table 1. Prediction of CTL epitopes using NetCTL v1.2 server and checking for antigenicity, toxicity, allergenicity, induction of (IL-4) and (IFN γ) cytokines.

Position	Peptide	Rank	Antigenicity	Toxicity	Allergenicity	Induction of IL4	IFN gamma
176	MDLEGKQGN	99.00	ANTIGEN 1.6834	Non-Toxin	allergen	IL4-inducer	POSITIVE
411	PGQTGKIAD	99.00	ANTIGEN 1.5384	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
460 ★	LKPFERDIS	99.00	ANTIGEN 0.4333	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
598	TPGTNTSNQ	99.00	ANTIGEN 0.5029	Non-Toxin	non-allergen	Non-IL4-inducer	NEGATIVE
632	WRVYSTGSN	99.00	ANTIGEN 0.4976	Non-Toxin	allergen	IL4-inducer	NEGATIVE
677	TNSPRRARS	99.00	NON-ANTIGEN -0.0598	Non-Toxin	non-allergen	Non-IL4-inducer	NEGATIVE
768 ★	GIAVEQDKN	99.00	ANTIGEN 1.0158	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
771	VEQDKNTQE	99.00	NON-ANTIGEN 0.0684	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
806	PDPSKPSKR	99.00	ANTIGEN 0.4780	Non-Toxin	allergen	IL4-inducer	POSITIVE
1065	TYVPAQEKN	99.00	ANTIGEN 0.7158	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
495	GFQPTNGVG	95.00	ANTIGEN 0.6424	Non-Toxin	non-allergen	Non-IL4-inducer	POSITIVE
549 ★	GVLTESNKK	90.00	ANTIGEN 0.8797	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
927 ★	NSAIGKIQD	90.00	ANTIGEN 0.5659	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
807	DPSKPSKRS	99.00	ANTIGEN 1.1039	Non-Toxin	allergen	IL4-inducer	POSITIVE
415 ★	GKIADYNYK	85.00	ANTIGEN 1.9241	Non-Toxin	non-allergen	IL4-inducer	POSITIVE

Table 2. Prediction of CTL epitopes using Rankpep MHC Class I server and checking for antigenicity, toxicity, allergenicity, induction of (IL-4) and (IFN γ) cytokines.

RANK	POS.	SEQUENCE	SCORE	Antigenicity	Non-Toxin	Allergenicity	Induction of IL4	IFN gamma
1	1121	**FVSGNCDVV**	25.32	NON-ANTIGEN -0.1537	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
2	797	**FGGFNFSQI**	24.762	ANTIGEN 1.2730	Non-Toxin	non-allergen	Non-IL4-inducer	NEGATIVE
3	133	**FQFCNDPFL**	23.518	NON-ANTIGEN 0.2493	Non-Toxin	allergen	IL4-inducer	POSITIVE
4	444	**KVGGNYNYL**	20.941	ANTIGEN 0.5994	Non-Toxin	allergen	IL4-inducer	POSITIVE
5	538	**CVNFNFNGL**	20.675	ANTIGEN 1.7985	Non-Toxin	allergen	IL4-inducer	POSITIVE
6	1042	**FCGKGYHLM**	20.628	ANTIGEN 0.5098	Non-Toxin	allergen	IL4-inducer	POSITIVE
7 ★	951	**VVNQNAQAL**	20.423	ANTIGEN 0.4749	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
8	915	**VLYENQKLI**	19.028	ANTIGEN 0.4361	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
9	145	**YHKNNKSWM**	18.321	NON-ANTIGEN 0.3388	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
10	1122	**VSGNCDVVI**	16.021	NON-ANTIGEN 0.0579	Non-Toxin	allergen	IL4-inducer	NEGATIVE

Table 3. The six selected CTL epitopes for vaccine design.

#	position	Peptide	Rank	Antigenicity	Toxicity	Allergenicity	Induction of IL4	IFN gamma
1	460	**LKPFERDIS**	99.00	**ANTIGEN** 0.4333	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
2	768	**GIAVEQDKN**	99.00	**ANTIGEN** 1.0158	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
3	549	**GVLTESNKK**	90.00	**ANTIGEN** 0.8797	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
4	927	**NSAIGKIQD**	90.00	**ANTIGEN** 0.5659	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
5	415	**GKIADYNYK**	85.00	**ANTIGEN** 1.9241	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
6	951	**VVNQNAQAL**	20.423	**ANTIGEN** 0.4749	Non-Toxin	non-allergen	IL4-inducer	POSITIVE

Prediction and assessment of B lymphocyte (BL) epitopes

B lymphocytes secrete specific antibodies in order to neutralize specific viral invaders. Through differentiation into long-lived plasma cells and memory B lymphocytes they ensure long term immunological protection (Dorner *et al*., 2007). Activation of B lymphocytes takes place through the binding of B cell receptor to either soluble or membrane bound epitopes (Travers *et al*., 2009). In vaccine design, reliable prediction of B lymphocyte (BL) epitopes through the use of various computational tools plays an important Part (Ahmad *et al*., 2016). In this study, we used Rankpep MHC Class II for prediction of 26 (9-mer) BL epitopes (Table 4) and ABCpred Prediction Server for prediction of 13 (16-mers) BL epitopes (Table 5). Only epitopes that showed a positive value for immunogenicity were kept for the next stage of evaluation. Toxicity and allergenicity of the immunogenic epitopes were checked using ToxinPred and AllerTOP 2.0 servers respectively. The non-toxic and non-allergenic epitopes were subjected to VaxiJen server for checking antigenicity. Inducibility of interferon γ (IFN γ) and interleukin-4 (IL-4) was checked using the IFNepitope and IL4pred servers respectively. Based on these criteria, only the antigenic, non-toxic, non-allergenic and all two cytokine inducing epitopes were selected for vaccine construction. Three epitopes selected from Table 4 are highlighted and marked by astrisks one of them at position 756 became IL-4 inducer by substituting the first mer (Y) by (C). We selected three other BL epitopes from Table 5 by ABCpred Prediction Server, one of them at position 245 became IL-4 inducer by substituting the first mer (H) by (C).

The six selected BL epitopes that are antigenic, non-toxic, non-allergenic, IL-4 and IFN gamma inducer are presented in table 6.

Table 4. Prediction of BL epitopes using Rankpep MHC Class II server and checking for antigenicity, toxicity, allergenicity, induction of (IL-4) and (IFN γ) cytokines.

RANK	POS	SEQUENCE	SCORE	Antigenicity	Toxicity	Allergenicity	Induction of IL4	IFN gamma
1	454	RLFRKSNLK	23.069	**NON-ANTIGEN**	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
2	831	AGFIKQYGD	18.653	**NON-ANTIGEN -0.4781**	Non-Toxin	allergen	IL4-inducer	NEGATIVE
3	323	TESIVRFPN	18.291	**NON-ANTIGEN -0.6508**	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
4 ★	970	FGAISSVLN	16.14	**ANTIGEN 0.5435**	Non-Toxin	non-allergen	IL4-inducer	POSITIVE
5	559	FLPFQQFGR	13.857	**ANTIGEN 0.4802**	Non-Toxin	allergen	Non-IL4-inducer	POSITIVE
6	815	RSFIEDLLF	13.657	**NON-ANTIGEN -0.5782**	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
7	335	LCPFGEVFN	13.143	**NON-ANTIGEN -0.0363**	Non-Toxin	allergen	Non-IL4-inducer	POSITIVE
8	1177	VNIQKEIDR	12.831	NON-ANTIGEN -0.3611	Non-Toxin	allergen	IL4-inducer	NEGATIVE
9 ★	41	KVFRSSVLH	10.707	NON-ANTIGEN -0.6913	Non-Toxin	non-allergen	Non-IL4-inducer	NEGATIVE
10	756	YGSFCTQLN	10.139	**ANTIGEN 0.9142**	Non-Toxin	**non-allergen**	Non-IL4-inducer By substituting Y by C became IL-4 inducer CGSFCTQLN	POSITIVE
11	352	AWNRKRISN	9.814	**ANTIGEN 0.7421**	Non-Toxin	allergen	IL4-inducer	POSITIVE
12	834	IKQYGDCLG	9.474	**NON-ANTIGEN -0.1337**	Non-Toxin	allergen	IL4-inducer	NEGATIVE
13	235	ITRFQTLLA	8.403	**NON-ANTIGEN -0.3233**	Non-Toxin	non-allergen	Non-IL4-inducer	NEGATIVE
14	987	VEAEVQIDR	8.338	**NON-ANTIGEN -0.0275**	Non-Toxin	allergen	IL4-inducer	NEGATIVE
15	794	IKDFGGFNF	8.07	**NON-ANTIGEN 0.2649**	Non-Toxin	**allergen**	IL4-inducer	NEGATIVE
16	296	LSETKCTLK	7.802	**ANTIGEN 0.6883**	Non-Toxin	**allergen**	IL4-inducer	NEGATIVE
17	1218	LGFIAGLIA	7.678	**ANTIGEN 0.5797**	Non-Toxin	**non-allergen**	Non-IL4-inducer	POSITIVE
18	890	AGAALQIPF	7.61	**ANTIGEN 0.4855**	Non-Toxin	**non-allergen**	Non-IL4-inducer	NEGATIVE

19	273	RTFLLKYNE	7.545	**NON-ANTIGEN -0.4873**	Non-Toxin	**non-allergen**	IL4-inducer	NEGATIVE
20 ★	928	NSAIGKIQD	6.889	**ANTIGEN 0.5659**	Non-Toxin	**non-allergen**	IL4-inducer	POSITIVE
21	720	ISVTTEILP	6.558	**ANTIGEN 0.9471**	Non-Toxin	**non-allergen**	IL4-inducer	NEGATIVE
22	449	YNYLYRLFR	6.164	**NON-ANTIGEN -0.8692**	Non-Toxin	**allergen**	IL4-inducer	POSITIVE
23	995	RLITGRLQS	6.13	**NON-ANTIGEN -0.7565**	Non-Toxin	**allergen**	IL4-inducer	POSITIVE
24	350	VYAWNRKRI	5.863	**ANTIGEN 0.5003**	Non-Toxin	allergen	IL4-inducer	POSITIVE
25	1197	LIDLQELGK	5.735	**ANTIGEN 0.9206**	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE
26	1180	QKEIDRLNE	5.608	**NON-ANTIGEN -0.1119**	Non-Toxin	non-allergen	IL4-inducer	NEGATIVE

Table 5. Prediction of BL epitopes using ABCpred Prediction Server and checking for antigenicity, toxicity, allergenicity, induction of (IL-4) and (IFN γ) cytokines.

#	Sequence	Start position	Score	Antigenicity	Toxicity	Allergenicity	Induction of IL4	IFN gamma
1	AGTITSGWTFGAGAAL	879	0.97	NON-ANTIGEN 0.2868	Non-Toxin	non-allergen	Non-IL4-inducer	POSITIVE
2	GVSVITPGTNTSNQVA	594	0.95	ANTIGEN 0.4651	Non-Toxin	non-allergen	Non-IL4-inducer	NEGATIVE
3	GWTAGAAAYYVGYLQP	257	0.95	ANTIGEN 0.6210	Non-Toxin	non-allergen	Non-IL4-inducer	POSITIVE
4	PQIITTDNTFVSGNCD	1112	0.95	NON-ANTIGEN 0.2404	Non-Toxin	allergen	Non-IL4-inducer	NEGATIVE
5 ★	HRSYLTPGDSSSGWTA	245	0.92	ANTIGEN 0.6017	Non-Toxin	non-allergen	Non-IL4-inducer by substituting H by C became IL-4 inducer CRSYLTPGDSSSGWTA	POSITIVE
6	QKEIDRLNEVAKNLNE	1180	0.92	NON-ANTIGEN 0.0684	Non-Toxin	allergen	Non-IL4-inducer	NEGATIVE
7	GSTPCNGVEGFNCYFP	476	0.91	NON-ANTIGEN 0.2489	Non-Toxin	allergen	IL4-inducer	NEGATIVE
8	TVEKGIYQTSNFRVQP	307	0.91	ANTIGEN 0.6733	Non-Toxin	allergen	IL4-inducer	POSITIVE
9	GCLIGAEHVNNSYECD	648	0.90	ANTIGEN 0.8480	Toxin	allergen	IL4-inducer	NEGATIVE
10	LQSYGFQPTNGVGYQP	492	0.90	ANTIGEN 0.5258	Non-Toxin	allergen	Non-IL4-inducer	POSITIVE
11	CCSCGSCCKFDEDDSE	1247	0.90	NON-ANTIGEN 0.1273	Toxin non toxic	allergen	IL4-inducer	NEGATIVE
12 ★	YEQYIKWPWYIWLGFI	1206	0.8	ANTIGEN 0.9510	Non-Toxin	Non-allergen	IL4-inducer	POSITIVE
13 ★	VVLSFELLHAPATVCG	511	0.66	ANTIGEN 0.6184	Non-Toxin	Non-allergen	IL4-inducer	POSITIVE

Table 6. The six selected BL epitopes for vaccine design.

#	POS.	SEQUENCE	SCORE	Antigenicity	Toxicity	Allergenicity	Induction of IL4	IFN gamma
1	970	FGAISSVLN	16.14	**ANTIGEN 0.5435**	Non-Toxin	Non-allergen	IL4-inducer	POSITIVE
2	756	CGSFCTQLN	10.139	**ANTIGEN 0.9142**	Non-Toxin	Non-allergen	IL-4 inducer	POSITIVE
3	928	NSAIGKIQD	6.889	**ANTIGEN 0.5659**	Non-Toxin	Non-allergen	IL4-inducer	POSITIVE
4	245	CRSYLTPGDSSSGWTA	0.92	**ANTIGEN 0.6017**	Non-Toxin	Non-allergen	IL-4 inducer	POSITIVE
5	1206	YEQYIKWPWYIWLGFI	0.89	**ANTIGEN 0.9510**	Non-Toxin	Non-allergen	IL4-inducer	POSITIVE
6	511	VVLSFELLHAPATVCG	0.66	**ANTIGEN 0.6184**	Non-Toxin	Non-allergen	IL4-inducer	POSITIVE

Prediction and assessment of HTL epitopes

After proteolytic cleavage of viral antigens, antigen presenting cells such as B cells, macrophages and dendritic cells present epitopes to HTLs or CD4+ T cells in the epitope-MHC class II complex form (Holling *et al*., 2004). Almost all immunoinformatics studies for vaccine construction incorporate the prediction of HTL epitopes as part of their work (Zahroh *et al*., 2016). In our study, 16-mer HTL epitope and its corresponding MHC class II alleles were predicted using the consensus algorithm of the IEDB MHC-II binding tool Table 7 (Wang *et al*., 2010). Antigenicity and toxicity of the epitope were then detected using the VaxiJen and ToxinPred servers respectively and AllerTOP 2.0 server was used in subsequent order to select the non-allergenic epitope. Inducibility of interferon γ (IFN γ) and interleukin-4 (IL-4) was checked using the IFNepitope and IL4pred servers respectively.

Table 7. Prediction of HTL epitope using IEDB MHC-II binding tool and checking for antigenicity, toxicity, allergenicity, induction of (IL-4) and (IFN γ) cytokines.

Sequence	POS	Binding with MHC class II	Antigenicity	Toxicity	Allergenicity	Induction of IL-4	IFN gamma
YEQYIKWPWYIWLGFI	1206	Good binder	ANTIGEN 0.9510	Non Toxin	Non-allergen	IL4-inducer	POSITIVE

Designing the vaccine construct

A highly-immunogenic mRNA vaccine construct requires the presence of five main elements in the Open Reading Frame (ORF) - 1. Kozak sequence 2. Epitopes 3. Adjuvant 4. Linkers (or spacers) and 5. Stop codon. The start codon should be part of a Kozak sequence (Kozak *et al*., 1986). The criteria used in this study for an epitope to be included in the vaccine construct are antigenicity, non-allergenicity, non-toxicity and cytokine-inducing properties. Along with appropriate antigenic epitopes, a strong adjuvant can boost the adaptive immune response (McKee *et al*., 2017). In our design, co-stimulatory molecule CD40 ligand (CD40L) has been added to the vaccine construct since its inclusion was previously reported to be of significant impact. CD40L is able to activate professional Antigen Presenting Cells (pAPCs) (Caux *et al*., 1994). Selection of an appropriate epitope-specific linker (e.g., flexible, rigid, cleavable) is an essential step in designing an immunogenic multi-epitope vaccine so that the domains can work independently avoiding interaction and interference between them (Chen *at al*., 2013).

The epitopes and the linkers have been combined in the following Ways:
- The adjuvant and the HTL epitopes were combined together by GPGPG linkers.
- HTL and BL epitopes were linked by $(EAAK)_2$.
- Intra-BL epitopes were spaced using $(EAAK)_2$.
- BL and CTL epitopes were linked by AAY.
- Intra-CTL epitopes were combined using AAY linkers.

The final entire construct of the vaccine contained the following elements in order from the N-terminal to C-terminal direction:
5′7GCap–5′UTR – Kozak sequence (for initiation of translation) – TPA (secretory Signal peptide) – CD40L(Adjuvant) – GPGPG linker – HTL epitope – EAAKEAAK linker – LBL epitope– EAAKEAAK linker – LBL epitope – EAAKEAAK linker – LBL epitope– EAAKEAAKE linker – LBL epitope – EAAKEAAKE linker – LBL epitope –EAAKEAAKE linker – LBL epitope – AAY linker – CTL epitope – AAY linker – CTL epitope – AAY linker – CTL epitope – AAY linker – CTL epitope– AAY linker – CTL epitope – AAY

linker – CTL epitope – AAY linker– MITD sequence – Stop codon – 3′UTR – Poly (A) tail (120-250 Adenine).

The proposed mechanism of synthesis, delivery and action of our vaccine has been presented in Fig. 8.

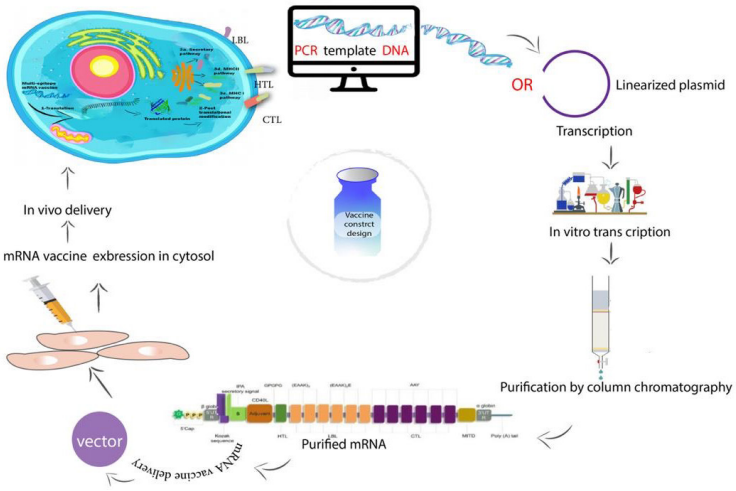

Figure 8. Proposed mechanism of synthesis, delivery and action of the mRNA vaccine against SARS-CoV-2. At first, the PCR template DNA or linearized plasmid DNA containing the designed vaccine sequences is transcribed in vitro in a media containing RNA polymerase and nucleotide phosphates. This results in a mixture of double stranded RNAs and other aberrant products. Therefore, chromatographic purification is carried out to obtain RNA with desired content and length. After vector-mediated delivery into the body, then RNA transits to the cytoplasm. In the cytosol, the cellular translation machinery synthesizes proteins which undergo post-translational modifications, resulting in properly folded, fully functional proteins. The secretory signal sequences direct the peptides to specific compartments of the endoplasmic reticulum and Golgi body for efficient secretion (BL) and presentation by MHC-I (CTL) and MHC-II (HTL).

Discussion

Vaccination has been undeniably very helpful in promoting a healthy global population. It has severally saved lives, reduced healthcare costs and raised man's quality of life (TERRY et al ,2015) greatly reduces disease burden, disability and death. However, newly emerging and re-emerging infectious diseases (ERID), infectious agents with complex lifecycle and antigenic variability and the need for personalized vaccination present additional challenges in vaccine development. (POLAND et al, 2016).

Since the first successful case of mRNA therapeutics in 1990, mRNA vaccines such as those against HIV-1, Zika, rabies, influenza virus etc. have represented a versatile and highly effective subset of vaccine candidates (GHANDI et al,2016, RICHNER et al, 2017 and BAHL et al, 2017). Despite hurdles like instability of mRNA due to degradation by ubiquitous RNases (TSUI et al, 2002) and inherent immunogenicity due to recognition by innate immune sensors (CHEN et al, 2017), mRNA vaccine technology has made great progress. Vaccines are capable of providing immunological memory that persists for several years to several decades (AMANNA et al., 2007). Stimulation of both B and T lymphocyte mediated immune responses is considered crucial for any successful vaccination strategy since it is able to provide a faster and more efficient immune response when the host encounters the target pathogen in the future (KALIA et al, 2006).

The goal of vaccination is to trick the body into thinking that it has been attacked by a pathogen and so generate an immune response that leads to the production of memory B and T lymphocytes (VETTER et al, 2017). Generation of effector and memory B and T lymphocytes depends upon the successful recognition of specific target antigens, more specifically, parts of the specific antigens called epitopes. Therefore, it is of great interest to predict B and T lymphocyte epitopes on target antigens when it comes to designing vaccines (Sanchez *et al*., 2017). CTL mediated cytotoxic activity is a crucial part of the immune response to viral infections. Virus infected cells degrade some of the viral proteins and present them to the CTLs in combination with MHC class I molecules. Recognition of degraded parts of viral proteins called epitopes by CTLs leads to the killing of infected cells through the release of cytotoxic granules (ROSENDAHL et al, 2014). In our study, 6 CTL epitopes have been selected using immunoinformatics program that fulfill certain immunogenic criteria (antigenicity, allergenicity, toxicity and induction of cytokines) for construction of a vaccine against SARS-CoV-2 (Table 3)

Antigen-presenting cells display viral particles to HTLs in combination with MHC class II molecules which leads to the activation of HTLs. Upon recognition of the epitopes, the HTLs secrete a wide range of cytokines and chemokines such as IFN, IL-4, IL-10 etc. which play diverse roles in the immune response against the invaders (DITTMER et

al, 2001). In our study, one HTL epitope was selected for vaccine construction (Table 7).

B lymphocytes bind to antigenic epitopes found on the surface of target cells and subsequently internalize, process and present them to T cells (BATISTA et al, 2001). The processed epitopes are presented in combination with MHC class II on the surface of B lymphocytes and are recognized by HTLs possessing a cognate T-cell Receptor (TCR). This leads to the differentiation of B lymphocytes into antibody secreting plasma cells (HODGKIN et al., 1994). These antibodies are extremely important in neutralizing pathogens (FORTHAL et al, 2014). In this study, 6 B lymphocytes epitopes have been identified as suitable for including in the vaccine construct that are antigenic, non allergenic, non toxic and inducer of both cytokines (Table 6).

While developing an mRNA-vaccine, important factors such as mRNA manufacturing, quality control, formulation, immunological and physicochemical properties of the vaccine as well as the translated form of the peptide come into question. Industrially, mRNA is obtained by in vitro transcription of DNA (linearized plasmid DNA or PCR template) containing a recognition site for RNA polymerase attachment (KRIEG et al, 1987). This mRNA preparation contains double-stranded RNA (dsRNA) contaminants (KARIKO et al., 2011), which in robust type I interferon production. This leads to the inhibition of translation (HARO et al, 1996) and the degradation of cellular mRNA and ribosomal RNA (LIANG et al, 2006). To solve this problem, purification of mRNA using High performance Liquid Chromatography (HPLC) has resulted in increase of translation up to 1000-fold.

mRNA formulation and administration are also crucial factors for antigen expression. Recently, Lipid Nanoparticle (LNP) has emerged as a promising vector for mRNA vaccines. Route of administration is also a contributing factor. Intramuscular and intradermal delivery of mRNA–LNPs has been shown to result in three-fold more persistent protein expression than intravenous delivery. As sustained antigen availability during vaccination drives higher antibody titres and pronounced immune responses, higher half-life can contribute to higher potency of the vaccine. (PARDI et al, 2017).

After incorporation into the cytosol, mRNA comes into contact of the cytosolic translation complex. After translation and post-translational modifications, the vaccine gets ready to generate immune response (Fig. 8). The translated form of our mRNA construct was predicted to be almost neutral, stable, highly antigenic, non-allergenic, non-toxic and inducer of both IFN γ and IL-4 cytokines thus inducing both humoral and cell mediated immunity.

Conclusion

Instead of beginning with the laboratory-based expensive and time-consuming methods, immunoinformatics provides the advantage of low cost and fast identification and screening of epitopes and designing a vaccine. In the face of worldwide transmission of the COVID-19 pandemic at an alarming rate, rapid and high-yield technologies like mRNA vaccine production is the one to meet the challenge. Our immunoinformatics based approach for designing a multi-epitope mRNA vaccine against the spike glycoprotein of SARS-CoV-2 have demonstrated that this novel vaccine candidate can be a useful tool in mankind's arsenal against the deadly virus. This can only come to fruition after further validation of its performance in vitro and in vivo.

References

1. ABDULLA F., U.K. ADHIKARI, M.K. UDDIN (2019), Exploring T & B-cell epitopes and designing multi-epitope subunit vaccine targeting integration step of HIV-1 lifecycle using immunoinformatics approach, Microb. Pathog. 137, 103791.
2. AHMAD T.A., A.E. EWEIDA, S.A. SHEWEITA (2016) B-cell epitope mapping for the design of vaccines and effective diagnostics, Trials in Vaccinology 5 71–83.
3. A.G-PRECIADO, E. MERINO (2009). Genome Sequence Databases: Types of Data and Bioinformatic Tools.
4. ANGUS NNAMDI OLI, WILSON OKECHUKWU OBIALOR,AND GORDON C IBEANU (2020) Immunoinformatics and Vaccine Development.
5. ALBERTS B., A. JOHNSON, J. LEWIS, et al., (2002) T cells and MHC proteins, Molecular biology of the cell, 4th ed Garland Science, New York, [Online] Available https://www.ncbi.nlm.nih.gov/books/NBK26926. (Accessed 26 June 2020).
6. ALI S.A., Y.A. ALMOFTI, K.A. ABD-ELRAHMAN (2019). Immunoinformatics approach for multiepitopes vaccine prediction against glycoprotein B of avian infectious laryngo tracheitis virus, Adv. Bioinforma. 1–23
7. AMANNA I.J., N.E. CARLSON, M.K. Slifka (2007) Duration of humoral immunity to common viral and vaccine antigens, N. Engl. J. Med. 357 (19) 1903–1915 Nov
8. ASQUITH B., A.R. MCLEAN (2007) In vivo CD8+T cell control of immunodeficiency virus infection in humans andmacaques, Proc. Natl. Acad. Sci. 104 (15) 6365–6370
9. BAHL K., et al (2017) Preclinical and clinical demonstration of immunogenicity by mRNA vaccines against H10N8 and H7N9 influenza viruses, Mol. Ther. 25 (6) 1316–1327 Jun.
10. BATISTA F.D., D. IBER, M.S. NEUBERGER (2001) B cells acquire antigen from target cells after synapse formation, Nature 411 (6836) 489–494 May
11. BOUNDS C.E., et al (2017) An immunoinformatics-derived DNA vaccine encoding human class II T cell epitopes of Ebola virus, Sudan virus, and Venezuelan equine encephalitis virus is immunogenic in HLA transgenic mice, Human Vaccines & Immunotherapeutics 13 (12) 2824–2836 Jun.
12. BLACKBURN LAURA (2018) RNA vaccines: an introduction. October www.phgfoundation.org
13. CAUX C., et al. (1994) Activation of human dendritic cells through CD40 cross-linking, J.Exp. Med. 180 (4) 1263–1272 Oct.

14. CANRONG WU et al (2020). Acta Pharm Sin B. MayAnalysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods.
15. CHEN X., J.L. ZARO,W.-C. SHEN (2013) Fusion protein linkers: property, design and function ality, Adv. Drug Deliv. Rev. 65 (10) 1357–1369 Oct
16. CHEN N., P. XIA, S. LI, T. ZHANG, T.T.WANG, J. ZHU (2017) RNA sensors of the innate immune system and their detection of pathogens, IUBMB Life 69 (5) 297–304 Apr.
17. CHEN LIANGJUN ET AL. EMERG (2020). Microbes Infect . RNA based mNGS approach identifies a novel human coronavirus from two individual pneumonia cases in 2019 Wuhan outbreak.
18. DE GROOT ANNE S., LEONARD MOISE,FRANCES TERRY, ANDRES H. GUTIERREZ,POOJA HINDOCHA, GUILHEM RICHARD, DANIEL FREDRIC HOFT, TED M. ROSS, AMY R. NOE, YOSHIMASA TAKAHASHI, VINAYAKA KOTRAIAH, SARAH E. SILK, CAROLYN M. NIELSEN, ANGELA M. MINASSIAN, REBECCA ASHFIELD, MATT ARDITO, SIMON J. DRAPER ANDWILLIAM D. MARTIN,(2020) Better Epitope Discovery, Precision Immune Engineering, and Accelerated Vaccine Design Using Immunoinformatics Tools.
19. DITTMER U., K.E. PETERSON, R. MESSER, I.M. STROMNES, B. RACE, K.J. HASENKRUG (2001) Role of interleukin-4 (IL-4), IL-12, and gamma interferon in primary and vaccine-primed immune responses to friend retrovirus infection, J. Virol. 75 (2) 654–660
20. DÖRNER T., A. RADBRUCH (2007), Antibodies and B cellmemory in viral immunity, Immunity 27 (3) 384–392 Sep.
21. FEHR ANTHONY R. AND PERLMAN STANLEY (2015). Coronaviruses: An Overview of Their Replication and Pathogenesis.
22. FORTHAL D.N., (2014) Functions of antibodies, Microbiology Spectrum 2 (4)
23. GHANDI R.T., et al. (2016) Immunization of HIV-1-infected persons with autologous den dritic cells transfected with mRNA encoding HIV-1 Gag and Nef, JAIDS Journal of Acquired Immune Deficiency Syndromes 71 (3) 246–253 Mar.
24. GUSTAVO R. VILLAS-BOAS et al (2020). The New Coronavirus (SARS-CoV-2): A Comprehensive Review on Immunity and the Application of Bioinformatics and Molecular Modeling to the Discovery of Potential Anti-SARS-CoV-2 Agents Sep
25. GRALINSKI, L. E. & MENACHERY, V. D. (2020) Return of the coronavirus: 2019-nCoV. Viruses 12, 135.
26. GORBALENYA AE, BAKER SC, BARIC RS, et al. (2020). Severe acute respiratory syndrome-related coronavirus: the species and its viruses—a statement of the Coronavirus Study Group. bioRxiv Feb 11.
27. HARO C. DE, R. MÉNDEZ, J. SANTOYO (1996), The eIF-2a kinases and the control of protein synthesis 1, FASEB J. 10 (12) 1378–1387 Oct.
28. HASAN M., et al., (2019) Contriving a chimeric polyvalent vaccine to prevent infections caused by herpes simplex virus (type-1 and type-2): an exploratory immunoinformatics approach, J. Biomol. Struct. Dyn. 1–18
29. HODGKIN P.D., B.E. CASTLE, M.R. KEHRY, B cell differentiation induced by helper T cell membranes: evidence for sequential isotype switching and a requirement for lymphokines during proliferation, Eur. J. Immunol. 24 (1) (1994) 239–246
30. HOLLING T.M., E. SCHOOTEN, P.J. VAN DEN ELSEN, Function and regulation of MHC class II molecules in T-lymphocytes: of mice and men, Hum. Immunol. 65 (4) (2004) 282–290
31. KHALILI S., A. JAHANGIRI, H. BORNA, K. AHMADI ZANOOS, J. AMANI, (2014) Computational vaccinology and epitope vaccine design by immunoinformatics, Acta Microbiol Immunol. Hung. 61 (3) 285-307.
32. KALIA V., S. SARKAR, T.S. GOURLEY, B.T. ROUSE, R. AHMED (2006) Differentiation of memory B and T cells, Curr. Opin. Immunol. 18 (3) 255–264 Jun.
33. KARIKÓ K., H. MURAMATSU, J. LUDWIG, D. WEISSMAN (2011) Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translationof nucleoside-modified, protein-encoding mRNA, Nucleic Acids Res. 39(21).
34. KOZAK M. (1986) Pointmutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44 (2) 283–292 Jan.
35. KRIEG P.A., D.A. MELTON (1987) In vitro RNA synthesis with SP6 RNA polymerase, Methods Enzymol. 397–415.
36. LETKO MICHAEL.,ANDREA MARZI (2020) Vincent Munster Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses.
37. LIANG S.-L., D. QUIRK, A. ZHOU (2006), RNase L: its biological roles and regulation, IUBMB Life (International Union of Biochemistry and Molecular Biology: Life) 58 (9) 508–514 .
38. IVAN DIMITROV et al. (2014) J Mol Model. Jun. AllerTOP v.2--a server for in silico prediction of allergens
39. MARÍA R.R., C.J. ARTURO, J.A. ALICIA, M.G. PAULINA, A.O. GERARDO (2017)., The impact of bioinformatics on vaccine design and development, Vaccines
40. MCKEE A.S., P. MARRACK, (2017Old and new adjuvants, Curr. Opin. Immunol. 47) 44–51
41. PARDI N., et al., (2017) Zika virus protection by a single low-dose nucleoside-modified mRNA vaccination, Nature 543 (7644) 248–251 Feb.
42. PASCOLO S, (2006) "Vaccination with messenger RNA," DNA Vaccines, pp. 23–40.

43. PROBST J, et al., (2007) Spontaneous cellular uptake of exogenous messenger RNA in vivo is nucleic acid-specific, saturable and ion dependent, Gene Ther. 14 (15) 1175–1180.
44. POLAND GA, WHITAKER JA, POLAND CM, OVSYANNIKOVA IG, KENNEDY RB. (2016) Vaccinology in the third millennium: scientific and social challenges. Curr Opin Virol. 17:116–125. doi:10.1016/j.coviro.2016.03.003
45. LU ROUJIAN, XIANG ZHAO, MD, WENJIE TAN, PROF, MD (2019) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding.
46. RAKIB AHMED, SAAD AHMED SAMI ,MD. ASHIQUL ISLAM , SHAHRIAR AHMED ,FARHANA BINTA FAIZ , BIBI HUMAYRA KHANAM , KAY KAY SHAIN MARMA 1, MAKSUDA RAHMAN , MIR MUHAMMAD NASIR UDDIN , FIRZAN NAINU , TALHA BIN EMRAN ,AND JESUS SIMAL-GANDARA (.2020). Epitope-Based Immunoinformatics Approach on Nucleocapsid Protein of Severe Acute Respiratory Syndrome-Coronavirus-2.
47. RICHNER J.M., ET AL., (2017) Modified mRNA vaccines protect against Zika virus infection, Cell 168 (6) 1114–1125.e10 Mar.
48. RUI JIANG,XUEGONG ZHANG, MICHAEL Q, ZHANG (2007) Basics of Bioinformatics Lecture Notes of the Graduate Summer School on Bioinformatics of China.
49. SANCHEZ-TRINCADO J.L., M. GOMEZ-PEROSANZ, P.A. RECHE (2017). Fundamentals and methods for T- and B-cell epitope prediction, J Immunol Res 2017 1–14.
50. SCHLAKE THOMAS,ANDREAS THESS, MARIOLA FOTIN-MLECZEK, KARL-JOSEF KALLEN (2012). Developing mRNA-vaccine technologies . Nov..
51. TERRY FE, MOISE L, MARTIN RF, ET AL. TIME FOR T. (2015) Immunoinformatics addresses vaccine design for neglected tropical and emerging infectious diseases. Expert Rev Vaccines.;14(1):21–35.
52. TSUI N.B., E.K. NG, Y.D. LO, Stability of endogenous and added RNA in blood specimens, serum, and plasma, Clin. Chem. 48 (10) (2002) 1647–1653 Oct.
53. TRAVERS, C. A, P, M. WALPORT, AND M. J. SHLOMCHIK, "B-cell activation by armed helper T cells," Nih.gov, 2009. [Online]. Available: https://www.ncbi.nlm.nih.gov/book NBK27142/
54. VETTER V., G. DENIZER, L.R. FRIEDLAND, J. KRISHNAN, M. SHAPIRO (2017), Understanding modern-day vaccines: what you need to know, Ann. Med. 50 (2) 110.–120 Nov
55. WANG P., et al., Peptide binding predictions for HLA DR, DP and DQ molecules, BMC Bioinformatics 11 (1) (2010) 568.
56. WOLFF J, et al., Direct gene transfer into mouse muscle in vivo, Science 247 (4949) (1990) 1465–1468.
57. Xingguang Li, Junjie Zai, and Qiang Zhao 2020 Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2.
58. ZHU, N. et al. A Novel Coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 382, 727–733 (2020).
59. ZHANG YUN,CHRISTIAN ZMASEK ,GUANGYU SUN,CHRISTOPHER N, LARSEN RICHARD H, SCHEUERMANN (2018) Hepatitis C Virus Database and Bioinformatics Analysis Tools in the Virus Pathogen Resource (ViPR).
60. ZHU NA., DINGYU ZHANG,. WENJIE TAN (2019) A Novel Coronavirus from Patients with Pneumonia in China.
61. ZAHROH H., A. MA'RUP, U.S.F. TAMBUNAN, A.A. PARIKESIT, Immunoinformatics approachin designing epitope-based vaccine against meningitis-inducing bacteria(Streptococcus pneumoniae, Neisseria meningitidis, and Haemophilus influenzatype b), Drug Target Insights 10 (2016) DTI.S38458