ejpmr, 2020,7(6), 29-44

SJIF Impact Factor 6.222

EUROPEAN JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.ejpmr.com

Research Article

ISSN 2394-3211

EJPMR

EVOLUTION OF COVID-19 VIRUS AND DESIGNING B EPITOPE VACCINE FOR SPIKE PROTEIN

Zahra M. Al-Khafaji*[1], Aaisha B. Mahmood[2] and Marium B. Mahmood[3]

[1]Institute of Genetic Engineering and Biotechnology for Postgraduate Studies, University of Baghdad, Iraq.
[2]Ministry of Agriculture, Veterinary Directorate, Baghdad Veterinary Hospital, Al-Dora Hospital, Iraq.
[3]Financial Affairs Dept., Computer Science, University of Baghdad, Iraq.

***Corresponding Author: Zahra M. Al-Khafaji**

Institute of Genetic Engineering and Biotechnology for Postgraduate Studies, University of Baghdad, Iraq.

Article Received on 24/03/2020 **Article Revised on 13/04/2020** **Article Accepted on 03/05/2020**

ABSTRACT

One of coronaviruses back again to attack human, COVID-19 is highly similar to SARS-CoV Middle East respiratory syndrome (MERS). This study aimed to find the possibility of changing/evolution of one of the most virulent factor 'spike protein' which facilities the attachment of the virus to cell receptor(s), and in silico designing and finding of B epitopes that induce production of antibodies to neutralize and block this attachment. Results show that this protein is continuously changing among the retrieved protein sequences from all over the world. It has been found that there is some B epitopes,177 - MDLEGKQGNFKNL-189, 555- SNKKFLPF-562, 656 -VNNSYECDIPI -666, 1035- GQSKRVDFC- 1043, derived from Cons sequence constructed form global protein sequences released from 11 Feb to 06 April, these satisfied most of the required criteria, these can be offered for real wet applications.

KEYWORDS: COVID-19, evolution, B epitopes. spike protein, Iraq.

INTRODUCTION

Recently, the world facing a pandemic outbreak of new coronavirus, initiated in China and spread all over the world, the viral disease causing a sever and even fatal respiratory infections. The virus is highly contagious and transmitted via airborne droplets, in addition to other routes.[1]

The virus and the disease were given more or less similar names by the authorities. In this work COVID-19 will be used. The virus belongs to corona viruses which are common human pathogens, causing in general mild respiratory illness. The zoonotic virus belongs to Coronaviridae family. The virus is enveloped with positive ssRNA genome and few structural and non-structural proteins.[2] It has been found to be with high similarity with SARS-CoV form different points of view.[3,4] The life cycle of COVID-19 virus starts by binding to its receptor Angiotensin-Converting Enzyme (ACE2)[5,6,7], which expressed on different types of cells of respiratory system and other susceptible cells in different parts of the body.[8] The attachment via S1 protein, then fusion of the virus membrane with endosomal membrane probably mediated by S2, followed by releasing the viral genome into the cell cytoplasm. The virus starts replication and assembly of the proteins and releases new infective particles to infect new target cells. These events accompanying by production and activation of proinflammatory chemokines and cytokines causing significant damage to lung tissues resulting in atypical pneumonia with rapid abnormalities and failure.[9,10,11]

As COVID-19 is a RNA virus, it tends to mutate more frequently than other viruses, in general corona virus mutate at highly rates e.g. Infectious bronchitis virus (IBV) which belongs to Gammacoronavirus, its spike protein mutate at rate recorded to be $2*10^{-2}$ to $4*10^{-4}$ substitute/site/year, or 10^{-4} to 10^{-5} substitute/ nucleotide/round of replication compared to 10^{-9} to 10^{-11} substitute /nucleotide/round of replication in *Escherichia coli*.[12,13] Some proteins of the virus mutate more frequently than others especially the outer membrane proteins such as S protein. This high mutation rate happens due to more than one reason, first of all ; RNA molecule is unstable in comparison with DNA molecule, the other reason is responding to altered selective pressures, and infidelity of RNA polymerase due to the absence of proof-reading activity.[4] These changes help the viruses to escape the cell-mediated and humoral immunity and might enable the virus to cross species barriers, so it is likely that COVID-19 emerged from the accumulation of mutations.[14,15] All these make traditional vaccines for viral diseases with limited success and put additional strains on developmental efforts, The spike (S) protein is the main molecule present at the surface of the virion and could represent the most effective virulence factor in all corona virus

groups due to its function and location. It is a multifunctional protein, contributes to host receptor binding, cell tropism and pathogenesis, as it binds to the receptor it induces endocytosis of virion particle, the fusion between host and viral membranes allowing penetration of the virus genome into the host cytoplasm, S2 participates in the latter activities.[16,17,18]

The known receptor for SARS-CoV which is highly related to COVID-19 at multiple levels is the ACE2, but it has been found that COVID-19 has some changes which resulted in increasing the receptors range such as integrins and may enhances animal-to-human and human-to-human transmission.[19,20,21]

Vaccines are the appropriate method to deal with viral infections, epitope/peptides vaccines might be the best, since the peptide vaccines are intended to present appropriate B- and T-cell epitopes to stimulate the adaptative and specific immune response and subsequently induce immune protection and long -lasting response.[22] Such approached opened a new field in immunology called immunoinformatics and more specifically Vaccinomics, which thanks to advance in computer science and its related field Bioinformatics.[23,24,25,26]

In silico epitope vaccine design was used to find epitope vaccine for both B- and T-cell epitopes using different approaches were reported, especially from S protein.[27,28]

The aim of this study (Part I) is to find the evolution and charges of surface proteins of COVID-19 new strains, and to design B epitopes for vaccine developing depending on S protein.

MATERIALS AND METHODS

Number of databases and software were used in this study

NCBI

Used for sequences retrieved of surface proteins, and protein alignment using BLASTp when required.

IEDB database

Used for design B epitopes, (BepiPred2) used to predict linear epitopes with default threshold values, and study of some its characters according to database recommendations.

VaxiJen server

Used for prediction the antigenicity of proteins and epitopes using 0.4 as threshold value for virus group.

EMBOSS Explorer

Used to get consensus sequence for proteins.

MEGA software

Used for alignment and constructing Neighbor joining trees.

AllerCatPro and AllerTop software

Used to estimate the allergenicity of selected appropriate epitopes.

ToxinPred

Used to find the toxicity of the epitopes.

Swiss-pdb viewer 4.1.0.

Used for Ramachandran plot estimation

RESULTS AND DISCUSSION

Surface and secreted proteins of pathogens are mostly antigenic[29,30] and responsible for pathogenicity, it has been identified that spike and membrane proteins as targets for the COVID-19 vaccine.[15] The precursor S protein consisting of about 1273 amino acids, it is cleaved by proteases into surface subunit S1(position 17-680) and the transmembrane subunit S2 (681-1095).[31] The first part (S1) recognizes and binds to host receptors. The spike glycoprotein structure and sequence become available.[32]

Sequence retrieve surface

Surface protein Sequences for COVID-19 were retrieved from NCBI, the release interval extended from 11 Feb till 06 April /2020, using FASTA format, sequences having abnormal character were eliminated. Table 1 shows the information about sequences.

Table 1: Surface protein (Spike) of COVID-19 strains from all over the world.

Acc No	Country	Date
QHR84449	Australia	FEB-11
QHZ00358	China	FEB-11
QHU36864	China	FEB-11
QHU36854	China	FEB-11
QHU36844	China	FEB-11
QHU36834	China	FEB-11
QHU36824	China	FEB-11
QHO62112	China	FEB-11
QHO62107	China	FEB-11
QHN73810	China	FEB-11
QHN73795	China	FEB-11
QHZ00379	Korea	FEB-11
QHZ00399	USA	FEB-11
QHZ00389	USA	FEB-11
QHW06059	USA	FEB-11
QHW06049	USA	FEB-11
QHW06039	USA	FEB-11
QHU79204	USA	FEB-11
QHU79194	USA	FEB-11
QHQ82464	USA	FEB-11
QHQ71973	USA	FEB-11
QHQ71963	USA	FEB-11
QHO62877	USA	FEB-11
QHZ87592	USA	FEB-12
QHZ87582	USA	FEB-12
QID21068	USA	FEB-24
QID21058	USA	FEB-24
QID21048	USA	FEB-24

QID98794	USA	FEB-27
BCA87371	Japan	FEB-29
BCA87361	Japan	FEB-29
BCB15090	Japan	MAR-07
QII57338	USA	MAR-09
QII57328	USA	MAR-09
QII57318	USA	MAR-09
QII57308	USA	MAR-09
QII57298	USA	MAR-09
QII57288	USA	MAR-09
QII57278	USA	MAR-09
QII57268	USA	MAR-09
QII57258	USA	MAR-09
QII57248	USA	MAR-09
QII57238	USA	MAR-09
QII57228	USA	MAR-09
QII57218	USA	MAR-09
QII57208	USA	MAR-09
QII57198	USA	MAR-09
QII57188	USA	MAR-09
QII57178	USA	MAR-09
QII57168	USA	MAR-09
QIJ96523	USA	MAR-12
QIJ96513	USA	MAR-12
QIJ96503	USA	MAR-12
QIJ96493	USA	MAR-12
QIJ96483	USA	MAR-12
QIJ96473	USA	MAR-12
QIJ96463	USA	MAR-12
QIK02964	USA	MAR-13
QIK02954	USA	MAR-13
QIK02944	USA	MAR-13
QHU79173	Finland	MAR-17
QHD43416	China	MAR-18
QHO60594	USA	MAR-27
YP_009724390	China	MAR-30
QIG55994	Brazil	APR-06
QIE07481	China	APR-06
QIE07471	China	APR-06
QIE07461	China	APR-06
QIE07451	China	APR-06
QIA20044	China	APR-06
QIA98583	India	APR-06
QHS34546	India	APR-06
QIA98554	Italy	APR-06
QIB84673	Nepal	APR-06
QIC53204	Sweden	APR-06
QIK50417	Taiwan	APR-06
QIA98606	Taiwan	APR-06
QIA98596	Taiwan	APR-06
QIK50427	USA	APR-06
QII87840	USA	APR-06
QII87830	USA	APR-06
QII87818	USA	APR-06
QII87806	USA	APR-06
QII87794	USA	APR-06
QII87782	USA	APR-06
QIH55221	USA	APR-06
QIK50448	Viet Nam	APR-06
QIK50438	Viet Nam	APR-06

Eighty eight sequences were retrieved, 53 from USA (60.23%), followed by Chinese strains 17(19.3%), the rest from all over the world countries.

The sequences were subjected to VaxiJen to estimate their antigenicity at 0.4 threshold value. The value ranged from 0.4614 as lowest antigenicity value for Sweden strain (Acc # : QIC53204), to highest value 0.4687 for the Indian strain (QHS34546), while most of the rest (85.23%) having the value of 0.4646.

Then the sequences were subjected to multiple alignments using ClastW in order to construct phylogeny relationship using MEGA software, in general they aligned except those with changed antigenicity due to amino acids substitution as shown in Fig 1.

Fig 1: Relationship among all spike protein sequences.

The relation of Chinese strains is shown in Fig 2, as China was the origin of the virus

Fig 2: Relationship among spike protein sequences from China.

***** The first Wuhan Strain

Fig 3 shows the phylogeny of the USA strains represented the largest number of deposited sequences

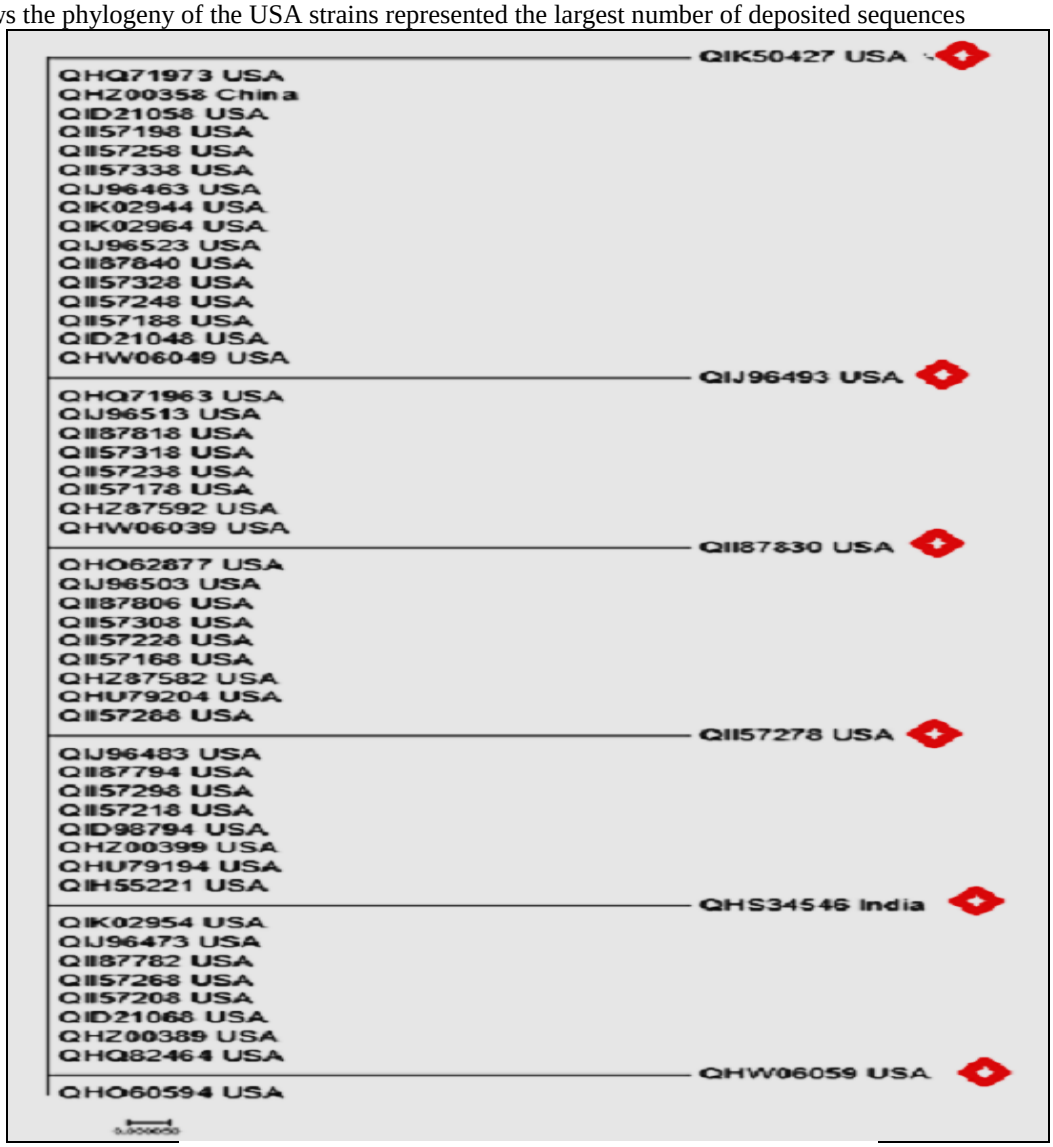
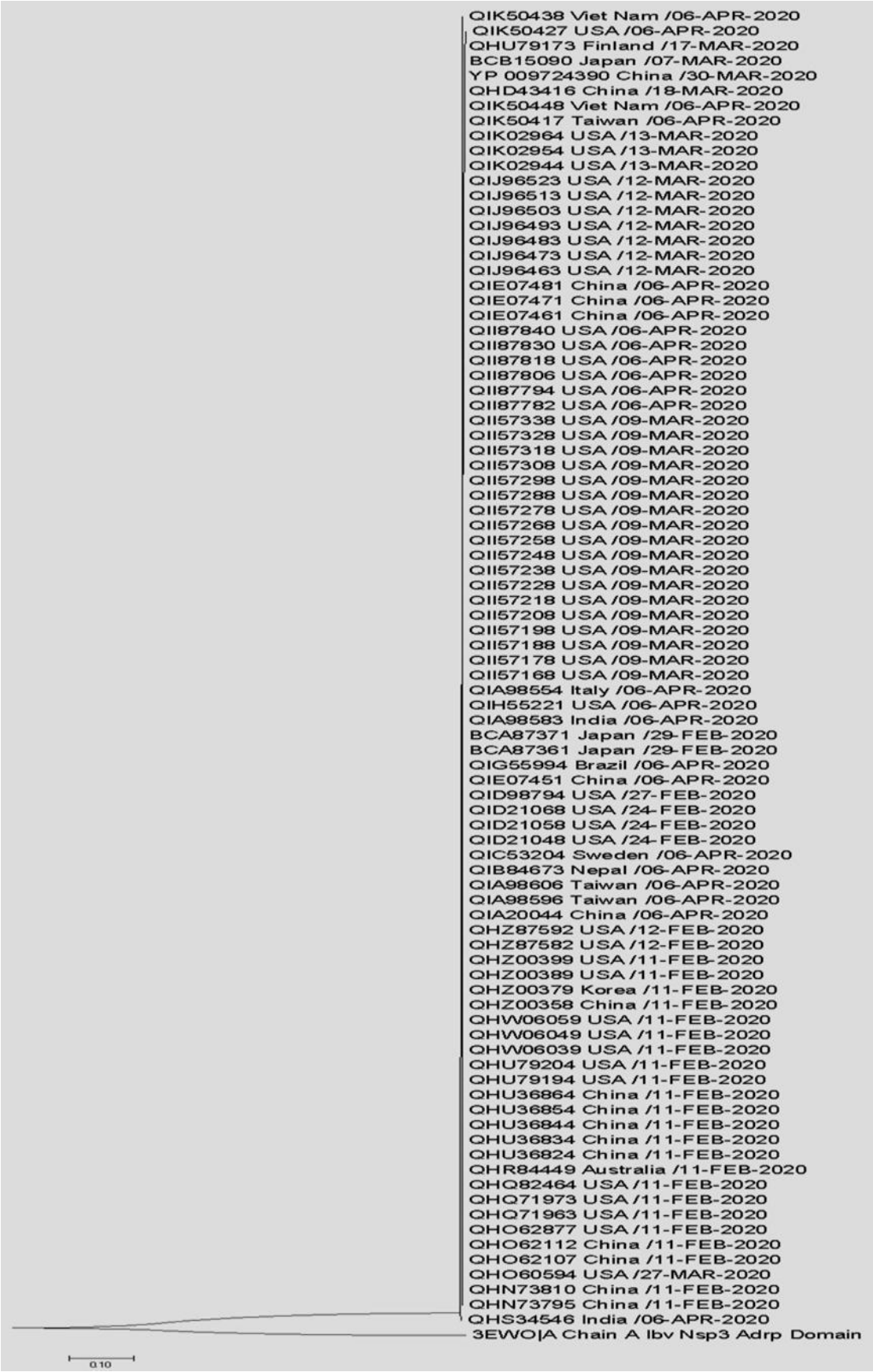


Fig 3: Relationship among spike protein sequences from USA

◇ Strains with altered antigenicity

In order to find the evolution of virus, the sequences were aligned with one IBV virus protein since the IBV belongs to corona virus / Gammacoronavirus as outgroup, results in Fig 4.



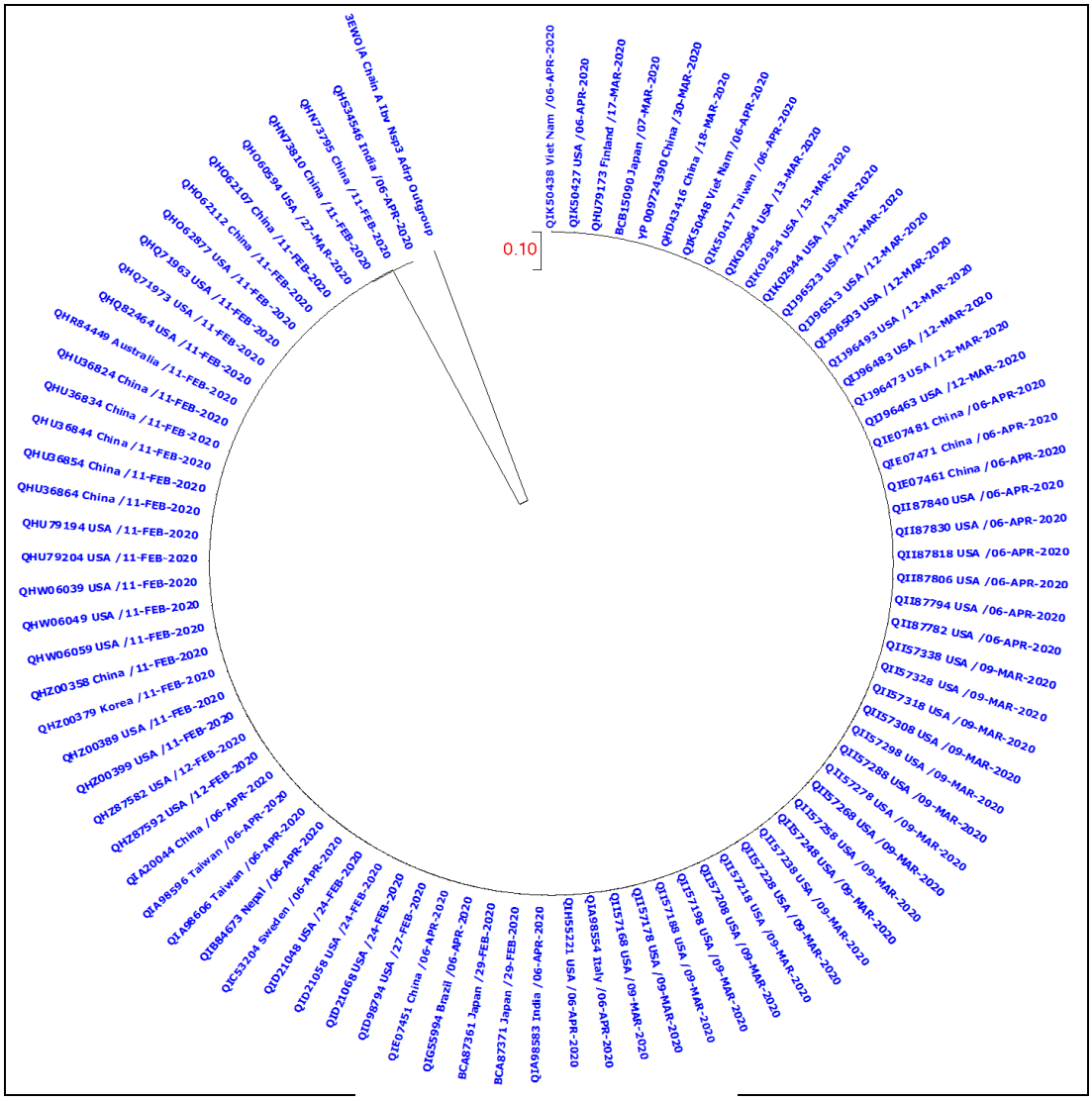


Fig 4: Evolution of spike proteins.

The results above show changes 10% of residues, These results in addition to above results clarify that there is changing in virus spike protein, although it is very slight (as in the Figure), this may be due to limited used sequences and short period of collection, this is with general trends about instability of COVID-19 genome, and most mutations occur in the outer membrane proteins such as spike protein.[33]

These mutations increase the sustainability and ability of the viruses to escape from immune system responses.[34] This virus i.e. COVID-19 diverse from other corona viruses by 75% at nucleotide sequences, and found that it has short insertions in the terminal domain and has 4 out of 5 key residues changes in the receptor binding motif in comparison with SARS-CoV.[21]

At this situation, the sequences were aligned to get consensus sequences to be used to derive different epitopes in an attempt to use bulk estimation. The consensus sequences obtained using EMBOSS Explorer, aligned with other sequences and subjected for phylogeny estimation to find its position among other sequences, the results shown in Fig 5.

Fig 5: Position of Cons sequence among other strains.

The antigenicity of this sequence is 0.4631 (Cons sequence), the resulted sequence was subjected to different studies, such as estimation of general properties using ExPASy ProtParam as shown in Table 2.

Table 2: Parameters of Cons sequence protein using ExPASy protParam server.

Character	Value
Molecular weight	141178.47
Theoretical pI	6.24
Total number of negatively charged residues (Asp + Glu)	110
Total number of positively charged residues (Arg + Lys)	103
instability index /stable	33.01
Aliphatic index	84.67
Grand average of hydropathicity (GRAVY):	-0.079

The isoelectric point (PI) was 6.24 suggests its negative existence, since value less than 7 means that the protein with negative charge, the protein is stable (instability index 33.01). The conserved domains/NCBI showed that the protein composed of two domains : Spike receptor binding domain (17-680) and S2 domain (681-1273), shown in Fig 6.

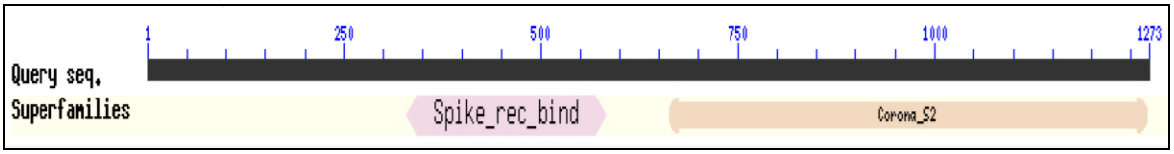


Fig 6: Domains of selected spike protein (Cons sequemce).

The resulted Cons sequence subjected to modelling using SwissModel, the best Template is 6vsb.1A (pdb ID) since sequence identity was 99.26%, followed by 6acc.1A (pdb ID), and 6acd.1A (pdb ID) at sequence identity 76.47%. Model 2 (6acc, pdb ID) was used for further studies depending on its correlation to the target protein (Spike protein), sequence similarity and model quality.[35] Secondary structure was estimated using PISPRED[36] shown in Fig 7.

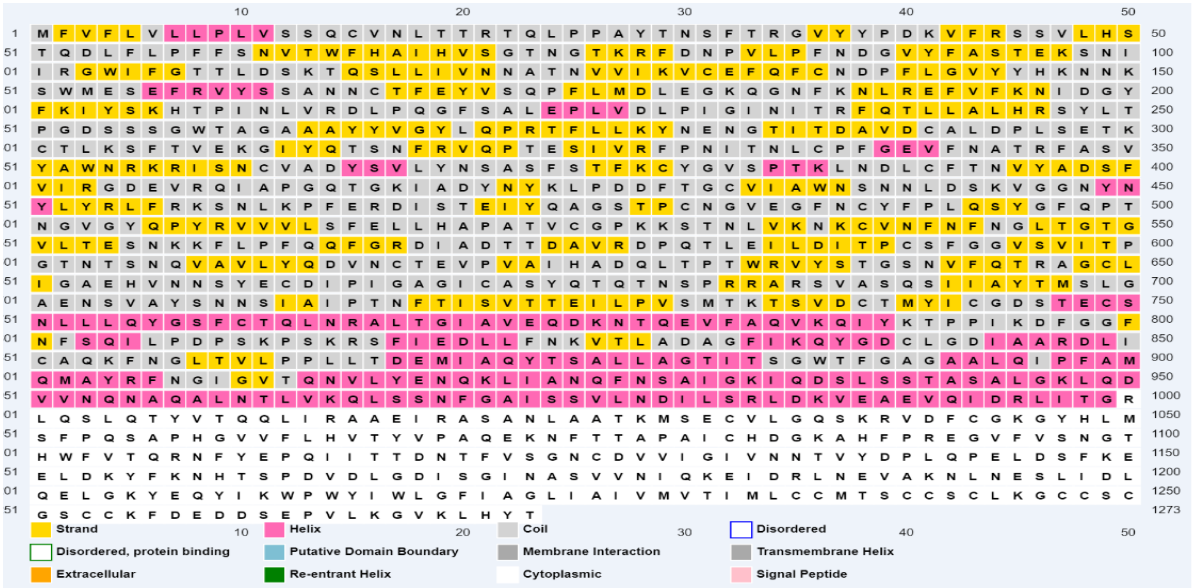


Fig 7: Secondary structure of Cons Sequence.

And 3D structure is shown in Fig 8

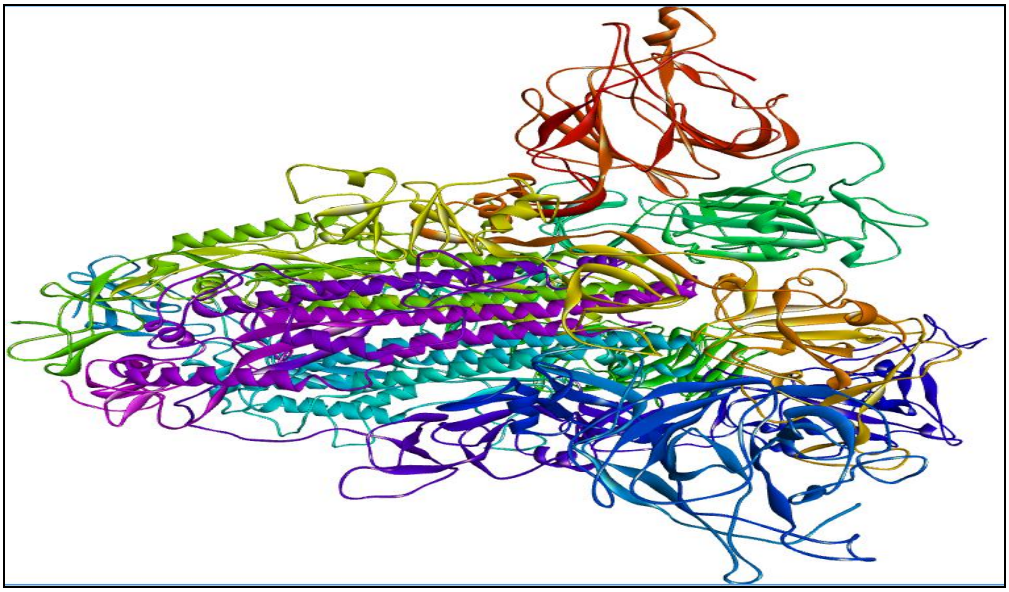


Fig 8: Tertiary (3D) structure of used protein (Cons sequence).

Ramachandran plot was estimated using Swiss-pdb viewer 4.1.0., the results shown in Fig 9

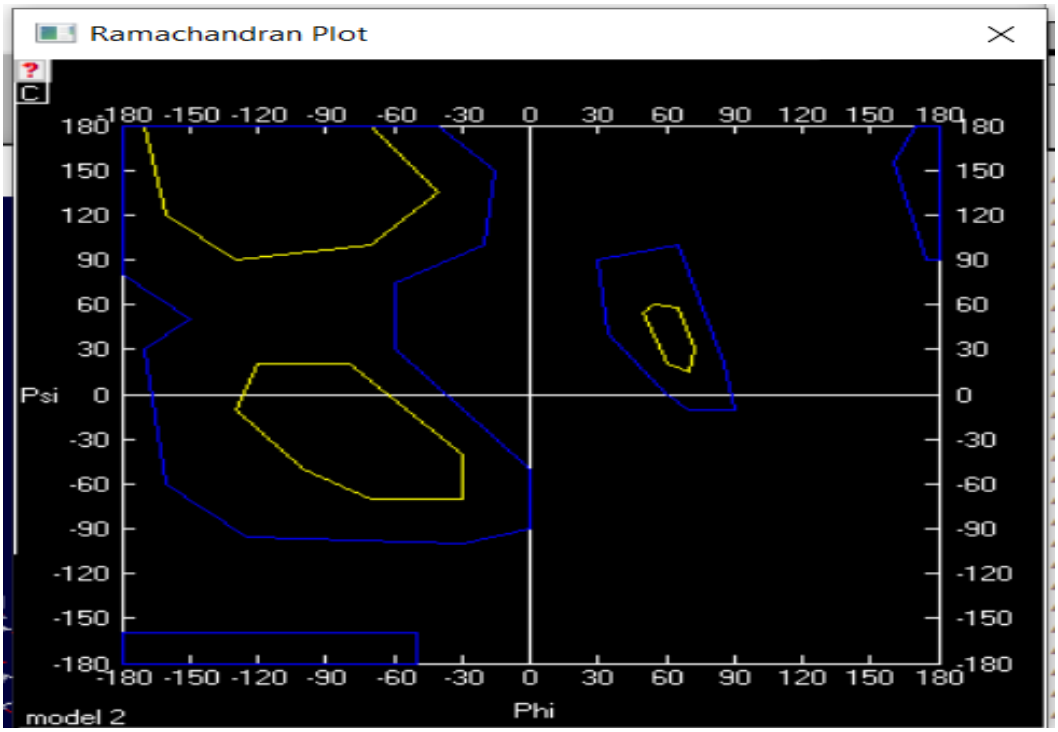
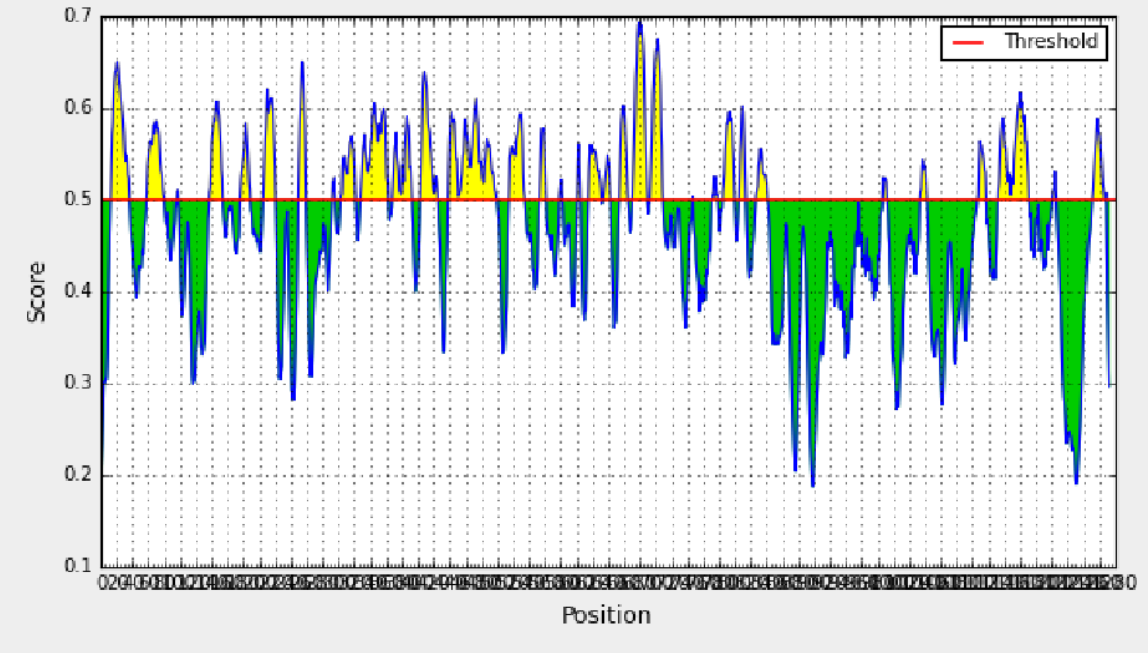


Fig 9: Ramachandran plot for residues distribution of Cons sequence.

Prediction of B-Epitopes

Peptide or epitope-based vaccines are specific and induce immunity against selected epitopes. Full-length S protein might considered as a good vaccine antigen as it could induce neutralizing antibodies preventing host cell attachment and infection, in addition, antiviral peptides against S2 may be considered as therapeutic candidates, the latter subunit of COVID-19 was found highly conserved and has 99% sequence similarity with other human coronaviruses.[37] Besides that B- and T-cell epitopes were found in viral nucleocapsid (N) protein as well.[3,38]

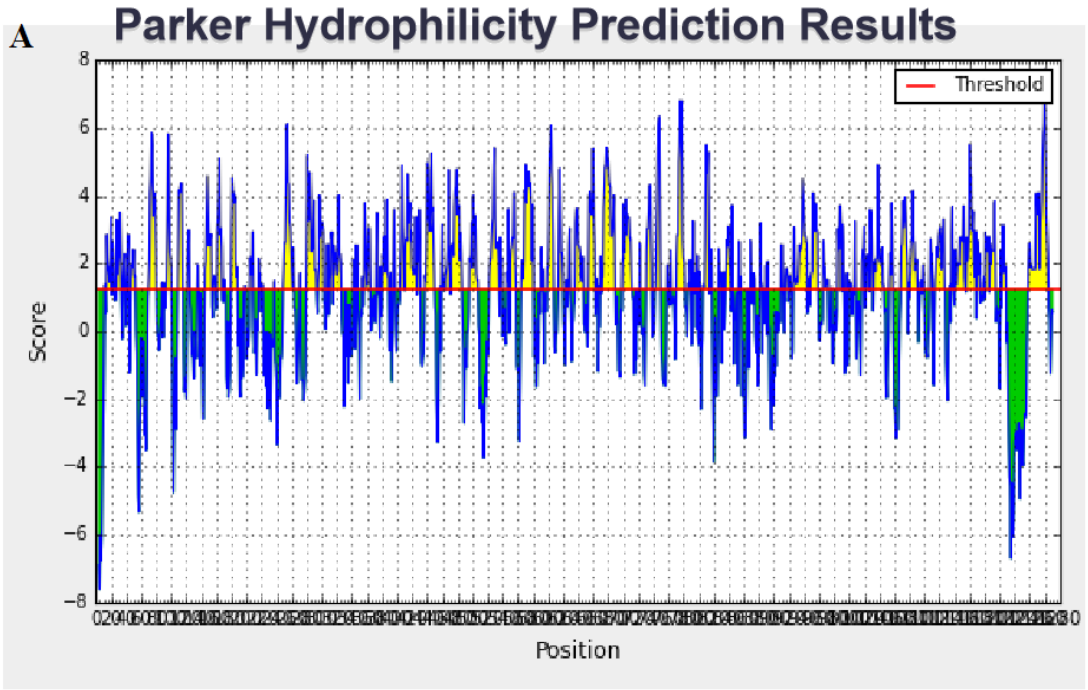
The aim of prediction of B cell epitopes was to find the potential antigens that would interact with B cells and initiate an immune response[39], linear B cell epitopes were predicted using Bepipred 2.0[40] tool from IEDB, the residues scoring above 0.5 were considered epitope, the results shown in Fig 10



Average: 0.470 **Minimum:** 0.183 **Maximum:** 0.695

Fig 10: Predicted B epitopes using Cons sequence in IEDB database.

The most important criteria for B epitopes are the hydrophilicity[16] and is usually in β turns regions[41], these were estimated using IEDB facilities at default parameters as shown in Fig 11A and 11B.



Average: 1.238 **Minimum:** -7.629 **Maximum:** 7.743

Fig 11 A: Hydrophilicity of predicted epitopes.

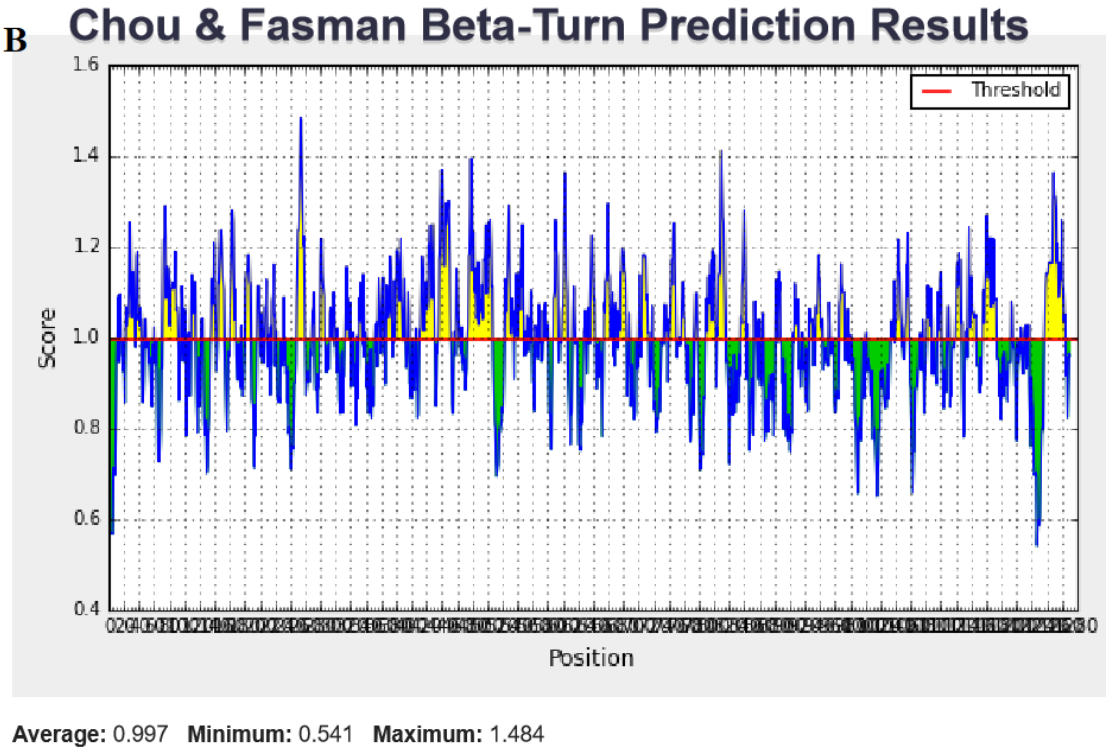


Fig 11 B: Chou and Fasman β – turns of predicted epitopes.

Several epitopes were resulted, each epitope was BLASTed with BLASTp using nr and RefSeq databases and expected value 0.05, those had similarity with human proteins were excluded to avoid autoimmune threats.[14] Toxicity was checked using ToxinPred Software, since certain peptides may induce toxic reactions when administrated by inhibiting certain biological process.[42] In addition, allergenicity was checked for epitopes and they found non-allergen using AllerCap, AllerTop.[43,44] The final accepted epitopes shown in Table 3, the epitopes were selected upon their length in addition to other criteria to avoid folding of long peptides which might exhibit reducing immune activity.[14]

Table 3: Selected B epitopes.

Epitope			Antigenicity	Allergenicity	Toxicity
Start	End	Amino acid Sequence			
177	189	MDLEGKQGNFKNL	1.2592	Non -Allergen	Non-Toxin
555	562	SNKKFLPF	1.3952	Non -Allergen	Non-Toxin
656	666	VNNSYECDIPI	0.6124	Non -Allergen	Non-Toxin
1035	1043	GQSKRVDFC	0.6792	Non -Allergen	Non-Toxin

The modeled structure of the selected epitopes was estimated using PEP-FOLD3 as in Fig 12

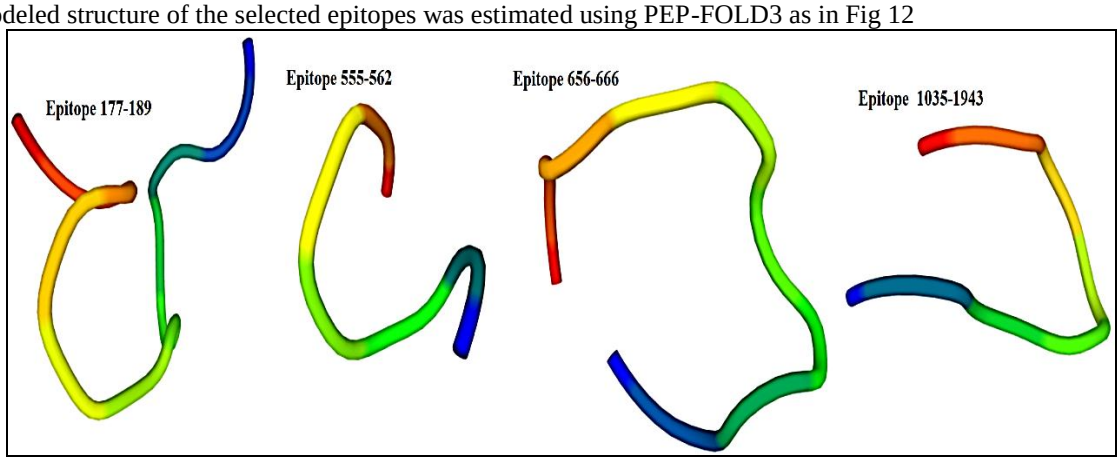


Fig 12: 3D structure of selected epitopes.

Fig 13 indicates the position of B epitopes on the protein surface to facilitates its accessibility.

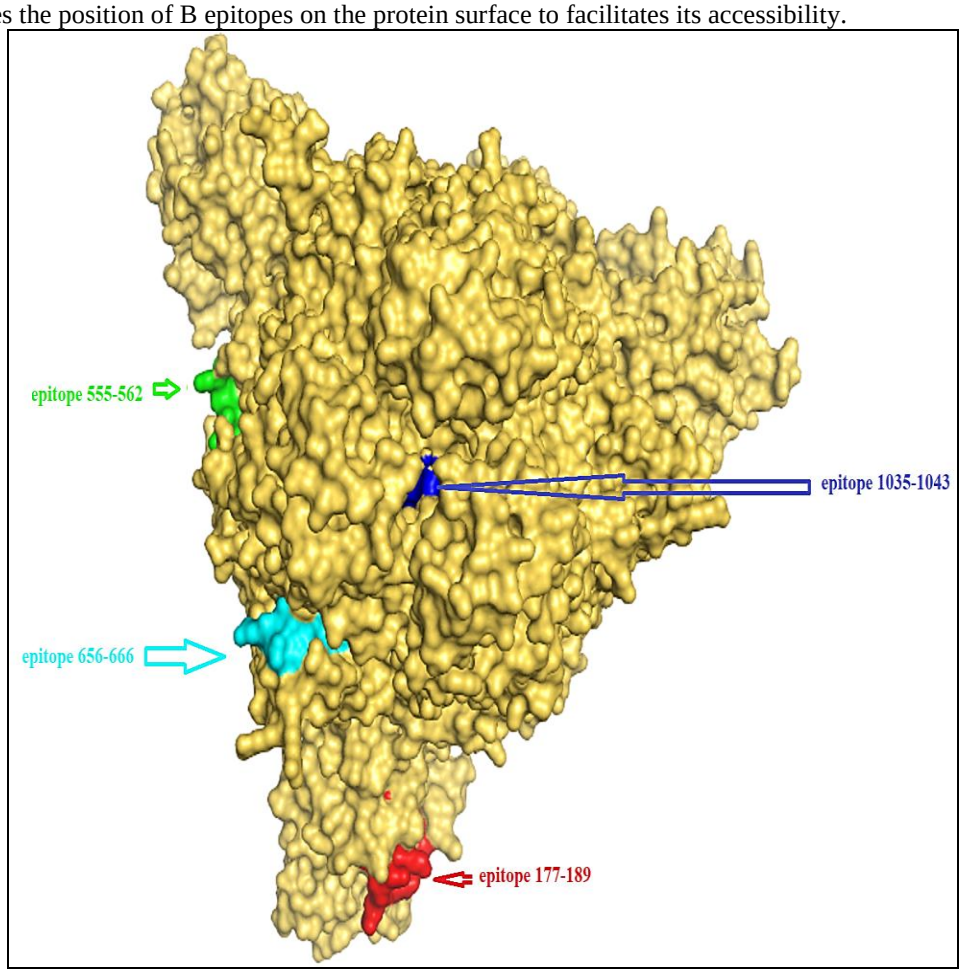


Fig 13: Position of selected epitopes on the surface of the protein.

Another important feature for the epitopes to be used for vaccination is their conservancy. IEDB was used for this task, the conservancy analysis was found to be 98.88-100% across all the sequences from different parts of the world in addition to the built cons sequence, the results in Table 4.

Table 4: Conservancy of selected epitopes across the all retrived protein sequences.

Epitope #	Epitope name	Epitope sequence	Epitope length	Percent of protein sequence matches at identity <= 100%	Minimum identity	Maximum identity
1	177-189	MDLEGKQGNFKNL	13	98.88% (88/89)	92.31%	100.00%
2	555-562	SNKKFLPF	8	100.00% (89/89)	100.00%	100.00%
3	656-666	VNNSYECDIPI	11	100.00% (89/89)	100.00%	100.00%
4	1035-1043	GQSKRVDFC	9	100.00% (89/89)	100.00%	100.00%

Generally speaking, epitope -based vaccine design is now become more popular and already has been established for some viruses and bacteria and tumors[45.46] with desired results. So it would be practical to identify the vaccine candidates through in silico approaches before being subjected to in vitro and in vivo confirmatory studies[47], i.e. the aspects of preclinical, clinical and post licensure vaccine enterprises.[3] It would be considered to be more effective than vaccines based on inactivated or live-attenuated viruses.[48] This simplified by vaccine informatics which is an emerging research area focusing on development and applications of Bioinformatics methods, this will ensure using only the parts of a protein that can elicit an immune response and unnecessary components that are potentially antigenic can be eliminated, thereby preventing complications.[49]

In addition the peptide - based vaccines have more advantages such as easy production and transportation[50], with high selectivity and can be used in multiple manner, this significantly reduce time, efforts

and is cost –effective.[16,35] IEDB offers multiple facilities in this aspect and offers selection of methods.[35]

In general there is a continuous emerging of new viruses which might attack human host with higher frequency, and this rapidly render existing drugs and vaccines ineffective.

Vaccine design strategies used to design vaccine against COVID-19 using conserved regions could generate immunity that perform cross protection to Betacoronaviruses and might be effective for ongoing virus evolution.

CONCLUSION

COVID-19 is changing and it will do in the future due to reasons explained previously, spike protein is the main part of the virus undergoes continuous changing in order to find appropriate hosts, and blocking this part i.e. (spike protein) using epitope vaccine could be one of the solutions. In this work, the whole protein sequence (S1 and S2)were used to derive epitopes depending on the fact that inactivation of one part or both of them could lead to inactivation of the virus, this could be safer and more effective than other types of vaccines, the idea beyond using full-length of spike protein, is the latter contains large number of antigenic epitopes, so it is the most potential target for vaccine design, since it has the ability to induce a faster and longer term mucosal immune response than other proteins.[51,52,53] There is a possibility to use multiepitope to cover arising problems. These results can be transferred from theoretical aspects into practical field, it might need some modifications such as using adjuvants to enhance the immune response, and using innovative methods for applications.

REFERENCES

1. Zhang W, Du RH, Li B, Zheng XS, Yang XL, Hu B. et al. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. Emerg Microb Infect, 2020; 9: 386–9.
2. Li G, Fan Y, Lai Y, Han T, Li Z, Zhou P. Coronavirus Infections and Immune Responses J Med Virol, 2020; 92: 424-32.
3. Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. A Sequence homology and Bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. Cell Host & Microbe, 2020; 27: 671–80.
4. Fahmi M, Kubota Y, Ito M. Nonstructural proteins NS7b and NS8 are likely to be phylogenetically associated with evolution of 2019-nCoV. Infect Genet Evol, 2020; 81.104272: 1-6.
5. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP.et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science, 2003; 300: 1394–99.
6. DimitrovDS. The secret life ofACE2 as a receptor for the SARSvirus. Cell, 2003; 115: 652–53.
7. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature, 2003; 426: 450–54.
8. Gu J, Korteweg C. Pathology and pathogenesis of severe acute respiratory syndrome. Am J Pathol, 2007; 170: 1136–47.
1. 9.Betacoronavirus ViralZone page. https: //viralzone.expasy.org/764?outline=all_by_species.
9. Ding Y, He L, Zhang Q, Huang Z, Che X, Hou J, et al. Organ distribution of severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV) in SARS patients: implications for pathogenesis and virus transmission pathways. J Pathol, 2004; 203: 622–30.
10. Jiang S, Lu L, Liu Q, Xu W, Du L. Receptor-binding domains of spike proteins of emerging or re-emerging viruses as targets for development of antiviral vaccines Emerg Microbes Infec, 2012; 1, e13: 1-9.
11. Lin S, Chen H. Infectious Bronchitis Virus variants: molecular analysis and pathogenicity investigation. Int. J. Mol. Sci, 2017; 18: 2-17.
12. Savić V. Emergence of a novel variant of avian infectious bronchitis virus in Croatia and subsequent recombination with 793B type. First Annual Meeting of COST Action FA1207: Towards Control of Avian Coronaviruses: Strategies for Diagnosis, Surveillance and Vaccination Potency Testing of Veterinary Vaccines: 19-21 N, 2013 Langen, Germany, 2013.
13. Nandy A, Basak S. A Brief review of computer-assisted approaches to rational design of peptide vaccines. Int. J. Mol. Sci, 2016; 17, 666: 1-11.
14. Bojin F, Gavriliuc O, Margineanu M, Paunescu V. Design of an epitope-based synthetic long peptide vaccine to counteract the novel China coronavirus (2019-nCoV). *Preprints*, 2020; 2020020102.
15. Oany AR, Emran AA, Jyoti TP. Design of an epitope-based peptide vaccine against spike protein of human corona virus: an in silico approach. Drug Des Devel Ther, 2014; 8: 1139–49.
16. Sigrist C, Bridge A, LeMercier P. A potential role for integrins in host cell entry by SARS-CoV-2. Antivir Res, 2020; 177: 104759.1-3.
17. Villanueva RA, Rouillé Y, Dubuisson J. Interactions between virus proteins and host cell membranes during the viral life cycle. Int Rev Cytol, 2005; 245: 171–244.
18. Belouzard, S. et al. Mechanisms of coronavirus cell entry mediated by the viral spike protein., Viruses, 2012; 4: 1011–33.
19. Li F, Li WH, Farzan M, Harrison SC. Structure of SARS coronavirus spike receptor binding domain complexed with receptor. Science, 2005; 309: 1864–8.
20. Zhou S, Wang Y, Zhu T, Xia L. CT Features of coronavirus disease 2019 (COVID-19) pneumonia

in 62 patients in Wuhan, China. Am J Roentgenol, 2020; 5: 1-8.
21. Jacob CO, Leitner M, Zamir A, Salomon D, Arnon R. Priming immunization against cholera toxin and *E. coli* heat-labile toxin by a cholera toxin short peptide-beta galactosidase hybrid synthesized in *E. coli*. EMBO J, 1985; 4: 3339-43.
22. Backert L, Kohlbacher O. Immunoinformatics and epitope prediction in the age of genomic medicine. Genome Med, 2015; 7, 119: 1-13.
23. Tripathy DN, Schnitzlein WM, Morris PJ, Janssen DL, Zuba JK, Massey G, et al. Characterization of poxviruses from forest birds in Hawaii. J Wildl Dis, 2000; 36: 225-30.
24. Bolte AL, Meurer J, Kaleta EF. Avian host spectrum of avipoxviruses. Avian Pathol, 1999; 28: 415-32.
25. Drexler I, Staib C, Kastenmuller W, Stevanovic S, Schmidt B, Lemonnier FA, et al. Identification of vaccinia virus epitope-specific HLA-A*0201-restricted T cells and comparative analysis of smallpox vaccines. Proc Natl Acad Sci U S A, 2003; 100: 217-22.
26. Abdelmageed M, Abdelmoneim A, Mustafa M, Elfadol NM, Naseem S. Murshed NS. et al. Design of multi epitope-based peptide vaccine against E protein of human COVID-19: An immunoinformatics approach. doi: bioRxiv preprint https: //doi.org/10.1101/2020.02.04.934232
27. Buchholz U, Bukreyev A, Yang L, Lamirande E, Murphy B, Subbarao K. et al. Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity. ProcNatlAcad Sci USA, 2004; 101: 9804–09.
28. Jagusztyn-Krynicka EK, Roszczenko P, Grabowska A. Impact of proteomics on anti-*Mycobacterium tuberculosis* (MTB) vaccine development, Pol J Microbiol, 2009; 58: 281–7.
29. He Y, Rappuoli R, DeGroot AS Chen RT. Emerging vaccine informatics, J Biomed Biotechnol, 2010; vol.2010:, ArticleID 218590, 26pages.
30. Du L, Kao RY, Zhou Y, He Y, Zhao G, Wong C. et al. Cleavage of spike protein of SARS coronavirus by protease factor Xa is associated with viral infectivity. Biochem Biophys Res Commun, 2007; 359: 174–9.
31. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science, 2020; 13, 367(6483): 1260-63.
32. Manzin A, Solforosi L, Petrelli E, Macarri G, Tosone G, Piazza M. et al. Evolution of hypervariable region 1 of hepatitis C virus in primary infection. J Virol, 1998; 72: 6271–76.
33. Christie JM, Chapel H, Chapman RW, Rosenberg WM. Immune selection and genetic sequence variation in core and envelope regions of hepatitis C virus. Hepatology, 1999; 30: 1037–44.
34. Prasasty V, Grazzolie L, Rosmalena R, Yazid F, Ivan F, Ernawati Sinaga E. Peptide-based subunit vaccine design of T- and B-cells multi-epitopes against Zika virus using immunoinformatics approaches. Microorganisms, 2019; 7, 226: 1-27.
35. McGuffin LJ[1], Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics, 2000; 16: 404-5.
36. Chan JF, Kok KH, Zhu Z, Chu H, To KK, Yuan S, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microbes Infect, 2020; 9: 221-36.
37. Ahmed A, Siman-Tov G, Hall G, Bhalla N, Narayanan A. Human antimicrobial peptides as therapeutics for viral infections. Viruses, 2019; 11, E704 .1-26.
38. Nair DT, Singh K, Siddiqui Z, Nayak BP, Rao KV, Salunke DM. Epitope recognition by diverse antibodies suggests conformational convergence in an antibody response. J Immunol, 2002; 168: 2371–82.
39. Jespersen MC, Peters B, Nielsen M, Marcatili P. Bepi Pred 2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res, 2017; 45: W24–W29.
40. Rose GD, Gierasch L, Smith JA. Turns in peptides and proteins. Adv Protein Chem, 1985; 37: 1–109.
41. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R. In silico approach for predicting toxicity of peptides and proteins. PLoS One, 2013; 13; 8(9): e73957.
42. Maurer-Stroh S, Krutz N, Kern NS, Gunalan V, Nguyen MN, Limviphuvadh V. et al. AllerCatPro—prediction of protein allergenicity potential from the protein sequence. *Bioinformatics*, 2019; 35: 3020–7.
43. Dimitrov I, Flower DR, Doytchinova I. AllerTOP - a server for *in silico* prediction of allergens. BMC Bioinformatics, 2013; 14(Suppl 6): S4 1-9.
44. Lapelosa M, Gallicchio E, Arnold GF, Arnold E, Levy RM. In silico vaccine design based on molecular simulations of rhinovirus chimeras presenting HIV-1 gp41 epitopes. J Mol Biol, 2009; 385: 675-91.
45. Chakraborty S1, Chakravorty R, Ahmed M, Rahman A, Waise TM, Hassan F, et al. A computational approach for identification of epitopes in dengue virus envelope protein: a step towards designing a universal dengue vaccine targeting endemic regions. In Silico Biol, 2010; 10: 235–46.
46. Ranjbar MM, Gupta SK, Ghorban K, Nabian S, Sazmand A, Taheri M, et al.(2015) Designing and modeling of complex DNA vaccine based on tropomyosin protein of Boophilus genus tick. Appl Biochem Biotechnol, 2015; 175: 323–39.
47. He YX, Li JJ, Heck S, Lustigman S, Jiang SB. Antigenic and immunogenic characterization of recombinant baculovirus-expressed severe acute respiratory syndrome coronavirus spike protein: implication for vaccine design. J Virol, 2006; 80: 5757–67.
48. Wang X, Xie G, Liao J, Yin D, Guan W, Pan M, et al.. Design and evaluation of a multiepitope

assembly peptide (MEAP) against herpes simplex virus type2 infection in BALB/c mice. Virol J, 2011; 8: 1-10.
49. Panahi HA, Bolhassani A, Javadi G, Noormohammadi Z. A comprehensive in silico analysis for identification of therapeutic epitopes in HPV16, 18, 31 and 45 oncoproteins. PLoS ONE, 2018; 13(10): e020593319
50. Ma C, Li Y, Wang L, Zhao G, Tao X, Tseng CT. et al. Intranasal vaccination with recombinant receptor-binding domain of MERS-CoV spike protein induces much stronger local mucosal immune responses than subcutaneous immunization: Implication for designing novel mucosal MERS vaccines. Vaccine, 2014; 32: 2100–08.
51. Yang ZY1, Kong WP, Huang Y, Roberts A, Murphy BR, Subbarao K.et al. A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. Nature, 2004; 428(6982): 561–64.
52. Agnihothram S, Gopal R, YountJr B, Donaldson E, Menachery V, Graham R. et al. Platform strategies for rapid response against emerging coronaviruses: MERS-CoV serologic and antigenic relationships in vaccine design. J Infect Dis, 2014; 209: 995-1006.