



## Prediction of B-cell and T-cell epitopes in the spike glycoprotein of SARS-CoV-2 in Bangladesh: An in-silico approach

Mehedi Hasan<sup>1</sup>, Md. Maksudur Rahman Shihab<sup>1</sup>, Mohammad Ariful Islam<sup>1\*</sup>

<sup>1</sup> Department of Microbiology, Faculty of Life and Earth Science, Jagannath University, Dhaka 1100, Bangladesh

### \*Corresponding author

Mohammad Ariful Islam, PhD  
Department of Microbiology,  
Faculty of Life and Earth Science  
Jagannath University, Dhaka 1100,  
Bangladesh  
Email: [ariful@mib.jnu.ac.bd](mailto:ariful@mib.jnu.ac.bd)

### Academic editor

Md Abdul Hannan, PhD  
Dongguk University, South Korea

### Article info

Received: 12 July 2020  
Accepted: 21 August 2020  
Published: 02 September 2020

### Keywords

COVID-19, SARS-CoV-2, B-cell epitope, T-cell epitope, Bangladesh

### ABSTRACT

To combat highly infectious Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), scientists and researchers are toiling hard globally to develop effective drugs and vaccines. By exploring the structural proteins of SARS-CoV-2 can be a feasible way to find an effective vaccine. In this study by using in-silico tools, we recommended B-cell and T-cell epitopes of spike protein from a Bangladeshi isolate which can be considered for incorporation into a vaccine against the SARS-CoV-2. Homology modelling, energy minimization process, and finally Ramachandran model was used for the prediction of a more stable conformation of the spike protein. The most important peptides were screened through the VaxiJen server followed by the IEDB server and CTLPred Score predicted and analysed the desired epitopes. In the final analysis, the peptide EVRQIAPGQTGKIADY (starting from 91) showed the highest antigenicity score (1.3837) as a B-cell epitope although GSTPCNGVEGFNCYFP, starting at 161, showed highest score (0.91) in an initial analysis. On the contrary, as a T-cell epitope, 71 KLNDLCFTNV- 80 was found with the highest antigenicity score (2.6927) which was also found as an epitope in further analysis. A combination of B-cell and T-cell epitopes may evoke a humoral and cell-mediated immune response which will possibly lead to an effective vaccine. Further, the various computational analyses will provide valuable information that will pave the way for modelling a novel vaccine against SARS-CoV-2.

### INTRODUCTION

Coronavirus disease 2019 (COVID-19) pandemic is the consequence of the respiratory tract infection caused by novel Coronavirus (2019-nCoV) which is later known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). It has emerged on the eve of 2020 from the Chinese city of Wuhan (Hubei province) [1]. As of June 15, 2020, the pandemic has globally claimed more than 436,900 lives and infected over 8.03 million individuals [2]. Therefore, the development of a vaccine is no longer merely a discussion or part of a debate about whether a vaccine is ultimately needed to suppress the viral spread and to control and prevent SARS-CoV-2. As of June 29, 2020, seventeen candidate vaccines were in clinical evaluation across the globe whereas other 132 candidates in preclinical stages [3].

SARS-CoV-2, which is a positive-sense single-stranded RNA virus, has a genome size of approximately 30 kilobases which encode for various structural and non-structural proteins. The structural proteins of SARS-CoV-2 are the spike (S) protein, the envelope (E) protein, the membrane (M) protein, and the nucleocapsid (N) protein and the non-structural proteins contain open reading frame 1ab (ORF1ab), ORF3a, ORF6, ORF7a, ORF8, and ORF10 [4, 5]. The S protein has two major domains: S1 and S2 [6]. S1 subunit has the receptor-binding domain (RBD) and binds to angiotensin-converting enzyme 2 (ACE2) while S2 mediates the fusion of viral and host cell membranes [7, 8].

The first laboratory-confirmed case in Bangladesh was identified on March 8, 2020 [9]. So far, 90,619 people have been tested positive and 1,209 have died in the

country (as of June 15, 2020) [10]. Genomic data of SARS-CoV-2 in Bangladesh is limited, however, efforts from scientists here are underway to make genome sequences available gradually [11]. On May 11, 2020, the first whole-genome sequence (EPI\_ISL\_437912) from Bangladesh was published. A few days later, the study sequence (EPI\_ISL\_447590) was made publicly available by the Genome Centre, Jashore University of Science and Technology, Jashore, Bangladesh.

To develop an effective vaccine, the identification of B-cell and T-cell epitopes for SARS-CoV-2 proteins are critical especially for S protein. Both humoral immunity and cellular immunity, provided by B-cell antibodies and T-cell respectively, are essential for effective vaccines [12, 13]. By analyzing an S protein of SARS-CoV-2 from Bangladesh, we identified and reported the B-cell and T-cell epitopes. Therefore, the epitopes presented in the study may evoke an effective immune response against SARS-CoV-2 for which further experimental analysis will be required. We tend to target the epitopes within the S protein as the protein gave a dominant and long-lasting immune response against SARS-CoV that was previously reported [4]. However, our analyses will probably slender down the rummage around for targets for designing a novel vaccine candidate against SARS-CoV-2.

## MATERIALS AND METHODS

### Nucleotide sequence retrieval and structure prediction

The nucleotide sequence (EPI\_ISL\_447590) of a Bangladeshi SARS-CoV-2 spike (S) glycoprotein was downloaded from the GISAID EpiCoV™ database. The nucleotide sequence was then translated into an amino acid sequence for further analysis. The protein sequence was then applied to form its 3-D model by using SWISS-MODEL [14]. The SMTL ID: 6vyb.1 (<https://swissmodel.expasy.org/templates/6vyb.1>), chain A, was used as a template that shared 100% sequence identity with the study sequence. The 6vyb.1, chain A, is a SARS-CoV-2 spike ectodomain structure (open state) solved experimentally using Electron Microscopy deposited at early this year in the protein data bank [15]. A PDB file was found after the modeling which was then used for further prediction. Moreover, we minimized energy after homology modeling to find out a stable conformation of proteins using FoldX (<http://foldxsuite.crg.eu/>). Ramachandran model was also used for finding stable conformation. Ramachandran plot that we used here to determine torsional angles which are permitted, so we could understand the insight into the structure of the peptides, ultimately stable conformation of the peptides. In the Ramachandran plot, the outlier with

percentage was measured by SWISS-PDB. From the outliers, the conformations of the phi ( $\Phi$ ) and psi ( $\Psi$ ) angles were predicted in the protein. Moreover, empirical distribution was observed in this single structure. Ultimately protein stability was confirmed by this value.

### Prediction of B-cell epitope

The linear and discontinuous B-cell epitopes based on protein's 3-D structure were predicted using Ellipro [16]. The prediction parameters were a minimum score and a maximum distance of 0.5 and 6, respectively. The number of residues of each chain was predicted by Ellipro. For identifying B-cell epitopes in an antigen sequence, ABCpred was used to identify the peptide sequence [17]. The protein sequence was submitted where the threshold was 0.51 with a length of 16 to be used for prediction. The antigenicity index of the predicted epitopes was examined and for predicting and designing Interferon-gamma (IFN- $\gamma$ ) inducing peptides, a combination of VaxiJen v2.0 with a threshold of 0.4 and IFN- $\gamma$  response were used [18, 19].

### Prediction of T-cell epitope

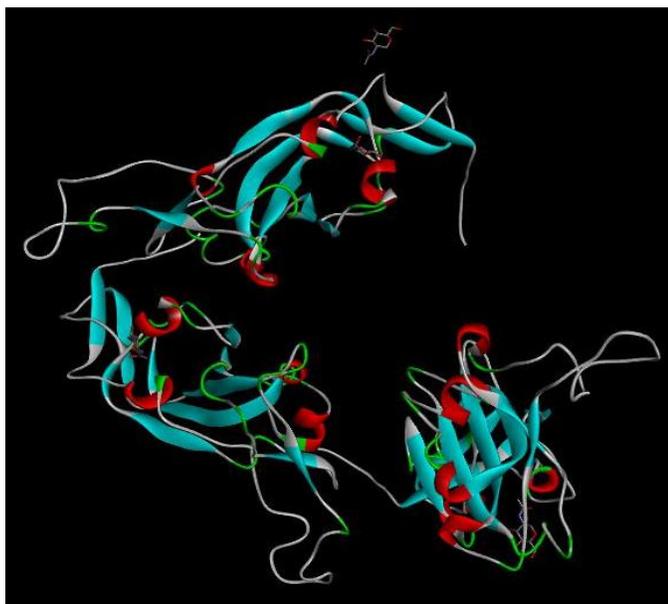
The T-cell epitope prediction was performed by using the IEDB tool where all MHC alleles were taken with a threshold of 0.7 for the study [20-22]. The output peptides were again examined by using a combination of Vixen, IFN- $\gamma$  response, and CTLPred which scores were based on using both support vector machine (SVM) and artificial neural network (ANN) score [23]. In CTLPred, ANN and SVM cut-off scores were used as of 0.51 and 0.36, respectively.

## RESULTS

### Nucleotide sequence retrieval and structure prediction

The retrieved nucleotide sequence from the GISAID EpiCoV™ database was found to have a length of 694 base pairs. The converted protein sequence was then used for modeling (Figure 1) which showed a 100% sequence identity with template sequence. After conducting the energy minimization process, we got energy minimized state of that protein (Table 1) and a new 3-D figure of the S protein (Figure 2).

From the Ramachandran model (Figure 3), the stable conformation of the PDB file was confirmed with Ramachandran Favoured, and Ramachandran Outliers of 92.58%, and 1.75%, respectively.



**Figure 1.** Homology modeling of spike protein generated by SWISS-MODEL.



**Figure 2.** 3-D model of the spike protein after energy optimization generated by using FoldX.

### B-cell epitope prediction

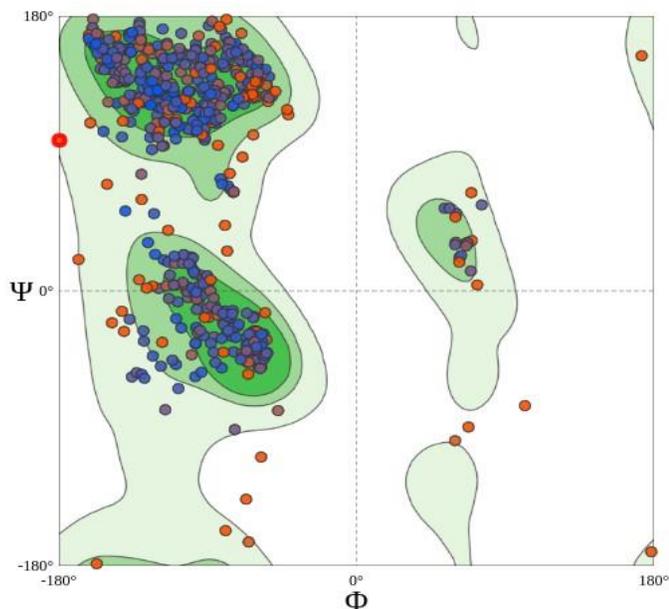
The PDB file was used to identify linear and discontinuous epitopes by using Ellipro. This revealed the presence of B-cell epitopes according to their chain ID. It showed 3 chain IDs with their number of residues, 231 equally (Figure 4).

The linear peptides sequence with a threshold of 0.51 provided 23 sequences at a specific length of 16. The antigenicity of peptides was confirmed by the score (Table 2).

For more confirmation, combined processes were done to identify antigen in the spike glycoprotein (Table 3). IFN- $\gamma$  responses can be either positive or negative depending on the cytokines. The probability of being an antigen of specific peptides was confirmed by VaxiJen with a threshold of 0.4.

**Table 1.** Energy minimization of spike protein sequence to find out a stable conformation of protein.

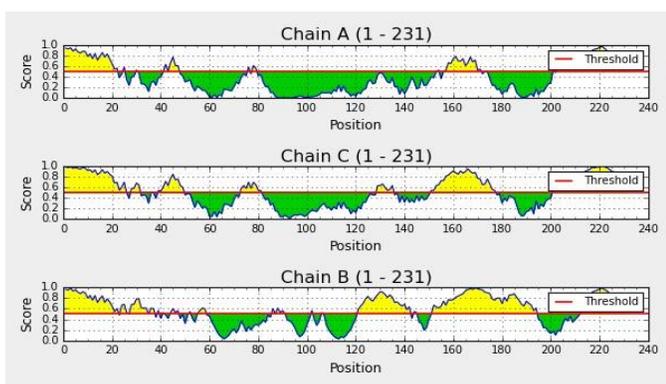
Parameters	Before Optimization	After Optimization
BackHbond	-292.19	-294.70
SideHbond	-96.10	-99.56
Energy_VdW	-703.44	-703.06
Electro	-27.33	-27.33
Energy_SolvP	962.96	960.58
Energy_SolvH	-908.69	-908.52
Energy_vdwclash	82.40	74.35
Energy_torsion	16.69	15.62
Backbone_vdwclash	235.57	235.46
Entropy_sidec	338.98	339.36
Entropy_mainc	1037.52	1035.73
Water bonds	0.00	0.00
Helix dipole	-2.17	-2.17
Loop_entropy	0.00	0.00
Cis_bond	0.00	0.00
Disulfide	-38.47	-38.47
Kn electrostatic	-0.25	-0.25
Partial covalent interactions	0.00	0.00
Energy_Ionisation	0.17	0.15
Entropy Complex	0.00	0.00
Total	370.09	351.73



**Figure 3.** Ramachandran plot using SWISS-MODEL.

### Prediction of T-cell epitope

In spike glycoprotein, a total of 27 MHC-I peptides were predicted at different amino acid positions. However, 12 peptide sequences were found to be antigenic in VaxiJen. IFN- $\gamma$  responses can be either positive or negative. MHC restriction refers that a particular MHC molecule bounded with the peptides that can be interacted with T cell epitopes. CTLPred score confirms the epitopes probability by combining both support vector machine (SVM) and artificial neural network (ANN) score (Table 4).



**Figure 4.** Predicted linear B-cell epitopes and their locations to the surface membrane (yellow) by Ellipro.

**Table 2.** Predicted antigenic B-cell linear epitopes found in the spike glycoprotein with their positions and antigenicity scores.

Sequence	Start Position	Score
GSTPCNGVEGFNCYFP	161	0.91
LQSYGFQPTNGVGYQP	177	0.90
TEIYQAGSTPCNGVEG	155	0.89
CGPKKSTNLVKNKCVN	210	0.86
FERDISTEIQAGSTP	149	0.86
EVRQIAPGQTGKIADY	91	0.85
CFTNVYADSFVIRGDE	76	0.85
TGKIADYNYKLPDDFT	100	0.84
FPNITNLCPFGEVFNA	14	0.82
FASVYAWNRKRISNCV	32	0.81
VQPTEISIVRFPNITNL	5	0.80
NCVADYSVLYNSASF	45	0.79
NGVGYQPYRVVLSFE	186	0.77
FSTFKCYGVSPKLN	59	0.76
SVLYNSASFSTFKCYG	51	0.74
EGFNCYFPLQSYGFQP	169	0.73
VGGNYNYLRLFRKSN	130	0.71
TGCVIAWNSNLDISKV	115	0.71
YKLPDDFTGCVIAWNS	108	0.71
FVIRGDEVRQIAPGQT	85	0.69
VVLSFELLHAPATVCG	196	0.66
RKSNLKPFERDISTEI	142	0.65
PTKLNLCFTNVYADS	69	0.62

**Table 3.** Predicted antigenic B-cell linear epitopes found in the spike glycoprotein with their antigenic scores, VaxiJen scores, and IFN- $\gamma$  responses

Sequence	Start Position	Score	VaxiJen Score	IFN- $\gamma$ Response
LQSYGFQPTNGVGYQP	177	0.90	Antigen (0.5258)	Positive (14)
EVRQIAPGQTGKIADY	91	0.85	Antigen (1.3837)	Negative (1)
TGKIADYNYKLPDDFT	100	0.84	Antigen (0.9642)	Negative (-0.334)
FPNITNLCPFGEVFNA	14	0.82	Antigen(0.6058)	Negative (-0.138)
VQPTESIVRFPNITNL	5	0.80	Antigen (0.4454)	Positive (7)
NGVGYQPYRVVLSFE	186	0.77	Antigen (0.9736)	Positive (0.371)
FSTFKCYGVSPKLNLD	59	0.76	Antigen (0.9650)	Positive (1)
EGFNCYFPLQSYGFQP	169	0.73	Antigen (0.6870)	Positive (0.0261)
TGCVIAWNSNNLDSKV	115	0.71	Antigen (0.5086)	Positive (0.101)
FVIRGDEVQRQIAPGQT	85	0.69	Antigen (0.5147)	Positive (2)
VVLSFELLHAPATVCG	196	0.66	Antigen (0.6184)	Positive (0.395)
RKSNLKPFERDISTEI	142	0.65	Antigen (0.4298)	Positive (0.332)
PTKLNLDLFTNVYADS	69	0.62	Antigen (1.2303)	Negative (-0.426)

**Table 4.** MHC-I associated antigenic peptide predicted on the spike glycoprotein.

Peptide	VaxiJen	IFN- $\gamma$ Response	CTLPred Score (ANN/SVM)	MHC Restriction
174- YFPLQSYGF-182	Antigen (0.5107)	Negative (-0.2288)	-	HLA-A*24:02, HLA-A*23:01
102- KIADYNYKL- 110	Antigen (1.6639)	Negative (-0.3740)	0.72/0.74005 (Epitope)	HLA-A*02:01, HLA-A*02:06, HLA-A*32:01, HLA-A*02:03
71- KLNLDLFTNV- 80	Antigen (2.6927)	Negative (-0.8044)	0.03/0.67438 (Epitope)	HLA-A*02:03, HLA-A*02:01
93-RQIAPGQTGK-102	Antigen (1.7893)	Negative (-0.1924)	0.48/0.83303 (Epitope)	HLA-A*03:01
125 -NLDSKVGNGY -134	Antigen (0.7882)	Negative (-0.1738)	0.02/0.1194 (Non-epitope)	HLA-A*01:01
143 -KSNLKPFER - 151	Antigen (0.9490)	Negative (-0.3191)	0.82/1.07575 (Epitope)	HLA-A*31:01
34 -SVYAWNRKR -42	Antigen (0.7650)	Positive (0.2458)	0.27/0.15625343 (Non-epitope)	HLA-A*31:01, HLA-A*68:01
192 -PYRVVLSF - 200	Antigen (1.0281)c	Negative (-0.018)	0.49/0.10768 (Non-epitope)	HLA-A*23:01, HLA-A*24:02

29-ATRFASVYA -38	Antigen (0.0785)	Negative (-0.020)	0.86/0.24604 (Epitope)	HLA-A*30:01, HLA-B*58:01, HLA-B*57:01
98- GQTGKIADY -108	Antigen (1.4019)	Negative (-0.206)	-	HLA-B*15:01, HLA-A*01:01
200- FELLHAPATV-209	Antigen (0.5982)	Negative (-0.135)	0.00/0.84172 (Epitope)	HLA-A*02:03
129 -KVGGNVNYLY- 138	Antigen (0.5338)	Positive (0.393)	1.00/1.07922 (Epitope)	HLA-A*30:02

## DISCUSSION

To control the current pandemic, the development of new drugs and vaccines is an urgent issue. Especially, effective vaccination or immunotherapy could play a remarkable role in suppressing the viral spread and eventually eliminating it from humans. In the case of vaccine development, a challenge is the absence of sufficient data concerning specific immune responses against SARS-CoV-2.

This study has recommended some epitopes that may have the ability to evoke sufficient response in the human body and to be incorporated into a novel vaccine. The 3-D structure of spike (S) protein derived from a Bangladeshi SARS-CoV-2 strain was firstly determined, minimized energy, and confirmed by the Ramachandran model followed by the probable antigenic B-cell, and T-cell epitopes prediction using the amino acid sequence of the same protein.

Recently, a report unveiled that T-cell epitopes are more conducive as a result of more long-lasting immune response mediated by CD8<sup>+</sup> T-cell and due to the antigenic drift, by which an antibody is not able to respond against an antibody [24]. In this study, we predicted 12 peptides as T-cell epitopes in a combined strategy (VaxiJen score and IFN- $\gamma$  responses). All of the peptides were antigenic whereas 71- KLNDLCFTNV- 80 showed the highest antigenicity score (2.6927) though the IFN- $\gamma$  response was negative which mainly depends on cytokines and can be either positive or negative. In another analysis (CTLPred score), the peptide was also found as an epitope with support vector machine (SVM) and artificial neural network (ANN) score was 0.03 and 0.67438, respectively. Besides these, all peptides were able to interact with the MHC class I alleles.

In the case of SARS-CoV, antibodies were generated in mice against spike protein that gave protection from the infection [25-27]. Besides, the B-cell epitope provides a strong immune response without causing any adverse effects [28]. Therefore, we also predicted 23 linear B-cell epitopes of 16-mer which could be effective. Among these peptides, the peptide GSTPCNGVEGFNCYFP

with start position 161 showed the highest antigenicity score (0.91). But further analysis (VaxiJen score and IFN- $\gamma$  responses) revealed only 13 peptides and among them, the peptide EVRQIAPGQTGKIADY, started from 91, showed the highest antigenicity score (1.3837) with negative (1) IFN- $\gamma$  responses.

The potential candidate epitopes, reported in this study, may contribute to the development of a novel vaccine against SARS-CoV-2. A large number of Bangladeshi individuals can be covered by these peptides. The limitation of the study is that all the analyses here were based on a single sequence of SARS-CoV-2 derived from Bangladesh. More mutations will be observed as the virus is evolving continuously which may affect our present analysis although most of the mutations are synonymous. However, further experimental analysis will also be required to prove the immunogenicity of the recommended peptides. Overall, the study provides insight that will eventually contribute to the development of a vaccine against SARS-CoV-2.

## ACKNOWLEDGEMENT

This research received no external funding. We are grateful to the authors and the laboratory from where the study sequence was originated and submitted followed by sharing through the GISAID EpiCoV<sup>TM</sup> database. Accession ID of the study sequence: EPI\_ISL\_447590 Originating and submitting laboratory: Genome Centre, Jashore University of Science and Technology, Jashore-7408, Bangladesh. Authors: A. S. M. Rubayet Ul Alam, M. Rafiul Islam, M. Shaminur Rahman, Md. Tanvir Islam, Md. Shazid Hasan, Pravas Chandra Roy, Habiba Ibnat, MD. Ali Ahasan Setu, Tanay Chakrovarty, Sourav Dutta Dip, Ruhul Amin, Md Nur Kabidul Azam, Ovinu Kibria Islam, Hassan M. Al-Emran, Shireen Nigar, Selina Akter, Md. Nazmul Hasan, Iqbal Kabir Jahid, M. Anwar Hossain.

## AUTHOR CONTRIBUTIONS

MAI conceived the idea and supervised the project. MH performed the database search. MMRS contributed to perform the experiments. MH analyzed data and wrote the manuscript. MH and MMRS illustrated the figures and prepared the tables. MAI critically revised the manuscript. All the authors proofread and approved the final manuscript.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## REFERENCES

- [1] Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *The Lancet*. 2020;395(10223):470-3.
- [2] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*. 2020;20(5):533-4.
- [3] World Health Organization (WHO). DRAFT landscape of COVID-19 candidate vaccines – 29 June 2020. Retrieved June 30, 2020, from <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines>.
- [4] Ahmed SF, Quadeer AA, McKay MR. Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies. *Viruses*. 2020;12(3):254.
- [5] Kiyotani K, Toyoshima Y, Nemoto K, Nakamura Y. Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. *Journal of Human Genetics*. 2020;65(7):569-75.
- [6] Ortega JT, Serrano ML, Pujol FH, Rangel HR. Role of changes in SARS-CoV-2 spike protein in the interaction with the human ACE2 receptor: An in silico analysis. *EXCLI journal*. 2020; 19:410.
- [7] Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is 278 Blocked by a Clinically Proven Protease Inhibitor. *Cell*. 2020. doi: 10.1016/j.cell.2020.02.052.
- [8] Verdecchia P, Cavallini C, Spanevello A, Angeli F. The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. *European Journal of Internal Medicine*. 2020. doi: 10.1016/j.ejim.2020.04.037.
- [9] Hossain I, Khan MH, Rahman MS, Mullick AR, Aktaruzzaman M. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in Bangladesh: A descriptive study. *Journal of Medical Science and Clinical Research*. 2020;8(04).
- [10] WHO Bangladesh COVID-19 Situation Report #16, 15 June 2020. Retrieved June 18, 2020, from <https://www.who.int/bangladesh/emergencies/coronaviru>s-disease-(covid-19)-update/coronavirus-disease-(covid-2019)-bangladesh-situation-reports.
- [11] Hasan S, Khan S, Ahsan GU, Hossain MM. Genome Analysis of SARS-CoV-2 Isolate from Bangladesh. *BioRxiv*. 2020. doi:10.1101/2020.05.13.094441.
- [12] Olsson S-E, Villa LL, Costa RL, Petta CA, Andrade RP, Malm C, et al. Induction of immune memory following administration of a prophylactic quadrivalent human papillomavirus (HPV) types 6/11/16/18 L1 virus-like particle (VLP) vaccine. *Vaccine*. 2007;25(26):4931-9.
- [13] Rappuoli R, Black S, Bloom DE. Vaccines and global health: In search of a sustainable model for vaccine development and delivery. *Science Translational Medicine*. 2019;11(497): eaaw2888.
- [14] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*. 2018;46(W1): W296-W303.
- [15] Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*. 2020. doi: 10.1016/j.cell.2020.02.058.
- [16] Ponomarenko J, Bui H-H, Li W, Fusseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC bioinformatics*. 2008;9(1):514.
- [17] Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins: Structure, Function, and Bioinformatics*. 2006;65(1):40-8.
- [18] Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. *Biology direct*. 2013;8(1):30.
- [19] Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28(12):1647-9.
- [20] Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*. 2016;32(4):511-7.
- [21] Lundegaard C, Lund O, Nielsen M. Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics*. 2008;24(11):1397-8.
- [22] Lundegaard C, Nielsen M, Lund O. The validity of predicted T-cell epitopes. *Trends in biotechnology*. 2006;24(12):537-8.
- [23] Bhasin M, Raghava G. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine*. 2004;22(23-24):3195-204.
- [24] Chiou S-S, Fan Y-C, Crill WD, Chang R-Y, Chang G-JJ. Mutation analysis of the cross-reactive epitopes of Japanese encephalitis virus envelope glycoprotein. *Journal of general virology*. 2012;93(6):1185-92.

- [25] Deming D, Sheahan T, Heise M, Yount B, Davis N, Sims A, et al. Vaccine efficacy in senescent mice challenged with recombinant SARS-CoV bearing epidemic and zoonotic spike variants. *PLoS medicine*. 2006;3(12).
- [26] Graham RL, Becker MM, Eckerle LD, Bolles M, Denison MR, Baric RS. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nature medicine*. 2012;18(12):1820.
- [27] Yang Z-y, Kong W-p, 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):561-4.
- [28] Rakib A, Sami SA, Islam MA, Ahmed S, Faiz FB, Khanam BH, et al. Epitope-Based Peptide Vaccine Against Severe Acute Respiratory Syndrome-Coronavirus-2 Nucleocapsid Protein: An in silico Approach. *BioRxiv*. 2020. doi: 10.1101/2020.05.16.100206.



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.