Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Immuninformatics-guided designing of epitope-based subunit vaccines against the SARS Coronavirus-2 (SARS-CoV-2)

Bishajit Sarkara,⁎, Md. Asad Ullaha, Fatema Tuz Johoraa, Masuma Afrin Taniyab, Yusha Arafc

⁎ Department of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Jahangirnagar University, Dhaka, Bangladesh
a Department of Microbiology, School of Life Sciences, Independent University, Dhaka, Bangladesh
b Department of Genetic Engineering and Biotechnology, School of Life Sciences, Shahjalal University of Science and Technology, Sylhet, Bangladesh

A R T I C L E   I N F O

Keywords:
COVID-19
Coronavirus
SARS-COV-2
Reverse vaccinology
Vaccine
Immunoinformatics

A B S T R A C T

SARS Coronavirus-2 (SARS-CoV-2) pandemic has become a global issue which has raised the concern of scientific community to design and discover a counter-measure against this deadly virus. So far, the pandemic has caused the death of hundreds of thousands of people upon infection and spreading. To date, no effective vaccine is available which can combat the infection caused by this virus. Therefore, this study was conducted to design possible epitope-based subunit vaccines against the SARS-CoV-2 virus using the approaches of reverse vaccinology and immunoinformatics. Upon continual computational experimentation, three possible vaccine constructs were designed and one vaccine construct was selected as the best vaccine based on molecular docking study which is supposed to effectively act against the SARS-CoV-2. Thereafter, the molecular dynamics simulation and in silico codon adaptation experiments were carried out in order to check biological stability and find effective mass production strategy of the selected vaccine. This study should contribute to uphold the present efforts of the researches to secure a definitive preventative measure against this lethal disease.

1. Introduction

Coronaviruses are a group of viruses that belong to the family, Coronaviridae and the order, Nidovirales. These viruses are enveloped, single stranded, positive sense RNA viruses with the genome size ranging from 26 to 32 kilobases in length. Coronaviruses infect humans as well as some other animals such as murine, porcine, feline, bovine, avian and are known to cause acquired acute upper respiratory tract infections and severe respiratory infections in children and adults (Su et al., 2016; Weiss and Navas-Martin, 2005; Masters and Perlman, 2013). Seven different human coronaviruses (HCoVs) have been identified so far. Among them, four HCoVs i.e., HCoV-OC43, HCoV-229E, HCoV-NL63, and HCoV-HKU1 cause common cold in immunocompromised individuals and two other HCoVs i.e., severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) cause severe respiratory diseases (van der Hock et al., 2004; Hamre and Procknow, 1966; Drosten et al., 2003; Zaki et al., 2012). The severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which is responsible for the recent pandemic all over the world, is the seventh strain known to infect human and causes the lethal coronavirus disease-2019 (COVID-19).

In December 2019, the COVID-19 was first identified in a cluster of patients with pneumonia in Wuhan, China (Peeri et al., 2020). First fatality case due to COVID-19 was reported on 11th January 2020 in Wuhan, China and first infected case outside China was reported in Thailand on 13th January 2020 (Wang et al., 2020). The most common symptoms at onset of COVID-19 are fever, cough, fatigue, diarrhoea and in severe conditions patients face difficulties in breathing (Huang et al., 2020). World Health Organization (WHO) declared COVID-19 as pandemic on 11th March 2020, as by the end of February 2020, the infected cases outside China increased 13 fold and more than 4000 fatality cases were reported globally (World Health Organization., 2020). At the time of writing, as of 29th March 2020, 652,079 infected cases, 30,313 death cases, 137,319 recovery cases were recorded globally in 177 countries (Hopkins, 2020).

To date, there is no effective vaccine that can combat the SARS-CoV-2 infections and hence the treatments are only supportive. Use of interferons in combination with Ribavirin is somewhat effective. However, the effectiveness of combined remedy needs to be further evaluated (Fehr and Perlman, 2015). This experiment was carried out to design novel epitope-based vaccine against four proteins of SARS-CoV-2 i.e., nucleocapsid phosphoprotein which is responsible for genome packaging and viral assembly (Chang et al., 2014); surface glycoprotein that is responsible for membrane fusion event during viral entry (Petit et al., 2005; Cavanagh, 1995); ORF3a protein that aids in the viral replication, characterized virulence, viral spreading and infection (Siu et al., 2019) and membrane glycoprotein.
which mediates the interaction of virions with cell receptors (Rottier, 1995), using the approaches of reverse vaccinology and immunoinformatics.

Reverse vaccinology and immunoinformatics refer to the processes of developing vaccines where the novel antigens of a virus or microorganism or a pathogenic organism are detected by analyzing the genomic and genetic information of that particular entity. In reverse vaccinology, the tools of bioinformatics are used for identifying and analyzing these novel antigens. These tools are used to dissect the genome and genetic makeup of a pathogen for developing a potential vaccine. Reverse vaccinology approach of vaccine development also allows the scientists to easily understand the antigenic segments of a virus or pathogen that should be given more emphasis during the vaccine development process. These methods of vaccine development are quick, cheap, efficient, easy and cost-effective way to design vaccines. These methods have been successfully used for developing vaccines to fight against many viruses i.e., the Zika virus, Chikungunya virus etc. (Chong and Khan, 2019; María et al., 2017).

2. Materials and methods

The current experiment was conducted to develop potential vaccines against the SARS-CoV-2, by exploiting the strategies of reverse vaccinology and immunoinformatics (Fig. 1). The materials and methods used in this experiment were taken and adapted from the works of Ullah et al. (2020a).

3. Results

3.1. Identification, selection and retrieval of viral protein sequences

The SARS-CoV-2 was identified from the NCBI database (https://www.ncbi.nlm.nih.gov/). Four protein sequences i.e., Nucleocapsid Phosphoprotein (accession no: QHD43423.2), Membrane Glycoprotein (accession no: QHD43419.1), ORF3a Protein (accession no: QHD43417.1) and Surface Glycoprotein (accession no: QHD43416.1) were selected for the possible vaccine construction and retrieved from

---

Fig. 1. Step-by-step strategies employed in the overall vaccine designing study.
the NCBI database in fasta format. Table 1 lists the proteins sequences with their NCBI accession numbers.

### 3.2. Antigenicity prediction and physicochemical property analysis of the protein sequences

Two proteins: nucleocapsid phosphoprotein and surface glycoprotein, were identified as potent antigens and used in the next phases of the experiment (Table 2). The physicochemical property analysis was conducted for these two selected proteins. Nucleocapsid phosphoprotein had the highest predicted theoretical pI of 10.07, however, surface glycoprotein had the highest predicted extinction co-efficient of 148960 M\(^{-1}\)cm\(^{-1}\). Both of them were found to have similar predicted half-life of 30 h. However, surface glycoprotein had the highest predicted aliphatic index and grand average of hydropathicity (GRAVY) values among the two proteins (Table 3).

### 3.3. T-cell and B-cell epitope prediction and their antigenicity, allergenicity and topology determination

The MHC class-I and MHC class-II epitopes were determined for potential vaccine construction. The IEDB (https://www.iedb.org/) server generated a good number of epitopes. The server contains experimental data on antibody and T-cell epitopes from studied conducted on humans, non-human primates and other animal species in the context of allergy, infectious disease, autoimmunity and transplantation. The server generates epitopes by analyzing these experimental data (Vita et al., 2018). However, based on the antigenicity scores, ten epitopes were selected from the top twenty epitopes because the epitopes generated almost similar AS and percentile scores. The percentile scores represent the predicted binding affinity and lower percentile scores correspond to higher binding affinity (Vita et al., 2018). Later, the epitopes with high antigenicity, non-allergenicity and non-toxicity were selected for vaccine construction. The B-cell epitopes were also selected based on their antigenicity, non-allergenicity and length (the sequences with more than 10 amino acids).

Table 4 lists the potential T-cell epitopes of nucleocapsid phosphoprotein and Table 5 lists the potential T-cell epitopes of surface glycoprotein. Table 6 lists the predicted B-cell epitopes of the two proteins and Table 7 lists the epitopes that followed the mentioned criteria and were selected for further analysis and vaccine construction.

### 3.4. Cluster analysis of the MHC alleles

The online tool MHCcluster 2.0 (http://www.cbs.dtu.dk/services/MHCcluster/), was used for the prediction or cluster analysis of the
Table 4
MHC class-I and MHC class-II epitope prediction and topology, antigenicity, allergenicity and toxicity analysis of the epitopes of nucleocapsid phosphoprotein. AS: Antigenic Score.

| MHC class | Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold = 0.4) | Allergenicity | Toxicity |
|-----------|---------|-------|-----|----------|----|-------------------|---------------------------------------------|--------------|---------|
| MHC class-I | AGLPYGANK | 119 | 127 | Outside | 0.547 | 0.28 | Antigenic | Non-allergic | Non-toxic |
|            | KTPPTEPK | 361 | 369 | Outside | 0.967 | 0.01 | Antigenic | Non-allergic | Non-toxic |
|            | AADLDFFSK | 397 | 405 | Outside | 0.235 | 0.96 | Antigenic | Allergic | Non-toxic |
|            | SSAAEASKK | 249 | 257 | Inside | 0.670 | 0.17 | Antigenic | Non-allergic | Non-toxic |
|            | TQPALPQRSK | 379 | 387 | Inside | 0.349 | 0.60 | Antigenic | Non-allergic | Non-toxic |
|            | SSRTSPAR | 201 | 209 | Inside | 0.166 | 1.40 | Antigenic | Non-allergic | Non-toxic |
|            | ASWFTALTQ | 50 | 58 | Inside | 0.115 | 1.80 | Antigenic | Allergic | Non-toxic |
|            | QLESKMSKG | 229 | 237 | Inside | 0.082 | 2.30 | Antigenic | Allergic | Non-toxic |
| MHC class-II | GTPPRKPY | 164 | 172 | Outside | 0.169 | 1.30 | Antigenic | Allergic | Non-toxic |
|            | QELRQCTTDYK | 289 | 300 | Inside | 2.90 | 3.60 | Antigenic | Non-allergic | Non-toxic |
|            | ELRQCTTDYK | 290 | 301 | Inside | 2.90 | 3.60 | Antigenic | Allergic | Non-toxic |
|            | QELRQCTTDYK | 289 | 300 | Inside | 2.90 | 3.60 | Antigenic | Allergic | Non-toxic |
|            | SRIGMEVTPST | 318 | 329 | Inside | 2.60 | 4.80 | Non-antigenic | Allergic | Non-toxic |
|            | LIRQTDYKHP | 291 | 302 | Inside | 2.90 | 3.60 | Antigenic | Non-allergic | Non-toxic |
|            | QRLESKMSKG | 226 | 237 | Inside | 2.00 | 8.10 | Antigenic | Non-allergic | Non-toxic |
|            | WFTALTQHG | 49 | 60 | Outside | 1.70 | 11.0 | Antigenic | Allergic | Non-toxic |

Table 5
MHC class-I and MHC Class-II epitope prediction and topology, antigenicity, allergenicity and toxicity analysis of the epitopes of surface glycoprotein. AS: Antigenic Score.

| MHC class | Epitope | Start | End | Topology | AS | Percentile scores | Antigenicity (tumor model, threshold = 0.4) | Allergenicity | Toxicity |
|-----------|---------|-------|-----|----------|----|-------------------|---------------------------------------------|--------------|---------|
| MHC class-I | GVYFASTEK | 89 | 97 | Inside | 0.938 | 0.01 | Non-antigenic | Non-allergic | Non-toxic |
|            | ASANLAATK | 1020 | 1028 | Inside | 0.911 | 0.01 | Non-antigenic | Allergic | Non-toxic |
|            | SVLNDLSR | 975 | 983 | Inside | 0.731 | 0.12 | Antigenic | Non-allergic | Non-toxic |
|            | GVLTESNKK | 550 | 558 | Inside | 0.721 | 0.14 | Antigenic | Allergic | Non-toxic |
|            | SITASALGGK | 939 | 947 | Outside | 0.779 | 0.09 | Antigenic | Non-allergic | Non-toxic |
|            | GTHWFPYTQR | 1099 | 1107 | Inside | 0.776 | 0.09 | Non-antigenic | Non-allergic | Non-toxic |
|            | ELIPVSMTK | 725 | 733 | Inside | 0.773 | 0.09 | Non-antigenic | Allergic | Non-toxic |
|            | ALDPLSETK | 292 | 300 | Outside | 0.679 | 0.16 | Antigenic | Non-allergic | Non-toxic |
|            | RLFRKSNLK | 454 | 462 | Inside | 0.677 | 0.16 | Antigenic | Non-allergic | Non-toxic |
| MHC class-II | QAPAQGTPK | 409 | 417 | Inside | 0.674 | 0.16 | Antigenic | Non-allergic | Non-toxic |
|            | FLGVYYHKNKSG | 140 | 151 | Inside | 4.40 | 0.45 | Non-antigenic | Non-allergic | Non-toxic |
|            | TFVYVYHKNKSG | 143 | 154 | Inside | 4.40 | 0.45 | Non-antigenic | Non-allergic | Non-toxic |
|            | NFRFQPTEISVR | 317 | 328 | Inside | 5.30 | 0.08 | Antigenic | Non-allergic | Non-toxic |
|            | GVFSVNGTHWVF | 1093 | 1104 | Outside | 4.10 | 0.78 | Non-antigenic | Non-allergic | Non-toxic |
|            | VYHYNKNSWSME | 143 | 154 | Inside | 4.40 | 0.45 | Non-antigenic | Non-allergic | Non-toxic |
|            | NFRFQPTEISVR | 317 | 328 | Inside | 5.30 | 0.08 | Antigenic | Non-allergic | Non-toxic |
|            | SNFPRQPTESVR | 316 | 327 | Inside | 5.30 | 0.08 | Antigenic | Non-allergic | Non-toxic |
|            | LLHYNATNIVI | 117 | 128 | Inside | 4.30 | 0.58 | Antigenic | Non-allergic | Non-toxic |
|            | EGVFVSNGTHWVF | 1092 | 1103 | Outside | 4.10 | 0.78 | Non-antigenic | Non-allergic | Non-toxic |
|            | VFPSGNTHWVT | 1094 | 1105 | Outside | 4.10 | 0.78 | Non-antigenic | Non-allergic | Non-toxic |
|            | IVNATNIVKVK | 119 | 130 | Inside | 4.30 | 0.58 | Antigenic | Non-allergic | Non-toxic |

Table 6
B-cell epitope prediction and antigenicity, allergenicity analysis of the epitopes of nucleocapsid phosphoprotein and surface glycoprotein.

| Nucleocapsid phosphoprotein | Epitope | Antigenicity (tumor model, threshold = 0.4) | Allergenicity |
|-----------------------------|---------|---------------------------------------------|--------------|
|                             | MSDNGOPQNPQRAPTRFGPGDSTGSNQNGERSGASKRRQRPQGPNNTAS | Antigenic | Non-allergic |
|                             | RIRGGDGKMKDL | Antigenic | Non-allergic |
|                             | TGPEAGPAPGANK | Antigenic | Non-allergic |
|                             | GTPPLPKGFGAYAEERGSGQASSRRSSRSRSSNSSRPGSRGSIPTSPARMAGNNGGD | Antigenic | Non-allergic |
|                             | QHGSKDELEKPPAGEQRGGQAPQVTNNTSPIDQIQ | Non-antigenic | Non-allergic |
|                             | SKMSGKGGQGGQGGQTEVYTSSASSEEKSPKQKRTATKAYN | Non-antigenic | Non-allergic |
|                             | KTPPTEPKKDKKDKKADGQALPQQRKKEQ | Non-antigenic | Allergic |
|                             | AFGRRGEPFTQGNGF | -- | -- |

| Surface glycoprotein | Epitope | Antigenicity (tumor model, threshold = 0.4) | Allergenicity |
|---------------------|---------|---------------------------------------------|--------------|
|                     | RTQLPPATYNS | Non-antigenic | Allergic |
|                     | LTIPGGSSSGWTAG | Antigenic | Non-allergic |
|                     | VRQAPQGPTQKIAD | Antigenic | Non-allergic |
|                     | SGTNGTKKFDN | Antigenic | Non-allergic |
|                     | YQAGSTPCNGV | Antigenic | Non-allergic |
|                     | QTQTNSPRRARSSV | Non-antigenic | Allergic |
|                     | ILPDPSKKPKRS | Non-antigenic | Allergic |
|                     | YGQPSTGNGVQYQ | Non-antigenic | Allergic |
|                     | RDIADTTDAVRDPQ | -- | -- |
|                     | VYDPLQPELDVF | -- | -- |
|                     | KNHTSPOVOLG | Non-antigenic | Non-allergic |
| Name of the proteins               | MHC class-I epitopes | MHC class-II epitopes | B-cell epitopes |
|-----------------------------------|----------------------|----------------------|----------------|
| Nucleocapsid phosphoprotein       | AGLPYGANK            | QELIRQGTDYKH         | MSDNGPQHQPQNATPRITFGPSDSTGSNQNGERSGARSKQRQRPQGLPNNTAS |
|                                   | AADLDDFSK            | LIRQGTDYKH           | RIRGGDICTKMDL  |
|                                   | QUEKMSGK            | RINQLEKSMKG          | TGPGEALPYGANK  |
|                                   | –                    | LNSQLESKSMKG         | –              |
|                                   | –                    | LDRLNQLESKSMG        | –              |
| Surface glycoprotein              | SVLNDILSR           | TSNFRVQPTESI         | –              |
|                                   | GVLTRSNK            | SNFRVQPTESIV         | LTPGDSGSGWTAG  |
|                                   | RLFRBNLKK           | LLLIVNNATNVI         | VRQIAAPQQTGIA  |
|                                   | QIAPOQTTGK          | –                    | –              |
|                                   | –                    | –                    | –              |

Fig. 2. The results of the MHC cluster analysis. Here, (a) is the heat map of MHC class-I cluster analysis, (b) is the tree map of MHC class-I cluster analysis, (c) is the heat map of MHC class-II cluster analysis, (d) is the tree map of MHC class-II cluster analysis.
possible MHC class-I and MHC class-II alleles that may interact with the selected epitopes during the immune responses. The tool illustrates the relationship of the clusters of the alleles in phylogenetic manner (Thomsen et al., 2013). Fig. 2 depicts the results of the cluster analysis where the red zone indicates strong interaction and the yellow zone corresponds to weaker interaction.

3.5. Generation of the 3D structures of the epitopes and peptide-protein docking

After 3D structure prediction of the selected epitopes, the peptide-protein docking was conducted to find out, whether all the epitopes had the ability to bind with the MHC class-I as well as MHC class-II molecules or not. The HLA-A*11-01 allele (PDB ID: 5WJL) was used as the receptor for docking with the MHC class-I epitopes and HLA-DRB1*04-01 (PDB ID: 5JLZ) was used as the receptor for docking with the MHC class-II epitopes. Among the MHC class-I epitopes of nucleocapsid phosphoprotein, QLESKMSGK showed the best result with the lowest global energy of -53.28. Among the MHC class-II epitopes of nucleocapsid phosphoprotein, LIRQGTDYKHWP generated the best global energy score of -16.44. GVLTESNKK generated the best global energy score of -34.60 of the MHC class-I epitopes of surface glycoprotein. And among the MHC class-II epitopes of surface glycoprotein, TSNFRVQPTESI generated the best global energy score of -2.28 (Table 8 & Fig. 3).

3.6. Vaccine construction

After successful docking, three vaccines were constructed using the selected epitopes which are supposed to be directed to fight against the SARS-CoV-2. To construct the vaccines, three different adjuvants were used i.e., beta defensin, L7/L12 ribosomal protein and HABA protein and different linkers i.e., EAAAK, GGGS, GPGPG and KK linkers were used at their appropriate positions. PADRE sequence is an important sequence which was used in vaccine construction. It has the capability to increase the potency of the vaccines with minimal toxicity. Moreover, PADRE sequence also improve the CTL response, thus ensuring potent immune response (Wu et al., 2010). The newly constructed vaccines were designated as: CV-1, CV-2 and CV-3 (Table 9).

3.7. Antigenicity, allergenicity and physicochemical property analysis of the vaccine constructs

The results of the antigenicity, allergenicity and physicochemical property analysis are listed in Table 10. All the three vaccine constructs were found to be antigenic as well as non-allergenic.

Table 8

| Name of the protein | Epitope       | MHC allele       | Global energy | Hydrogen bond energy |
|---------------------|---------------|------------------|---------------|----------------------|
| Nucleocapsid        | AGLPYGAN      | HLA-A*11-01      | -38.28        | -1.53                |
| phosphoprotein      | AADLDDFS      | allele (PDB ID: 5WJL) | -15.60        | -2.69                |
|                     | QLESKMSGK     | SWJL             | -53.28        | -4.32                |
|                     |               |                  |               |                      |
| Surface glycoprotein| SVLNDLISR     |                  | -27.72        | -3.74                |
|                     | GVLTESNNK     |                  | -34.60        | -3.64                |
|                     | RLFRKSNLK     |                  | -27.48        | -4.77                |
|                     | QIAPQGTGK     |                  | -26.86        | -0.89                |

Fig. 3. The best poses of predicted interactions between the selected epitopes from the two proteins and their respective receptors. Here, (a) is the interaction between QLESKMSGK and MHC class-I, (b) is the interaction between GVLTESNNK and MHC class-I, (c) is the interaction between LIRQGTDYKHWP and MHC class-II, (d) is the interaction between TSNFRVQPTESI and MHC class-II. The interactions were visualized by Discovery Studio Visualizer.
highest predicted molecular weight, extinction co-efficient and amino acids in the disallowed regions. The CV-3 vaccine showed the worst performance in the 3D structure generation experiment. Moreover, three different templates were used for generating 3D structures of the three different vaccines. The RaptorX server used these templates for generating the 3D structures of the query vaccine constructs (Källberg et al., 2012). The results of the 3D structure analysis are listed in Table 12 and illustrated in Fig. 5.

### 3.9. 3D structure refinement and validation

The three vaccine constructs were refined and then validated in the 3D structure refinement and validation step. The PROCHECK server (https://services.mbi.ucla.edu/PROCHECK/) divides the Ramachandran plot into four regions: the most favored region (represented by red color), the additional allowed region (represented by yellow color), the generously allowed region (represented by light yellow color) and the disallowed region (represented by white color). According to the server, a valid protein (the best quality protein) should have over 90 % of its amino acids in the most favored region. The additional allowed region and generously allowed region might also contain some percentage of the amino acids of the protein. However, no amino acid should reside within the disallowed region (Sateesh et al., 2010; Laskowski et al., 1993; Zobayer, 2018).

The 3D protein structures generated in the previous step were refined for further analysis and validation. The refined structures were validated with the aid of the Ramachandran Plots. The analysis showed that CV-1 vaccine had excellent percentage of 94.3 % of the amino acids in the most favored region, 4.4 % of the amino acids in the additional allowed regions, 0.0 % of the amino acids in the generously allowed regions and 1.3 % of the amino acids in the disallowed regions. The CV-2 vaccine had 90.0 % of the amino acids in the most favored regions, 8.3 % of the amino acids in the additional allowed regions, 0.6 % of the amino acids in the generously allowed regions and 1.1 % of the amino acids in the disallowed regions. The CV-3 vaccine showed the worst result with 77.4 % of the amino acids in the most favored region, 20.9 % of the amino acids in the additional allowed regions, 1.4 % of the amino acids in the generously allowed regions and 0.3 % of the amino acids in the disallowed regions (Fig. 6).

### 3.10. Vaccine protein disulfide engineering

In protein disulfide engineering, disulfide bonds were generated within the 3D structures of the vaccine constructs. In the experiment, the amino acid pairs that had bond energy value less than 2.00 kcal/mol, were selected. Since about 90 % of the native disulfide bonds in proteins have energy value of less than 2.2 kcal/mol, the bond energy value of 2.00 kcal/mol was selected as the cut-off value for the experiment for better prediction (Craig and Dombkowski, 2013). The CV-1 generated 10 amino acid pairs that had the capability to form disulfide bonds. However, only one pair was selected because they had the
bond energy, less than 2.00 kcal/mol: 276 Ser-311 Arg. Although, CV-2 and CV-3 generated 04 and 05 pairs of amino acids, respectively, that might form disulfide bonds but no pair of amino acids showed bond energy less than 2.00 Kcal/mol. The selected amino acid pairs of CV-1 formed the mutant version of the original vaccines (Fig. 7).

3.11. Protein-protein docking study

The protein-protein docking study was carried out to find out the best constructed COVID-19 vaccine. The vaccine construct with the best result in the molecular docking, was considered as the best vaccine construct. According to docking results, it was found that CV-1 was the best constructed vaccine. CV-1 showed the best and lowest scores in the docking as well as in the MM-GBSA study by HawkDock server. However, CV-2 showed the best binding affinity (ΔG scores) with DRB3*0202 (-18.9 kcal/mol) and DRB1*0301 (-18.5 kcal/mol) when analyzed with ClusPro 2.0 and the PRODIGY tool of HADDOCK server. Moreover, when analyzed with PatchDock and FireDock servers, CV-3 showed best global energy scores with the MHC alleles i.e., DRB5*0101 (-10.70), DRB5*0101 (-19.59), DRB1*0101 (-17.46) and DRB3*0101 (-12.32). Since CV-1 showed the best results in the protein-protein docking study with almost all the targets by all the servers and also with the TLR-8, it was considered as the best vaccine construct among the three constructed vaccines (Fig. 8 and Table 13). Later, the molecular dynamics simulation and in silico codon adaptation studies were conducted only on the CV-1 vaccine.

3.12. Molecular dynamics simulation study

The results of molecular dynamics simulation of CV-1-TLR-8 docked complex is illustrated in Fig. 9. Dynamic simulation of proteins gives easy determination of the stability and physical movements of their atoms and molecules (Chauhan et al., 2019). So, the simulation was carried out to determine the relative stability of the vaccine protein. The deformability graph of the complex illustrates the peaks representing the regions of the protein with moderate degree of deformability (Fig. 9b). The B-factor graph of the complex gives easy visualization and comparison between the NMA and the PDB field of the docked complex (Fig. 9c). The eigenvalue of the docked complex is depicted in Fig. 9d. CV-1 and TLR8 docked complex generated quite good eigenvalue of 3.817339e-06. The variance graph illustrates the individual variance by red colored bars and cumulative variance by green colored bars (Fig. 9e). Fig. 10f depicts the co-variance map of the complex, where red color represents the correlated motion between a pair of residues, uncorrelated motion is indicated by white color as well as the anti-correlated motion is marked by blue color. The elastic map of the complex refers to the connection between the atoms and darker gray regions indicate stiffer regions (Fig. 9g) (López-Blanco et al., 2014; López-Blanco et al., 2011; Kovacs et al., 2004).

3.13. Codon adaptation and in silico cloning study

Since the CV-1 protein had 596 amino acids, after reverse translation, the number nucleotides of the probable DNA sequence of CV-1 would be 1788. The codon adaptation index (CAI) value of 1.0 of CV-1 indicated that the DNA sequences contained higher proportion of the codons that should be used by the cellular machinery of the target organism E. coli strain K12 (codon bias). For this reason, the production of the CV-1 vaccine should be carried out efficiently (Solanki and Tiwari, 2018; Carbone et al., 2003). The GC content of the improved sequence was 51.34 % (Fig. 10). The predicted DNA sequence of CV-1 was inserted into the pET-19b vector plasmid between the SgrA1 and Shl restriction sites and since the vaccine DNA sequence did not have restriction sites for SgrA1 and Shl restriction enzymes, SgrA1 and Shl restriction sites were conjugated at the N-terminal and C-terminal sites, respectively. The newly constructed vector is illustrated in Fig. 11.
4. Discussion

The current study was designed to construct possible vaccines against the SARS-CoV-2, which is the cause of the recent pandemic of the deadly viral disease, COVID-19 around the world. The pneumonia has already caused the death of thousands of people worldwide. For this reason, possible vaccines were predicted in this study to fight against this lethal virus. To carry out the vaccine construction, four candidate proteins of the virus were identified and selected from the NCBI database. Only highly antigenic sequences were selected for further analysis since the highly antigenic proteins can induce better immunogenic response (Demkowicz et al., 1992). Because the nucleocapsid phosphoprotein and surface glycoprotein were found to be antigenic, they were taken into consideration for vaccine construction.

The physicochemical property analysis was conducted for the two predicted antigenic proteins. The extinction coefficient can be defined as the amount of light that is absorbed by a particular compound at a certain wavelength. Surface glycoprotein had the highest predicted extinction coefficient of 148960 M$^{-1}$ cm$^{-1}$. The aliphatic index of a protein corresponds to the relative volume occupied by the aliphatic amino acids in the side chains of the protein, for example: alanine, valine etc. (Pace et al., 1995; Gill and Von Hippel, 1989; Ikai, 1980; Ullah et al., 2020b). Surface glycoprotein also had the highest predicted aliphatic index among the two proteins (84.67). Therefore, surface glycoprotein had greater amount of aliphatic amino acids in its side chain than the nucleocapsid phosphoprotein. The grand average of hydropathicity value (GRAVY) for a protein is calculated as the sum of hydropathy values of all the amino acids of the protein, divided by the number of residues in its sequence. The negative GRAVY value represents hydrophilic characteristic and the positive GRAVY value represents hydrophobic characteristic of a compound (Kyte and Doolittle, 1982; Chang and Yang, 2013). Surface glycoprotein had the highest predicted GRAVY value of -0.079 among the two proteins. Since both proteins had the predicted negative GRAVY value, both of them were considered to be hydrophilic. Moreover, both of them had the predicted in vivo half-life of 30 h and nucleocapsid phosphoprotein had the highest theoretical pl of 10.07. Both the proteins showed quite good results in the physicochemical property analysis.

After the physicochemical property analysis of the protein sequences, the T-cell and B-cell epitope prediction was conducted. T-cell and B-cell are the two main types of cells that function in immunity. When an antigen is encountered in the body by the immune system, the antigen presenting cells or APC like macrophage, dendritic cell etc. present the antigen to the T-helper cell, through the MHC class-II molecules on their surface. The helper T-cell contains CD4+ molecule on its surface, for this reason, it is also known as CD4+ T-cell. On the other hand, the other type of T-cell, cytotoxic T-cell contains CD8+ molecule on its surface, for which, they are called CD8+ T-cell. MHC class-I molecules present antigens to cytotoxic T-lymphocytes. After activation by the antigen, the T-helper cell activates the B-cell, which starts to produce large amount of antibodies. Macrophage and CD8+ cytotoxic T cell are also activated by the T-helper cell that cause the final destruction of the target antigen (Goerdt and Orfanos, 1999; Tanchot and Rocha, 2003; Pavli et al., 1993; Arpin et al., 1995; Cano and Lopera, 2013). The possible T-cell and B-cell epitopes of the selected proteins were determined by the IEDB (https://www.iedb.org/)

Fig. 4. Results of the secondary structure prediction of the three vaccine constructs. Here, (a) is the CV-1 vaccine, (b) is the CV-2 vaccine, (c) is the CV-3 vaccine.
Fig. 4. (continued)
The epitopes with high antigenicity, non-allergenicity and non-toxicity were selected to construct the vaccines. The B-cell epitopes (predicted by the server) that were more than ten amino acids long were taken into consideration and the antigenic and non-allergenic epitopes were selected for vaccine construction. However, most of the epitopes were found to be located within the cell membrane.

The cluster analysis of the MHC alleles which may interact with the server. The epitopes with high antigenicity, non-allergenicity and non-toxicity were selected to construct the vaccines. The B-cell epitopes (predicted by the server) that were more than ten amino acids long were taken into consideration and the antigenic and non-allergenic epitopes were selected for vaccine construction. However, most of the epitopes were found to be located within the cell membrane.

The cluster analysis of the MHC alleles which may interact with the server.

Table 11
Results of the secondary structure analysis of the vaccine constructs.

| Name of the vaccine | Alpha helix (percentage of amino acids) | Beta sheet (percentage of amino acids) | Coil structure (percentage of amino acids) |
|---------------------|----------------------------------------|---------------------------------------|--------------------------------------------|
| CV-1                | 25%                                    | 8%                                    | 67.1%                                      |
| CV-2                | 31.6%                                  | 6.8%                                  | 61.6%                                      |
| CV-3                | 37.8%                                  | 5%                                    | 57.2%                                      |

Table 12
Results of the tertiary structure analysis of the vaccine constructs.

| Name of the vaccine | Number of the domains | p-value | PDB Id of the est. matched template |
|---------------------|-----------------------|---------|------------------------------------|
| CV-1                | 02                    | 2.37e-04| 1kj6A                              |
| CV-2                | 02                    | 6.35e-05| 1dd3A                              |
| CV-3                | 01                    | 2.36e-04| 6cfeA                              |

Fig. 5. 3D structures of the three predicted vaccine constructs. Here, (a) is CV-1, (b) is CV-2, (c) is CV-3.
Fig. 6. The results of the Ramachandran plot analysis of the three coronavirus vaccine constructs. Here, 01. CV-1 vaccine, 02. CV-2 vaccine, 03. CV-3 vaccine.

Fig. 7. The disulfide engineering of CV-1. The original form is illustrated in the left side and the mutant form is illustrated in the right side.
Since all the epitopes generated quite good docking scores, it can be
bind with their respective MHC class-I and MHC class-II alleles or not.

Results of the docking study of all the vaccine constructs.

| Name of the vaccines | Name of the targets | PDB IDs of the targets | Binding affinity, ΔG (kcal mol⁻¹) | Global energy (kcal mol⁻¹) | HawkDock score (the lowest score) | MM-GBSA (binding free energy, in kcal mol⁻¹) |
|----------------------|---------------------|------------------------|-----------------------------------|---------------------------|-----------------------------------|-------------------------------------------|
| CV-1                 | DRB3*0202           | 1A6A                   | −17.2                             | −4.22                     | −6436.60                          | −55.56                                    |
|                      | DRB*0101            | 1HI5                   | −19.9                             | −4.92                     | −6689.84                          | −141.66                                   |
|                      | DRB1*0101           | 2FSE                   | −19.1                             | 4.31                      | −7297.17                          | −148.58                                   |
|                      | DRB3*0101           | 2Q6W                   | −19.2                             | −7.20                     | −7581.70                          | −138.4                                    |
|                      | DRB1*0401           | 2EB                    | −21.4                             | −11.58                    | −6758.33                          | −98.56                                    |
|                      | DRB1*0301           | 3C5J                   | −17.7                             | −9.09                     | −5201.43                          | −114.35                                   |
|                      | TLR8                | 3W3M                   | −23.2                             | −23.12                    | −6514.36                          | −52.06                                    |
| CV-2                 | DRB3*0202           | 1A6A                   | −18.9                             | −10.32                    | −3477.55                          | 1.01                                      |
|                      | DRB5*0101           | 1HI5                   | −17.7                             | −10.46                    | −3761.37                          | −106.17                                   |
|                      | DRB1*0101           | 2FSE                   | −16.9                             | 1.58                      | −3531.12                          | −106.13                                   |
|                      | DRB3*0101           | 2Q6W                   | −19.1                             | 16.68                     | −3707.86                          | −90.89                                    |
|                      | DRB1*0401           | 2EB                    | −20.0                             | −10.01                    | −4766.17                          | −45.76                                    |
|                      | DRB1*0301           | 3C5J                   | −18.5                             | −1.87                     | −3561.16                          | −18.72                                    |
|                      | TLR8                | 3W3M                   | −21.1                             | −18.91                    | −2945.44                          | −54.79                                    |
| CV-3                 | DRB3*0202           | 1A6A                   | −16.9                             | −10.70                    | −4023.68                          | −9.2                                      |
|                      | DRB5*0101           | 1HI5                   | −18.9                             | −19.59                    | −4556.87                          | −12.38                                    |
|                      | DRB1*0101           | 2FSE                   | −17.1                             | −17.46                    | −4602.08                          | −10.54                                    |
|                      | DRB3*0101           | 2Q6W                   | −18.4                             | −12.32                    | −4767.21                          | −27.71                                    |
|                      | DRB1*0401           | 2EB                    | −20.1                             | 6.55                      | −3571.79                          | −8.74                                     |
|                      | DRB1*0301           | 3C5J                   | −17.8                             | 5.35                      | −4001.56                          | −12.38                                    |
|                      | TLR8                | 3W3M                   | −22.8                             | −10.92                    | −5008.23                          | −19.83                                    |

Fig. 8. The interaction between TLR-8 (in green color) and CV-1 vaccine construct (in light blue color). The interaction was visualized with PyMol.

The secondary structure prediction of the vaccine constructs determined that CV-1 had the lowest number of amino acids in alpha-helix formation, with 25 % of the amino acids in the alpha-helix formation and 67.1 % of the amino acids in coil formation. For this reason, most of the amino acids of CV-1 vaccine were predicted to be in coil structure, which was also the highest percentage of amino acids in coil structure among the three vaccines. On the other hand, CV-3 had the highest amount of amino acids in the alpha-helix formation (37.8 %), according to the prediction of the study. However, all the three vaccine constructs had most of their amino acids in their coil structures. In the tertiary structure prediction, all the three vaccine constructs showed quite satisfactory results. Thereafter, in the tertiary structure refinement and validation, CV-1 vaccine construct generated the best result with 94.3 % of the amino acids in the most favored region and 4.4 % of the amino acids in the additional allowed regions. CV-2 also showed good result with 90.0 % of the amino acids in the most favoured region. In the disulfide bond engineering experiment, only CV-1 was found to follow the selection criteria for disulfide bond formation. With the lowest and best results generated by the MM-GBSA study, HawkDock server and ClusPro 2.0 server, CV-1 was considered as the best vaccine construct among the three vaccines. Therefore, CV-1 was selected for molecular dynamics simulation study, codon adaptation and in silico coding study.

The molecular dynamics simulation study, conducted by the online tool iMODS (http://imods.chaconlab.org/) revealed that the TLR-8-CV-1 docked complex should be quite stable with a good eigenvalue of 3.8173396-06. The complex had less chance of deformation and for this reason, the complex should be quite stable in the biological environment. The Fig. 9 shows that a good number of amino acids were in the correlated motion that were marked by red color. Finally, codon adaptation and in silico cloning experiments were performed and with the predicted CAI value of 1.0, it could be concluded that the DNA sequence might contain very high amount of favorable codons that should be able to express the desired amino acids in the target microorganism, E. coli strain K12. The DNA sequence also had quite high and good amount of GC content of 51.34 %. Finally, the pET-19b vector, containing the CV-1 vaccine insert was constructed which should efficiently encode the vaccine protein in the E. coli cells.
The vaccine development using genome based technologies provides scientists the opportunity to develop vaccines by optimizing the target antigens. Conventional vaccines, like the attenuated vaccines or the inactivated vaccines sometimes fail to provide potential immunity towards a target antigen. Moreover, the conventional approach of vaccine development has raised many safety concerns in the pre-clinical and clinical trials. The subunit vaccines like the vaccines predicted in the study could overcome such difficulties (Tameris et al., 2013; Merten, 2002; Hasson et al., 2015; Kaufmann et al., 2014; Stratton et al., 2002). Finally, this study recommends CV-1 as the best vaccine to be an effective countermeasure based on the strategies employed in the study to be triggered against SARS-CoV-2 infection. However, further in vivo and in vitro experiments are suggested to strengthen the findings of this study.

Fig. 9. The results of molecular dynamics simulation study of CV-1 and TLR-8 docked complex. Here, (a) NMA mobility, (b) deformability, (c) B-factor, (d) eigenvalues, (e) variance (red color indicates individual variances and green color indicates cumulative variances), (f) co-variance map (correlated (red), uncorrelated (white) or anti-correlated (blue) motions) and (g) elastic network (darker gray regions indicate more stiffer regions).
Fig. 10. The results of the codon adaptation study of the best constructed vaccine, CV-1.

Fig. 11. Constructed pET-19b vector with the CV-1 insert (marked in red color). In the plasmid, the larger purple colored arrow represents the lacI gene (from 2500 bp to 3582 bp), the smaller purple colored arrow represents the rop gene (from 4896 bp to 5085 bp), yellow colored arrow represents the origin of replication (from 5517 bp to 6103 bp), the light green colored arrow represents the AmpR (ampicillin resistance) gene (from 6274 bp to 7134 bp), the white rectangle represents the T7 terminator (from 195 bp to 242 bp), the light blue colored arrow represents the multiple cloning site (from 301 bd to 317 bp) and the desired gene has been inserted (marked by red color) between the 485 bp and 2128 bp nucleotide. Various restriction enzyme sites are mentioned in the plasmid structure.
5. Conclusion

The SARS-CoV-2 has caused one of the deadliest pandemics in the recent times. Prevention of the newly emerging infection is very challenging as well as mandatory. The potentiality of recent times. Prevention of the newly emerging infection is very challenging as well as mandatory. The potentiality of recent times. Prevention of the newly emerging infection is very challenging as well as mandatory. The potentiality of recent times. Prevention of the newly emerging infection is very challenging as well as mandatory.

Craig, D.B., Dombkowski, A.A., 2013. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinformatics 14 (346). https://doi.org/10.1186/1471-2105-14-346.

Demkowicz, W.E., Maa, J.S., Esteban, M., 1992. Identification and characterization of vaccines virus genes encoding proteins that are highly antigenic in animals and are immunodominant in vaccinated humans. J. Virol. 66 (January (1)), 388–398.

Goerdt, S., Orfanos, C.E., 1999. Other functions, other genes: alternative activation of antigen-presenting cells. Immunology 10, 137–142. https://doi.org/10.1111/j.1749-7718.1999.00301.x.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. 157 (May (1)), 105–122.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395 (February (10223)), 497–506.

SARS-CoV-2.

Goerdt, S., Orfanos, C.E., 1999. Other functions, other genes: alternative activation of antigen-presenting cells. Immunology 10, 137–142. https://doi.org/10.1111/j.1749-7718.1999.00301.x.

Hamre, D., Procknow, J.J., 1966. A new virus isolated from the human respiratory tract. Proc. Soc. Exp. Biol. Med. 121 (January (1)), 190–193.

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., Xu, J., 2012. Template-based protein structure modeling using the RaptorX web server. Nat. Protoc. 7, 1511. https://doi.org/10.1038/nprot.2012.85.

Kaufmann, S.H., McElrath, M.J., Lewis, D.J., Del Giudice, G., 2014. Challenges and responses in human vaccine development. Curr. Opin. Immunol. 28, 18–26. https://doi.org/10.1016/j.coi.2014.01.009.

Kovacs, J.A., Chacón, P., Abagyan, R., 2004. Predictions of protein flexibility: first-order measures. Proteins: Struct. Funct. Bioinf. 56, 661–668. https://doi.org/10.1002/prot.20151.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. 157 (May (1)), 105–122.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., 1993. Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291. https://doi.org/10.1107/s0021889892009944.

López-Blanco, J.R., Garzón, J.I., Chacón, P., 2011. iMod: multipurpose normal mode analysis in internal coordinates. Bioinformatics 27, 2843–2850. https://doi.org/10.1093/bioinformatics/btr497.

López-Blanco, J.R., Aliaga, J.J., Quintana-Oriti, E.S., Chacón, P., 2014. iMDS: internal coordinates normal mode analysis server. Nucleic Acids Res. 42, W271–6. https://doi.org/10.1093/nar/gku339.

María, R.R., Arturo, C.J., Alicia, J.A., Paulina, M.G., Gerardo, A.O., 2017. The Impact of reverse vaccinology for the development of chimeric vaccine against Acinetobacter baumannii. Virology 341 (October (2)), 215–230.

Kousoulas, K.G., 2005. Genetic analysis of the SARS-coronavirus spike glycoprotein antigenic epitopes of the SARS coronavirus ORF3a protein activates the NLRP3 inflammasome by pro- lucing 1074–1078.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., Gray, T., 1995. How to measure and predict protein structure modeling using the RaptorX web server. Nat. Protoc. 7, 1511. https://doi.org/10.1038/nprot.2012.85.

Pavli, P., Hume, D.A., VanDePol, E., Doe, W.F., 1993. Dendritic cells, the major antigen-presenting cells. Immunity 10, 137–142. https://doi.org/10.1016/s1074-1696(00)80123-4.

Siu, K.L., Yuen, K.S., Castaño-Rodriguez, C., Ye, Z.W., Yeung, M.L., Fung, S.Y., Yuan, S., Siu, K.L., Yuen, K.S., Castaño-Rodriguez, C., Ye, Z.W., Yeung, M.L., Fung, S.Y., Yuan, S., Sateesh, P., Rao, A., Sangeeta, S.K., Babu, M.N., Grandhi, R.S., 2010. Homology modeling of the novel coronavirus spike glycoprotein. Cytotechnology 39, 91–116. https://doi.org/10.1007/s10527-010-9261-8.

Siu, K.L., Yuen, K.S., Castaño-Rodriguez, C., Ye, Z.W., Yeung, M.L., Fung, S.Y., Yuan, S., Siu, K.L., Yuen, K.S., Castaño-Rodriguez, C., Ye, Z.W., Yeung, M.L., Fung, S.Y., Yuan, S., Sateesh, P., Rao, A., Sangeeta, S.K., Babu, M.N., Grandhi, R.S., 2010. Homology modeling of the novel coronavirus spike glycoprotein. Cytotechnology 39, 91–116. https://doi.org/10.1007/s10527-010-9261-8.

Kousoulas, K.G., 2005. Genetic analysis of the SARS-coronavirus spike glycoprotein antigenic epitopes of the SARS coronavirus ORF3a protein activates the NLRP3 inflammasome by producing TRAF3-dependent ubiquitination of ASC. FASEB J. 33 (August (8)), 8865–8877.

Solanki, V., Tiwari, V., 2018. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against Acinetobacter baumannii. Sci. Rep. 8, 9044. https://doi.org/10.1038/s41598-018-26689-7.

Stratton, K., Almario, D.A., McCormick, M., 2002. Institute of Medicine (US) Immunization Safety Review Committee. Immunization safety review: SV40
contamination of polio vaccine and cancer. Immunization Safety Review: SV40 Contamination of Polio Vaccine and Cancer. National Academies Press, US.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A.C., Zhou, J., Liu, W., Bi, Y., Gao, G.F., 2016. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends Microbiol. 24 (June (6)), 490–502.

Tameris, M.D., Hatherill, M., Landry, B.S., Scriba, T.J., Snowden, M.A., Lockhart, S., Shea, J.E., McClain, J.B., Hussey, G.D., Hanekom, W.A., Mahomed, H., 2013. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. Lancet 381, 1021–1028. https://doi.org/10.1016/s0140-6736(13)60177-4.

Tanchot, C., Rocha, B., 2003. CD8 and B cell memory: same strategy, same signals. Nat. Immunol. 4, 431. https://doi.org/10.1038/nimm-0403-431.

Thomsen, M., Lundegaard, C., Buu, S., Lund, O., Nielsen, M., 2013. MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 65, 655–665. https://doi.org/10.1007/s00251-013-0714-0719.

Ullah, M.A., Sarkar, B., Islam, S.S., 2020a. Exploiting the Reverse Vaccinology Approach to Design Novel Subunit Vaccine against Ebola Virus. Manuscript submitted for publication.

Ullah, M.A., Sarkar, B., Islam, S.S., 2020b. Exploiting the reverse vaccinology approach to design novel subunit vaccine against ebola virus. medRxiv. https://doi.org/10.1016/j.imbio.2020.151949.

van der Hoek, L., Pyrc, K., Jubbink, M.F., Vermeulen-Oost, W., Berkhour, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J., Berkhour, B., 2004. Identification of a new human coronavirus. Nat. Med. 10 (April (4)), 368–373.

Vita, R., Mahajan, S., Overton, J.A., Dhanda, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., Peters, B., 2018. The immune epitope database (IEDB): 2018 update. Nucleic Acids Res. (47), D339–343. https://doi.org/10.1093/nar/gky1006.

Thomsen, M., Lundegaard, C., Buu, S., Lund, O., Nielsen, M., 2013. MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 65, 655–665. https://doi.org/10.1007/s00251-013-0714-0719.

Zaki, A.M., Van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., Fouchier, R.A., 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N. Engl. J. Med. 367 (November (19)), 1814–1820.

Zobayer, M., 2018. In Silico Characterization and Homology Modeling of Histamine Receptors (Doctoral dissertation). Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh. https://doi.org/10.3923/jbs.2018.178.191.