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.

Journal of Infection and Public Health 14 (2021) 938–946

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: http://www.elsevier.com/locate/jiph

Original Research Article

Design of a novel multiple epitope-based vaccine: An immunoinformatics approach to combat SARS-CoV-2 strains

Muhammad Naveed[a,*], Sana Tehreem[b], Sundas Arshad[c], Syeda Aniqa Bukhari[d], Muhammad Aqib Shabbir[a], Ramsha Essa[e], Nouman Ali[a], Sumera Zaib[f], Ajmal Khan[g], Ahmed Al-Harrasi[g,*], Imtiaz Khan[h,*]

[a] Department of Biotechnology, Faculty of Life Sciences, University of Central Punjab, Lahore 54590, Pakistan
[b] Department of Biochemistry and Molecular Biology, School of Life Sciences, Hubei University, Wuhan 430062, China
[c] Faculty of Science, Technology and Medicine, University of Luxembourg, Luxembourg
[d] Research Center for Modeling and Simulation, National University of Science and Technology, Islamabad, Pakistan
[e] Center of Excellence in Molecular Biology (CEMB), Punjab University, Lahore, Pakistan
[f] Department of Biochemistry, Faculty of Life Sciences, University of Central Punjab, Lahore 54590, Pakistan
[g] Natural and Medical Sciences Research Center, University of Nizwa, P.O Box 33, Postal Code 616, Birkat Al Mauz, Nizwa, Oman
[h] Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom

ARTICLE INFO

Article history:
Received 16 August 2020
Received in revised form 12 April 2021
Accepted 22 April 2021

Keywords:
B and T-cells
Molecular docking
Dynamic simulations
PET28a (+) vector
SARS-CoV-2
Vaccine design

ABSTRACT

Background: Since the SARS-CoV-2 outbreak in December 2019 in Wuhan, China, the virus has infected more than 153 million individuals across the world due to its human-to-human transmission. The USA is the most affected country having more than 32-million cases till date. Sudden high fever, pneumonia and organ failure have been observed in infected individuals.
Objectives: In the current situation of emerging viral disease, there is no specific vaccine, or any therapeutics available for SARS-CoV-2, thus there is a dire need to design a potential vaccine to combat the virus by developing immunity in the population. The purpose of present study was to develop a potential vaccine by targeting B and T-cell epitopes using bioinformatics approaches.
Methods: B- and T-cell epitopes are predicted from novel M protein-SARS-CoV-2 for the development of a unique multiple epitope vaccine by applying bioinformatics approaches. These epitopes were analyzed and selected for their immunogenicity, antigenicity scores, and toxicity in correspondence to their ability to trigger immune response. In combination to epitopes, best multi-epitope of potential immunogenic property was constructed. The epitopes were joined using EAAAK, AAY and GPGPG linkers.
Results: The constructed vaccine showed good results of worldwide population coverage and promising immune response. This constructed vaccine was subjected to *in-silico* immune simulations by C-ImmSim. Chimeric protein construct was cloned into *PET28a* (+) vector for expression study in *Escherichia coli* using snapgene.
Conclusion: This vaccine design proved effective in various computer-based immune response analysis as well as showed good population coverage. This study is solely dependent on developing M protein-based vaccine, and these *in silico* findings would be a breakthrough in the development of an effective vaccine to eradicate SARS-CoV-2 globally.

Introduction

Coronaviruses are enveloped RNA viruses that are known to cause various diseases including respiratory, enteric, hepatic and neurological in humans, birds and other mammals [1]. There are six species of coronavirus having capability to cause disease in humans and four viruses HKU1 OC43, NL63, and 229E are known to cause cold like symptoms in humans [2]. In past, though an unknown coronavirus has already caused an outbreak of SARS which is also known as Severe Acute Respiratory Syndrome in humans in 2003 [3,4]. In December 2019, an enormous number of pneumonia cases were observed in Wuhan, China. On 2 January, 2020, 41 patients were diagnosed with infection of a novel coronavirus (SARS-CoV-2) in China [5]. Later, on March 11, 2020, WHO announced the COVID-

* Corresponding authors.
E-mail addresses: dr.naveed@ucp.edu.pk (M. Naveed), aharrasi@unizwa.edu.om (A. Al-Harrasi), imtiaz.khan@manchester.ac.uk (I. Khan).

https://doi.org/10.1016/j.jiph.2021.04.010

19 outbreak as pandemic. Most of the infected people were exposed to the wet animal market of Wuhan, China where the animal meat was sold routinely. So, it was suggested that the virus may belong to zoonotic origin [6]. However, cases were also reported in the people who were not exposed to wet animal market. On the basis of this, new reports suggested that the actual route of virus transmission is from person to person [7]. The detailed analysis showed that the virus may spread from person to person through contact and the droplet transmission i,e., from the infected individuals by coughing and sneezing [8].

The WHO stated to get extensive measures in order to control the spread. For this purpose, isolation of infected individuals was recommended, and traveler screening is getting mandatory to avoid further spread of the infection. Although, early death cases were observed in elder people because of their weaker immunity allows the viral replication hence the progression of viral infection occurs [9]. The latency period of virus is almost 3–7 days according to epidemiological studies with a maximum of 14 days [10]. The novel coronavirus (SARS-CoV-2) is contagious during the latency period and promised to be emerging continuously at rapid rate. This property of the novel virus makes it more lethal and difficult to combat. The COVID-19 infection shows common symptoms like flu including fatigue, fever, and dry cough. In some cases, lymphopenia as well as acute respiratory symptom are observed. However, about 30% of severe cases of the disease are held in intensive care, with median mortality rate of 4.3% [11].

The positive sensed, non segmented single stranded RNA genome of SARS-CoV-2 is responsible for making many proteins for its survival and attack. It has four main structural proteins such as Envelope glycoprotein (E), Spikes glycoprotein (S), Matrix/membrane glycoprotein (M) and nucleocapsid (N). For the selection and entry into the target cell, the novel coronavirus SARS-CoV-2 uses its spike proteins. Crown-like shape of the spikes protein forms homotrimer onto the viral surface and helps in binding of viral body onto the ACE2 receptor of the host cell. When co-expressed, the instinctive envelope trimeric S glycoprotein is incorporated onto virus like particle in order to cause infection [12]. According to recent reports, a protein on host surface, cellular protease TMPRSS2 is used for SARS-CoV-2 priming [13]. Therefore, the diagnosis of lethal pathogen SARS-CoV-2 can be done by various methods including RT-PCR and CT-scans [14]. CT scan findings are considered better and authentic than RT-PCR diagnostics due to its good sensitivity power. The sensitivity of CT-scan was observed as 97.2% as compared to RT-PCR (83.3%). So patients with promising CT-scan findings and negatively reported on PCR should be isolated [15].

Although at present, there is no authentic treatment method available for the novel coronavirus disease (COVID-19). But some reported drugs are having a noticeable effect on the patients [16]. Remdesivir is observed to be effective on coronavirus patients due to promising effects. Remdesivir is an adenosine correspondent which causes pre-mature termination by integrating into nascent viral RNA chains. Hydroxychloroquine, another drug, is recommended to be used in the updated version of guidelines due to its acceptable safety and efficacy [16]. However, a parasitic drug ivermectin, is now declared FDA-approved drug to be an inhibitor of COVID-19 by showing almost 5000 folds reduction in viral replication in many *in-vitro* studies [17]. Since there is a large work in the pharmaceutical field, still not a single drug cures this disease completely. This arises the need of vaccine development in order to target virus at its molecular level. However, some advancement in the field of vaccine development has been made. Nano-based vaccination and homology based vaccination against coronavirus are new talks of the era. On the other hand, recent advancement at vaccine strategies like live-attenuated vaccine, protein subunits vaccine, inactivated vaccine, viral vector vaccine also facilitates in encouraging advancements [18]. The role of promising vaccine will be the activation of initiate herd immunity. Herein, we chose M protein or vaccine design, due to its high antigenic and conserved property, for the purpose of multi-epitope vaccine development of the novel coronavirus (SARS-CoV-2). As M protein is a main component of viral envelope, it contributes in viral morphogenesis and interaction during assembly with other cell bodies. Hence, it plays a central role in viral immune envasion. This novel vaccine shows potential results by accomplishing all-natural criteria in the artificial immune environment. The vaccine will be produced on an industrial scale in a suitable expression host. Therefore, the peptide is optimized according to the *Escherichia coli*, as it is the most suitable and affordable expression system for vaccine development. Hence, it is cloned in *PET28a* (+) vector for further expression analysis. After testing *in vivo* and *in vitro*, this vaccine may be considered as a future therapeutic.

Material and methods

This section has been added in Supporting information.

Results

Sequence and structure analysis

For the construction of potential vaccine, 95 M protein sequences of SARS-CoV-2 were retrieved from ViPR database. M protein is one of the structural proteins which is the utmost abundant constituent of the coronaviruses [19,20] enabling it to conquer and assemble viral particles into the host cell. M protein with UniprotKB ID: P0DTC5 (VME1_SARS2) was found to be the most antigenic viral protein. The antigenicity estimation of viral protein was done using Vaxijen 2.0 online server [21]. For a higher specificity, the threshold value was set at $\geq$0.4. The antigenicity for M protein was found to be 0.5102 by the full-length protein antigenicity analysis that indicated it as a probable antigen. Then, the protein was used for further investigation.

The calculation of physiochemical properties of SARS-CoV-2 M protein *via* ProtParam determined that it contains 222 amino acids and the molecular weight was found to be 25146.62 Da [22]. The theoretical isoelectric point (PI) was obtained as 9.51 that specifies it as positive in nature as isoelectric point above 7.0 shows positively charged proteins [23]. Protparam calculated instability-index (II) of 39.14; which categorized our protein as a stable one. Aliphatic index of 120.86 designates that protein is thermo-stable over a wide temperature assortment [24,25]. Total number of carbon (C), hydrogen (H), nitrogen (N), oxygen (O), and sulfur (S) were entitled by formula $C_{1165}H_{1823}N_{303}O_{301}S_8$. The details of physiochemical properties calculated by Protparam are given in Fig. S1.

Prediction of secondary structure *via* PSIPRED [26] and 3D structure *via* I-TASSER [27] showed that M protein exhibited 10% α helixes, 50% β sheets, and 40% loops (Fig. S2a). Moreover, the prediction of 6 disulfide bonds (S—S) and the allocation of scores in Table 1 was done *via* DiANNA1.1 tool [28]. Prediction of transmembrane topology was completed using an online tool TMHMM (Fig.

Table 1
Disulfide bonds prediction by DiANNA1.1.

Cysteine sequence position	Distance	Bond[a]	Score
33–64	31	FLTWICLLQFA-PVTLACFVLAA	0.01039
33–86	53	FLTWICLLQFA-AIAMACLVGLM	0.01063
33–159	126	FLTWICLLQFA-HHLGRCDIKDL	0.012
64–86	22	PVTLACFVLAA-AIAMACLVGLM	0.01037
64–159	95	PVTLACFVLAA-HHLGRCDIKDL	0.0107
86–159	73	AIAMACLVGLM-HHLGRCDIKDL	0.01043

[a] All are weak bonds.

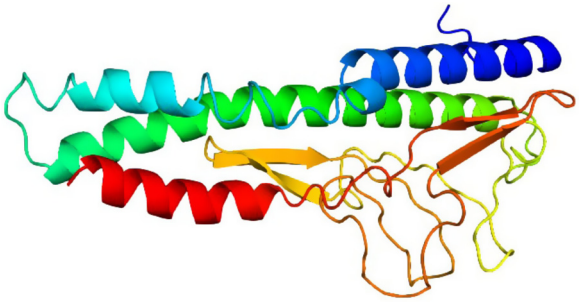

Fig. 1. 3D structure of M-protein.

Table 2
List of Bepipred linear epitope predicted *via* IEDB analysis resource.

Start	End	Peptide	Length
1	8	MADSNGTI	8
112	113	FN	2
163	163	D	1
169	171	TVA	3
185	194	QRVAGDSGFA	10
206	215	LNTDHSSSSD	10

S2b). The residues from 1 to 19 and 74 to 77 were found to be surface exposed, while residues from 40 to 50 and 101 to 122 were found inside the transmembrane region. Residues from 20 to 39 and 78 to 100 were found to be buried within the core region of the M protein (Fig. 1).

Recognition of B-cell epitopes

B-cell epitopes are important for resistance against viral infection. Potential B-cell epitopes have altered features that direct B-cell to recognize as well as activate the rich immune responses against distinctive viral infection. Here, we anticipated amino acid screened based methods for the analysis of potential B-cell epitopes. We used a consensus-based approach to predict potential B-cell epitopes.

Total 6 linear epitopes were predicted by Bepipred linear epitope prediction with threshold score of 0.350. The tool uses Hidden Markov model-based algorithm which is one of the best linear epitopes calculating methods. The minimum score for linear epitope prediction was 0.010 and the maximum score was 1.623. Moreover, the average prediction score was found to be –0.779. After analyzing all the data, we found that the peptide sequences from 206 to 215 amino acids can speed up the desired immune response and declared as B-cell epitopes. The results obtained from this method are explained in Fig. S3a and Table 2. Antigenicity of experimentally known amino acid epitopes was examined by the Kolaskar and Tongaonkar method [29]. The maximum tendency of antigenicity was found to be 1.235 and the minimum value was 0.904. The average observed value was 1.053. Threshold value was adjusted at 1.00 and all the values greater than 1.00 were candidates of antigenic factors. Total 7 epitopes were found satisfying the threshold value and selected for further analysis. They were found to have potential to initiate B-cell response. The results of antigenicity analysis are shown in Fig. S3b and Table 3.

Good surface accessibility is an important factor for an effective B-cell epitope. Therefore, Emini surface accessibility tool was utilized for this prediction. The default threshold value was found as 1.000 and total 7 epitopes were selected passing the default value of threshold. The more accessible region was 88–94 amino acid residues and average observed value was 1.00. Although the minimum value was found to be 0.078 and maximum value was 5.199. The result are shown in Fig. S3c and Table 4.

Table 3
The Kolaskar and Tongaonkar method for the prediction of antigenicity.

Start	End	Peptide	Length
10	17	VEELKKLL	8
19	39	QWNLVIGFLFLTWICLLQFAY	21
45	73	FLYIIKLIFLWLLWPVTLACFVLAAVYRI	29
79	103	GIAIAMACLVGLMWLSYFIASFRLF	25
119	159	LLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRC	41
161	188	IKDLPKEITVATSRTLSYYKLGASQRVA	28
193	199	FAAYSRY	7

Table 4
Emini surface accessibility prediction for the accessible region.

Start	End	Peptide	Length
11	20	EELKKLLEQW	10
38	44	AYANRNR	7
103	109	FARTRSM	7
162	167	KDLPKE	6
171	188	ATSRTLSYYKLGASQRVA	18
195	202	AYSRYRIG	8
204	215	YKLNTDHSSSSD	12

Beta-turn is surface exposed and hydrophilic in nature. It plays a dynamic part in the commencement of immune system defense response. For the prediction of beta-turn in M protein, Chou and Fasman beta-turn evaluating algorithm was used [30]. Threshold level of the tool was adjusted at 0.915 and the computed results suggested various values which are 0.600 (minimum) and 1.384 (maximum). The average value was 0.915. Graphical representation of Chou and Fasman's results is shown in Figure S3d. The region from 204 to 208 are found to be more likely to persuade beta turns in the peptide structure. The experimental evidences suggested that the flexibility of the peptides is associated with the antigenicity of the protein. Hence, the Karplus and Schulz [31] method was instigated for this purpose. The region from 203 to 216 was found to be more flexible by this prediction method as shown in Fig. S3e. Threshold value of the tool was adjusted to 0.970 and the computed results are 0.873 (minimum) and 1.115 (maximum). Average computed value was 0.915.17. The epitopes were further screened out based on antigenicity, allergenicity and toxicity. Threshold value was set at 0.4. Non-antigenic, allergic and toxic epitopes were excluded, and non-toxic, non-allergic and antigenic epitopes were selected. 5 out of 17 epitopes were assumed as effective B-cell epitopes that can evoke B lymphocyte (QWNLVIGFLFLTWICLLQFAY, FLYIIKLIFLWLLWPVTLACFVLAAVYRI, GIAIAMACLVGLMWLSYFIASFRLF, FAAYSRY, QRVAGDSGFA).

B-cell epitopes were further tested for conservancy analysis using IEDB conservancy analysis tool. Total 5 selected epitopes were used in conservancy analysis that were chosen to be utilized in vaccine construction. These epitopes were found to be conserved with maximum conservancy at 100% coverage and identity.

T-cell epitope identification

Prediction of MHC Class-I binding profile for conserved epitopes

We selected homo sapiens as MHC source and the SMM method to investigate diverse set of MHC HLA alleles in humans. This tool provides an output interface for HLA-binding affinity of epitopes according to IC_{50} nM unit. A lower IC_{50} value designates a higher binding affinity of epitopes with MHC Class-I molecule. Total 756 number of epitopes were selected based on IC_{50} values less than 200 for certifying a higher affinity estimated to interact with many MHC Class-1 alleles. Out of 756 epitopes, 50 were selected based on maximum interaction of MHC Class-1 alleles with the epitopes. Based on antigenicity, allergenicity and toxicity, 50 epitopes were further screened out. Toxic and allergic epitopes having less than

0.4 antigenic scores were excluded. 16 epitopes of MHC Class-1 were finalized for further analyses. The core epitopes AGDSGFAAY, ATSRTLSYY, TSRTLSYYK, were perceived to be the dominant binder with 7 alleles (HLA-C*12:03, HLA-C*05:01, HLA-B*15:02, HLA-C*14:02, HLA-A*30:02, HLA-B*35:01, HLA-C*03:03), (HLA-C*12:03, HLA-C*03:03, HLA-A*30:02, HLA-A*11:01, HLA-A*29:02, HLA-A*01:01, HLA-C*14:02), (HLA-A*30:01, HLA-C*15:02, HLA-A*11:01, HLA-C*12:03, HLA-A*31:01, HLA-C*07:01, HLA-A*68:01) respectively. The core epitope FLFLTWICL showed highest antigenic score, 1.4835 (Table S1).

Prediction of MHC Class-II binding profile for conserved epitopes

630 conserved predicted epitopes with the IC_{50} less than 200 were found to interact with MHC Class-II alleles. Out of 630 epitopes, 38 were selected that were interacting with more than 6 MHC Cass-II alleles. 18 epitopes were finalized for further analyses because of allergenicity, toxicity and antigenicity. The core epitopes IKLIFLWLL is considered to be the top binder as it interacts with 14 alleles; (HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DRB1*01:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*07:01, HLA-DRB1*15:01, HLA-DRB4*01:01) followed by LYIIKLIFL and LFLTWICLL are supposed to bind with 10 and 9 alleles, respectively (Table S2).

Multi-epitopic subunit vaccine assembly

For multi-epitopic vaccine chimera construction, a total of 6 B-cell epitopes, 15 MHC Class-I epitopes and 18 MHC Class-II epitopes were used. 50S ribosomal protein L7/L12 with UniProt ID: P9WHE3 was used as an adjuvant in order to build the vaccine construct. The adjuvant was linked to the first B-cell epitope using an EAAAK linker at the amino (N) terminus to generate a specific immune response. Furthermore, the linkage of B-cell and MHC Class-I epitopes, GPGPG linkers were utilized. MHC Class-II epitopes were connected using AAY linkers. The overlapping regions of B-cell epitopes, CTL and HLT epitopes were merged to reduce the size of vaccine (Table S3). For protein identification and purification process, 6x His tag was incorporated at the C-terminus of the vaccine sequence. The final vaccine construct sequence was constructed with a molecular weight of 32928.31 Da with 309 amino acids (Fig. 2).

Population coverage and epitope conservancy analysis

In order to figure out the worldwide coverage of MHC Class-I and MHC Class-II allele interacting epitopes, population coverage analysis was performed. By utilizing the IEDB population coverage analysis tool, the most predominant potential epitopes were recognized for each coverage method. The fluctuation of the distribution of MHC HLA allele occurs among diverse geographical regions around the globe. Therefore, it is mandatory to engage population coverage for the design of a potential operative vaccine.

The highest population exposure for MHC Class-II allele was computed in the South Africa (98.2%) which was followed by Europe, North America and North Africa, with population coverage of 98.14, 97.18 and 97.01%, respectively. The lowest population coverage was computed in Central America (7.76%). The highest population coverage in case of MHC Class-II allele was computed in the North Africa (100%) which was closely followed by Europe, West Indies and Central America (both), East Africa and South Asia with population coverage of 99.99, 99.98, 99.96 and 99.92%, respectively. However, lowest population coverage was computed in South Africa (1.79%). When considering MHC Class-I, three epitopes are responsible for most of the interactions MHC Class-I alleles (TRSMWSFNP, FVLAAVYRI, FLFLTWICL); characterize a significant coverage in contradiction to the whole world inhabitants. The percentage of concentrated population coverage of these epitopes in the world was predicted to be 62.98% for TRSMWSFNP. In the case of MHC Class-II, 3 epitopes were anticipated to interact with frequent MHC Class-II alleles (IKLIFLWLL, LFLTWICLL, NLVIGFLFL); concluding an enormous global coverage. The amplest population coverage percentage of these epitopes in the world was given to IKLIFLWLL with percentage of 97.96%.

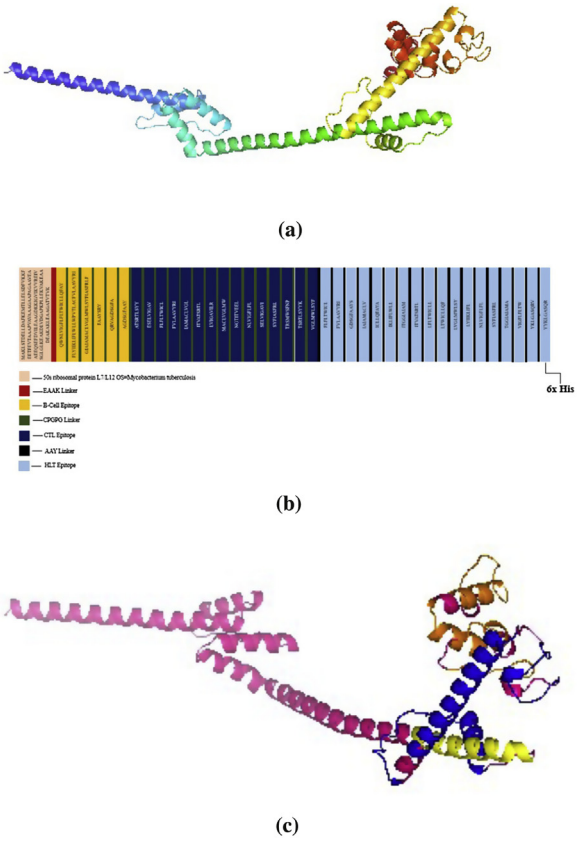

Fig. 2. (a) Vaccine assemblage by M-protein. The 3D structure of designed vaccine is shown. (b) Graphical view of constructed vaccine. All epitopes are represented in different colors. All linkers, 50 s ribosomal protein and histadine tag are also shown in different colors. (c) represents the presence of different epitopes on the surface of M-protein. Yellow highlighted area represents the B-cell epitopes present on M-protein surface. CTL epitopes are highlighted in blue color and HTL epitopes are indicated in orange color.

The most fascinating part in this assessment is the population coverage investigation results for the abundant binders to MHC Class-I and MHC Class-II alleles, respectively in combined form; demonstrate a remarkable coverage with the percent of 96.66% and 99.75%, respectively (Figs. S4 and S5). The IEDB conservancy analysis tool evaluated the conservancy of the predicted epitopes (Tables S1 & S2).

Antigenicity and allergenicity evaluation of the vaccine protein

The antigenicity of the vaccine protein along with adjuvant was calculated using VaxiJen 2.0 web server and it was predicted to be 0.5430. The antigenicity of vaccine construct without adjuvant was found to be 0.6604. It was concluded from the results that the vaccine construct is antigenic in nature either it is attached with adjuvant or not. The vaccine was found to be non-allergenic according to the results of AllerTOP v2 either the adjuvant is attached or not. The toxicity of the protein was found to be non-toxic by Toxinpred with or without adjuvant.

Analysis of solubility and physiochemical properties of multi-epitope subunit

Physiochemical properties were predicted using ExPASY ProtParam tool and the results provided various properties related to the protein nature. The molecular weight (MW) of the multi-epitope subunit was 32928.31 Da. The calculated pI of the protein was 5.74. This value suggested that the protein is of basic nature. The instability index (II) was 26.63. The instability index calculations ranked it as a stable protein as the value greater than 40 indicated the instability of the protein. It was confirmed as a thermostable protein because of the aliphatic index values that were found to be 104.34 [32] and GRAVY index was 0.417. The positive value indicates that the protein is not hydrophilic in nature [33]. Our vaccine construct showed a higher solubility rate calculated through SOLpro server with the score 0.500000.

Secondary structure extrapolation

Extrapolation of secondary structure was performed using Raptor X tool that analyzed the actual nature of the protein. Results provided various aspects of the protein and it was concluded that the protein contains 48% helix, 9% beta strands and 41% coils. Total 46% protein content was found to be exposed, 25% was found to be medium exposed and 17% buried. There were total 15% residues found to be present in disordered domain.

Tertiary structure assessment of the protein

1 best tertiary structure model of chimeric vaccine construct were constructed by 3D PRO by Scratch Protein Predictor. On the basis of high coverage values, the models were predicted using top 10 threading templates. In this study, the model having highest coverage-score was selected for refinement procedures [34].

Refinement of the tertiary structure

GalaxyRefine tool provided total five models of the vaccine chimera after refinement. Various parameters were considered in the process of refinement such as GDT-HA (0.8875), RMSD (0.568), and MolProbity (2.000). The calculation of clash score was found to be 17.8, score of poor rotamers was 1.3 and the Ramachandran score was predicted as 97.1%. For latter investigations, model 3 was selected as it was found to be the most authentic one.

Validation of 3D structure

The refined tertiary structure was validated by RAMPAGE server. The structure was analyzed, and Ramachandran plot was produced for the structure of protein (Fig. S6). There was total 92% region that was in favored region of the plot before refinement and only 8% structural region was found to be in allowed region. 0% structure was in the outlier region. After performing the refinement, better results were generated by RAMPAGE. According to these results, total 94.3% residues in favoured region, 5.3% residues in allowed region and 0.4% residues in outlier region.

Molecular docking with ligand binding domain of TLR3

Molecular docking was performed in order to estimate the interaction between refined vaccine model and ligand binding domain of immune receptor TLR3 by using online server Cluspro2.0 for protein-protein docking [35]. Docking gives multiple models for analysis. After analyzing all 30 docked poses, model number 9 was proved to be the best-docked model having 26 hydrogen bonds and 9 electrostatic interactions with score of −310.8 as shown in Fig. 3. The interacting residues of vaccine and receptors are A:LYS74:HZ2–LX0:ASP97:OD1, A:LYS165:HZ1–LX0:GLU90:OE1, A:LYS165:HZ3–LX0:GLU90:OE2, A:LYS165:HZ2–LX0:ASP83:OD1, A:LYS286:HZ1–LX0:GLU50:OE1 A:LYS286:HZ3–LX0:GLU50:OE2, A:LYS371:HZ1–LX0:GLU29:OE1, A:LYS371:HZ2–LX0:GLU29:OE2, A:LYS447:HZ1–LX0:GLU14:OE1, A:LYS447:HZ2–LX0:GLU14:OE2, LX0:LYS60:HZ1–A:GLU145:OE1, LX0:LYS60:HZ3–A:GLU145:OE2, LX0:LYS60:HZ3–A:ASP53:OD2, LX0:LYS91:HZ1–A:GLU99:OE2, LX0:LYS91:HZ2–A:GLU145:OE1, LX0:LYS91:HZ3–A:GLU99:OE1, A:LYS74:NZ–LX0:GLU95:OE2, LX0:LYS60:NZ–A:GLU99:OE1, A:LYS74:HZ1–LX0:LYS94:O, A:LYS74:HZ3–LX0:GLU95:O, A:SER147:HG–LX0:GLU56:O, A:ASN148:HN–LX0:GLU56:OE1, A:ASN148:HD21–LX0:GLU56:OE1, A:SER169:HG–LX0:GLU56:OE2, A:LYS371:HZ3–LX0:GLN48:OE1, LX0:ALA57:H–A:ASP124:OD2, LX0:LYS60:HZ2–A:ASN77:OD1, LX0:LYS91:HZ2–A:THR122:OG1, LX0:LYS94:HZ2–A:GLN144:OE1, LX0:LYS100:HZ1–A:ASN96:OD1, A:SER147:CB–LX0:GLU56:OE1, LX0:GLY58:CA–A:GLN79:OE1, A:GLN79:HE21–LX0:LYS60:HZ2, LX0:ALA57–A:MET103, LX0:LYS94–A:ILE121.

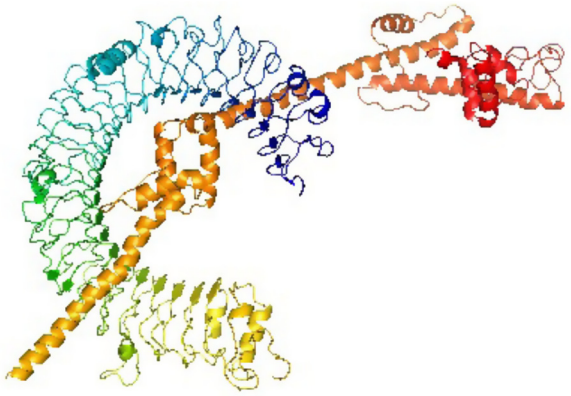

Fig. 3. Docked complex of vaccine construct and TLR-3.

Molecular dynamics simulation

IModS does the critical analysis of the structure by adjusting force field of the complex with respect to different time interval. The resultant model shows less deformation at each residues capacity level. The eigon value of the complex is 2.930209e − 07. Low RMSD and highly co-related region in heat maps revealed better interactions of the individual residues (Fig. 4).

The figure explains the detailed result explanation from IModS molecular dynamics simulation results. In Fig. 4(a) the MNA mobility is shown in the given protein structure and the (b) part shows deformability which indicates low level of deformation at all the residues. In Fig. 4(c) the B-factor is indicated. In (d), eigon values are indicated as 2.930209e − 07 and (e) indicates the variance explained in both red and green color. The other (f) and (g) figures show the co-variance and elastic network of the complex as well.

Codon optimization of designed vaccine peptide for expression analysis

Java codon adaptation tool (JCat) was employed for codon optimization to allow maximum protein expression. The optimized codon had a length of 927 nucleotides and it displayed that codon adaptation index (CAI) for optimized sequence was 0.96 with GC content of 51.88%. These values indicate stable vector expression in *E. coli*, as GC content ranges between 30–70% and is considered as optimum (Fig. S7a). Finally, the optimized sequence was cloned in pET28 (+) vector to make a recombinant plasmid after being

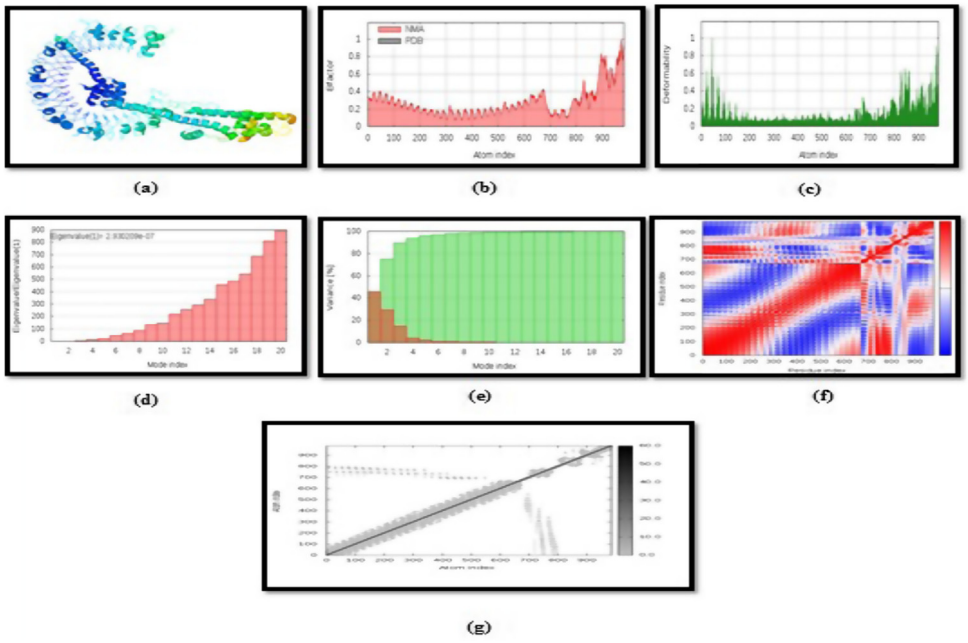

Fig. 4. The results of molecular dynamic simulation study of vaccine construct and TLR-3 docked complex. (a) MNA mobility, (b) deformability, (c) B-factor, (d) Eigon values, (e) variance (red color indicates individual variances and green color indicated cumulative variances), (f) co-variance map (correlated (red), uncorrelated (white), or anti-correlated (blue) motions) and (g) elastic network (darker grey regions indicate more stiffer regions).

amplified through *in-silico* PCR by using SnapGene software (Fig. S7b–c).

Immune simulation

Immune simulation was carried by C-ImmSim server. This exhibits immune response as that of real immune response. The primary response was marked by the increased level of IgM. High B-cell population characterized secondary and tertiary responses. As the concentration of antigen decreased; the concentration of IgM, IgG_1 + IgG_2 and IgG + IgM increased (Fig. S8a). The results also showed the production of memory cell upon subsequent exposure (Fig. S8b). Along with this, there was also an increased concentration of helper (T_H) and cytokines (Fig. S8c–d).

Discussion

With the rise in global crises of this deadly virus, the need of the hour is to take a step forward in order to find the cure against this novel disease. With the advancement of computational aided sequence-based technology, bioinformatics tools provide a vital approach in peptide-based vaccine designing. The peptide based vaccine design in other viruses like dengue virus, chikungunya, rhino virus, and SLE virus have already proved a potential pipeline of viral targeting [36]. RNA based SARS-CoV-2 shows multiple resistance due to high mutation in its genome. Because of this reason, the outer protein or membrane protein (M Protein) covering the virus is the main focus of this study [37]. However, the physicochemical and secondary structure analysis of the targeted protein reveals it to be highly antigenic helical candidate to design vaccine. Moreover, the PDB structure of the related protein was also designed by using bioinformatics software due to its unavailability in different protein data banks. Overall, this study demonstrates the design of multi-epitope-based subunit peptide vaccine by means of *in-silico* approaches to increase the both humoral and cell-mediated immune response.

In the past, B-cells were the only source for the design of a potential vaccine. However, with the innovation in computational biology, targeting major histocompatibility complex (MHC) T-cells with most interacting human leukocyte antigen (HLA) scheme is the new field of clinical research [38]. With interval, the antigenic drift may liberate the antigen from the antibody memorial response. Although, the T-cell immunity gives long-lasting immune response [39], the journey of epitope to become vaccine involves some strict criteria. First, we develop the database of all possible epitopes of M protein. For B-cell epitope prediction, five methods from IEDB database have been chosen for the prediction of antigenicity. Parameters such as turns, hydrophilicity, flexibility, polarity exposed surface, accessibility, and antigenic propensity of polypeptide chains have been associated with the position of an epitope. The computational analysis of Kolaskar and Tongaonkar, Emini surface, Chou and Fasman beta-turn, Karplus and Schulz and Bepipred linear epitope give the graphical idea of each residue which have the potency to contribute in the epitope development [40]. However, three of them provide the peptide sequence of epitopes for further analysis. On the other hand, IEDB also comes in handy when analysis was made on binding and processing prediction of T cells [40]. Here, we screened out potential T-cell epitopes with IC_{50} value less than 200 making it highly active against its targeted allele. Considering the consensus approach, epitopes interacting with more than five MHC Class-I and MHC Class-II get separated for further screening.

The epitopes get strengthen debate after proving best among different criteria. Foremost of them is the checking antigenic property of both B and T-cells. Using Vaxijen v2.0, the ACC calculation of the peptide based on its physicochemical properties was performed. Peptides with a value higher than the threshold declared as antigen in nature. The antigenic epitope should be nontoxic in nature in order to deliver a successful immune response. Toxinpred uses SVM classifier to predict the peptide toxicity. Allergenicity is the prominent obstacle in vaccine development [40]. Currently, most vaccines develop the allergic reaction in order to stimulate the immune system. The AllerTop v.2.0 works on the quantitative structural activity relationship (QSAR) to give score value which is the probability of allergen [41]. The method is constructed on auto cross covariance (ACC) proposed epitope's with no allergenicity screened

out to be as a non-allergen. However, according to the FAO/WHO scheme of predicting allergenicity, a sequence is possibly allergenic if it has an at least six contiguous amino acids when compared to known allergens amino acid database [42]. Hence, our selected epitopes according to AllerTop v.2.0 did not accomplish the criteria for the FAO/WHO assessment scheme of allergenicity prediction and thus characterized as a non-allergen. The putative antigenic epitopes with no allergenicity and toxicity making immunoreactive peptides for our further studies [43].

Conservancy of screened epitopes, among all the strains of HCoV, make them potential candidate towards vaccine development. IEDB screened out epitope shows good conservancy in fraction of protein sequences and identityis the degree of correspondence among strains. Moreover, population coverage analysis for T-cell was done by IEDB as MHC molecules are enormously polymorphic and found in thousand diverse human MHC (HLA) alleles [44]. For this purpose, selected multiple T-cell peptides with different HLA binding were analyzed. Both classes of MHC show high conservation throughout the globe. This will have increased coverage of HCoV patient's population embattled by peptide-based vaccines. After accomplishing all criteria, 5 epitopes of B-cell, 16 epitopes of MHC Class-I T-cells and 18 epitopes of MHC Class-II T-cells were selected as subunits for the vaccine construction process.

A study on 50S ribosomal protein L7/L12 shows that it is considered to be involved in increasing the response of vaccine in the pathogen recognition and immune system activation, therefore, selected as adjuvant to increase the immunoreactive property [45]. Initially selected B and T-cell epitopes were fused using suitable linkers in order to develop the multiepitope subunit vaccine construct. Due to their optimum effect, spacer sequences are considered to be central in the processes of vaccine development. From the previous experimentation, GPGPG and AAY linkers were integrated between the predicted epitopes to develop a potential vaccine with optimum antigenicity. In order to link the adjuvant with first predicted B-cell epitope, an EAAAK linker adjusted in the sequence design. The entanglement of this linker has been reported in designing bi-functional peptides which boosts the fused protein. The 6× His tag at the C terminal of sequence, also acknowledged as polyhistidine tag, is motif of at least 6 histidine residues fused to the carboxyl (C-) terminal. Histidine residues bind to immobilized ions, making easy for the sequence to work in a buffer condition [46].

The bioinformatic analysis together with immunologic investigation directed that the constructed protein sequence has absence of allergenic as well as toxic properties. Few documented studies reveal the low value of antigenicity of vaccine construct. Although, this constructed vaccine chimera expressed satisfactory antigenic score either bound to adjuvant or not. The molecular weight of designed vaccine protein was found to be 32 kDa and it was further analyzed for its solubility according to its stimulated antigenicity. Approving the basic nature of vaccine protein, the theoretical pI of the vaccine is 5.71. The vaccine shows low instability index having the value of 26.63, which means the designed vaccine protein, is stable and could be presumed as a vaccine model when expressed. The chimeric vaccine construct was predicted to be thermostable as per aliphatic index calculations. The secondary and tertiary structures are considered vital in the vaccine design. The secondary structure analysis of the chimeric vaccine protein was found to be consisting 49% helix, 13% beta-turn and 37% coil. The refinement of 3D structure of chimeric vaccine construct was performed by GalaxyRefine and desirable properties were obtained. RAMPAGE (Ramachandran Plot) shows the suitable properties that are required for a potential vaccine candidate. The results verified that most of the residues are present in favored areas and only a few residues in outlier region. It indicates the acceptable intended model quality.

Next step in the vaccine construction is crucial to achieve promising impact. Immunoinformatics involving computational methods such as molecular docking has been recognized as a potent tool when it comes to predict the protein-protein interaction. Various studies have reported the involvement of TLR particularly TLR3 in generating immune response against SARS-CoV-2. A research has described the specificity of TLR3 in the action of innate immunity [47]. Docking is done by Clustpro2.0 showing that the model 9 of the vaccine-receptor complex exhibits promising interaction proximal to reference structure. Least energy value of the docked complex shows the stable interaction and less RMSD from the original conformation. The presence of strong electrostatic, van der Waals, hydrogen bonds and hydrophobic interaction making ligand in stable configuration inside the binding pocket of the receptor [37]. Docking gives us single snapshot of complex physiological movement. Therefore, there is a need to study protein-protein interaction in a more flexible environment [33]. For this purpose, molecular dynamics simulation was performed mimicking the natural behavior of our dynamic system. The results indicated various factors that are explained in Fig. 4. The figure explains the MNA mobility of the complex structure, deformability values, B-factor, eigenvalues, c-variance and elasticity of the complex. Based on maximum eigenvalue, the constructed complex is shown to be stable and reveals less chances of deformation during immune response. Immune simulation of the designed construct was discovered by co-variance matrix analysis and its output was in accordance with the immune responses. The humoral response was anticipated to be produced as a result of vaccine introduction in the body [48,49].

Typical immune response was obtained by immune simulation. The immune response is enhanced as the antigen exposure is increased repeatedly. It was evident that memory B-cells were developed that are consistent for several months. There were also the development of memory T-cells along with helper T-cells simulation. There was an indication of increased level of IL-2 after first injection. It had been studied that the T-cell response is higher for structural protein [37]. In order to validate the designed vaccine, it was tested for immunoreactivity [50], which was done by expressing it in *E. coli* [51]. For maximum expression the codon was optimized according to host which results in CAI of 0.96 and GC content of 51.8%.

However, many vaccine candidates have been evaluated using *in-silico* approaches that could be possible vaccine candidates but no vaccine has been efficiently designed using M-protein of the novel coronavirus. Furthermore, in previously designed vaccines, immune simulation and vaccine cloning were not performed. Some vaccine candidates have a low population coverage as compared to our designed vaccine chimera. Therefore, it would prove an efficient possible vaccine candidate if it is carried out *in vitro* and *in vivo* studies.

Conclusions

In the present study, computational approaches were used to develop an effective vaccine against SARS-CoV-2 infections. Providence of artificial environment favors the illustrated approach to design promising vaccine as a time saving and cost effective strategy due to minimal chances of failure. *In silico* vaccine with a good immune response and population coverage is the potential candidate for clinical trials. *In silico* immune simulation showed immune response in accordance with the clearing of antigen. Computational cloning in *PET28a* (+) plasmid using snapgene showed good protein expression. However, to ensure the efficacy of vaccine construct against COVID-19, experimental validation is essential. In many trials, peptide vaccines had shown good results with a better immune response, so this vaccine construct may be considered. This study would be definitely helpful in eradicating the global threat caused by COVID-19 by developing an effective vaccine.

Funding

The project was supported by grant from The Oman Research Council (TRC) through the funded project (BFP/RGP/HSS/19/198).

Competing interests

There are no conflicts to declare.

Ethics approval

Not applicable because there are no animals and human used in this study.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Code availability

Not applicable.

Authors' contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgements

We acknowledge the support of Department of Biotechnology, University of Central Punjab, Lahore. The authors would like to thank the University of Nizwa for the generous support of this project. We thank analytical and technical staff for assistance.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jiph.2021.04.010.

References

[1] Weiss SR, Leibowitz JL. Coronavirus pathogenesis. Adv Virus Res 2011;81:85–164. Elsevier.
[2] Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends Microbiol 2016;24:490–502.
[3] Peiris JSM, Lai ST, Poon LLM, Guan Y, Yam LYC, Lim W, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 2003;361:1319–25.
[4] Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 2012;367:1814–20.
[5] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395:497–506.
[6] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 2020;382:1199–207.
[7] Wu P, Hao X, Lau EHY, Wong JY, Leung KSM, Wu JT, et al. Real-time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan, China, as at 22 January 2020. Eurosurveillance 2020;25.
[8] Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun 2020;109:102433.
[9] Jin Y-H, Cai L, Cheng Z-S, Cheng H, Deng T, Fan Y-P, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (SARS-CoV-2) infected pneumonia (standard version). Mil Med Res 2020;7:4.
[10] Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. bioRxiv 2020.
[11] Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. JAMA 2020;323:1061–9.
[12] Indwiani Astuti Y. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): an overview of viral structure and host response. Diabetes Metab Syndr 2020.
[13] Hoffmann M, Kleine-Weber H, Krüger N, Muller M, Drosten C, Pöhlmann S. The novel coronavirus 2019 (SARS-CoV-2) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. bioRxiv 2020.
[14] Li X, Geng M, Peng Y, Meng L, Lu S. Molecular immune pathogenesis and diagnosis of COVID-19. J Pharm Anal 2020;10:102–8.
[15] Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, et al. Diagnosis of the coronavirus disease (COVID-19): rRT-PCR or CT? Eur J Radiol 2020;126:108961.
[16] Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (SARS-CoV-2) in vitro. Cell Res 2020;30:269–71.
[17] Caly L, Druce JD, Catton MG, Jans DA, Wagstaff KM. The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro. Antiviral Res 2020;178:104787.
[18] Badgujar KC, Badgujar VC, Badgujar SB. Vaccine development against coronavirus (2003 to present): an overview, recent advances, current scenario, opportunities and challenges. Diabetes Metab Syndr Clin Res Rev 2020;14(September (5)):1361–76.
[19] Sturman LS, Holmes K, Behnke J. Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. J Virol 1980;33:449–62.
[20] Sturman LS. Characterization of a coronavirus: I. structural proteins: effects of preparative conditions on the migration of protein in polyacrylamide gels. Virology 1977;77:637–49.
[21] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007;8:4.
[22] Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Springer; 2005. p. 571–607.
[23] Novák P, Havlíček V. Protein extraction and precipitation. In: Proteomic profiling and analytical chemistry. Elsevier; 2016. p. 51–62.
[24] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M, Appel R, Bairoch A, Walker JM (ed): The proteomics protocols handbook. Humana Press: 2005.
[25] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–32.
[26] Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED protein analysis workbench. Nucleic Acids Res 2013;41:W349–57.
[27] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: protein structure and function prediction. Nat Methods 2015;12:7.
[28] Ferrè F, Clote P. DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. Nucleic Acids Res 2006;34:W182–5.
[29] Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 1990;276:172–4.
[30] Rose GD, Glerasch LM, Smith JA. Turns in peptides and proteins. Adv Protein Chem 1985;37:1–109. Elsevier.
[31] Karplus P, Schulz G. Prediction of chain flexibility in proteins. Naturwissenschaften 1985;72:212–3.
[32] Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem 1980;88:1895–8.
[33] Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, Prajapati VK. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. Sci Rep 2017;7:1–13.
[34] Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. Proteins Struct Funct Bioinform 2004;57:702–10.
[35] Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 2005;33:W363–7.
[36] Ahmad I, Ali SS, Shah I, Khan S, Khan M, Ullah S, et al. Computational vaccinology based development of multi-epitope subunit vaccine for protection against the Norovirus infections. bioRxiv 2020;73:390–400.
[37] Li G, Fan Y, Lai Y, Han T, Li Z, Zhou P, et al. Coronavirus infections and immune responses. J Med Virol 2020;92:424–32.
[38] Basu A, Sarkar A, Maulik U. Strategies for vaccine design for corona virus using Immunoinformatics techniques. bioRxiv 2020.
[39] Phan AT, Goldrath AW, Glass CK. Metabolic and epigenetic coordination of T cell and macrophage immunity. Immunity 2017;46:714–29.
[40] Oany AR, Emran A-A, Jyoti TP. Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. Drug Des Devel Ther 2014;8:1139.
[41] Khan S, Khan A, Rehman AU, Ahmad I, Ullah S, Khan A, et al. Immunoinformatics and structural vaccinology driven prediction of multi-epitope vaccine against

Mayaro virus and validation through in-silico expression. Infect Genet Evol 2019;73:390–400.
[42] Herman RA, Song P, ThirumalaiswamySekhar AJC, Allergy M. Value of eight-amino-acid matches in predicting the allergenicity status of proteins: an empirical bioinformatic investigation. Clin Mol Allergy 2009;7:9.
[43] Khan MK, Zaman S, Chakraborty S, Chakravorty R, Alam MM, Bhuiyan TR, et al. In silico predicted mycobacterial epitope elicits in vitro T-cell responses. Mol Immunol 2014;61:16–22.
[44] Chiu SS, Chan KH, Chu KW, Kwan SW, Guan Y, Poon LLM, et al. Human coronavirus NL63 infection and other coronavirus infections in children hospitalized with acute respiratory disease in Hong Kong, China. Clin Infect Dis 2005;40:1721–9.
[45] Oliveira S, Zhu Y, Splitter G. Recombinant L7/L12 ribosomal protein and gamma-irradiated Brucella abortus induce a T-helper 1 subset response from murine CD4+ T cells. Immunology 1994;83:659.
[46] Majid M, Andleeb S. Designing a multi-epitopic vaccine against the enterotoxigenic Bacteroides fragilis based on immunoinformatics approach. Sci Rep 2019;9:1–15.
[47] Garcia MM, Goicoechea C, Molina-Álvarez M, Pascual D. Toll-like receptor 4: a promising crossroads in the diagnosis and treatment of several pathologies. Eur J Pharmacol 2020:172975.
[48] Shey RA, Ghogomu SM, Esoh KK, Nebangwa ND, Shintouo CM, Nongley NF, et al. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. Sci Rep 2019;9:1–18.
[49] Chauhan V, Rungta T, Rawat M, Goyal K, Gupta Y, Singh MP. Excavating SARS-coronavirus 2 genome for epitope-based subunit vaccine synthesis using immunoinformatics approach. J Cell Physiol 2020;(July).
[50] Gori A, Longhi R, Peri C, Colombo G. Peptides for immunological purposes: design, strategies and applications. Amino Acids 2013;45:257–68.
[51] Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 2014;5:172.