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.

Informatics in Medicine Unlocked 21 (2020) 100438

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

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

Exploration of surface glycoprotein to design multi-epitope vaccine for the prevention of Covid-19

Elijah Kolawole Oladipo [a,h,*], Ayodeji Folorunsho Ajayi [b], Olumuyiwa Elijah Ariyo [c], Samson Olugbenga Onile [d], Esther Moradeyo Jimah [e,h], Louis Odinakaose Ezediuno [f,h], Oluwadunsin Iyanuoluwa Adebayo [b,h], Emmanuel Tayo Adebayo [b,h], Aduragbemi Noah Odeyemi [b,h], Marvellous Oluwaseun Oyeleke [b], Moyosoluwa Precious Oyewole [g], Ayomide Samuel Oguntomi [h], Olawumi Elizabeth Akindiya [i], Bunmi Olayemi Olamoyegun [h], Victoria Oyetayo Aremu [b,h], Abiola O. Arowosaye [j], Dorcas Olubunmi Aboderin [k], Habibat Bolanle Bello [l], Tosin Yetunde Senbadejo [m], Elukunbi Hilda Awoyelu [n], Adio Abayomi Oladipo [o], Bukola Bisola Oladipo [p], Lydia Oluwatoyin Ajayi [q], Olusola Nathaniel Majolagbe [r], Olubukola Monisola Oyawoye [a], Julius Kola Oloke [n]

[a] Department of Microbiology, Laboratory of Molecular Biology, Immunology and Bioinformatics, Adeleke University, Ede, Osun State, Nigeria
[b] Reproduction and Bioinformatics Unit, Department of Medical Physiology, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria
[c] Department of Medicine, Infectious Diseases and Tropical Medicine Unit, Federal Teaching Hospital, Ido-Ekiti, Ekiti State, Nigeria
[d] Biotechnology Programme, Department of Biological Sciences, Elizade University, Nigeria
[e] Department of Medical Microbiology and Parasitology, University of Ilorin, Kwara State, Nigeria
[f] Department of Microbiology and Parasitology, University of Ilorin, Kwara State, Nigeria
[g] Department of Biochemistry, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria
[h] Genomics Unit, Helix Biogen Consult, Ogbomoso, Oyo state, Nigeria
[i] Microbiology Programme, Department of Biological Science, Olusegun Agagu University of Science and Technology, Okitipupa, Ondo State, Nigeria
[j] Department of Virology, University of Ibadan, Ibadan, Oyo State, Nigeria
[k] Department of Medical Microbiology and Parasitology, University of Lagos, Lagos State, Nigeria
[l] Department of Microbiology, Kebbi State University, Aliero, Kebbi State, Nigeria
[m] Department of Microbiology, Fountain University, Osogbo, Osun State, Nigeria
[n] Department of Natural Sciences, Precious Conerstone University, Ibadan, Oyo State, Nigeria
[o] Department of Haematology and Blood Grouping Serology, Obafemi Awolowo Teaching Hospital Complex, Ile-Ife Wesley Guild Hospital Wing, Osun State, Nigeria
[p] Department of Clinical Nursing, Bowen University Teaching Hospital, Ogbomoso, Oyo State, Nigeria
[q] Department of Biochemistry, Adeleke University, Ede, Osun State, Nigeria
[r] Department of Pure and Applied Biology, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

ARTICLE INFO

Keywords:
COVID-19
SARS-CoV-2
Subunit vaccine
TLRs
Immunity

ABSTRACT

Stimulation and generation of T and B cell-mediated long-term immune response are essential for the curbing of a deadly virus such as SAR-CoV-2 (Severe Acute Respiratory Corona Virus 2). Immunoinformatics approach in vaccine design takes advantage of antigenic and non-allergenic epitopes present on the spike glycoprotein of SARS-CoV-2 to elicit immune responses. T cells and B cells epitopes were predicted, and the selected residues were subjected to allergenicity, antigenicity and toxicity screening which were linked by appropriate linkers to form a multi-epitope subunit vaccine. The physiochemical properties of the vaccine construct were analyzed, and the molecular weight, molecular formula, theoretical isoelectric point value, half-life, solubility score, instability index, aliphatic index and GRAVY were predicted. The vaccine structure was constructed, refined, validated, and disulfide engineered to get the best model. Molecular binding simulation and molecular dynamics simulation

* Corresponding author. Department of Microbiology, Laboratory of Molecular Biology, Immunology and Bioinformatics, Adeleke University, Ede, Osun State, Nigeria. Tel.: 2348035369472.
E-mail address: koladipo2k3@yahoo.co.uk (E.K. Oladipo).
URL: http://drkoladipo2k3@helixbiogenconsult.org (E.K. Oladipo).

https://doi.org/10.1016/j.imu.2020.100438
Received 30 July 2020; Received in revised form 11 September 2020; Accepted 27 September 2020
Available online 4 October 2020

were carried out to predict the stability and binding affinity of the vaccine construct with TLRs. Codon acclimatization and *in silico* cloning were performed to confirm the vaccine expression and potency. Results obtained indicated that this novel vaccine candidate is non-toxic, capable of initiating the immunogenic response and will not induce an allergic reaction. The highest binding energy was observed in TLR4 (Toll-like Receptor 4) (−1398.1), and the least is TLR 2 (−1479.6). The steady rise in Th (T-helper) cell population with memory development was noticed, and IFN-g (Interferon gamma) was provoked after simulation. At this point, the vaccine candidate awaits animal trial to validate its efficacy and safety for use in the prevention of the novel COVID-19 (Coronavirus Disease 2019) infections.

1. Introduction

COVID-19 is a respiratory infection that is caused by a beta-coronavirus known as SARS-COV-2. It is an RNA virus which has 14 receptors binding residue on it glycoprotein that interacts with the angiotensin-converting enzyme 2 (ACE2) receptor [1]. Spike (S) glycoprotein, which protrudes from the viral surface mediates the entry of coronaviruses into the host cells [1,2]. On mature viruses, the spike glycoprotein comprises of two functional subunits which are S_1 head and S_2 stalk they function in the binding of the virus to host cell receptor and fusion of host cell membrane with the invading virus respectively. The S_1 head of SARS-CoV-2 has a receptor-binding domain (RBD) which recognizes and binds with human angiotensin-converting enzyme 2 (hACE2) with an affinity more significant than that of SARS-CoV [3,4]. Since the SARS-COV-2 glycoproteins (S) facilitate the viral entry into the host cell, they become the main target of antibodies. The spike glycoprotein of coronavirus is found to be antigenic and immunogenic with a good number of antigenic sites [5–7]. Therefore SARS-COV-2 spike glycoproteins are one of the essential proteins for multi-epitope vaccine construct. Consequently, the antigenicity and immunogenicity of the spike glycoprotein was exploited in this study.

A multi-epitope vaccine consists of epitope peptides (series or overlapping) retrieved from more than two immunodominant epitopes. These immunodominant epitopes are selected from appropriate candidate antigens [8]. The epitopes can be selected based on binding assays, evaluation of immunogenicity and analysis of motifs. Also, the potency of vaccine is assessed by antigenicity and allergenicity assays [9]. Linkers like AAY and GPGPG are added between two epitopes for effective separation needed for the efficiency of the epitope. At the N-terminal of the vaccine construct, an adjuvant is added to improve immunogenicity and vaccine delivery in the host [10]. Immunoinformatics tools have become a novel instrument in the design of a potential multi-epitope vaccine candidate, this tool have been used successfully in the design of multi-epitope vaccine against Hepatitis C virus infection [10]. The vaccine candidate successfully stimulated innate, humoral and cellular immune responses [10]. The design of a multi-epitope vaccine against dengue virus involved CTL (Cytotoxic T Lymphocyte) and HTL (Helper T Lymphocyte) epitopes that were common with epitopes of B-cell. The designed vaccine model was also subjected to the prediction of IFN-γ inducing epitopes using various informatics tools [11]. In addition, the primary amino acid sequence of the salivary protein was used to design a multi-epitope sub-unit vaccine against malaria parasite plasmodium, this vaccine composed of CTL, HTL and B-cell epitopes that were antigenic but non-allergenic [12]. The efficacy of this method has been established for the design of a vaccine to prevent different infections.

Presently there is no accepted vaccine or drug for the prevention or treatment of COVID-19. However, a lot of vaccine and drug candidate is currently in development, whereas, the WHO has stressed its concern on the use of antibiotics but recommended weakly the use of antiviral agent such as remdesivir, rather than not using it [13]. Beigel et al., also concluded that remdesivir shortens the time of recovery and lower respiratory tract infection in an adult patient with COVID-19 disease [14]. Hydroxychloroquine has been shown to interfere with SARS-CoV-2 binding with ACE 2 receptor, which makes it a possible remedy for the treatment of COVID-19 [15]. Although there are many potential therapeutic and prophylactic candidates for the treatment and prevention of COVID-19 but none has been approved as a suitable agent for the control of the virus.

Vaccination is the most efficient approach towards the prevention of infectious diseases, especially influenza, which is a highly contagious respiratory infection. Human vaccines that are licensed are mostly based on their capability to elicit humoral antibodies that block infection and possibly reduce pathogen loads [16].

Therefore, this study proposed the design of a multi-epitope protein subunit based vaccine, which comprised of immune-stimulating epitopes that can evoke a humoral and cell-mediated immune response to prevent COVID-19 infection.

2. Methodology

2.1. Retrieval of sequences

The whole-genome nucleotide sequences were retrieved from the National Center for Biotechnology Information (NCBI) and Global Initiative on Sharing All Influenza Data (GISAID). Thirty whole-genome nucleotides were selected in GISAID from six African countries (Congo-16, South Africa – 3, Senegal −7, Gambia-2, Egypt – 1, Ghana – 5). The whole-genome nucleotide sequences were selected based on their submission date (between 15 March and 3 May 2020) the date was picked to avoid imported cases considering community transmission (https://www.epicov.org) [17]. Five available sequences for three African countries, Nigeria – 1, Tunisia – 3 and South – Africa – 1 as at 5 May 2020, were retrieved from NCBI. A reference sequence of an isolate from Wuhan with accession number NC_045512.2 deposited on 17 January 2020, was recovered from NCBI (https://www.ncbi.nlm.nih.gov) [18]. Therefore, total of thirty-five sequences from eight African countries were used for this study.

2.2. Annotation of retrieved sequences

The retrieved African SARS-CoV-2 sequences from GISAID were annotated using the retrieved Wuhan isolate with accession number NC_045512.2 as a reference sequence using NCBI tools (https://www.ncbi.nlm.nih.gov/guide/data-software/) [19]. The whole-genome sequences were annotated to identify the position of previously established antigenic spike glycoprotein of SARS-COV-2 using pairwise alignment (https://www.ebi.ac.uk/Tools/psa/emboss needle/) [6,20], and the aligned nucleotide sequences of the spike glycoprotein were translated to protein with protein transeq (https://www.ebi.ac.uk/Tools/st/emboss_transeq/) [21].

2.3. Antigenicity prediction of spike glycoprotein

The antigenicity of the spike glycoprotein was predicted using ANTIGENpro on the scratch protein (http://scratch.proteomics.ics.uci.edu/) [22] and VaxiJen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) [23]. Scratch protein, sequences above the threshold of 0.8 were selected [22], while protein sequences above the limit of 0.4 were chosen on VaxiJen [24]. The spike glycoproteins that

Table 1
Selected epitopes: (a) selected CTL epitopes, (b) selected HTL epitopes, (c) selected B-cell epitopes.

S/N	CTL Epitopes	Score
	FTLPDWWLY	3.1911
	WTAGAAAYY	3.1128
	TSNQVAVLY	3.0758
	ATSRTLSYY	2.6146
	TSVDCTMY	2.3795
	STECSNLLL	2.3492
	KLDHRWNCY	2.1760
	ITSTSLKIY	2.1542
	GAEHVNNSY	1.9960
S/N	HTL Epitopes	Score
	EFLIFWSKRTKYYI	0.03
	QQEVFVYNVNFPLAV	0.05
	RLFARTRS	0.22
	ILFALLQRY	0.52
	HQMLIVT	1.43
	WWLYKMGIWS	1.45
	RARSVASQSIIAYT	2.20
	MAYRFNGIGVTQNVL	2.51
	MIAQYTSA	2.80
	ELLHAPATV	2.90
	DLPQGFSA	6.05
	GYFKIYSKHTPINLV	6.90
	FNDGVYFA	7.00
	YKLGASQRVA	8.35
S/N	B-Cell Epitopes	
	FTVEKGIYQTSNFRVQPT	
	LADAGFIKQYGDC	
	SNNLDSKVGGNYNYLYRLFRK	
	LQDWQLIKHRPFQQ	
	TVCGPKKSTNLVK	
	IHVSGTNGTKRFDN	
	SRNHSSQRATPWHYSDQTA	
	SIIAYTMSLGAENSVAYSN	
	FSTFKCYGVSPTKLNDLCF	
	VNNTVYDPLQPELDSFKEELDKY	
	RYYYRRAKAPTMEPS	
	NLCPFGEVFNATRFASVY	
	TGKIADYNYKLP	
	YHKNNKSWMESEFRVYSSANN	
	SWTSSYCWTPSRTL	
	FKNHTSPDVDLGDISGINA	

met the criteria of at least one of the predictive analyses were further subjected to the next phase of the study.

2.4. Prediction of cytotoxic T lymphocytes (CTL) epitopes

NetCTL 1.2 server accessed through http://www.cbs.dtu.dk/services/NetCTL/was utilized in the prediction of CTL epitopes of SARS-CoV-2 spike glycoprotein sequences that passed either ANTIGENpro (http://scratch.proteomics.ics.uci.edu/) or VaxiJen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html). The MHC-1 binding peptides, Proteasomal C-terminal cleavages and Transporter Associated with Antigen Processing (TAP) transport efficiency were the three major components that the prediction was based on. The threshold value used for the Epitopes identification during the CTL Epitopes prediction was 0.75, while the weights on C-terminal cleavage and TAP transport efficiency used during the prediction were 0.15 and 0.05, respectively [25].

2.5. Projection of helper T-Cell (HTL) epitope

The HTL Epitopes on SARS COV-2 surface glycoprotein were projected using the Immune Epitope Database (IEDB) (tools.IEDB.org/mhcii/) [26]. Three mouse alleles which are H2-IAb, H2-IEd and H2-IAd were selected for the Major Histocompatibility Class II (MHC II). The predicted Epitopes were chosen by making use of their MHC-II affinity and percentile rank (which has an inverse relationship with the affinity of the epitopes predicted in the database) [27]. The six lowest scored epitopes were selected for each allele.

2.6. Projection of B-Cell epitopes

The linear B-cell Epitopes of SARS COV-2 surface glycoproteins were projected using B-cell Epitope prediction server; BepiPred (http://www.cbs.dtu.dk/services/BepiPred) and BCPREDS (http://ailab-projects1.ist.psu.edu:8080/bcpred/index.html). All epitopes above 0.5 thresholds (BepiPred) and justified by BCPREDS were selected and subjected to further analysis [28,29]. In totality, eight B-cell Epitopes were chosen for the vaccine construct.

2.7. Construction of multi-epitope vaccine sequence

Subsequent to the predictions made above using different immunoinformatics tools, a multi-epitope vaccine was built using the predicted CTL, HTL and B-Cell epitopes. The predicted epitopes for CTL and HTL were subjected to toxicity test using ToxinPred (http://crdd.osdd.net/raghava/toxinpred). RS09 motif was used as an adjuvant to improve the immunogenicity of the vaccine [30]. The CTL Epitopes were linked using AAY, while HTL and B cell epitopes were linked using GPGPG [31].

2.8. Allergenicity prediction of constructed vaccine

AllerTOP (https://www.ddg-pharmfac.net/AllerTOP/) was employed in predicting the allergenicity of the final vaccine construct; the server uses k nearest neighbours (kNN) model with 85.3% accuracy [32]. Also, AlgPred was used to predict multi-epitope (http://www.imtech.res.in/raghava/algpred/submission.html) vaccine allergenicity with high precision; the different algorithm was combined (SVMc + MAST + ARPs BLAST + IgEepitope) and exploited using AlgPred. The different approaches used can be employed in the high accuracy prediction of allergenic proteins [33].

2.9. Antigenicity projection of constructed vaccine

VaxiJen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) was employed in the prediction of antigenicity of the vaccine;

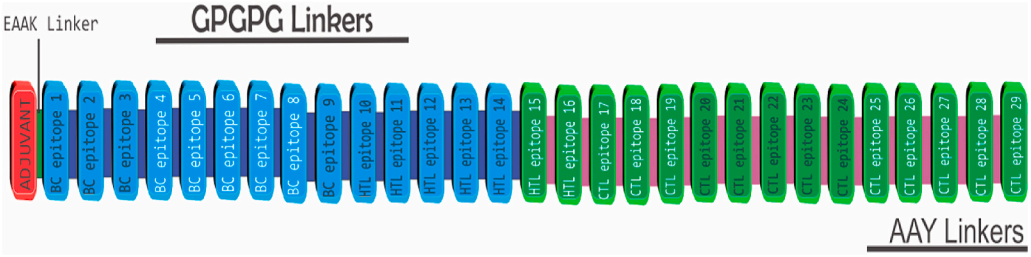
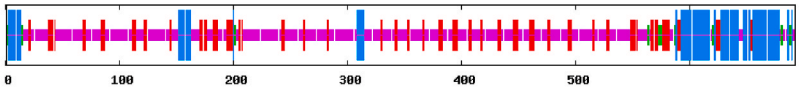


Fig. 1. The schematic representation of vaccine construct.

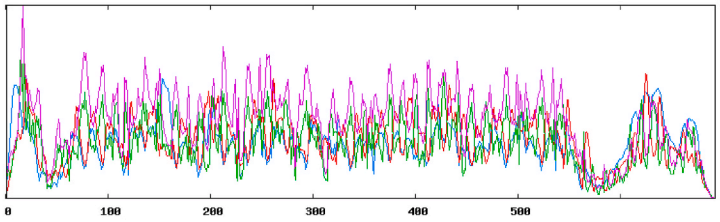
```
 10        20        30        40        50        60        70
         |         |         |         |         |         |         |
APPHALSEAAAKFKNHTSPDVDLGDISGINAGPGPGSWTSSYCWTPSRTLGPGPGYHKNNKSWMESEFRV
ccthhhhhhhhhhhttccccceeccccccccccccccccceeeeeeccccccccccccccccccccccee
YSSANNGPGPGTGKIADYNYKLPGPGPGNLCPFGEVFNATRFASVYGPGPGRYYYRRAKAPTMEPSGPGP
ecccccccccccceeeeccccccccccccccccccccccceeeeccccccceeeccccccccccccc
GVNNTVYDPLQPELDSFKEELDKYGPGPGFSTFKCYGVSPTKLNDLCFGPGPGSIIAYTMSLGAENSVAY
ccceeccccccchhhhhhhhhhhccccccceeeeeeccccceeeeeccccccceeeeeehttcceeecc
SNGPGPGSRNHSSQRATPWHYSDQTAGPGPGIHVSGTNGTKRFDNGPGPGTVCGPKKSTNLVKGPGPGLQ
ccccccccccccccccccccccccccccccceeeccccccccccccccceeccccccccccccccc
DWQLIKHRPFQQGPGPGSNNLDSKVGGNYNYLYRLFRKGPGPGLADAGFIKQYGDCGPGPGFTVEKGIYQ
cceeccccccccccccccccccccchhhhhhhccccccccccccceeccccccccceeeeccccc
TSNFRVQPTGPGPGILFALLQRYGPGPGRLFARTRSGPGPGHQMLIVTGPGPGDLPQGFSAGPGPGYKLG
ccceeccccccccccceeccccccccccceeeeccccccceeeeeccccccccceeeccccccccc
ASQRVAGPGPGELLHAPATVGPGPGMIAQYTSAGPGPGWWLYKMGIWSGPGPGRARSVASQSIIAYTGPG
ccccccccccceeeccccccccccceeeeeccccccccceeeeeccccccccccceeeccccccccccc
PGMAYRFNGIGVTQNVLGPGPGFNDGVYFAGPGPGGYFKIYSKHTPINLVGPGPGQQEVFVYNVNFPLAV
ccceeeeccccccccccccccccceeccccccccceeeccccccccccccccceeeeeeccccc
GPGPGEFLIFWSKRTKYYIAAYFTLPDWWLYAAYATSRTLSYYAAYWTAGAAAYYAAYTSNQVAVLYAAY
cccttceeeeeettteeeeeeccccthheeehhhhhhhhhhhhhhhhhhhhhhhhhctthheeeehhh
STECSNLLLAAYITSTSLKIYAAYTSVDCTMYAAYGAEHVNNSYAAYKLDHRWNCYHHHHHH
hhhhhhhhhhhhhccchhhhhheehhhhhhhhhhhhhhhhhhhhhhhcttccchhtthcc
Sequence length: 692
```

SOPMA :	Alpha helix (Hh) : 109 is 15.75%	3_{10} helix (Gg) : 0 is 0.00%		Pi helix (Ii) : 0 is 0.00%	
Beta bridge (Bb) :	0 is 0.00%	Extended strand (Ee) : 127 is 18.35%		Beta turn (Tt) : 18 is 2.60%	
Bend region (Ss) :	0 is 0.00%	Random coil (Cc) : 438 is 63.29%		Ambiguous states (?) : 0 is 0.00%	
Other states :	0 is 0.00%				

A



B1



B2

Fig. 2. Prediction of the secondary structure of the vaccine construct. (A) cartoon secondary structure of vaccine constructs showing alpha helix, extended strand, random coil and beta structure. (B1, B2) OMPL prediction of the secondary structure represented by different colours. Blue is alpha-helix, and green is Beta strands, red is extended strand and yellow random coil. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

this server calculates antigenicity based on two sets of data, 100 known antigenic peptides and 100 non-antigenic peptides which were modelled with an accuracy of 70%–89% [23,24]. ANTIGENpro located on http://scratch.proteomics.ics.uci.edu/is a sequenced-based, pathogen-independent and alignment-free predictor. It was also employed in the projection of the antigenicity of the vaccine construct. The antigenicity prediction was based on five machine learning algorithms and multiple representations of the primary sequence, which gives result based on protein microarray data [22].

2.10. Physiochemical properties and domain identification

Protparam (https://web.expasy.org/cgi-bin/protparam/) was employed in the determination of various physiochemical parameters of the constructed vaccine, the following parameters among others were accessed: molecular weight, theoretical Protrusion Index (PI),

Fig. 3. Tertiary structure of vaccine construct.

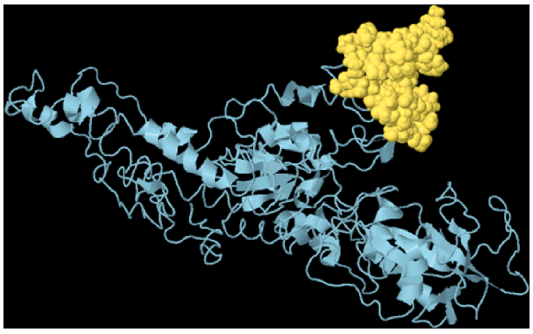


Fig. 4. The conformational B-cell epitopes of final vaccine construct.

hydropathicity (GRAVY), Aliphatic index, Instability index, Extinction coefficients, Atomic Composition, Charged Residues, *in vitro* and *in vivo* half-life [34].

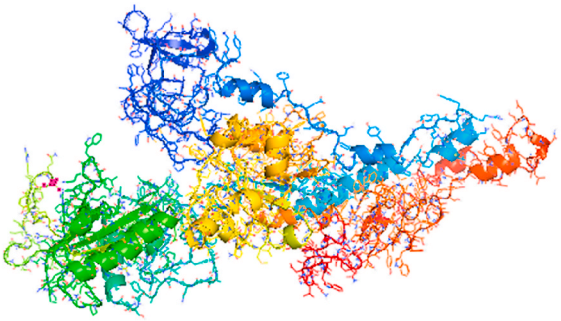
2.11. Prediction of secondary structure

The Self-Optimized Prediction Method (SOPMA) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) was employed for predicting the secondary structure of the vaccine construct. SOPMA predict parameters that include the helices, sheets, turns and coils [35,36].

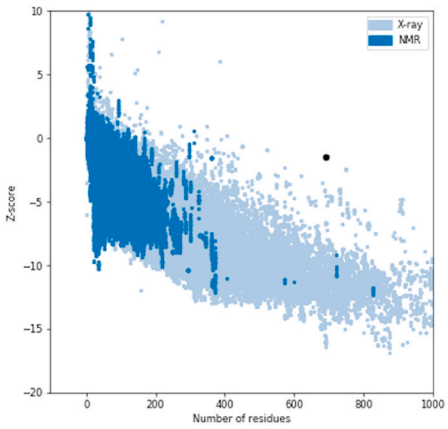
2.12. Prediction of 3D configuration and discontinuous B-Cell epitopes

I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER) was used to predict the tertiary structure of the multi-epitope vaccine construct. I-TASSER tertiary structure prediction server is designed to generate a tertiary protein structure using a quantitative assessment scoring system to obtain models. I-TASSER server gives up to five full-length output for each query: the estimated TM-score, confidence score, standard deviation and Root Mean Square Deviation (RMSD). These estimations were also predicted [37].

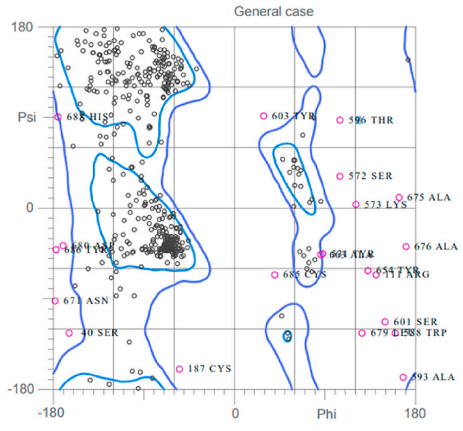
The tertiary composition of the vaccine construct was used for the prediction of B-Cell 3D conformational structure using Ellipro (http://tools.iedb.org/ellipro), this assigned each epitope a Protrusion Index (PI). Therefore, ellipro was used to determine the conformational 3D structure of the linear B-cell epitope predicted. Jmol viewer was used in visualizing the antibody epitope predicted. Ellipro results contain the number of residue in each of the epitopes, the higher the residue, the greater it's solvent availabilities [38].



a



b



c

Fig. 5. Protein validation: (a) Z-Score scatter graph (b & c) Ramachandran plot showing favoured allowed and disallowed region.

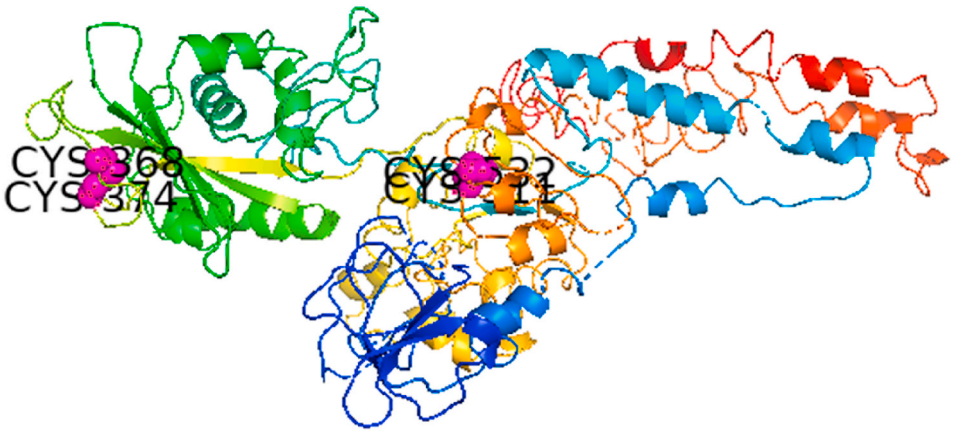


Fig. 6. Disulfide engineering showing disulfide bond as predicted by Disulfide by Design 2.

2.13. Refinement of the tertiary structure

In order to refine the multi-epitope vaccine construct, Galaxy Refine web (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) which is based on CASP 10 version, refinement method was employed for protein side-chain reconstruction which was followed by a relaxation of the structure through repacking and molecular dynamic simulation. Galaxy Refine is a proven performing algorithm for the enhancement of the quality of the local structure [39,40].

2.14. Validation of tertiary structure

ProSA-web server with the URL: https://prosa.services.came.sbg.ac.at/prosa.php is used for validation of the projected and refined 3D configuration of proteins as this is an essential step in modeling of sequences. ProSA server was employed in the validation of the 3D configuration of the vaccine construct; ProSA provided an overall magnitude of quality for the input structure. Outliers outside the characterized range of the native protein indicate a probability of occurrence of error. Ramachandran plot was also obtained by inputting the PDB file of the vaccine structure on PROCHEK server; this also validates the vaccine and authenticates their potential functionality [34,41–44].

2.15. Molecular docking of vaccine with toll-like receptors

Stimulating immune response is the target for vaccine design; this necessitates the importance of evaluating the interactions between an antigen and TLRs. ClusPro 2.0 server (http://cluspro.org/login.php?redir=/qeueue.php) was used to predict the binding pocket among TLR-2 (5d3i), TLR-3 (3ciy), TLR-4 (2z64) and TLR-9 (5zln) in the most stable complex form. Molecular docking also predicts the binding energy between the ligand and receptor by their scoring function. This was achieved by uploading the PDB file of TRLs, and the vaccine constructs onto cluspro 2.0 sever for processing and prediction [45,46].

2.16. Stability enhancement of final vaccine construct through disulfide engineering

Disulfide by design (http://cptweb.cpt.wayne.edu/DbD2/) was employed in the *in silico* process of Disulfide engineering, which provides strength to the 3D configuration of the Vaccine construct by using disulfide. The method of examining protein dynamics and interactions enhances the stability of the protein structure. Conformational entropy of unfolded protein slate is reduced due to the strength of protein after the increase. The quantification of the dynamic mobility of atoms in a protein known as B-factor is the criteria for selection [47].

2.17. Molecular dynamics simulation of receptor-ligand complex

Predicting the stability of the protein-ligand complex is essential, and molecular dynamics simulation is a well-accepted method to achieve this. iMOD server (http://inods.chaconlab.org) was used to carry out simulation; this approach studies the physical basis, the structure and function attributed to the biological molecules for determining the stability of the complex. This can be deduced from the result obtained as deformability, eigenvalues, and covariance [48,49].

2.18. Immune response simulation

In silico tool, C-ImmSim (http://150.146.2.1/C-IMMSIM/index.php) was used in profiling the immune response of the multi-epitope vaccine construct both humoral and cellular response against the predicted vaccine was illustrated. One shot of the vaccine product was administered *in silico*. One step of the simulation is equivalent to Eight hours of real-time; the simulation was performed using default parameters. The volume simulated and simulation steps of 1000 were used, (random seed = 12345 with an injection of vaccine containing no LPS).

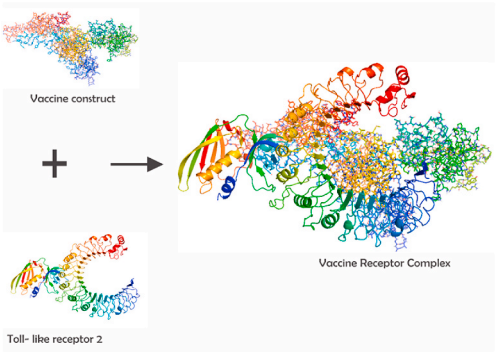
2.19. The in silico cloning and optimization of the vaccine protein

The vaccine construct was codon-optimized JCAT Java tool. JCAT tool expresses protein sequences into expression systems of another biological host to adapt the codon for the new host, JCAT provides GC and CAI value of the adapted codon, the tool also back-translate protein sequences to DNA sequences which are further used for *in silico* cloning. In this study, JCAT was used to adjust the final vaccine sequences into the *E. coli* K12 strain expression system by inputting the construct of the final vaccine, which is processed for adaptation [50]. DNA sequences obtained from back-translation was cloned into the *E. coli* K12 pET-28a (+) vector expression system at specific restriction enzymes with the aid of SnapGene software [51].

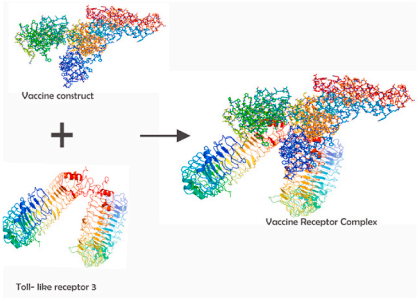
3. Results

3.1. Antigenicity prediction of spike glycoprotein

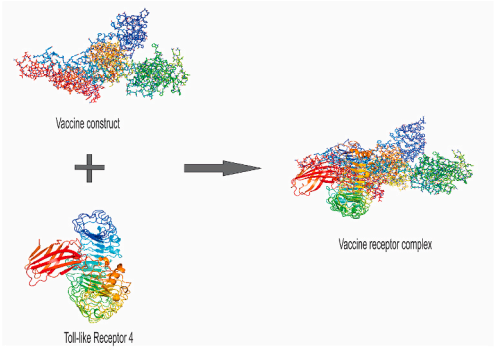
The antigenicity of spike glycoprotein annotated from each of the sequences was predicted, and 25.7% of the sequences passed AntigenPRO at a threshold 0.8. In comparison, 100% of the selected available sequences passed VaxiJen at a threshold of 0.4. The sequences that passed were then subjected to further analysis.



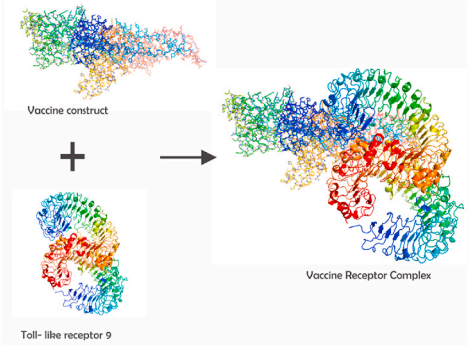
(a)



(b)



(c)



(d)

Fig. 7. Molecular docking of vaccine with TLR: (a) Molecular docking of vaccine with TLR2 (b) Molecular docking of vaccine with TLR3(c) Molecular docking of vaccine with TLR4 (d) Molecular docking of vaccine with TLR9.

Table 2
The binding energy weight of the selected docked model.

Interaction	Vaccine-TLR2	Vaccine-TLR3	Vaccine-TLR4	Vaccine-TLR9
Binding Energy Weight	−1479.6	−1414.6	−1398.1	−1477.3

3.2. Projection of novel cytotoxic T lymphocytes (CTL), helper T lymphocytes (HTL, MHC-II) and B-cells epitopes

CTL, HTL and B-cell epitopes were projected using a different server with the selected glycoprotein sequences that passed antigenicity screening. 9 CTL with 9-mer was predicted based on their high score, which was within the threshold of 0.75 (Table 1a). T Helper Lymphocytes epitopes were also predicted; those with low percentile rank were selected (Table 1b). B-Cell epitopes within the threshold of 0.5 were selected (Table 1c) and used in vaccine construction along with the HTL and CTL.

3.3. Construction of novel multiple epitope subunit vaccine

A multi-epitope vaccine was constructed from the predicted CTL, HTL and B-cell binding epitopes, linkers were employed in the linking of the epitopes, and an adjuvant was linked to the construct to aid the potency of the vaccine. AAY was used to link CTL together, GPGPG was used to connect HTL and B-cells, and EAAK was used to attach the adjuvant [30] to the whole construct. This construct is the final vaccine construct as shown in Fig. 1.

3.4. Allergenicity, antigenicity and toxicity prediction of constructed vaccine

The final vaccine construct was subjected to allergenicity, antigenicity and toxicity screening; the vaccine was non-allergenic as reported on the AllerTOP and Algpred server, VaxiJen 2.0 and ANTIGENpro predicted that the vaccine could elicit an antigenic reaction and tonxinPred predicted its non-toxicity. A comparative analysis of the epitopes in the vaccine with the human proteome was done using BLASTp server, which shows that no significant similarity exists between the vaccine and human proteome.

3.5. Physiochemical properties and domain identification

Physiochemical analysis done on protparam showed that the molecular weight of the construct was 73253.02. The construct has a molecular formula of $C_{3330}H_{4905}N_{895}O_{952}S_{17}$; the theoretical isoelectric point value was predicted at 9.22. The half-life was expected to be 4.4 h in mammalian reticulocytes *in vitro*, >20 h for yeast and >10 h in *E. coli*. Instability index (II) was 25.82; aliphatic index prediction was 56.60. The GRAVY was at-0.425. Alanine, Glycine and serine were observed to be in abundance.

3.6. Projection of secondary structure

SOPMA online server was utilized in the prediction of the secondary structure of the constructed vaccine. The server provided in addition to the secondary structure information on the vaccine construct, which includes, Alpha helix 15.75%, extended strand 19.35%, Random coil 63.29% and Beta turn structure 2.60%. These were all represented in Fig. 2. The high percentage of the random coil, as seen from the figure, indicates the high concentration of epitope at that point [52] (see Fig. 3).

3.7. Prediction of 3D configuration and discontinuous B-Cell epitopes

I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER) was

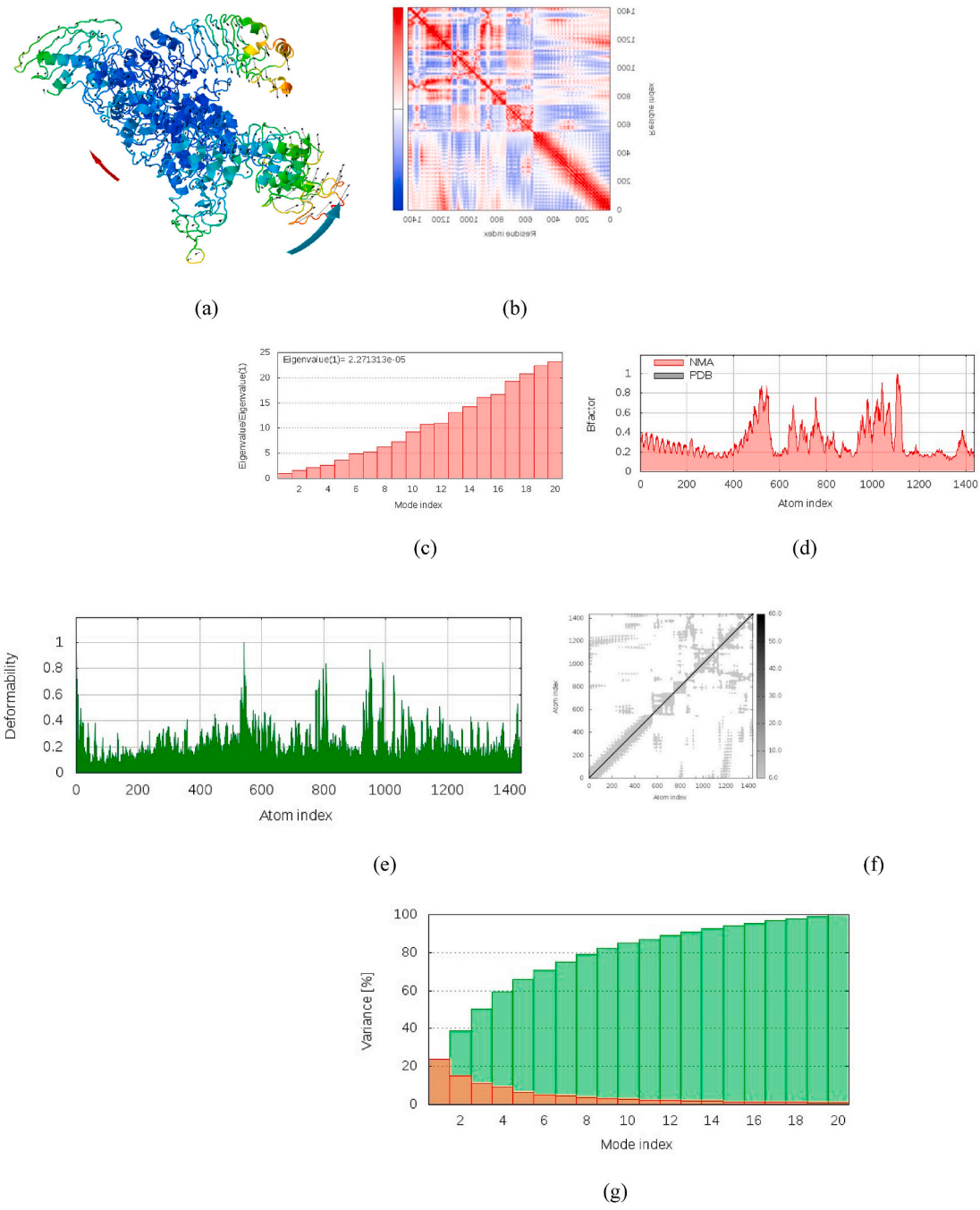


Fig. 8. Molecular dynamics simulation for TLR2: (a) Spin prediction of the ligand-receptor interaction (b) Covariance map of the ligand-receptor interaction (c) Eigenvalues of the ligand-receptor interaction (d) Mobility B-factor of the ligand-protein interaction (e) Deformability B-factor region of the ligand-protein interaction (f) Elastic network of the ligand-protein interaction (g) Variance of the ligand-protein interaction.

employed in predicting the 3D configuration of the multi-epitope vaccine construct. Five models were predicted, but model 1 was chosen based on the confidence score of −2.14, the estimated TM-score, 0.46 ± 0.15 and Root Mean Square Deviation (RMSD) at 13.3 ± 4.1 Å. The B-cell conformational epitope for the vaccine construct was identified using Ellipro server (Fig. 4.)

3.8. Refinement of the tertiary structure

Refinement of the 3D configuration involves reconstruction of side chains of the protein, molecular dynamic simulation and repacking; this is to refine the tertiary structure of the vaccine. Galaxy Refine webserver was employed to refine the configuration of the vaccine, five refined models were predicted, and model 1 was selected on the criteria which include 0.9187 GDT-HA, 0.501 RMSD, 2.814 MolProbity, Clash score of 38.1, Poor rotamers of 1.1, Rama favoured 82.0.

3.9. Validation of tertiary structure

The Ramachandran map showed (Fig. 5) distribution for the protein model for which 83.3% was in the favoured region, 11.3% in the allowed area and 5.4% found in the disallowed region. ProSA web predicted a Z-score of −1.46 for the inputted vaccine model; this is lying outside the score range that is frequently found in the case of native proteins of comparable size, and ERRAT online server revealed a quality score of 68.64%.

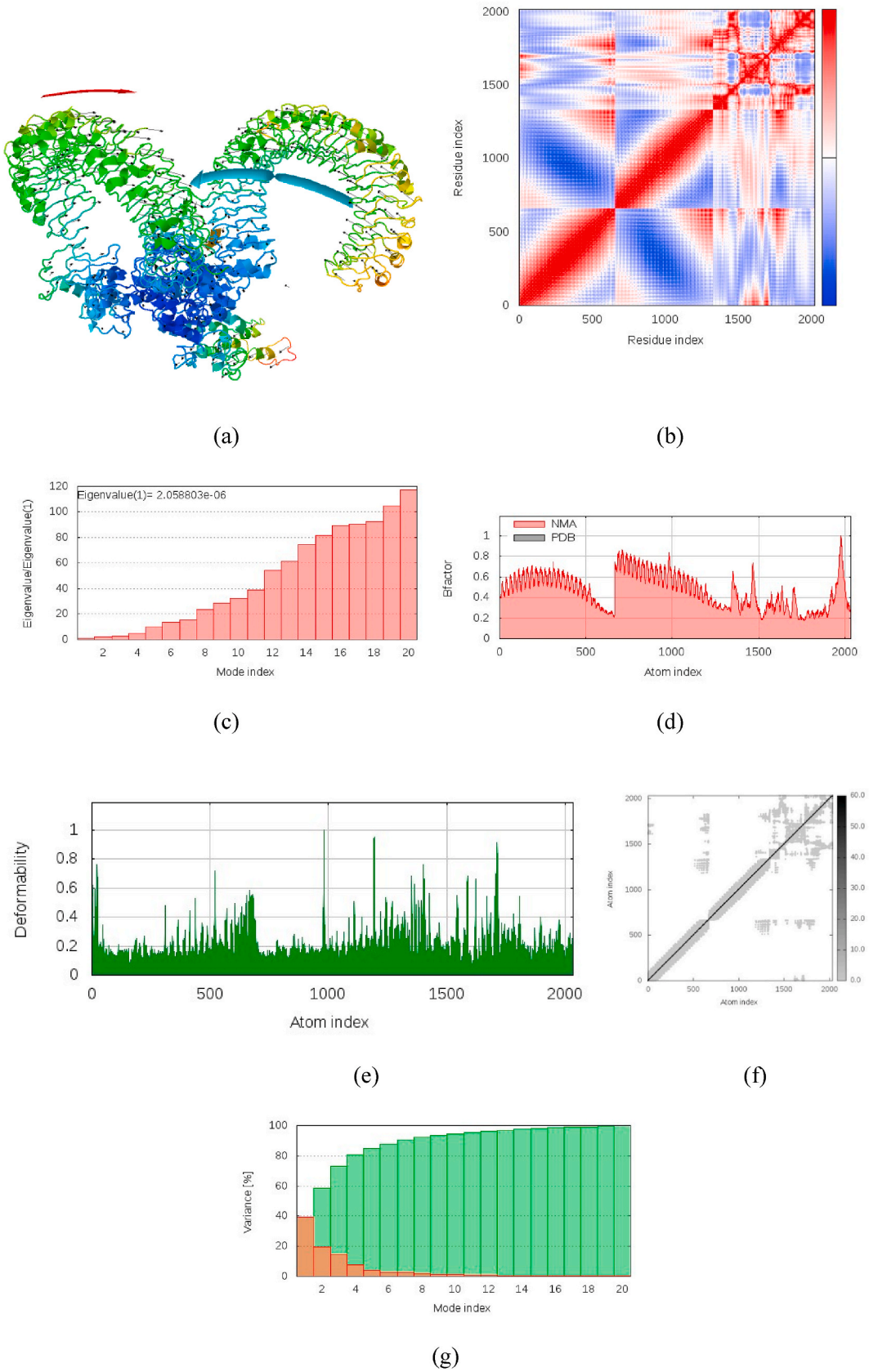


Fig. 9. Molecular dynamics simulation for TLR3: (a) Spin prediction of the ligand-receptor interaction (b) Covariance map of the ligand-receptor interaction (c) Eigenvalues of the ligand-receptor interaction (d) Mobility B-factor of the ligand-protein interaction (e) Deformability B-factor region of the ligand-protein interaction (f) Elastic network of the ligand-protein interaction (g) Variance of the ligand-protein interaction.

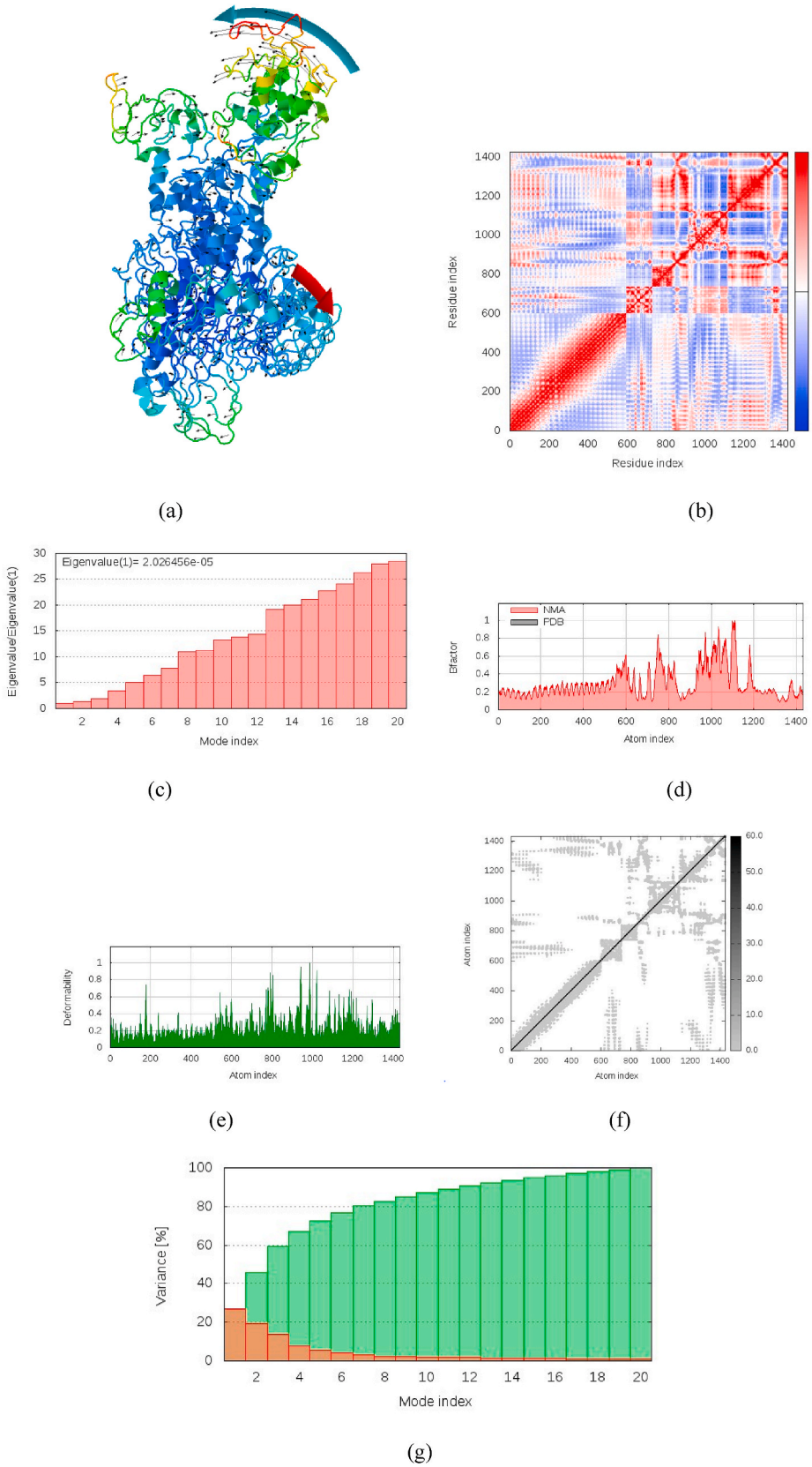


Fig. 10. Molecular dynamics Simulation for TLR4: (a) Spin prediction of the ligand-receptor interaction (b) Covariance map of the ligand-receptor interaction (c) Eigenvalues of the ligand-receptor interaction (d) Mobility B-factor of the ligand-protein interaction (e) Deformability B-factor region of the ligand-protein interaction (f) Elastic network of the ligand-protein interaction (g) Variance of the ligand-protein interaction.

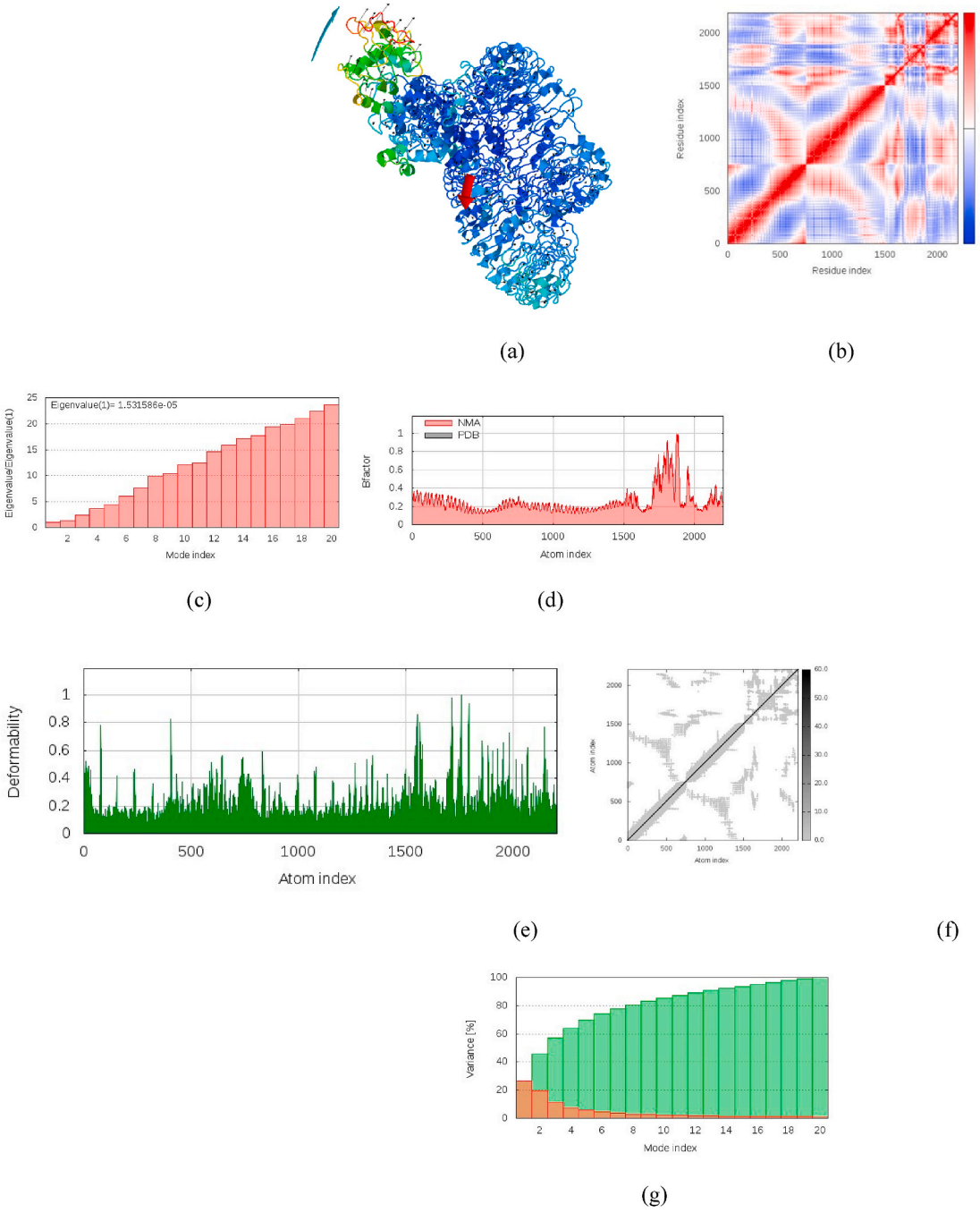


Fig. 11. Molecular dynamics simulation for TLR9: (a) Spin prediction of the ligand-receptor interaction (b) Covariance map of the ligand-receptor interaction (c) Eigenvalues of the ligand-receptor interaction (d) Mobility B-factor of the ligand-protein interaction (e) Deformability B-factor region of the ligand-protein interaction (f) Elastic network of the ligand-protein interaction (g) Variance of the ligand-protein interaction.

3.10. Disulfide engineering

Disulfide engineering was done via the disulfide by Design v2.0 server; this is necessary to enhance and aid the stability of the structure of the refined vaccine. A total of 98 pair of residue useful for disulfide engineering was found, from which ten pairs (PRO511-SER532 ARG574-TYR577 ARG48-GLY77 ASP587-TYR602 SER597-SER601 VAL106-GLY543 ALA593-ARG598 PRO168-GLU566 GLY51-ASN514 MET64-ARG69) were selected on the ground of their energy, Chi3 value and high B-factor as shown in Fig. 6.

3.11. Molecular docking

Docking is the *in silico* process of simulating the binding of the vaccine with TLRs to predict how the vaccine would bind to TLRs *In vivo*. Molecular docking for this vaccine was done using ClusPro 2.0, TLR2, TLR3, TLR4 and TLR9 were used individually as the receptor (Fig. 7). The projection of Molecular interaction of the vaccine with the receptors generated models from which the best were selected for each receptor-ligand complex based on their binding energy weight, as indicated in Table 2.

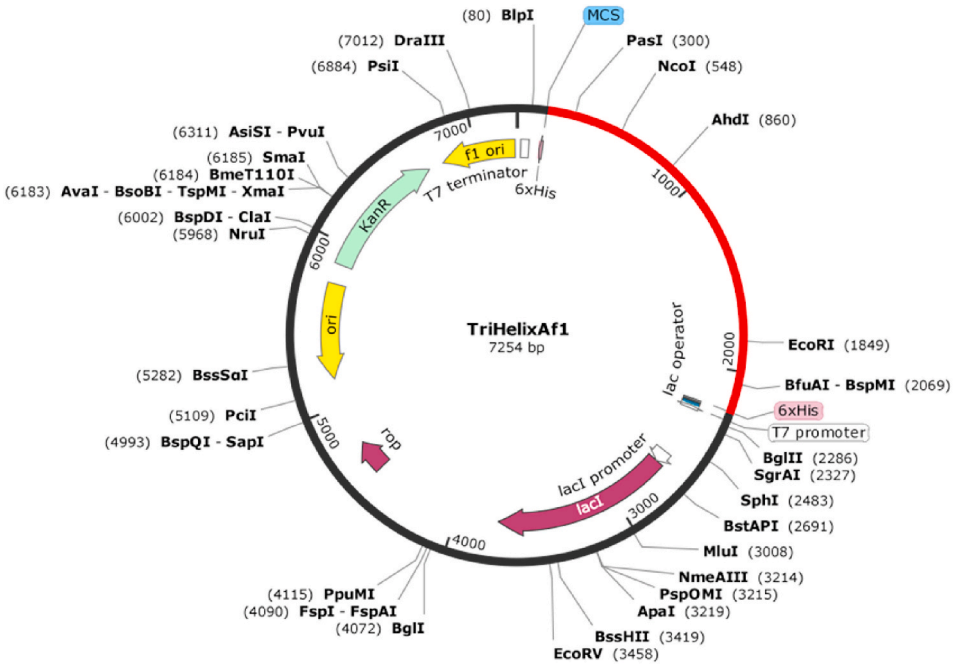
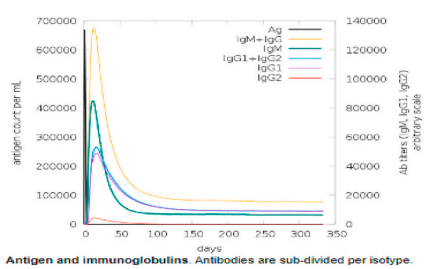
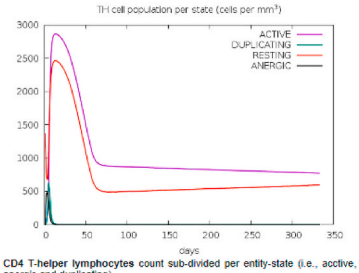


Fig. 12. *In silico* cloning for adapted vaccine into pET28a (+) vector showing the region of choice in red. The restriction enzyme XhoI (158) and XbaI (335) was used as the cloning site. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



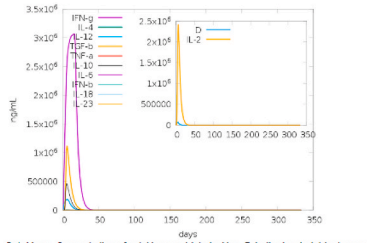
Antigen and immunoglobulins. Antibodies are sub-divided per isotype.

(a)



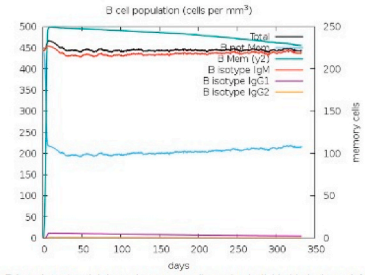
CD4 T-helper lymphocytes count sub-divided per entity-state (i.e., acctive, resting, anergic and duplicating).

(b)



Cytokines. Concentration of cytokines and interleukins. D in the inset plot is danger signal.

(c)



B lymphocytes: total count, memory cells, and sub-divided in isotypes IgM, IgG1 and IgG2.

(d)

Fig. 13. Immune simulation server (C-ImmSim) prediction results of immune response after administration of vaccine construct; (a) Antigen and immunoglobulins; (b) CD4þ helper T cells population per state; (c) Induced levels of the cytokine and Simpson index, D (d) B-lymphocytes cell population.

3.12. Molecular dynamic simulation

Normal Model Analysis (NMA) for the selected docked vaccine-receptor complex was carried out to study the stability and mobility employing the iMODs server. The vaccine protein and its receptor were predicted to spin towards each other; the relationship is represented by small black arrows as shown in Figs. 8a, 9a and 10a and 11a for TLR2, TLR3, TLR4 and TLR9, respectively. The hinges in the high deformability region indicate the deformability of the vaccine-receptor complex as shown in Figs. 8e, 9e and 10e and 11e for TLR2, TLR3, TLR4 and TLR9 respectively. The B-factor is directly proportional to the value of RMS as inferred through NMA (Figs. 8d, 9d and 10d, and 11d).

Eigenvalue for the vaccine-receptor complexes as obtained from iMODs server were 2.27e-05, 2.06e-06 and 1.53e-05 for TLR2, TLR3, TLR4 and TLR9 respectively (Figs. 8c, 9c and 10c and 11c). Variance is reported to be inversely proportional to the Eigen's value. The graphs of the variance can be seen from Figs. 8g, 9g and 10g, and 11g for TLR2, TLR3, TLR4 and TLR9, respectively. From the graphs of the residual index, the colour codes red, blue and white represent the correlated, anti-correlated and uncorrelated pairs of residues in the variance matrix (Fig. 8b 9b, 10b and 11b). The map was observed to be different for each of the vaccine-TLR docking. The elastic network model generated by iMODs (Figs. 8f, 9f and 10f and 11f) represents the pair of atoms connected via springs. Each dot represents one spring between the corresponding pairs of atoms; stiffer springs are represented by the dark grey areas in the elastic graph.

3.13. Codon adaptation and In silico cloning

The integration of COVID-19 vaccine construct into the *E. coli* expression system requires the use of JCAT and SnapGene server. Adapting the vaccine into *E. coli* k12 strain predicted a GC content of 57.03%, Codon Adaptation Index (CAI) of 0.92 and back-translate the protein sequence to an *E. coli* codon compatible nucleotide. The back-translated nucleotide was adapted into the expression system of *E. coli* using restriction enzyme XhoI (158) and XbaI (335) as a cloning site (Fig. 12).

3.14. Immune response model

The C-ImmSim server linked on http://150.146.2.1/C-IMMSIM/index.php revealed the system of successful immune response and the increased half-life of the vaccine candidate. The specific immunoglobulin and interleukin concentrate was shown (Fig. 13a and c). Activity of the CD4, T-helper lymphocytes count (Fig. 13b) as well as B lymphocytes count was uncovered.

4. Discussion

The recent outbreak of the COVID-19 pandemic placed the world in a state of emergency that resulted in the need for the solution against the viral scourge by the world of science and research. Prevention is far better than cure is a fact that can never be over-emphasized; hence it demands a quick response of scientists all over the world.

A vaccine is known as the surest way to prevent the spread of infectious diseases. Immunoinformatics has made vaccine development possible in a shorter period, and subunit vaccines which are a type of vaccine-derived from a fragment of the whole genome of the organism. Having the ability to induce a humoral and cellular mediated response, without reverting to broad type strain and eliminating the risk of pathogen handling, subunit vaccines has impacted positively vaccine design and development. As Fauci once said: "the global need for a vaccine and the wide geographical diversity of the pandemic requires more than one effective vaccine approach."

For this study, the amino acid sequence of SARS-CoV-2 spike glycoproteins was gotten from NCBI and GISAID. SARS-CoV-2 spike glycoproteins were selected based on the report that they possess receptor-binding domain (RBD) which is specific to ACE-II receptor, this mediates viral entry into the host. Spike glycoproteins are also known to be antigenic with relatively conserved protein residue [1] hence, making it a good target for inducing immunogenic response. The selected sequences were subjected to antigenic and allergenic screening and were found to be antigenic and non-allergenic; which is a crucial aspect of vaccinology. CTL, HTL and B-Cell epitope prediction was carried out because multi-epitope vaccine requires CTL, HTL and B-Cell epitopes [53]. This is important since T-cells recognize the surface antigen presented by the MHC molecules. MHC class II molecules present surface antigen to T-helper cells and B-cell; recognize and bind to B-cell epitopes to elicit antibody and memory cells. The result obtained from these analyses revealed the spike glycoproteins comprise a large quantity of MHC-II and B-cell binding epitope; these were employed in building the vaccine candidate using appropriate linkers. The vaccine requires an adjuvant for a robust immune response, attaching an adjuvant to a vaccine implies a linker was also used; hence the complete vaccine construct.

Understanding the physiochemistry of a vaccine is necessary to estimate its stability, solubility and other physical properties of the protein *in vitro* and *in vivo*. The properties as predicted by the physical and chemical predicting server (protparam) suggests that the vaccine has a molecular weight of 73.5 kDa, this is within the range for an accepted vaccine candidate [34]. The theoretical pI value of 9.22 showed an essential vaccine depicting a good structural fold, the aliphatic index was predicted at a score of 56.60 showing that the vaccine has aliphatic side chains, the instability index (25.82) predict the vaccine to be stable. The hydrophobicity of the vaccine was revealed by the GRAVY result of −0.425. The half-live was predicted to be 4.4 h in mammalian reticulocyte *in vitro*, >20 h inside yeast and 10 h in *E. coli in vivo*.

The tertiary structure of the vaccine candidate was predicted using I-TASSER server, and the best configuration was selected based on its high c-score (−2.14). Tertiary structure prediction provided information that includes TM-score (0.46 ± 0.15) and RMSD at 13.3 ± 4.1 Å, the arrangement of the residue that is important in the protein study of the vaccine was also provided. The structure was further used for various studies which include, molecular docking for a receptor-ligand relationship, dynamics study which at the microscopic level gives an insight into the stability of the receptor-ligand complex. Disulfide prediction is required for vaccine stability that helps in analyzing how the vaccine will react when the vaccine is exposed to biochemical stress within the host system as earlier described by Pandey et al. [53].

Structural prediction expresses the structure of vaccine candidate, but the system requires refinement and validation [54]. Refinement provides a refined 3D vaccine model having a higher number of residues in the favoured region (82.0%) of Ramachandran plot. Validation is often carried out to recognize errors within the structure of the final vaccine model, ProSA and PROCHECK servers provided a Z-score value of −1.46, which indicates the stability of the model. Ramachandran plot specified that a high percentage of residues (83.3%) were clustered tightly into the favoured and allowed regions of the map. Structural validation scores obtained from ERRAT and ProSA tools proved that the overall quality of the vaccine construct meets requirement which has been reported in a similar study by Wiederstein & Sippl [42].

Pattern-recognition receptors (PRRs) are means by which the immune system recognizes pathogen-associated molecular pattern (PAMPs). Vertebrate host encompasses a different class of PRRs such as RIG-I-like receptor (RLRs), C-type lectin receptor, NOD-like receptor and Toll-like receptors (TLRs) [55]. TLRs ligands activate the maturation of dendritic cells and consequently activate their ability to trigger adaptive immune responses and in effect, tag antigens as pathogenic. Another significance of TLR-induced DC maturation is the amplification of a variety of cytokines and chemokines [55]. Malcolm et al. [56] pointed out that TLR signaling is a fundamental link between innate and adaptive immunity; TLRs expression varies among different host cell. Several TLRs have been associated in antiviral immunity, TLR 2 is responsible for recognition of structural and non-structural viral protein which induce inflammatory cytokine production. TLR 2, 3, 4 and 9 identify specific forms of viral nucleic acids and are sacrosanct in recognizing viral genetic materials in the endolysosomal compartment hence responsible for antiviral signaling [57].

The binding energy and molecular relationship of the multi-epitope subunit vaccine with TLR (2, 3, 4 and 9) were probed by molecular docking. One model was selected from each of the docked complexes based on their proper receptor interactions, low binding energy and center energy scores [12]. Their Van Der Waals interactions, partial electrostatics and hydrophobic-favour was also considered (Table 2).

Dynamics result showed a positive eigenvalue (2.27e-05; 2.06e-06; 2.03e-05; 1.53e-05) which is significant in the vaccine and rotation of the constructed vaccine which has been reported from several studies [58–60].

How soluble the overexpressed recombinant protein in the *E. coli* host occur is critical for many biochemical and functional investigations. Thus adapting the vaccine model into an *E. coli* expression system is an essential step in vaccine design and codon adaptation is one of the preferred ways to express the rate of foreign genes in a host. This is because when the codon used by the host differs from that of the organisms gene, minor expression rate occur when the genes are not adapted. Therefore we adapted the final vaccine protein sequences to *E. coli* strain K12 utilizing JCAT server and obtained results that were satisfactory. JCAT also back-translated the protein sequences to nucleotides; the nucleotide was cloned into *E. coli* pET28a(+) vector XhoI and XbaI restriction sites, the total length of the clone was 7.2Kbp and is represented in red. The target sequence was encoded between 6-histidine residues; this would prove useful in purification purpose.

The memory of the immune cells patterned by a mechanism that increases cell half-life and a successful immune response is reflected by C-ImmSim. The pathway enables some cells to live longer than others by improving their half-life. The primary response is designated by IgM, and the antigen clearance is a result of immunoglobulins secretion; a property of an increased B-cell population. The steady rise in Th (helper) cell population with memory development was noticed, and IFN-g was provoked after simulation (Fig. 13a–d). B cells and T cells population were highly reactive with consistency in other immune cells. The IFN-g, which is capable of inducing and modulating an array of immune response that was observed in this study, is different from the IFN-c reported by Peele et al. [61]. Also, the difference in data output could be due to the administration of successive dose at intervals compared to the single-dose administered in our study. However, it was observed that the B cell lymphocyte curve was maintained with time (which is an indication of a stable/prolonged half-life and invariably the efficacy of the vaccine prototype within the host cell with time), contrary to the curve in the report of Peele et al. [61]. Transforming growth factor (TGF) is responsible for sustaining tolerance against self and mild antigen and in controlling immune responses to pathogen. Interleukin (IL 2, 4, 10 & 12) maintains the propagation, enhancement of cytotoxicity, IFN-g secretion, and antibody production; Immune suppression, decreases antigen presentation and MHC class II expression of dendritic cells, down regulates pathogenic Th (1,2, 17) responses. IL (4&10) has specific anti-inflammatory properties by suppressing the expression of inflammatory cytokines. IL12 stimulates cytokine secretion and boosts cytotoxic activity enhancing T and NK cells for IFN-g production, which promotes Th1 differentiation and plays functional interaction between the innate and adaptive immunity [62]. Immune stimulation results in authenticate consistency with actual immune responses.

5. Conclusion

Eradicating SARS-COV-2 will be unachievable unless novel control methods are implemented. Vaccines should be taken as very important because the number of infected people is still very few compared to the uninfected at the risk of being infected, due to the high rate of transmission resulting from it being extremely contagious. In this research, immunoinformatics approach was employed to construct a potential vaccine using the spike glycoproteins coding for multiple T-cell (HTL and CTL) and B-cell epitopes. The vaccine peptide has the potential to confer prophylactic benefits. Howbeit, this *a priori* computational work requires experimental validation to substantiate this study, and this may include the development of the designed subunit vaccine followed by *in vitro* and *in vivo* study to ascertain the immunogenicity and safety of the potential vaccine.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to acknowledge Helix Biogen Consult, Ogbomoso and TrimLabs, Ogbomoso, Oyo State, Nigeria for resources and technical support. We would also like to appreciate Odeyemi Ayodeji Emmanuel and Akintola Deborah Adenike for their significant contribution towards the success of this study.

References

[1] Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 2020;181(2): 281–92. https://doi.org/10.1016/j.cell.2020.02.058. e6.

[2] Tortorici MA, Veesler D. Structural insights into coronavirus entry. Adv Virus Res 2019;105:93–116. https://doi.org/10.1016/bs.aivir.2019.08.002.

[3] Walls AC, Tortorici MA, Bosch BJ, Frenz B, Rottier P, DiMaio F, Rey FA, Veesler D. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. Nature 2016;531(7592):114–7. https://doi.org/10.1038/nature16988.

[4] Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li Fl. Structural basis of receptor recognition by SARS-CoV-2. Nature 2020;581(7807):221–4. https://doi.org/10.1038/s41586-020-2179-y.

[5] Kumar A, Sharma BP. In Silico study to predict and characterize of SARS CoV 2 Surface glycoprotein. Vaccine Res 2020:10–6.

[6] Kumar S, Maurya VK, Prasad AK, Bhatt M, Saxena SK. Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). Virus disease 2020;31(1):13–21. https://doi.org/10.1007/s13337-020-00571-5.

[7] Yadav PD, Potdar VA, Choudhary ML, Nyayanit DA, Agrawal M, Jadhav SM, Majumdar TD, Shete-Aich A, Basu A, Abraham P, Cherian SS. Full-genome sequences of the first two SARS-CoV-2 viruses from India. Indian J Med Res 2020; 151(2 & 3):200–9. https://doi.org/10.4103/ijmr.IJMR_663_20.

[8] Zhang L. multi-epitope vaccines: a promising strategy against tumours and viral infections. Cell Mol Immunol 2018;15(2):182–4. https://doi.org/10.1038/cmi.2017.92.

[9] Sette A, Livingston B, McKinney D, Appella E, Fikes J, Sidney J, Newman M, Chesnut R. The development of multi-epitope vaccines: epitope identification, vaccine design and clinical evaluation. Biologicals: J Int Assoc Buddhist Stud 2001; 29(3–4):271–6. https://doi.org/10.1006/biol.2001.0297.

[10] Ikram A, Zaheer T, Awan FM, Obaid A, Naz A, Hanif R, Paracha RZ, Ali A, Naveed AK, Janjua HA. Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches. Sci Rep 2018;8(1):16107. https://doi.org/10.1038/s41598-018-34254-5.

[11] Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, Prajapati VK. Exploring the dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to the battle against dengue infection. Sci Rep 2017; 7(1):9232. https://doi.org/10.1038/s41598-017-09199-w.

[12] Pandey RK, Bhatt TK, Prajapati VK. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating Anopheles salivary protein. Sci Rep 2018;8(1):1125. https://doi.org/10.1038/s41598-018-19456-1.

[13] Rochwerg B, Agarwal A, Zeng L, Leo YS, Appiah JA, Agoritsas T, Bartoszko J, Brignardello-Petersen R, Ergan B, Ge L, Geduld H, Gershengorn HB, Manai H, Huang M, Lamontagne F, Kanda S, Kawano-Dourado L, Kurian L, Kwizera A, Murthy S, Guyatt G. Remdesivir for severe COVID-19: a clinical practice guideline. Br Med J (Clin Res Ed.) 2020;370:m2924. https://doi.org/10.1136/bmj.m2924.

[14] Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, Hohmann E, Chu HY, Luetkemeyer A, Kline S, Lopez de Castilla D, Feinberg RW, Dierberg K, Tapson V, Hsieh L, Patterson TF, Paredes R, Sweeney DA, Short WR, Touloumi G. ACTT-1 study group members remdesivir for the treatment of covid-19 - preliminary report. The New England journal of medicine. 2020. https://doi.org/10.1056/NEJMoa2007764. NEJMoa2007764. Advance online publication.

[15] Li X, Wang Y, Agostinis P, Rabson A, Melino G, Carafoli E, Shi Y, Sun E. Is hydroxychloroquine beneficial for COVID-19 patients? Cell Death Dis 2020;11(7): 512. https://doi.org/10.1038/s41419-020-2721-8.

[16] Eunju O, Lee Young-Tae, Eun-JuKo, Kim Ki-Hye, Yu-Na Lee Jae-Min Song, Kwon Young-Man, Kim Min-Chul, Perez Daniel R, Kang Sang-Moo. Roles of major histocompatibility complex class II in inducing protective immune responses to influenza vaccination. J Virol 2014;88(14):7764–75. https://doi.org/10.1128/JVI.00748-14.

[17] Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data–from vision to reality. Euro Surveill 2017;22(13):30494. https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494.

[18] Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Feolo M. Database resources of the national centre for biotechnology information. Nucleic Acids Res 2007;36(1):D13–21. https://doi.org/10.1093/nar/gkm1000.

[19] Jenuth JP. The NCBI. In: Bioinformatics methods and protocols. Totowa, NJ: Humana Press; 2000. p. 301–12.

[20] Lee H, Shin J, Chung YJ, Baek JY, Chung DR, Peck KR, Ko KS. Evolution of Klebsiellapneumoniae with mucoid and non-mucoid type colonies within a single patient. Int. J. Med. Microbiol. 2019;309(3–4):194–8. https://doi.org/10.1016/j.ijmm.2019.03.003.

[21] Amat-ur-Rasool H, Saghir A, Idrees M. Computational prediction and analysis of enveloping glycoprotein epitopes of DENV-2 and DENV-3 Pakistani isolates: the first step towards Dengue vaccine development. PloS One 2015;10(3):e0119854. https://doi.org/10.1371/journal.pone.0119854.

[22] Magnan Christophe N, Zeller Michael, Kayala Matthew A, Adam Vigil, Randall Arlo, Felgner Philip L, Baldi Pierre. High-throughput prediction of protein antigenicity using protein microarray data. Bioinformatics 2010;26(23):2936–43. https://doi.org/10.1093/bioinformatics/btq551.

[23] Pandey K, Sharma M, Saarav I, Singh S, Dutta P, Bhardwaj A, Sharma S. Analysis of the DosRregulon genes to select cytotoxic T lymphocyte epitope, specific vaccine candidates, using a reverse vaccinology approach. Int. J. Mycobacteriol. 2016;5(1):34–43. https://doi.org/10.1016/j.ijmyco.2015.10.005.

[24] Doytchinova IA, Flower DR VaxiJen. A server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinf 2007;8:4. https://doi.org/10.1186/1471-2105-8-4.

[25] Zhao JW, Yan M, Shi G, Zhang SL, Ming L. In silico identification of cytotoxic T lymphocyte epitopes encoded by RD5 region of Mycobacterium tuberculosis. J Inf Dev Countries 2017;11(10):806–10. https://doi.org/10.3855/jidc.7207.

[26] Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Peters B. The immune epitope database (IEDB) 3.0. Nucleic Acids Res 2015;43(D1):D405–12. https://doi.org/10.1093/nar/gku938.

[27] Zhang W, Liu J, Niu YQ, Wang L, Hu X. A Bayesian regression approach to the prediction of MHC-II binding affinity. Comput Methods Progr Biomed 2008;92(1):1–7. https://doi.org/10.1016/j.cmpb.2008.05.002.

[28] Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res 2017;45(W1):W24–9. https://doi.org/10.1093/nar/gkx346.

[29] Carter C J Schizophrenia. A pathogenetic autoimmune disease caused by viruses and pathogens and dependent on genes. J Pathogens 2011:128318. https://doi.org/10.4061/2011/128318.

[30] Ito M, Hayashi K, Minamisawa T, Homma S, Koido S, Shiba K. Encryption of agonistic motifs for TLR4 into artificial antigens augmented the maturation of antigen-presenting cells. PloS One 2017;12(11). https://doi.org/10.1371/journal.pone.0188934.

[31] Kalita P, Lyngdoh DL, Padhi AK, Shukla H, Tripathi T. Development of multi-epitope driven subunit vaccine against Fasciolagigantica using immunoinformatics approach. Int J Biol Macromol 2019;138:224–33. https://doi.org/10.1016/j.ijbiomac.2019.07.024.

[32] Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v. 2—a server for in silico prediction of allergens. J Mol Model 2014;20(6):2278. https://doi.org/10.1007/s00894-014-2278-5.

[33] Saha S, Raghava G P S AlgPred. Prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res 2006;34(suppl_2):W202–9. https://doi.org/10.1093/nar/gkl343.

[34] Garg VK, Avashthi H, Tiwari A, Jain PA, Ramkete PW, Kayastha AM, Singh VK. MFPPI–Multi FASTA ProtParam interface. Bioinformation 2016;12(2):74. https://doi.org/10.6026/97320630012074.

[35] Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics 1995;11(6):681–4. https://doi.org/10.1093/bioinformatics/11.6.681.

[36] Lee HW, Kim MJ, Park MY, Han KH, Kim J. The conserved proline residue in the LOB domain of LBD18 is critical for DNA-binding and biological function. Mol Plant 2013;6(5):1722–5. https://doi.org/10.1093/mp/sst037.

[37] Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinf 2008;9(1):40. https://doi.org/10.1186/1471-2105-9-40.

[38] Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinf 2008;9(1):514. https://doi.org/10.1186/1471-2105-9-514.

[39] Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement is driven by side-chain repacking. Nucleic Acids Res 2013;41(W1):W384–8. https://doi.org/10.1093/nar/gkt458.

[40] Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res 2012;40(W1):W294–7. https://doi.org/10.1093/nar/gks493.

[41] Laskowski RA, MacArthur MW, Thornton JM. PROCHECK: Validation of protein-structure coordinates, 25.2; 2006. p. 722–5. doi/10.1107/9780955360206000724/full, https://onlinelibrary.wiley.com/.

[42] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007;35(suppl_2):W407–10. https://doi.org/10.1093/nar/gkm290.

[43] Prajapat R, Marwal A, Gaur RK. Recognition of errors in the refinement and validation of three-dimensional structures of AC1 proteins of begomovirus strains by using ProSA-Web. J Viruses 2014;35. https://doi.org/10.1155/2014/752656. Web Server issue), W407–W410.

[44] Gholami A, Shahin S, Mohkam M, Nezafat N, Ghasemi Y. Cloning, characterization and bioinformatics analysis of novel cytosine deaminase from Escherichia coli AGH09. Int J Pept Res Therapeut 2015;21(3):365–74. https://doi.org/10.1007/s10989-015-9465-9.

[45] Yueh C, Hall DR, Xia B, Padhorny D, Kozakov D, Vajda S. ClusPro-DC: Dimer classification by the CLUSPRO server for protein-protein docking. J Mol Biol 2017;429(3):372–81. https://doi.org/10.1016/j.jmb.2016.10.019.

[46] Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Vajda S. The ClusPro web server for protein-protein docking. Nat Protoc 2017;12(2):255. https://doi.org/10.1038/nprot.2016.169.

[47] Craig DB, Dombkowski AA. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinf 2013;14(1):1–7. https://doi.org/10.1186/1471-2105-14-346.

[48] López-Blanco JR, Aliaga JI, Quintana-Ortí ES, Chacón P iMODS. Internal coordinates normal mode analysis server. Nucleic Acids Res 2014;42(W1):W271–6. https://doi.org/10.1093/nar/gku339.

[49] Bansal S, Dalrymple O, Gaffar A. Design, development, and implementation of the instructional module development system (IMODs). In: American society for engineering education annual conference (ASEE). Indianapolis; 2015. #13987.

[50] Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, Jahn D JCat. A novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 2005;33(2):W526–31. https://doi.org/10.1093/nar/gki376.

[51] Li M, Li M, Lin H, Wang J, Jin Y, Han F. Characterization of the novel T4-like Salmonella enterica bacteriophage STP4-a and its endolysin. Arch Virol 2016;161(2):377–84. https://doi.org/10.1007/s00705-015-2647-0.

[52] Tahmoorespur M, Nazifi N, Pirkhezranian Z. In silico prediction of B-cell and T-cell epitopes of protective antigen of *Bacillus anthracis* in development of vaccines against anthrax. Iran J Appl Anim Sci 2017;7(3):429–36.

[53] Chauhan V, Rungta T, Goyal K, Singh MP. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. Sci Rep 2019;9(1):2517. https://doi.org/10.1038/s41598-019-39299-8.

[54] Pandey RK, Ojha R, Aathmanathan VS, Krishnan M, Prajapati VK. Immunoinformatics approaches to design a novel multi-epitope subunit vaccine against HIV infection. Vaccine 2018;36(17):2262–72. https://doi.org/10.1016/j.vaccine.2018.03.042.

[55] Pasare C, Medzhitov R. Toll-like receptors: linking innate and adaptive immunity. n.d. Adv Exp Med Biol 2005:11–8. https://doi.org/10.1007/0-387-24180-9_2.

[56] Duthie Malcolm S, PlessnerWindish Hillarie, Fox Christopher B, Steven G, Reed. Use of defined TLR ligands as adjuvants within human vaccines. Immunol Rev 2018;239(1):178–96. https://doi.org/10.1111/j.1600-065X.2010.00978.x.

[57] Lester, Li. Toll-like receptors in innate antiviral immunity. J Mol Biol 2014;426(6):1246–64. https://doi.org/10.1016/j.jmb.2013.11.024.

[58] Hasan M, Islam S, Chakraborty S, Mustafa AH, Azim KF, Joy ZF, Hossain MN, Foysal SH, Hasan MN. Contriving a chimeric polyvalent vaccine to prevent infections caused by herpes simplex virus (type-1 and type-2): an exploratory immunoinformatic approach. J Biomol Struct Dynam 2019:1–18. https://doi.org/10.1080/07391102.2019.1647286. Advance online publication.

[59] Chauhan V, Singh MP. Immuno-informatics approach to design a multi-epitope vaccine to combat cytomegalovirus infection. Eur J Pharmaceut Sci: Off J Eur Fed Pharmaceut Sci 2020;147:105279. https://doi.org/10.1016/j.ejps.2020.105279.

[60] Ullah MA, Sarkar B, Islam SS. Exploiting the reverse vaccinology approach to design novel subunit vaccines against Ebola virus. Immunobiology 2020;225(3):151949. https://doi.org/10.1016/j.imbio.2020.151949.

[61] Peele Abraham K, Srihansa T, Krupanidhi S, Vijaya Sai A, Venkateswarulu TC. Design of multi-epitope vaccine candidate against SARS-COV-2: an *in-silico* study. J Biomol Struct Dyn 2020:1–9. https://doi.org/10.1080/07391102.2020.1770127.

[62] Mirjačić Martinović K, Babović N, Džodić R, Jurišić V, Matković S, Konjević G. Favourable in vitro effects of combined IL-12 and IL-18 treatment on NK cell cytotoxicity and CD25 receptor expression in metastatic melanoma patients. J Transl Med 2015;13:120. https://doi.org/10.1186/s12967-015-0479-z.