ORIGINAL RESEARCH

Design an Efficient Multi-Epitope Peptide Vaccine Candidate Against SARS-CoV-2: An in silico Analysis

This article was published in the following Dove Press journal:
Infection and Drug Resistance

Zahra Yazdani[1]
Alireza Rafiei[1]
Mohammadreza Yazdani[2]
Reza Valadan[1]

[1]Department of Immunology, Molecular and Cell Biology Research Center, School of Medicine, Mazandaran University of Medical Sciences, Sari, Iran; [2]Department of Chemistry, Isfahan University of Technology, Isfahan 84156-83111, Iran

Background: To date, no specific vaccine or drug has been proven to be effective against SARS-CoV-2 infection. Therefore, we implemented an immunoinformatic approach to design an efficient multi-epitopes vaccine against SARS-CoV-2.

Results: The designed-vaccine construct consists of several immunodominant epitopes from structural proteins of spike, nucleocapsid, membrane, and envelope. These peptides promote cellular and humoral immunity and interferon-gamma responses. Also, these epitopes have a high antigenic capacity and are not likely to cause allergies. To enhance the vaccine immunogenicity, we used three potent adjuvants: Flagellin of *Salmonella enterica* subsp. enterica serovar Dublin, a driven peptide from high mobility group box 1 as HP-91, and human beta-defensin 3 protein. The physicochemical and immunological properties of the vaccine structure were evaluated. The tertiary structure of the vaccine protein was predicted and refined by Phyre2 and Galaxi refine and validated using RAMPAGE and ERRAT. Results of ElliPro showed 246 sresidues from vaccine might be conformational B-cell epitopes. Docking of the vaccine with toll-like receptors (TLR) 3, 5, 8, and angiotensin-converting enzyme 2 approved an appropriate interaction between the vaccine and receptors. Prediction of mRNA secondary structure and in silico cloning demonstrated that the vaccine can be efficiently expressed in *Escherichia coli*.

Conclusion: Our results demonstrated that the multi-epitope vaccine might be potentially antigenic and induce humoral and cellular immune responses against SARS-CoV-2. This vaccine can interact appropriately with the TLR3, 5, and 8. Also, it has a high-quality structure and suitable characteristics such as high stability and potential for expression in *Escherichia coli* .

Keywords: SARS-CoV-2, multi-epitope vaccine, structural proteins, humoral immunity, cellular immunity, adjuvant

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which the cause of COVID-19 disease was first reported as a pneumonia epidemic in the Chinese city of Wuhan (Hubei province) on December 31, 2019, belongs to the Beta coronavirus genus.[1–3] Coronaviruses are positive-sense single-stranded RNA viruses. They belong to the order of Nidovirales and superfamily of Orthocoronaviridae. This superfamily has four genus including alpha, beta, gamma, and delta coronaviruses. Gamma and delta coronaviruses infect birds, while alpha and beta coronaviruses generally infect mammals such as humans. In immunocompetent individuals, they generally cause mild respiratory infections, such as common cold, while in some individuals, coronavirus infections cause serious diseases, such as SARS (Severe

Correspondence: Alireza Rafiei
Department of Immunology, Molecular and Cell Biology Research Center, School of Medicine, Mazandaran University of Medical Sciences, KM 18 Khazarabad Road, Khazar Sq, Sari, Iran
Tel +981133543614
Fax +98-11-3354-3087
Email rafiei1710@gmail.com

Acute Respiratory Syndrome) and MERS (Middle East Respiratory Syndrome) epidemics.

SARS-CoV-2 has a large mRNA genome, 26 to 32 kb in length, meaning a 5′ cap cover structure and a 3′ polyadenylated. This encodes several structural and nonstructural proteins. Among structural proteins, spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins are important in inducing immune responses.[3–5] Spike protein is the main tool for virus entry into the cells, which interacts with host cell receptor, angiotensinconverting enzyme 2 (ACE2). ACE2 is a metallopeptidase expressed in variant tissues, including alveolar epithelial cells, fibroblasts, endothelial cells, and enterocytes.[6–10] The N protein is an essential RNA-binding protein for transcription and replication of the viral RNA. It has significant roles in forming of the helical ribonucleoproteins during packaging the RNA genome, regulating of viral RNA synthesis during replication, transcription, and modulating of metabolism in the infected cells.[11–14] T cell responses against the S and N proteins of SARS-CoV virus are most dominant and long-lived than other structural proteins.[15] E protein has an important role in the assembly of the viral genome.[16,17] The M protein plays a pivotal role in virus assembly, budding, and replication of virus particles in the host cells.[18]

SARS-CoV-2 is transmitted quickly and causes a considerable fatality rate, so that World Health Organization (WHO) reported over 14,562 550 cases globally and at least 9898 deaths because of this disease until July 21, 2020.[19] It can be controlled when medical resources are sufficient but, there is currently no vaccine or approved treatment for this disease. One of the major challenges for scientists in dealing with the new coronavirus pandemic will be to get a useful and effective vaccine. For this reason, trying to develop an effective vaccine to control the virus has been the subject of research by many scientists around the world.[4,20-24] Previous studies on vaccine production against infectious diseases proved that achieving such a vaccine with conventional methods is time consuming and very expensive. As such, in silico prediction of the vaccine targets is very important because the targets can be selected with an open eye in a limited time and resources. Bioinformatic approaches are very helpful to identify the effective epitopes and developing vaccines. Several vaccine candidates for viral diseases have been reported based on this approach that include effective vaccines in Human papilloma viridea (HPV),[25] Ebola,[26] Zika[27] and MERS[28,29] viruses. There are few reports of COVID-19 vaccine.[30–32] They contained multiple cytotoxic T lymphocytes (CTL) and B-cell epitopes against several proteins of SARS-CoV-2 virus. However, those vaccines did not cover most of the immunodominant viral proteins. Therefore, to overwhelm these limitations, we designed a multi-epitope peptide vaccine consisted of S, M, N, and E viral proteins. The vaccine has appropriate physicochemical properties such as stability at room temperature, more antigenic capability, and no possibility of allergy. It consists of three epitopes from S protein and one epitope from each of M, N, and E proteins, respectively. Each of the epitopes has a high ability to stimulate humoral and cellular immune responses, especially the proper production of Interferon gamma (IFN-γ). The selected epitopes aligned between SARS-CoV-2 and SARS species. We used appropriate adjuvants in the vaccine structure to potentiate the immunogenicity of the antigens. Therefore, we also incorporated the potent adjuvants of N and C terminus of Flagellin of *Salmonella enterica* subsp. enterica serovar *Dublin* as a toll-like receptor (TLR) 5, a driven peptide from high mobility group box 1 (HMGB1) as HP-91, and human beta-defensin 3 (HBD-3) in the construct of the vaccine. The vaccine segments were connected to each other by appropriate linkers. Physicochemical properties, structural stability, and immunological characterizations of the vaccine were evaluated. Modeling, refinement, and validation were performed to access a high-quality three-dimensional (3D) structure of the vaccine protein. Docking evaluation showed an appropriate interaction between the vaccine and TLRs 3, 5, 8, and ACE-2. In silico cloning showed that the vaccine could be effectively expressed in *E. coli*. Totally, a potential vaccine candidate with proper immunological and stable physicochemical properties against SARS-CoV-2 was designed. It is expected the vaccine could be capable to protect humans from COVID-19 disease.

Methods
Protein Sequence Retrieval
The protein sequences of spike glycoprotein (accession number of QIC53213.1), nucleocapsid phosphoprotein (QHU79211.1), membrane glycoprotein (QIC53207.1) and envelope protein (QIC53215.1) were driven using NCBI (https://www.ncbi.nlm.nih.gov/) databases and saved in FASTA format for subsequent analysis.

Screening of Potential Epitopes

Screening of Major Histocompatibility Complex Class I (MHC-I) Epitopes

MHC-I humans alleles with 9 mer length were selected using IEDB (www.iedb.org) database through IEDB recommended 2.22 method and Net-MHC 4.0 online server (http://www.cbs.dtu.dk/services/NetMHC/). IEDB (Instructor/Evaluator Database) is a repository of web-based tools which predicts and analysis of immune epitopes. It uses a consensus method based on artificial neural network (ANN), stabilized matrix method (SMM), and Combinatorial Peptide Libraries (CombLib), if any corresponding predictor is available for the peptide sequence.[33–51] NetMHC 4.0 software is a free server for the prediction of peptide-MHCI binding affinity by gapped sequence alignment. Prediction methods of this server are based on alignments that include insertions and deletions and have significantly higher performance than those based on peptides of single lengths methods.[35]

Screening of Cytotoxic T Lymphocytes (CTL) Epitopes

The prediction of CTL epitopes was done by an online server CTL Pred. This method is based on quantitative matrix (QM) and machine learning techniques, for example, artificial neural network (ANN) and support vector machine (SVM). The server also uses the consensus and combined prediction approaches. The consensus and combined prediction approaches are more specific and sensitive, respectively, in comparison with individual approaches such as ANN and SVM.[52]

Screening of Human Leukocyte Antigen (HLA)-II Epitopes

IEDB database (IEDB recommended 2.22 method)[4353–56] (www.iedb.org) and RANKPEP online server (http://imed.med.ucm.es/Tools/rankpep.html) were employed for screening of HLA-II epitope. The database of IEDB was described previously. The RANKPEP server predicts MHC-II binding epitope by position-specific scoring matrices (PSSMs) that are structurally consistent with the binding mode of MHC-II ligands.[57]

Screening of Linear B-Cell Epitopes

Prediction of linear B-cell epitope was done by IEDB[58] (Emini surface accessibility method) and BepiPred software. BepiPred 2.0 predicts the location of linear B-cell epitopes by a combination of a hidden Markov model and a propensity scale method. The epitope Threshold of this server was selected 0.5.[59]

Selection of the Epitope Segments

The results of all predictions were collected and compared together, and the regions with the highest overlaps were determined. These immunodominant regions were employed for future analyses to finally select the most appropriate epitope domains for vaccine construct.

Alignment of Selected Epitopes Between SARS-CoV-2 and SARS

The alignment of the selected epitopes in SARS-CoV-2 and SARS was performed to understand the selected epitopes that are conserved in two viruses. The amino acid sequences of structural proteins of SARS were retrieved from the NCBI (https://www.ncbi.nlm.nih.gov/) database and their epitopes blasted using CLC sequence viewer 8.0.[60]

In silico Analyzing of IFN-γ Inducing Epitopes

IFN epitope server (http://crdd.osdd.net/raghava/ifnepitope) was employed to determine the ability of IFN-γ production in the selected epitopes. This web server classifies MHC binder epitopes into IFN-γ inducing (positive numbers) and non-inducing IFN-γ (negative numbers) using several methods such as; motifs-based search, machine learning technique, and a hybrid approach. Accuracy of Best prediction based on hybrid approach in this software is 82.10%.[61]

Evaluation of Antigenicity

Antigenic properties of the predicted-epitopes were assessed by ANTIGENpro (http://scratch.proteomics.ics.uci.edu/) and VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) servers. Web server of ANTIGENpro is the first sequence-based, alignment-free, and pathogen independent predictor, using protein antigenicity microarray data for predicting the protein antigenicity.[62] VaxiJen is also the first server for the prediction of antigens. This server applies a new alignment-independent approach that is based on auto-cross covariance (ACC) transformation of protein sequences into uniform vectors of principal amino acid properties. Depending on the target organism (bacteria, virus, or tumor) accuracy of this server varies from 70% to 89%. The Threshold of this server was selected at 0.5.[63,64]

Allergenicity Assessment of Predicted Epitopes

To accurately predict the possibility of allergenicity, the software of AllergenFP v.1.0 (http://ddg-pharmfac.net/AllergenFP/) was used. This is an online server that identifies allergens based on amino acid principle properties such as hydrophobicity, helix, and β-strand forming with accuracy of about 88%.[40]

Vaccine Engineering, Evaluation of Physicochemical Properties, Antigenicity, and Allergenicity

According to the results of the previous steps, three epitopes from S protein and one epitope from each protein of N, E, and M were selected to be incorporated in the vaccine construction. The multi-epitope vaccine construct was joined to the carboxyl and amino terminals of the flagellin adjuvant. In addition, HP-91 and HBD-3 synthetic peptide adjuvants were respectively incorporated at the N- and C-terminal of the epitopes. The felagellin molecule was derived from Salmonella enterica subsp. enterica serovar Dublin. Various segments of the designed-vaccine connected by suitable linkers.

After designing the vaccine, several physicochemical parameters including molecular weight, theoretical isoelectric point (pI), total number of positive and negative residues, extinction coefficient, Instability index, half-life, aliphatic index, and grand average hydropathy (GRAVY) were computed using Expasy's ProtParam at http://us.expasy.org/tools/protparam.html.[65] Antigenicity of the final vaccine construct was evaluated using vaxijen v2.0 and ANTIGENpro and allergenicity were assayed by AllergenFP v.1.0 server.

Tertiary Structure Development

Three-dimensional (3D) structural modeling of the designed construct was predicted using Phyre2 server (http://www.sbg.bio.ic.ac.uk/phyre2). It is a suite of a free web server for prediction and analysis of protein structure, function, and mutations.[66] The predicted model by Phyre2 was refined using Galaxy Refine web services. It successfully tested in CASP10 (Critical assessment of techniques for protein structure prediction) experiments. This server refines the whole protein with gentle and aggressive relaxation methods. This first reconstructs all side-chain conformations and repeatedly relaxes the structure with short molecular dynamics simulations after side-chain repacking perturbations.[67] Finally, Galaxy Refine showed five structures as refined models of vaccine structure. The

RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php), and ERRAT (http://services.mbi.ucla.edu/ERRAT/) servers were used to validate the refined 3D structures obtained from Galaxy Refine web service.[68,69] Finally, the refined and high-quality 3D structure of the vaccine was observed by PyMOL software v2.1.1.PyMOL is an open-source that widely used for bimolecular function.[70]

Prediction of Discontinuous B-Cell Epitope

Discontinuous B-cell epitopes were predicted from the 3D vaccine structure using ElliPro in IEDB database (http://tools.immuneepitope.org/tools/ElliPro). ElliPro is a useful research tool for identifying immune epitopes in protein antigens which implements Thornton's method and with a residue clustering algorithm. The MODELLER program and the Jmol viewer allow the prediction and visualization of immune epitopes in a given protein sequence or structure. In comparison with six other structure-based methods that are using for epitope prediction, ElliPro performs the best prediction and gave an AUC value of 0.732, when the most significant prediction is considered for each protein.[71]

Molecular Docking with TLR 3, 5, 8, and ACE-2

The first, tertiary structure of the human TLR 3, 5, 8 and ACE-2 were obtained from PDB database (www.rcsb.org) with codes of 3J0A, 5GS0, 3w3g, and 1R42, respectively. Next, protein–protein docking of the vaccine structure (as a ligand) and each receptor was performed by CLUSPRO 2.0 online server (cluspro.bu.edu/login.php). CLUSPRO 2.0 is a fully automated and fast rigid-body protein–protein docking server. This server evaluates docking of two interacting proteins based on three techniques: first, the Fast Fourier Transform (FFT) correlation approach, second, clustering of the best energy conformations, third, refining the obtained model using short Monte Carlo simulations and the medium-range optimization method SDU.[72–74] The interaction between molecules was observed using Discovery studio 4.5.

Codon Optimization, in silico Cloning and Prediction of the mRNA Secondary Structure

The reverse translation of the designed-vaccine gene sequence was performed by Sequence Manipulation Suite (https://www.bioinformatics.org/sms2/rev_trans.html) to prepare a suitable

Table 1 The Final Immunodominant Epitopes Were Selected from S, M, N, and E Proteins of SARS-CoV-2

Protein	Sequence	Number of Aminoacids	HLA-I (Netmhc4-IEDB)	HLA-II (RANKPEP/IED)	CTL (CTLPRED)	Linear B Cell (BEPIPRED/IEDB)	Antigenicity	IFN stimulation	Allergenicity
Spike glycoprotein [SARS-CoV-2] (QIC53213.1)	QCVNLTTRTQLPPAYTNSFTRGVYYPDKVF	20–49	✓	✓	✓	✓	✓ 0. 6331	+1	Probable non-allergen
	KIADYNYKLPDDFTGCVIAWNSNNLDSKVG	418–446	✓	✓	✓	✓	0.7009	+5.0777656	Probable non-allergen
	YQPYRVVVLSFELLHAPATVCGPKKSTNLV	505–534	✓	✓	✓	✓	✓0.5147	+4.9606125	Probable non-allergen
Envelope protein [SARS-CoV-2] (QIC53215.1)	LVKPSFYV YSRVKNLNSSRVP	51–71	✓	✓	✓	✓	✓ 0.5468	+1.3230951	Probable non-allergen
Nucleocapsid phosphoprotein [SARS-CoV-2] (QHU79211.1)	YKTFPPTEPKKDKKKKADET QALPQRQKK	360–388	✓	✓	✓	✓	✓0.5844	+3.3604363	Probable non-allergen
Membrane glycoprotein [SARS-CoV-2] (QIC53207.1)	TVATSRTLSYYKLGASQRVA GDSGFAAYSR	169–198	✓	✓	✓	✓	✓ 0.5727	+ 4.8238494	Probable non-allergen

vaccine sequence for cloning and expression in an appropriate expression system. The properties of sequence genes such as Codon Adaptation Index (CAI), GC content, and Codon Frequency and Distribution (CFD) have the key roles in attaining a high level of protein expression in the host were evaluated using GenScript online server (https://www.genscript.com/tools/rare-codon-analysis).[24] Finally, the restriction sites of NcoI and XhoI were, respectively, added to the N- and C-terminus of the vaccine DNA sequence using CLC Sequence viewer v8.0 (http://www.cacbio.com) to facilitate the cloning in *E. coli* expression system.

To predict the mRNA folding and secondary structure of the vaccine, the optimized DNA sequence was converted to possible RNA sequence using DNA<->RNA<->Protein tool (http://biomodel.uah.es/en/lab/cybertory/analysis/trans.htm). Then, the mRNA secondary structure was predicted using RNAfold (http://rna.tbi.univie.ac.at/RNAWebSuite/help.html) server. This server predicts the mRNA secondary structure thermodynamically and provides minimum free energy (ΔG Kcal/mol) for structure.[75]

Results
Immunoinformatic Analysis
Protein S, M, N, and E were used to predict HLA-I binding epitopes (HLA-A, B, and C) using IEDB and NetMHC 4. CTL Pred server was used to predict CTL epitopes of SARS-CoV-2 structural proteins. HLA-II binding epitope (DP, DQ, and DR) from these proteins were predicted using RANKEP and IEDB servers. Prediction of linear B-cell epitopes was performed by BepiPred and IEDB. Finally, the obtained epitopes from the comparison of the above analyses were applied to future selection (Table 1).

Antigenic Ability Screening and Avoid Allergenicity
All selected peptides were submitted to the IFN epitope server for evaluation of their ability to induce IFN-γ. Allergenicity probability of epitopes evaluated by AllergenFPE valuation of epitope antigenicity was performed by ANTIGEN pro and VaxiJen v2.0.The six epitopes (three from S and one from each of E, M,
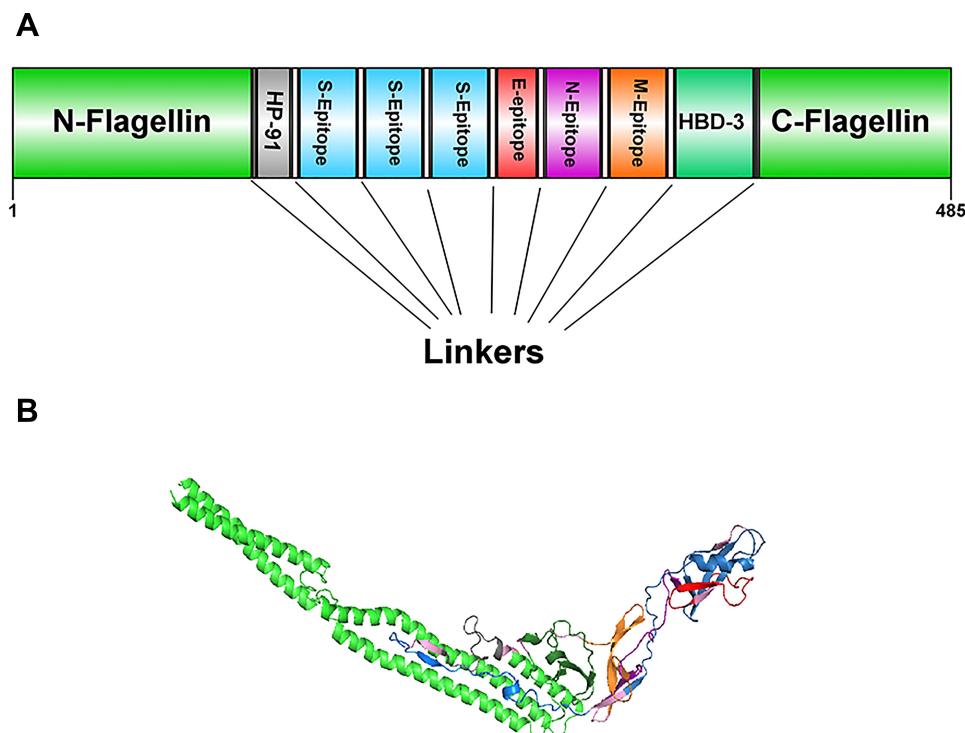

Figure 1 Schematic representation of the designed multi-epitope peptide based-vaccine. The vaccine consist of ten parts: Epitopes from structural proteins S, E, N, and M, adjuvants: Flagellin (in N-and C-terminus), HP-91 and HBD-3 that join to each other by linkers of repeat sequence of LE (**A**). Tertiary structure of the modeled multi-epitope vaccine construct (**B**). The 3D structure of the designed vaccine was predicted via homology modeling by Phyre2, then the best-predicted model was refined by Galaxy Refine and visualized using Discovery studio 4.5 software. N-and C-terminus of Flagellin is shown in green, HP-91 in gray, S epitopes in blue, E epitopes in red, N epitope in purple, M epitope in orange, HBD-3 in green, and linkers are shown in light pink.

Table 2 Analysis of the Physicochemical and Immunological Properties of the SARS-CoV-2 Vaccine

Physicochemical Properties	Value
Molecular weight (Da)	53102.91
Instability index	36.41
Gravy	-0.409
Aliphatic index	86.93
Theoretical pI	7.92
No. amino acids	485
Total no. of negatively charged residues (Asp+Glu)	53
Total no. of positively charged residues (Arg+Lys)	55
No. of atoms	7479
Antigenicity	
Antigenicity/ANTIGENpro	0.927820
Antigenicity/vaxijen	0.5585 (Probable antigen)
Allergenicity/AllergenFP v.1.0	Probable non-allergen

N proteins) that had the highest ability in IFN-γ induction, low allergenicity, and potent antigenic ability were chosen to be used in the vaccine structure (Table 1).

Alignment of the Selected Epitopes Between SARS-CoV-2 and SARS Viruses

The alignment of the selected epitopes between SARS-CoV-2 and SARS viruses showed the selected epitopes were conserved between two species (Table S1).

Designing a Multi-Epitopes Vaccine Construct

Six epitopes with high scores which selected to be incorporated in the final vaccine construct were the sequences of 20–49, 418–446, and 505–534 from S, 51–71 from E, 360–388 from N and 169–198 from M protein. Two epitopes, sequences 418–446 and 505–534, of S protein were located in RBD portion. Besides, three immunopotent adjuvants Flagellin of *Salmonella enterica* subsp. enterica serovar *Dublin*, HP-91, and HBD-3 were also added to the vaccine structure. The vaccine pieces were linked to each other using a short linker sequence of LE dipeptide repeats. The final multi-epitope peptide vaccine was 485 amino acid residues (Figure 1). Figure S1 .

Table 3 Conformational B-Cell Epitopes Identified in the Refined Tertiary Structure of the Multi-Epitope Vaccine Using ElliPro

	Residues	Number of Residues	Score
1	A:M1, A:A2, A:Q3, A:V4, A:I5, A:N6, A:T7, A:N8, A:S9, A:L10, A:S11, A:L12, A:L13, A:T14, A:Q15, A:N16, A:N17, A:L18, A:N19, A:K20, A:S21, A:Q22, A:S23, A:S24, A:L25, A:S26, A:S27, A:A28, A:I29, A:E30, A:R31, A:L32, A:S33, A:S34, A:G35, A:L36, A:D435, A:A436, A:D437, A:Y438, A:A439, A:T440, A:E441, A:V442, A:S443, A:N444, A:M445, A:S446, A:K447, A:A448, A:Q449, A:I450, A:L451, A:Q452, A:Q453, A:A454, A:G455, A:T456, A:S457, A:V458, A:L459, A:A460, A:Q461, A:A462, A:N463, A:Q464, A:V465, A:P466, A:Q467, A:N468, A:V469, A:L470, A:S471, A:L472, A:L473	75	0.801
2	A:N207, A:L208, A:D209, A:S210, A:K211, A:V212, A:G213, A:L214, A:E215, A:L216, A:E217, A:Y218, A:Q219, A:P220, A:Y221, A:R222, A:V223, A:V224, A:V225, A:L226, A:S227, A:F228, A:E229, A:L230, A:L231, A:H232, A:A233, A:P234, A:A235, A:T236, A:V237, A:C238, A:G239, A:P240, A:K241, A:K242, A:S243, A:T244, A:N245, A:L246, A:V247, A:L248, A:E249, A:L250, A:E251, A:L252, A:V253, A:K254, A:P255, A:S256, A:F257, A:Y258, A:V259, A:Y260, A:S261, A:R262, A:V263, A:K264, A:N265, A:L266, A:N267, A:S268, A:S269, A:R270, A:V271, A:P272, A:L273, A:E274, A:L275, A:E276, A:Y277, A:K278, A:T279, A:F280, A:P281, A:P282, A:T283, A:E284, A:P285, A:K286, A:K287, A:D288, A:K289, A:K291, A:K292, A:A293, A:D294, A:E295, A:T296, A:Q297	90	0.767
3	A:Q90, A:V92, A:R93, A:E94, A:L95, A:S96, A:V97, A:Q98, A:A99, A:T100, A:N101, A:G102, A:T103, A:N104, A:S105, A:D106, A:S107, A:D108, A:L109, A:K110, A:S111, A:I112, A:E115, A:E181, A:L182, A:K184, A:A186, A:D187, A:Y188, A:N189, A:Y190, A:K191, A:L192, A:P193, A:D194, A:D195, A:F196, A:T197, A:L318, A:N319, A:T320, A:D321, A:H322, A:S323, A:S324, A:S325, A:S326, A:G333, A:D334, A:P335, A:A379, A:A380, A:K381, A:K382, A:S383, A:T384, A:A385, A:N386, A:P387, A:L388, A:A389, A:S390, A:I391, A:D392, A:S393, A:A394, A:L395, A:S396, A:K397, A:D399, A:A400, A:S4	72	0.619
4	A:R37, A:I38, A:N39, A:S40, A:A41, A:K42, A:D43, A:D44, A:A45, A:A46	10	0.575

Immunological and Physicochemical Properties of the Vaccine Structure

To evaluate the antigenicity of the whole protein vaccine, the "virus" option was chosen as a target organism. The antigenicity of 0.5585 (probable antigen) was estimated at 0.5% threshold for the virus model. Assessment of antigenicity by Antigen pro showed this vaccine is high antigenic with antigenicity of 0.927820. AllergenFP showed that the T-cell epitopes in the vaccine protein are non-allergen. The molecular weight (MW) and theoretical isoelectric point (pI) of the vaccine protein were computed as

53.102 kDa and 7.92, respectively. The predicted half-lives were calculated as 30 hours (h) (mammalian reticulocytes, in vitro), 20 h (yeast, in vivo), and 10 h (*E. coli*, in vivo). The instability index was 36.41, indicating that the protein vaccine was stable enough (Table 2)

Tertiary Structure Modeling, Refinement, and Validation

The primary 3D model of the protein vaccine construct was predicted using Phyre2. Then, the selected Phyre2 model was refined by GalaxyRefine software. The quality
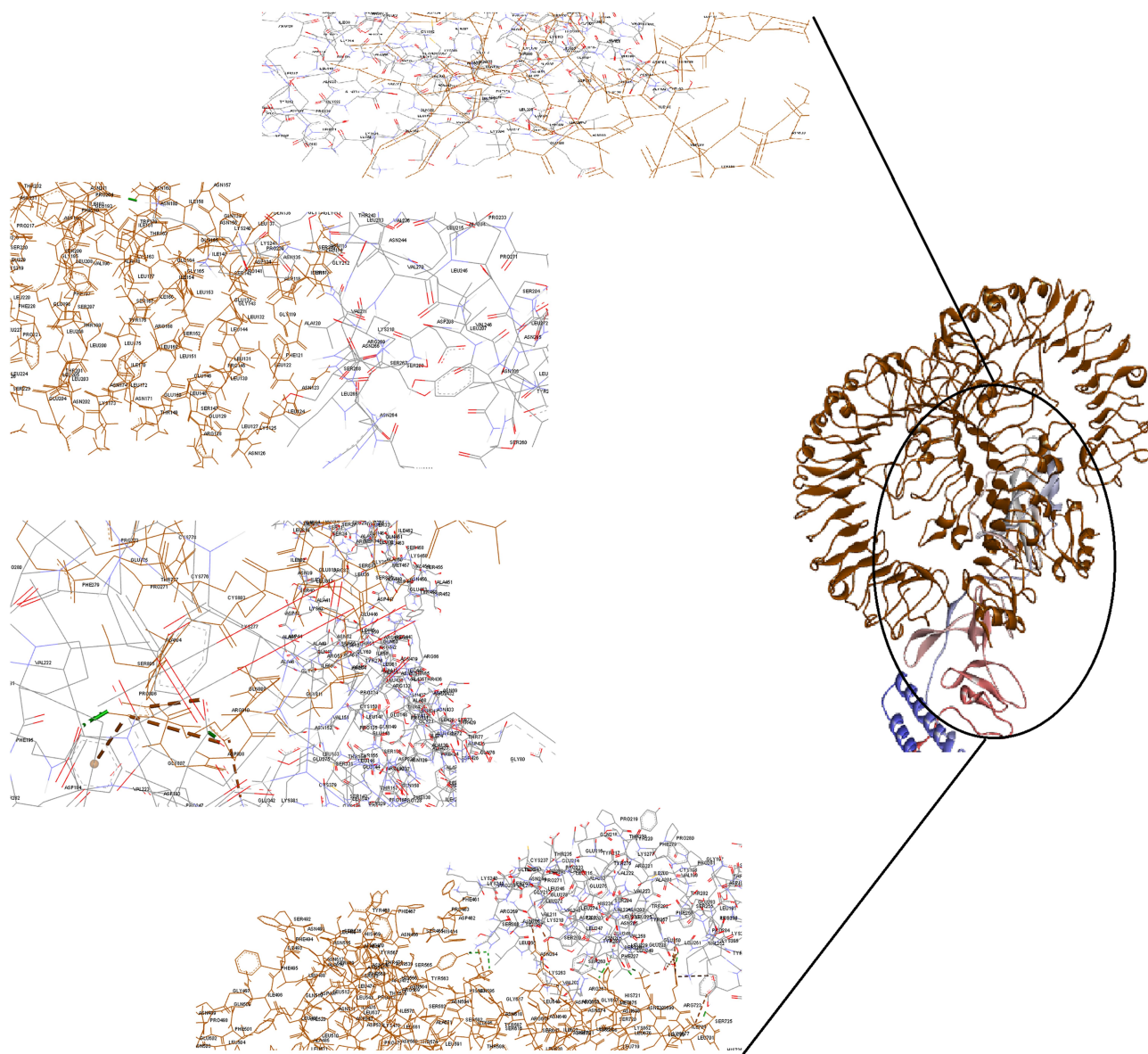

Figure 2 Docking model (cartoon representation) of human TLR3 in complex with the vaccine obtained using Cluspro server. TLR3 protein is shown in chocolate and vaccine was colored from N-to C-terminals (N-flagellin is shown in blue and C-flagellin is shown in red). As it shows some parts of two S, M, N, and E epitopes, HBD-3, and HP-91 interacted with TLR3. To more visualized interaction points, some of the interacting residues of the vaccine and TLR3 are magnified in 20 Angstrom. Docked model was visualized via Discovery studio 4.5 software.
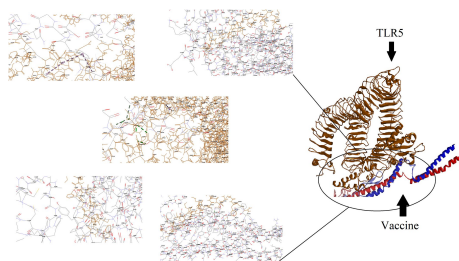
Figure 3 Docking model (cartoon representation) of human TLR5 in complex with the vaccine obtained using Cluspro server. TLR5 is shown in chocolate and vaccine was colored from N-to C-terminals (N-flagellin is shown in blue and C-flagellin is shown in red). As the figure shows some parts of Flagellin, S-epitope and HP-91 interacted with TLR5. In order to more visualized interaction points, some of the interacting residues of the vaccine and TLR5 are magnified in 20 Angstrom. Docked model was visualized by Discovery studio 4.5.

of the designed-vaccine was evaluated using the Ramachandran plot in the RAMPAGE server and characteristic atomic interaction in the ERRAT server. The RAMPAGE results of the final model showed that the majority of residues (95.9%) are located in the favored region and 2.1% are allowed, and only 2.1% of residues are the outlier. The ERRAT result showed that the refined

model has the ERRAT score 89.362 (Figure S2). The outputs of the Ramachandran plot and ERRAT indicated that the refined 3D structure is good and therefore, can be utilized as a reliable model for further evaluations.

Conformational B-Cell Epitope Identification

The tertiary structure of the designed vaccine was used as an input for conformational (discontinuous) B-cell epitope prediction via ElliPro in IEDB. From 485 amino acid residues, 246 were defined as discontinued B-cell epitope (Table 3).

Identifying Binding Sites and Protein–Protein Docking

ClusPro server has been used to perform the docking of the vaccine with TLR3, 5, 8, and ACE-2 molecules. Cluspro generated 30 models for each interaction. The docked model was selected based on most atoms interacted in the vaccine and each of the receptors (Figures 2–5).
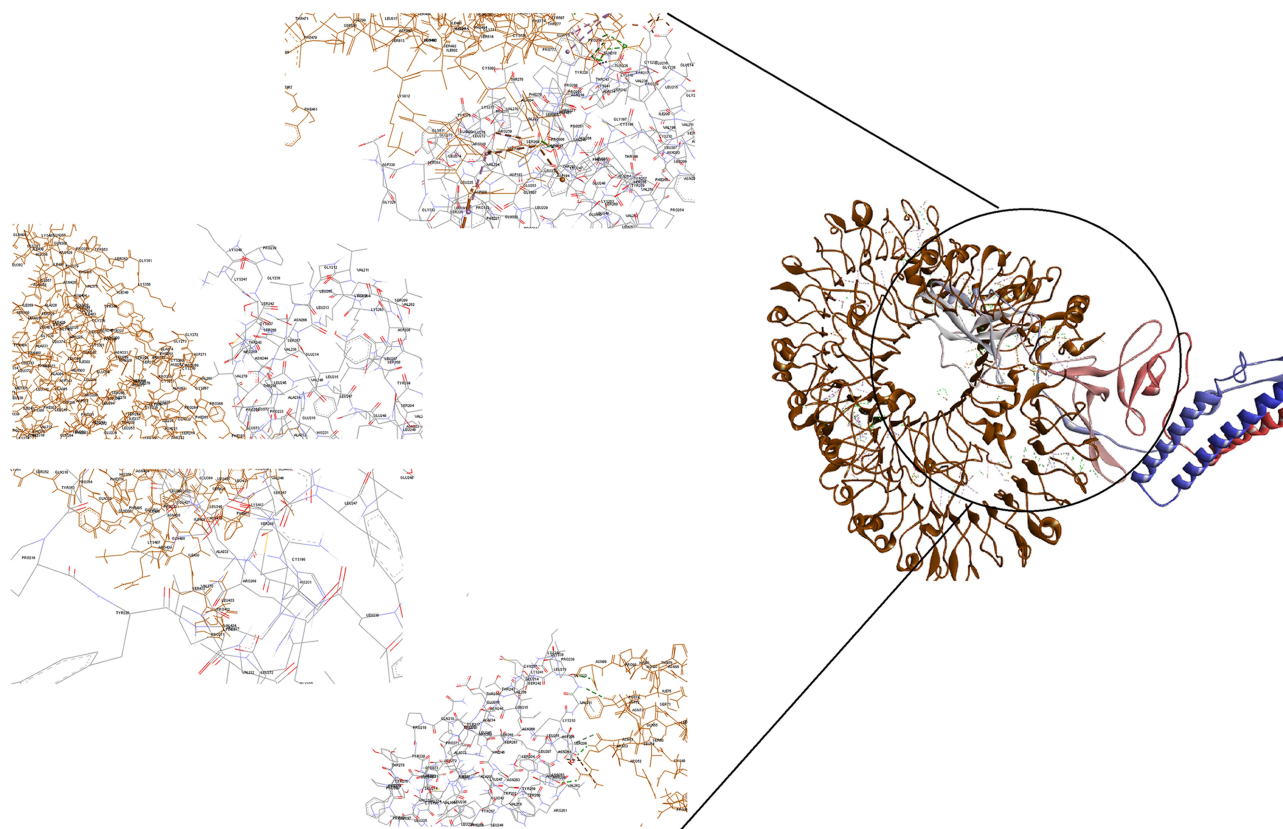

Figure 4 Docking model (cartoon representation) of human TLR8 in complex with the vaccine obtained using Cluspro server. TLR8 protein is shown in chocolate and vaccine was colored from N-to C-terminals (N-flagellin is shown in blue and C-flagellin is shown in red. As the figure shows some parts of S, M, N, and E epitopes, HBD-3 and HP-91 interacted with TLR8). In order to more visualized interaction points, some of the interacting residues of the vaccine and TLR8 are magnified in 20 Angstrom. Docked model was visualized using Discovery studio 4.5.

In silico Codon Optimization of the Vaccine Construct and Prediction of the mRNA Secondary Structure

The reverse translation of the protein vaccine into a nucleotide sequence was performed simultaneously using the Sequence Manipulation Suite server to express high-level protein in *E. coli*. Codon Adaptation Index (CAI), GC content, and Codon Frequency and Distribution (CFD) were evaluated using GenScript online server. A CAI of the vaccine optimized nucleotide sequence was 1. The GC content of the structure was in the ideal range of 58.73% (30–70%), and CFD 100 indicating the effective expression of the protein in the host. To clone the gene into *E. coli* vectors, NcoI and XhoI restriction sites were added into the N- and C-terminals of the sequence using CLC Sequence viewer v.8. Finally, the vaccine gene was inserted into the PET-28 vector

(Figure 6). Minimum free energy of the vaccine was determined using RNA fold server, the optimized mRNA construct showed ΔG value of −520 kcal/mol. Figure S2 illustrates the predicted secondary structure of the vaccine.

Discussion

Today, there is an urgent need to design effective vaccines to stop the indescribable spread of COVID-19. In light of recent advances in bioinformatic approaches, we can identify immunodominant T- and B-cell epitopes and design a potential vaccine to prevent the disease. We used an in silico analysis to design a potent multi-epitope peptide vaccine against SARS-CoV-2. The vaccine contains 485 amino acids which constructed from three specific epitopes from S (two epitopes from RBD section (and one epitope from each of the structural proteins of E, M, and
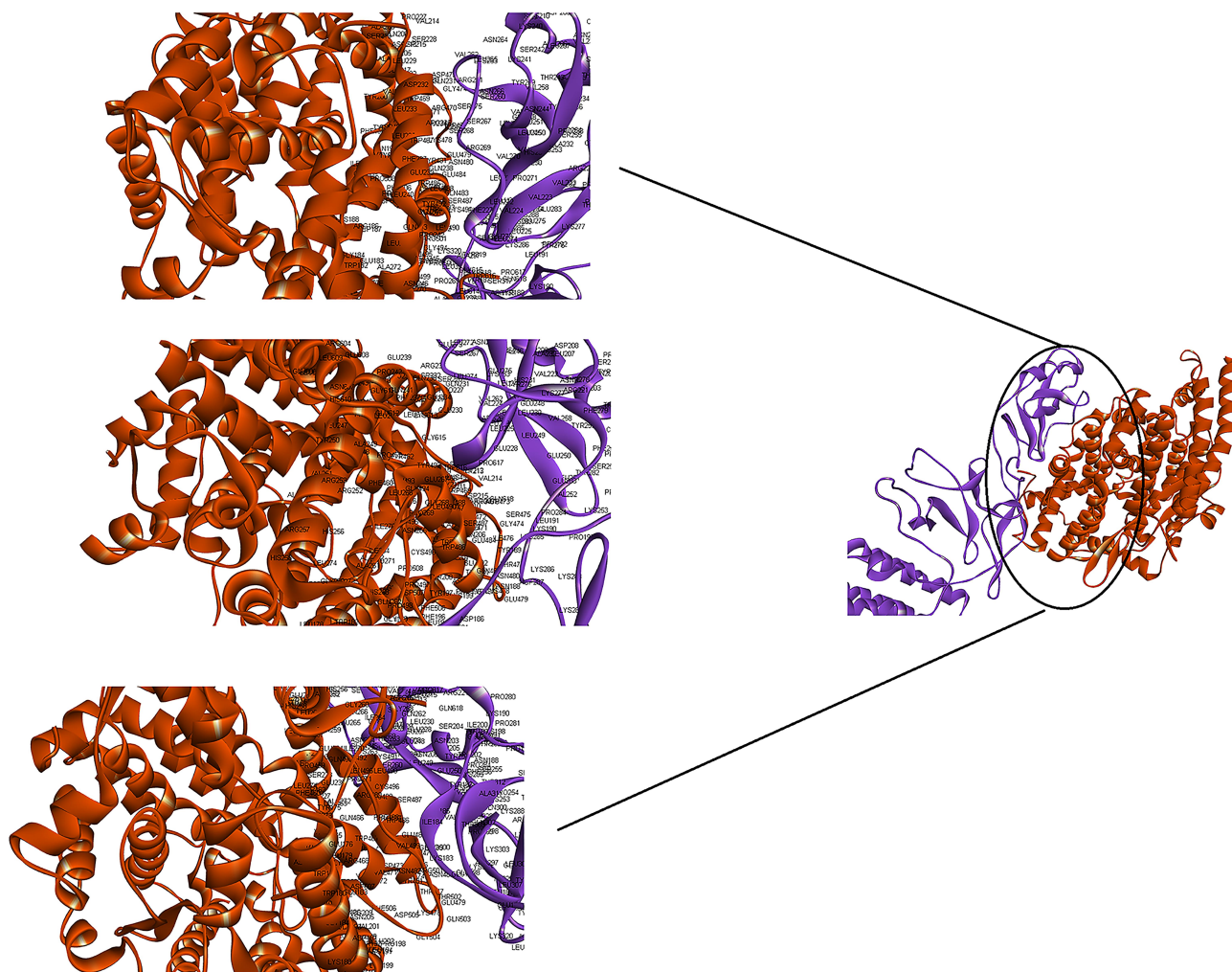

Figure 5 Docking model (cartoon representation) of human ACE-2 in complex with the vaccine obtained using Cluspro server. ACE-2 protein is shown in chocolate and vaccine was shown in purple. As the figure shows RBD epitopes of vaccine interacted with ACE-2. In order to more visualized interaction points, some of the interacting residues of the vaccine and ACE-2 are magnified in 20 Angstrom. Docked model was visualized using Discovery studio 4.5.

N)). These proteins have essential roles in the infection of host cells and host immune modulation. Therefore, it is expected that the vaccine will have a high ability to induce the production of neutralizing antibodies by B lymphocytes and the production of IFN-γ by (T helper) Th and CTL cells.

Adjuvants are essential components of multi-epitope vaccines because they increase the immunologic properties of the vaccine structures. We used Flagellin of *Salmonella enterica* subsp. enterica serovar *Dublin*, a TLR5 agonist, in two N and C terminus of the vaccine construct. Flagellin induces innate immune effectors such as cytokines and nitric oxide,[76,77] activates TLR5-positive dendritic cells (DCs), neutrophils, and epithelial cells[78–81] and stimulates the activation of adaptive immune responses mainly Th2-type and IgA production.[78,79,82–85] Intranasal administration of Flagellin stimulates the signaling of TLR5 in lung epithelial cells and pneumonocytes.[86] It was used broadly as an adjuvant in vaccine structure designed against viral infections such as influenza[85,87,88] and HPV.[25,89] In addition, we incorporated HP-91 and HBD-3 in the final vaccine construct. The peptide of HP-91 derived from B-box domain amino acids of 89–108 from HMGB1 protein. This peptide induces high levels of IL-2 and IL-15, increases secretion of IL-12 and IFN-α in human DCs and augmented T cell activation.[90] Also, it activates DCs independent of TLR2, 4, and 9, and MyD88

pathways.[91] This is a potent immune adjuvant for inducing cellular and humoral immune responses and has been used in vaccines against viral infections such as HIV and HPV.[90–93] HBD-3 is the third adjuvant in our vaccine, used as adjuvant to viruses such as influenza and MERS-CoV.[28,94] This peptide blocks viral fusion using creating a protective barricade of immobilized surface proteins.[94] It activates APCs via TLR1 and TLR2[95] and induces IL-22,[96] TGF-α[97,98] and IFN-γ.[99,100] It has been implicated in the chemotaxis of immature DCs and T cells via its interaction with chemokine receptor 6 (CCR6) and the chemotaxis of monocytes via interaction with CCR2,[101] As well as this peptide activates natural killer (NK), promotes and activates myeloid DCs directly and dependent NK cells activity.[95,100] To link different pieces of the vaccine, a repeated of a dipeptide linker of LE was used. This linker improves the expression of vaccine protein via the decreasing of a pI.[102,103]

An efficient vaccine should not only have stimulating ability immune response but also possess good physicochemical properties during production, formulation, storage, and consumption. According to the results of bioinformatic predictions, the designed vaccine was stable with a stability index of 36.41 and had a pI of 7.92. Conformational B-cell epitopes have a central role in induction humoral responses. The accessibility of a significant number of B-cell epitopes in the vaccine
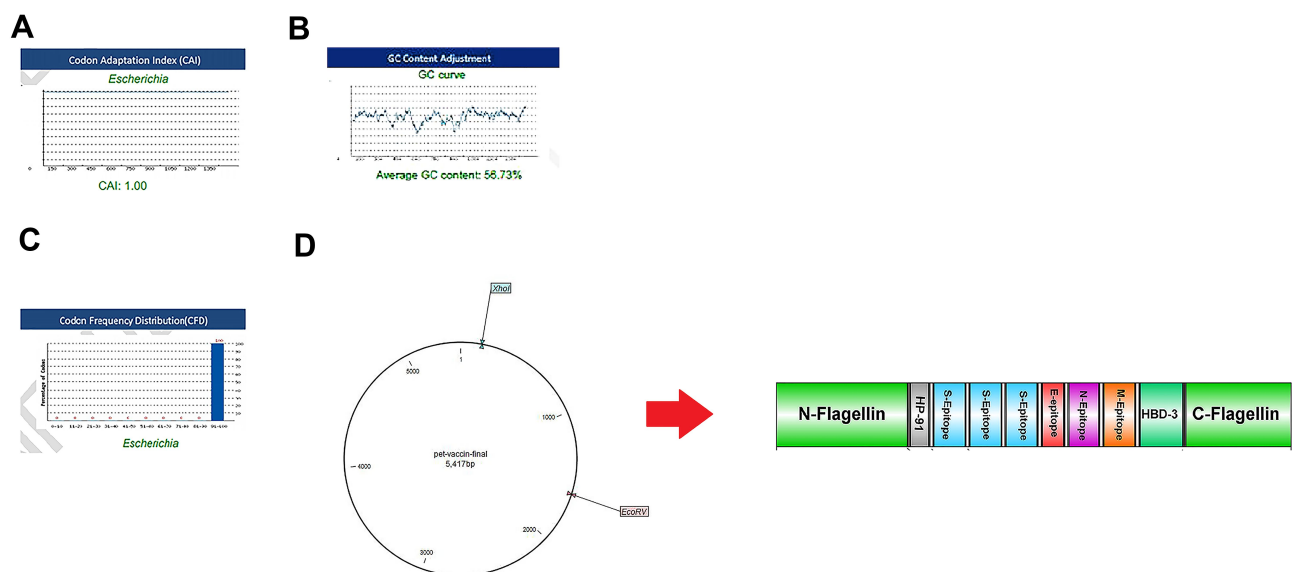

Figure 6 Evaluation of the three important parameters of the codon-optimized gene to express high-level protein in *E. coli*. (**A**) CAI of the gene sequence was 1. It is noted that a CAI of > 0.8 will be considered as good for expression in the selected host. (**B**) Average GC content of the sequence was 58.73%. (**C**) Codon frequency distribution (CFD) value of the gene sequence is 100. A CFD equal to100 supports maximum protein expression in the desired host. (**D**) Insertion of the vaccine gene in the PET28 vector by restriction enzymes of NcoI and XhoI.

molecule indicates the high ability of this structure to stimulate B lymphocytes`. The binding affinity of the vaccine with the immune receptors (TLR3, 5, and 8) is necessary to effectively transport vaccine protein into antigen-presenting cells.

Moreover, it has been reported that TLR3 induces a strong type I IFN-dependent antiviral response in murine Coronavirus Infection[104] and TLR8 has the inhibitory roles against viruses.[105,106] These three TLRs were used in designing in previous bioinformatic-based vaccines against SARS-CoV-2.[30,107,108] The results of docking analysis showed the vaccine protein properly interacts with TLR3, 5, and 8. In addition, docking results of the vaccine and ACE-2 reveal the RBD epitopes of the vaccine can efficiently interact with ACE-2. This interaction indicates the vaccine will be attached to ACE-2 and therefore, induces humoral and cellular responses against SARS-CoV-2.

The codon optimization and in silico cloning illustrated the vaccine can efficiently transcript and translate in *E. coli*. Alignment of the selected epitopes of structural proteins showed these epitopes are conserved between SARS-CoV-2 and SARS, so it is also possible that the vaccine may induce neutralizing antibodies against SARS coronavirus.

Conclusion

In this study, using a variety of bioinformatic methods, we developed a multi-epitope subunit vaccine against the SARS-Cov-2 virus. The results of this study showed that it could be possible to predict vaccine candidates against new emerging viral diseases such as COVID-19 with the help of reverse vaccinology. The present study, with its effective vaccine design against SARS-CoV-2, showed that bioinformatic approaches could help to develop effective treatments for other emerging infectious diseases in a short time and at low cost. However, the in silico results of this study need to be verified using laboratory and animal models.

Abbreviations

ACE2, angiotensin-converting enzyme 2; ANN, artificial neural network; CAI, Codon Adaptation Index; CASP10, Critical assessment of techniques for protein structure prediction; CCR, C-C Motif Chemokine Receptor; CFD, Codon Frequency and Distribution; CombLib, Combinatorial Peptide Libraries; CTL, Cytotoxic T lymphocytes; DC, Dendritic cells; E, envelope; *E. coli*, *Escherichia coli*; h, hour; HBD-3, human beta-defensin 3; HMGB1, high mobility group box1; HLA, Human leukocyte antigen; KDa, Kilo Dalton; INF, Interferon; M, membrane; MHC, Major histocompatibility complex class; MW, molecular weight; N, nucleocapsid; NK, Natural killer; pI, isoelectric point; PSSMs, position specific scoring matrices; QM, Quantitative matrix; S, spike; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SMM, stabilized matrix method; SVM, support vector machine; Th, T helper; TLR, Toll-Like Receptor; 3D, Three dimensional.

Ethics Approval and Consent to Participate

The protocol of this study approved by the Research Ethics committee of Mazandaran University of Medical Sciences IR.MAZUMS.REC.1398.1430.

Author Contributions

A.R conceived and designed this study; supervised the project and collected the data. Z.Y and M.Y contributed to performing the analysis and writing the draft. A.R and R.V edited the manuscript. All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors report no conflicts of interest in this work.

References

1. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395 (10223):497–506. doi:10.1016/S0140-6736(20)30183-5
2. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet*. 2020;395(10223):470–473. doi:10.1016/S0140-6736(20)30185-9
3. Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565–574. doi:10.1016/S0140-6736(20)30251-8
4. Ahmed SF, Quadeer AA, McKay MR. Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies. *Viruses*. 2020;12 (3):254. doi:10.3390/v12030254

5. Zhou P, Yang X-L, Wang X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270–3.

6. Stein RA. The 2019 Coronavirus: learning curves, lessons, and the weakest link. *Int J Clin Pract*. 2020;74:e13488.

7. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS. *J. Virol*. 2020;94. doi:10.1128/JVI.00127-20

8. Li F. Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol*. 2016;3:237–261. doi:10.1146/annurev-virology-110615-042301

9. Bosch BJ, van der Zee R, de Haan CA, Rottier PJ. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol*. 2003;77(16):8801–8811. doi:10.1128/JVI.77.16.8801-8811.2003

10. Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020;367(6483):1260–1263. doi:10.1126/science.abb2507

11. Nelson GW, Stohlman SA, Tahara SM. High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. *J Gen Virol*. 2000;81(1):181–188. doi:10.1099/0022-1317-81-1-181

12. Stohlman S, Baric R, Nelson G, Soe L, Welter L, Deans R. Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J Virol*. 1988;62(11):4288–4295. doi:10.1128/JVI.62.11.4288-4295.1988

13. Cong Y, Ulasli M, Schepers H, et al. Nucleocapsid protein recruitment to replication-transcription complexes plays a crucial role in coronaviral life cycle. *J Virol*. 2020;94(4).

14. Kang S, Yang M, Hong Z, et al. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *bioRxiv*. 2020.

15. Li CK-F, Wu H, Yan H, et al. T cell responses to whole SARS coronavirus in humans. *J Immunol*. 2008;181(8):5490–5500. doi:10.4049/jimmunol.181.8.5490

16. Westerbeck JW, Machamer CE. The infectious bronchitis coronavirus envelope protein alters Golgi pH to protect the spike protein and promote the release of infectious virus. *J Virol*. 2019;93(11):e00015–00019. doi:10.1128/JVI.00015-19

17. Gupta MK, Vemula S, Donde R, Gouda G, Behera L, Vadde R. In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel. *J Biomol Struct Dyn*. 2020;1–17.

18. Song Z, Yang Y, Wang L, et al. EIF4A2 interacts with the membrane protein of transmissible gastroenteritis coronavirus and plays a role in virus replication. *Res Vet Sci*. 2019;123:39–46. doi:10.1016/j.rvsc.2018.12.005

19. Control ECfDPa. COVID-19; 2020. Available from: http://mdanderson.libanswers.com/faq/26219. Accessed July 21,2020.

20. Pang J, Wang MX, Ang IYH, et al. Potential rapid diagnostics, vaccine and therapeutics for 2019 novel coronavirus (2019-nCoV): a systematic review. *J Clin Med*. 2020;9(3):623. doi:10.3390/jcm9030623

21. ul Qamar MT, Alqahtani SM, Alamri MA, Chen -L-L. Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. *J Pharm Anal*. 2020.

22. Robson B. Computers and viral diseases. Preliminary bioinformatics studies on the design of a synthetic vaccine and a preventative peptidomimetic antagonist against the SARS-CoV-2 (2019-nCoV, COVID-19) coronavirus. *Comput Biol Med*. 2020;103670.

23. Amanat F, Krammer F. SARS-CoV-2 vaccines: status report. *Immunity*. 2020;52(4):583–589. doi:10.1016/j.immuni.2020.03.007

24. Prompetchara E, Ketloy C, Palaga T. Immune responses in COVID-19 and potential vaccines: lessons learned from SARS and MERS epidemic. *Asian Pac J Allergy Immunol*. 2020;38(1):1–9. doi:10.12932/AP-200220-0772

25. Yazdani Z, Rafiei A, Valadan R, Ashrafi H, Pasandi M, Kardan M. Designing a potent L1 protein-based HPV peptide vaccine; a bioinformatics approach. *Comput Biol Chem*. 2020;85:107209. doi:10.1016/j.compbiolchem.2020.107209

26. Ahmad B, Ashfaq UA, Rahman M-U, Masoud MS, Yousaf MZ. Conserved B and T cell epitopes prediction of ebola virus glycoprotein for vaccine development: an immuno-informatics approach. *Microb Pathog*. 2019;132:243–253. doi:10.1016/j.micpath.2019.05.010

27. Shahid F, Ashfaq UA, Javaid A, Khalid H. Immunoinformatics guided rational design of a next generation multi epitope based peptide (MEBP) vaccine by exploring Zika virus proteome. *Infect Genet Evol*. 2020;80:104199. doi:10.1016/j.meegid.2020.104199

28. Srivastava S, Kamthania M, Singh S, Saxena AK, Sharma N. Structural basis of development of multi-epitope vaccine against middle east respiratory syndrome using in silico approach. *Infect Drug Resist*. 2018;11:2377. doi:10.2147/IDR.S175114

29. Ul Qamar MT, Saleem S, Ashfaq UA, Bari A, Anwar F, Alqahtani S. Epitope-based peptide vaccine design and target site depiction against Middle East Respiratory Syndrome Coronavirus: an immune-informatics study. *J Transl Med*. 2019;17(1):362. doi:10.1186/s12967-019-2116-8

30. Ul Qamar MT, Rehman A, Ashfaq UA, et al. Designing of a next generation multiepitope based vaccine (MEV) against SARS-COV-2: immunoinformatics and in silico approaches. *BioRxiv*. 2020.

31. Behbahani M. In silico Design of novel Multi-epitope recombinant Vaccine based on Coronavirus surface glycoprotein. *BioRxiv*. 2020.

32. Kalita P, Padhi A, Zhang KY, Tripathi T. Design of a peptide-based subunit vaccine against novel coronavirus SARS-CoV-2. *Microb Pathog*. 2020;145:104236. doi:10.1016/j.micpath.2020.104236

33. Zhang Q, Wang P, Kim Y, et al. Immune epitope database analysis resource (IEDB-AR). *Nucleic Acids Res*. 2008;36(suppl_2):W513–W518. doi:10.1093/nar/gkn254

34. Vita R, Mahajan S, Overton JA, et al. The immune epitope database (IEDB): 2018 update. *Nucleic Acids Res*. 2019;47(D1):D339–D343. doi:10.1093/nar/gky1006

35. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*. 2016;32(4):511–517. doi:10.1093/bioinformatics/btv639

36. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. *Nucleic Acids Res*. 2008;36(suppl_2):W509–W512. doi:10.1093/nar/gkn202

37. Lundegaard C, Nielsen M, Lund O. The validity of predicted T-cell epitopes. *Trends Biotechnol*. 2006;24(12):537–538. doi:10.1016/j.tibtech.2006.10.001

38. Lundegaard C, Lund O, Nielsen M. Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics*. 2008;24(11):1397–1398. doi:10.1093/bioinformatics/btn128

39. Nielsen M, Lundegaard C, Worning P, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci*. 2003;12(5):1007–1017. doi:10.1110/ps.0239403

40. Buus S, Lauemøller S, Worning P, et al. Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee'artificial neural network approach. *Tissue Antigens*. 2003;62(5):378–384. doi:10.1034/j.1399-0039.2003.00112.x

41. Peters B, Sette A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinform*. 2005;6(1):132. doi:10.1186/1471-2105-6-132

42. Kim Y, Sidney J, Pinilla C, Sette A, Peters B. Derivation of an amino acid similarity matrix for peptide: MHC binding and its application as a Bayesian prior. *BMC Bioinform*. 2009;10(1):394. doi:10.1186/1471-2105-10-394

43. Sidney J, Assarsson E, Moore C, et al. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome Res*. 2008;4(1):2. doi:10.1186/1745-7580-4-2

44. Moutaftsi M, Peters B, Pasquetto V, et al. A consensus epitope prediction approach identifies the breadth of murine T CD8+-cell responses to vaccinia virus. *Nat Biotechnol*. 2006;24(7):817–819. doi:10.1038/nbt1215

45. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol*. 2017;199(9):3360–3368. doi:10.4049/jimmunol.1700893

46. Nielsen M, Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med*. 2016;8(1):33. doi:10.1186/s13073-016-0288-x

47. Hoof I, Peters B, Sidney J, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics*. 2009;61(1):1. doi:10.1007/s00251-008-0341-z

48. Nielsen M, Lundegaard C, Blicher T, et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and-B locus protein of known sequence. *PLoS One*. 2007;2(8):e796. doi:10.1371/journal.pone.0000796

49. Karosiene E, Lundegaard C, Lund O, Nielsen M. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics*. 2012;64(3):177–186. doi:10.1007/s00251-011-0579-8

50. Zhang H, Lund O, Nielsen M. The PickPocket method for predicting binding specificities for receptors based on receptor pocket similarities: application to MHC-peptide binding. *Bioinformatics*. 2009;25 (10):1293–1299. doi:10.1093/bioinformatics/btp137

51. Rasmussen M, Fenoy E, Harndahl M, et al. Pan-specific prediction of peptide–MHC class I complex stability, a correlate of T cell immunogenicity. *J Immunol*. 2016;197(4):1517–1524. doi:10.4049/jimmunol.1600582

52. Bhasin M, Raghava G. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine*. 2004;22(23–24):3195–3204. doi:10.1016/j.vaccine.2004.02.005

53. Nielsen M, Lund O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinform*. 2009;10(1):296. doi:10.1186/1471-2105-10-296

54. Nielsen M, Lundegaard C, Lund O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinform*. 2007;8(1):238. doi:10.1186/1471-2105-8-238

55. Sturniolo T, Bono E, Ding J, et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol*. 1999;17 (6):555–561. doi:10.1038/9858

56. Andreatta M, Karosiene E, Rasmussen M, Stryhn A, Buus S, Nielsen M. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. *Immunogenetics*. 2015;67(11–12):641–650. doi:10.1007/s00251-015-0873-y

57. Reche PA, Glutting J-P, Zhang H, Reinherz EL. Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. *Immunogenetics*. 2004;56 (6):405–419. doi:10.1007/s00251-004-0709-7

58. Emini EA, Hughes JV, Perlow D, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol*. 1985;55(3):836–839. doi:10.1128/JVI.55.3.836-839.1985

59. 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–W29. doi:10.1093/nar/gkx346

60. Sarwar MA, Rehman A, Ferzund J. Database Search, Alignment Viewer and Genomics Analysis Tools: big Data for Bioinformatics. *Int J Comput Sci Netw*. 2016;14(12):317.

61. Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct*. 2013;8(1):30. doi:10.1186/1745-6150-8-30

62. Magnan CN, Baldi P. SSpro/ACCpro 5: almost perfect prediction of protein secondary structure and relative solvent accessibility using profiles, machine learning and structural similarity. *Bioinformatics*. 2014;30(18):2592–2597. doi:10.1093/bioinformatics/btu352

63. Agostini F, Cirillo D, Livi CM, Ponti RD, Tartaglia GG. ccSOL omics: a webserver for large-scale prediction of endogenous and heterologous solubility in E. coli. *Bioinformatics*. 2014;30 (20):2975–2977. doi:10.1093/bioinformatics/btu420

64. Doytchinova IA, Flower DR. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine*. 2007;25(5):856–866. doi:10.1016/j.vaccine.2006.09.032

65. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: Walker J.M. (eds) *The Proteomics Protocols Handbook*. Springer; 2005:571–607.

66. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10(6):845. doi:10.1038/nprot.2015.053

67. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res*. 2012;40 (W1):W294–W297. doi:10.1093/nar/gks493

68. Lovell SC, Davis IW, Arendall WB III, et al. Structure validation by Cα geometry: φ, ψ and Cβ deviation. *Proteins*. 2003;50 (3):437–450. doi:10.1002/prot.10286

69. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*. 1993;2 (9):1511–1519. doi:10.1002/pro.5560020916

70. Lill MA, Danielson ML. Computer-aided drug design platform using PyMOL. *J Comput Aided Mol*. 2011;25(1):13–19. doi:10.1007/s10822-010-9395-8

71. Ponomarenko J, Bui -H-H, Li W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform*. 2008;9(1):514. doi:10.1186/1471-2105-9-514

72. Vajda S, Yueh C, Beglov D, et al. New additions to the ClusPro server motivated by CAPRI. *Proteins*. 2017;85(3):435–444. doi:10.1002/prot.25219

73. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein–protein docking. *Nat Protoc*. 2017;12(2):255. doi:10.1038/nprot.2016.169

74. Kozakov D, Beglov D, Bohnuud T, et al. How good is automated protein docking? *Proteins*. 2013;81(12):2159–2166. doi:10.1002/prot.24403

75. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci*. 2004;101(19):7287–7292. doi:10.1073/pnas.0401799101

76. Mizel SB, Honko AN, Moors MA, Smith PS, West AP. Induction of macrophage nitric oxide production by Gram-negative flagellin involves signaling via heteromeric Toll-like receptor 5/Toll-like receptor 4 complexes. *J Immunol*. 2003;170(12):6217–6223. doi:10.4049/jimmunol.170.12.6217

77. Moors MA, Li L, Mizel SB. Activation of interleukin-1 receptor-associated kinase by gram-negative flagellin. *Infect Immun*. 2001;69(7):4424–4429. doi:10.1128/IAI.69.7.4424-4429.2001

78. Didierlaurent A, Ferrero I, Otten LA, et al. Flagellin promotes myeloid differentiation factor 88-dependent development of Th2-type response. *J Immunol*. 2004;172(11):6922–6930. doi:10.4049/jimmunol.172.11.6922

79. Lee SE, Kim SY, Jeong BC, et al. A bacterial flagellin, Vibrio vulnificus FlaB, has a strong mucosal adjuvant activity to induce protective immunity. *Infect Immun*. 2006;74(1):694–702. doi:10.1128/IAI.74.1.694-702.2006

80. Han S, Yang K, Ozen Z, et al. Enhanced differentiation of splenic plasma cells but diminished long-lived high-affinity bone marrow plasma cells in aged mice. *J Immunol*. 2003;170(3):1267–1273. doi:10.4049/jimmunol.170.3.1267

81. Pino O, Martin M, Michalek SM. Cellular mechanisms of the adjuvant activity of the flagellin component FljB of Salmonella enterica Serovar Typhimurium to potentiate mucosal and systemic responses. *Infect Immun*. 2005;73(10):6763–6770. doi:10.1128/IAI.73.10.6763-6770.2005

82. Honko AN, Mizel SB. Mucosal administration of flagellin induces innate immunity in the mouse lung. *Infect Immun*. 2004;72(11):6676–6679. doi:10.1128/IAI.72.11.6676-6679.2004

83. Cunningham AF, Khan M, Ball J, et al. Responses to the soluble flagellar protein FliC are Th2, while those to FliC on Salmonella are Th1. *Eur J Immunol*. 2004;34(11):2986–2995. doi:10.1002/eji.200425403

84. McSorley SJ, Ehst BD, Yu Y, Gewirtz AT. Bacterial flagellin is an effective adjuvant for CD4+ T cells in vivo. *J Immunol*. 2002;169(7):3914–3919. doi:10.4049/jimmunol.169.7.3914

85. Wang B-Z, Quan F-S, Kang S-M, Bozja J, Skountzou I, Compans RW. Incorporation of membrane-anchored flagellin into influenza virus-like particles enhances the breadth of immune responses. *J Virol*. 2008;82(23):11813–11823. doi:10.1128/JVI.01076-08

86. López-Boado YS, Cobb LM, Deora R. Bordetella bronchiseptica flagellin is a proinflammatory determinant for airway epithelial cells. *Infect Immun*. 2005;73(11):7525–7534. doi:10.1128/IAI.73.11.7525-7534.2005

87. Huleatt JW, Nakaar V, Desai P, et al. Potent immunogenicity and efficacy of a universal influenza vaccine candidate comprising a recombinant fusion protein linking influenza M2e to the TLR5 ligand flagellin. *Vaccine*. 2008;26(2):201–214. doi:10.1016/j.vaccine.2007.10.062

88. Skountzou I, Del Pilar Martin M, Wang B, et al. Salmonella flagellins are potent adjuvants for intranasally administered whole inactivated influenza vaccine. *Vaccine*. 2010;28(24):4103–4112. doi:10.1016/j.vaccine.2009.07.058

89. Negahdaripour M, Eslami M, Nezafat N, et al. A novel HPV prophylactic peptide vaccine, designed by immunoinformatics and structural vaccinology approaches. *Infect Genet Evol*. 2017;54:402–416. doi:10.1016/j.meegid.2017.08.002

90. Telusma G, Datta S, Mihajlov I, et al. Dendritic cell activating peptides induce distinct cytokine profiles. *Int Immunol*. 2006;18(11):1563–1573. doi:10.1093/intimm/dxl089

91. Saenz R, Futalan D, Leutenez L, et al. TLR4-dependent activation of dendritic cells by an HMGB1-derived peptide adjuvant. *J Transl Med*. 2014;12(1):211. doi:10.1186/1479-5876-12-211

92. Talebi S, Bolhassani A, Sadat SM, Vahabpour R, Agi E, Shahbazi S. Hp91 immunoadjuvant: an HMGB1-derived peptide for development of therapeutic HPV vaccines. *Biomed Pharmacother*. 2017;85:148–154. doi:10.1016/j.biopha.2016.11.115

93. Milani A, Bolhassani A, Shahbazi S, Motevalli F, Sadat SM, Soleymani S. Small heat shock protein 27: an effective adjuvant for enhancement of HIV-1 Nef antigen-specific immunity. *Immunol Lett*. 2017;191:16–22. doi:10.1016/j.imlet.2017.09.005

94. Leikina E, Delanoe-Ayari H, Melikov K, et al. Carbohydrate-binding molecules inhibit viral fusion and entry by crosslinking membrane glycoproteins. *Nat Immunol*. 2005;6(10):995–1001. doi:10.1038/ni1248

95. Funderburg N, Lederman MM, Feng Z, et al. Human β-defensin-3 activates professional antigen-presenting cells via Toll-like receptors 1 and 2. *Proc Natl Acad Sci*. 2007;104(47):18631–18635. doi:10.1073/pnas.0702130104

96. Wolk K, Kunz S, Witte E, Friedrich M, Asadullah K, Sabat R. IL-22 increases the innate immunity of tissues. *Immunity*. 2004;21(2):241–254. doi:10.1016/j.immuni.2004.07.007

97. Sørensen OE, Cowland JB, Theilgaard-Mönch K, Liu L, Ganz T, Borregaard N. Wound healing and expression of antimicrobial peptides/polypeptides in human keratinocytes, a consequence of common growth factors. *J Immunol*. 2003;170(11):5583–5589.

98. Ferris LK, Mburu YK, Mathers AR, et al. Human beta-defensin 3 induces maturation of human Langerhans cell–like dendritic cells: an antimicrobial peptide that functions as an endogenous adjuvant. *J Invest Dermatol*. 2013;133(2):460–468. doi:10.1038/jid.2012.319

99. Joly S, Organ CC, Johnson GK, McCray PB Jr, Guthmiller JM. Correlation between β-defensin expression and induction profiles in gingival keratinocytes. *Mol Immunol*. 2005;42(9):1073–1084.

100. Judge CJ, Reyes-Aviles E, Conry SJ, et al. HBD-3 induces NK cell activation, IFN-γ secretion and mDC dependent cytolytic function. *Cell Immunol*. 2015;297(2):61–68. doi:10.1016/j.cellimm.2015.06.004

101. Röhrl J, Yang D, Oppenheim JJ, Hehlgans T. Human β-defensin 2 and 3 and their mouse orthologs induce chemotaxis through interaction with CCR2. *J Immunol*. 2010;184(12):6688–6694. doi:10.4049/jimmunol.0903984

102. Chen X, Zaro JL, Shen W-C. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev*. 2013;65(10):1357–1369. doi:10.1016/j.addr.2012.09.039

103. Chen X, Lee H-F, Zaro JL, Shen W-C. Effects of receptor binding on plasma half-life of bifunctional transferrin fusion proteins. *Mol Pharm*. 2011;8(2):457–465. doi:10.1021/mp1003064

104. Mazaleuskaya L, Veltrop R, Ikpeze N, Martin-Garcia J, Navas-Martin S. Protective role of Toll-like receptor 3-induced type I interferon in murine coronavirus infection of macrophages. *Viruses*. 2012;4(5):901–923. doi:10.3390/v4050901

105. Campbell GR, Spector SA. Toll-like receptor 8 ligands activate a vitamin D mediated autophagic response that inhibits human immunodeficiency virus type 1. *PLoS Pathog*. 2012;8(11):e1003017. doi:10.1371/journal.ppat.1003017

106. Hammerbeck DM, Burleson GR, Schuller CJ, et al. Administration of a dual toll-like receptor 7 and toll-like receptor 8 agonist protects against influenza in rats. *Antiviral Res*. 2007;73(1):1–11. doi:10.1016/j.antiviral.2006.07.011

107. Bhatnager R, Bhasin M, Arora J, Dang AS. Epitope based peptide vaccine against SARS-COV2: an immune-informatics approach. *J Biomol Struct Dyn*. 2020;1–16. doi:10.1080/07391102.2020.1787227

108. Bhattacharya M, Sharma AR, Patra P, et al. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immunoinformatics approach. *J Med Virol*. 2020;92(6):618–631. doi:10.1002/jmv.25736