Vaccine Research
ISSN: 2423-4923
eISSN: 2383-2819

Assembling the Most Antigenic Peptides of COVID-19 Immunogenic Proteins Along with a Molecular Adjuvant to Develop a Novel Polyepitope Vaccine: a Bioinformatics Investigation

Nemat Shams[1]*, Narges Nazifi[1], Ali Forouharmehr[2], Amin Jaydari[1], Ehsan Rashidian[1]

[1]Department of pathobiology, Faculty of Veterinary Medicine, Lorestan University, Khorramabad, Iran. [2]Department of Animal Science, Faculty of Agriculture, Lorestan University, Khorramabad, Iran

A R T I C L E I N F O

Research Article

VacRes, 2021
Vol. 8, No.1, 36- 46
Received: October 10, 2021
Accepted: November 03, 2021
Pasteur Institute of Iran

***Corresponding Author:**
Nemat Shams
Department of pathobiology,
Faculty of Veterinary Medicine,
Lorestan University, Khorramabad, Iran.**Email:** shams.n@lu.ac.ir
Tel/Fax: +986633120330

KEYWORDS: SARS-CoV-2, COVID-19, Epitope Prediction, HBHA, Poly-Epitope Vaccine

A B S T R A C T

Introduction: Many countries are presently concerned about providing a safe vaccine with minimal side-effects against COVID 19. Here, we aimed to develop a multiepitope vaccine by utilization of spike, envelope, nucleocapsid and membrane proteins of SARS-CoV-2 virus. **Methods:** Online servers were employed for forecasting the most robust B-cell, T-cell and IFN-γ epitopes to stimulate the immune system. Then the top selected epitopes alongside the sequence of Heparin-Binding Hemagglutinin Adhesin (HBHA) protein were applied to design a novel multiepitope vaccine, bioinformatically. The physicochemical characteristics and the protein structures of the proposed vaccine were defined using online tools. The docking process between Toll-like Receptor 4/Myeloid Differentiation Factor 2 (TLR4/MD2 receptor) and the designed recombinant structure was also investigated. **Results:** The designed construct had -0.210 GRAVY and 36.39 instability indices which make it theoretically stable. The designed construct was predicted to be soluble and non-allergenic. The approximate half-life of the proposed structure was computed 30 hours in mammalian reticulocytes and more than 10 hours in *Escherichia coli*. In its tertiary structure, 93% of the residues were in the core region and had a score of 52.73 for 3D verification and -5.55 for Z-score. Protein-protein docking of HBHA and TLR4/MD2 receptor was successful with the lowest energy of -1310.6 kcal/mol. **Conclusion:** The bioinformatics evaluations indicate that the designed structure is stable and immunogenic for development of a protein-based subunit vaccine against COVID-19.
Citation:

INTRODUCTION

The emergence of COVID-19 and the declaration of pandemic due to a rapidly evolving Coronavirus (SARS-CoV-2) by the WHO, made an urgent need to develop safe vaccines with high immunogenicity for the global public health. Coronaviruses are single-stranded RNA-positive sense viruses that are classified into four groups, namely α, β, γ and δ [1]. COVID-19 is the newest isolated coronavirus, belonging to the CoV-β class [2], which has 88% genomic similarity to two MERS-derived coronaviruses (bat-SL-CoVZC45 and bat-SL-CoVZXC21) and has also shown 50% similarity to the MERS-CoV sequence [2]. The new CoV-β was named SARS-CoV-2 by the global virus classification committee.

In case of recombinant vaccine production, some researchers have demonstrated promising results against SARS-CoV-2. For instance, a recombinant protein composed of SARS-CoV receptor-binding domain (RBD; named as RBD219-N1) was able to show minimal immunopathology after viral challenge with SARS-CoV-2 virus in the animal model. In addition, it produced high levels of neutralizing antibodies and has been introduced as a heterologous vaccine for COVID-19 [3]. In another study, it has been reported that using a DNA vaccine against SARS-CoV-2 could induce cellular and humoral immunity successfully and can prevent interaction between the spike protein and the ACE2 receptor of the virus [4]. In order to control COVID-19, Tazehkand and Hajipour (2020) have designed a recombinant vaccines, comprised of B-cell and MHC-I epitopes of spike protein and RNA-dependent RNA along with the envelope and the nucleocapsid protein of SARS-CoV-2 [3]. The genome sequence of SARS-CoV-2 like other typical CoVs contains at least 10 ORFs. Other SARS-CoV-2 ORFs encode four important structural proteins in the remaining one-third of the genome including the spike glycoprotein (S), the small envelope glycoprotein (E), the nucleocapsid protein (N), and

[DOI: 10.52547/vacres.8.1.36]

the membrane protein (M). It also has a number of other proteins that are involved in viral replication [4]. Studies have shown that S protein with molecular weight of ~180 kDa is a glycoprotein of coronaviruses which has an important role in entry into the host cell [5]. N protein which has several functions including nuclear-import signals, interfering cell process, virus replication and RNA package, plays a crucial role in the host infection. Moreover, this protein is recognized as the most conserved antigen among the other isolates, and it shows a high expression rate during the infection inside the host [6, 7]. E protein is responsible for envelope forming in CoV which in cooperation with M, is involved in creating the virus-like particle [8, 9].

Heparin-Binding Hemagglutinin Adhesin (HBHA) is highly regarded as a molecular adjuvant in the design of recombinant safe vaccines [10-12]. Inhibitors such as heparin have been able to prevent hemagglutination and the binding of Tuberculosis bacteria to the epithelial cells; therefore, HBHA can be identified as a bacterial adhesion agent [12] and has been considered as an immunostimulatory adjuvant. Through docking of HBHA molecule and the Toll-like receptor 4 (TLR4)/Myeloid Differentiation Factor 2 (TLR4/MD2 receptor), cell migration to the lymph nodes and the consequent immune responses increases [13]. Herein, well-defined online tools were applied to design a novel and safe conjugated protein-based vaccine, based on robust immunogenic epitopes of SARS-CoV-2 and HBHA protein as a putative COVID-19 vaccine.

MATERIALS AND METHODS

Retrieving the Protein Amino Acid Sequences and their Signal Peptide Identification

Amino acid sequences of N, S, M and E of SARS-CoV-2 with NCBI GenBank accession numbers of P0DTC2, QJR96107, QJR85133 and YP_009724393, respectively were obtained from the database (https://www.ncbi.nlm.nih.gov/). Since for each protein, several sequences were recorded in the database, after reviewing and aligning those using CLC Main Workbench software, the sequences without any unknown amino acids (X), were used for further analyses. Also, amino acid sequence of HBHA protein (UniProtKB P9WIP9) was obtained from the server (https://www.uniprot.org/). The amino acid lengths of each of the above-mentioned proteins were 1273, 75, 419, 222 and 199 residues, respectively. Signal peptide identification for each protein was performed using SignalP server version 4.1 (http://www.cbs.dtu.dk/services/SignalP-4.1/). Since viruses rely exclusively on host mechanism expression systems [13], "eukaryotes" was selected as the “organism group” in this analysis. In general, signal peptides are cleaved during protein folding process by signal peptidase [14], consequently, signal peptides cannot be presented to the immune system; hence, their presence should be eliminated.

T Cell and CTL Epitope Prediction

To identify epitopes of MHC-I (A-0101, A0201 and B-2705 alleles), MHC-II (DRB1-0101 alleles DRB1-0401 alleles) classes of T-cell and CTL of S, E, N and M proteins, the following online tools were used (Table 1).

Table 1. List of bioinformatics tools to predict the epitopes.

servers	The method used to predict epitopes	Link
T-cell epitopes prediction		
IEDB	SVM (stabilized matrix method) and ANN-based (Artificial neural network) method [15]	https://www.iedb.org/
NetMHC	ANN-based method [16]	http://www.cbs.dtu.dk/services/NetMHC/
NetCTL	ANN-based method [17]	http://www.cbs.dtu.dk/services/NetCTL/
Propred	quantitative matrix [18, 19]	http://www.imtech.res.in/raghava/propred/
B-cell epitopes prediction		
IEDB	scales-based and HMMs methods[20, 21]	http://tools.iedb.org/bcell/
Bcepred	based on Physiochemical or combination properties of amino acids with 58.7% accuracy[22]	http://www.imtech.res.in/raghava/bcepred/
ABCpred	ANN-based method [23]	http://www.imtech.res.in/raghava/abcpred/
BepiPred	a random forest algorithm trained on epitopes annotated from antibody-antigen protein structures [24]	http://www.cbs.dtu.dk/services/BepiPred

The applied servers had been previously used and confirmed in vaccine design studies. In general, these servers use specific algorithmes and methods such as SVM, HMN and ANN methods [25, 26]. Since the antigen-binding groove of MHC-I is closed, it would be matched with epitopes having a shorter length while antigen-binding groove of MHC-II is open and it could be matched with an epitope with a longer length [27]. Hence, tools were adjusted to predict epitopes with 9-12-mer lengths for MHC-I and 15-mer lengths for MHC-II. The process of selecting the best epitopes based on the high affinity (i.e., peptides with IC_{50} values < 50 nM) was done as mentioned in the previous studies about epitope prediction [26, 28].

B-Cell and IFN-γ Epitope Prediction

B-cell epitope prediction of N, S, M, and E proteins were conducted using online tools specified in Table 1. The lengths of the epitopes were adjusted for 14-20-mer. The final B-cell epitopes were selected based on common outputs of different servers. Detection of the most probable IFN-γ epitopes of the mentioned proteins was accomplished by IFN epitope server (http://crdd.osdd.net/raghava/ifnepitope/). Predicted positive epitopes with 15- mer lengths and the highest scores were selected.

Designing and Determining the Antigenicity, Physicochemical Properties, Solubility and Allergenicity of the Immunodominant Structure

[DOI: 10.52547/vacres.8.1.36]

To design a new structure, all the selected epitopes from the previous steps with high affinity, were assessed by Vaxigen server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) and their probability of antigenicity were determined. In each category (MHC-I, MHC-II, B-cell and IFN-γ), epitopes with high antigenicity probability were selected for designing the final structure. Different arrangements of the selected epitopes were assessed by determining the antigenicity of the multiepitope structure using Vaxigen server. Finally, the best arrangement of the categories of the epitopes was selected as follows: MHC-I epitopes, MHC-II epitopes, B-cell epitopes and IFN-γ epitopes. In order to prevent any interaction between linear epitopes and also to prevent communication between an epitope and the adjuvant domains, "KP" and "EAAAK" rigid linkers were used, respectively. The number of replicates of each linker was determined based on the physicochemical properties of the structure. To increase the half-life of the structure as well as to enhance the stimulation of the immune system, HBHA as a molecular adjuvant, was conjugated to N-terminal of the developed structure using "EAAAK" rigid linker. Determination of physicochemical properties of the designed construct was performed using ProtParam Server (https://web.expasy.org/protparam/). In this regard, molecular weight, half-life, GRAVY, aliphatic index, instability index and isoelectric point (PI) of the final designed conjugated structure were evaluated. Solubility of the designed immunodominant structure was investigated by SOlpro procedure of Scratch Proteomics server (http://scratch.proteomics.ics.uci.edu/). Allergencity prediction of this structure was done using AllerTOP v.2.0 (https://www.ddg-pharmfac.net/AllerTOP/) and AllerCatPro (https://allercatpro.bii.a-star.edu.sg/) online tools.

Secondary and Tertiary Structures of the Multiepitope Construct fused with HBHA

To show the distribution of different states of the multiepitope construct fused to HBHA consisting of alpha helices, extend strands, beta turns and random coils, SOPMA server (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) was applied. The 3D structure of this recombinant structure was modeled using the I-TASSER server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) which is a hierarchical approach to protein structure and function prediction. The recommended models by I-TASSER server were compared and the best one was selected according to C-score and estimated RMSD values. Refinement of the best model of the 3D structure was accomplished using GalaxyRefine server (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) and pdb file provided by I-TASSER server. Among the models that were presented, the best model was selected based on the analysis of Ramachandran plot through the VADAR server (http://vadar.wishartlab.com/). Assessment of the compatibility of an atomic model (3D) with its own amino acid sequence (1D) of the best refined model was done byVerify3D server (https://servicesn.mbi.ucla.edu/Verify3D/). The overall model quality of the best refined model which contained the z-scores was evaluated through ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) servers.

Investigation of the Interaction of the Designed Structure and its Receptor through the Docking Process

The effects of the designed construct, incorporating HBHA molecule and its interaction with TLR4/MD2 receptor through molecular docking process, was performed using ClusPro server (https://cluspro.org/login.php?redir=/queue.php). In this process, pdb file of TLR4/MD2 receptor with reference number 2Z64 was mined from RCSB server (https://www.rcsb.org/) and the refined model was selected.

RESULTS

Data Mining and Epitope Prediction

The consensus amino acid sequences of N, S, M, and E as well as HBHA protein and pdb file of TLR4/MD2 receptor were applied for epitope prediction procedure, after omitting their signal peptides. The most frequent epitopes observed in the results of several IC50-based online tools were selected and the most desired peptides which indicate high affinity among different alleles and different online tools were ranked based on their antigenicity probability for MHC I and MHC II classes as indicated in Tables 2 and 3.

Table 2. List of best MHC-I binding epitopes of N, S, M and E proteins which were ranked by VaxiJen antigenicity probability.

Antigen	Rank	Sequence	VaxiJen antigencity probability
Nucleocapsid	1	DLSPRWYFY	1.7645 (Probable ANTIGEN)
Nucleocapsid	2	GRRGPEQTQGNF	0.8980 (Probable ANTIGEN)
Nucleocapsid	3	NSSPDDQIGYY	0.5757 (Probable ANTIGEN)
Nucleocapsid	4	LLLLDRLNQL	0.5158 (Probable ANTIGEN)
Spike	1	KLNDLCFTNV	2.6927 (Probable ANTIGEN)
Spike	2	RQIAPGQTGK	1.7893 (Probable ANTIGEN)
Spike	3	QTGKIADYNYKL	1.4851 (Probable ANTIGEN)
Spike	4	KRVDFCGKGYHL	1.3047 (Probable ANTIGEN)
Spike	5	IGAEHVNNSY	1.2671 (Probable ANTIGEN
Spike	6	YRVVVLSFEL	1.1494 (Probable ANTIGEN)
Spike	7	NLDSKVGGNYNY	1.0950 (Probable ANTIGEN)
Spike	8	NLDSKVGGNY	0.7882 (Probable ANTIGEN)
Spike	9	NLDSKVGGNY	0.7882 (Probable ANTIGEN)
Spike	10	GYLQPRTFLL	0.7535 (Probable ANTIGEN)
Spike	11	IVRFPNITNLC	0.6785 (Probable ANTIGEN)
Spike	12	NLNESLIDLQELGK	0.6631 (Probable ANTIGEN)
Spike	13	YIWLGFIAGLIAIV	0.6100 (Probable ANTIGEN)

[DOI: 10.52547/vacres.8.1.36]

Membrane	1	IGFLFLTWICLL	1.4857 (Probable ANTIGEN)
Membrane	2	LAAVYRINWI	1.4192 (Probable ANTIGEN)
Membrane	3	YRINWITGGI	1.3250 (Probable ANTIGEN)
Membrane	4	TLACFVLAAV	1.2318 (Probable ANTIGEN)
Membrane	5	VTLACFVLAAVY	1.1958 (Probable ANTIGEN)
Membrane	6	HLGRCDIKDL	1.1136 (Probable ANTIGEN)
Membrane	7	GIAIAMACLV	1.0674 (Probable ANTIGEN)
Membrane	8	FLWLLWPVTL	0.9162 (Probable ANTIGEN
Membrane	9	VATSRTLSYY	0.7297 (Probable ANTIGEN)
Membrane	10	VAGDSGFAAYS	0.6661 (Probable ANTIGEN)
Envelope	1	CAYCCNIVNV	1.1752 (Probable ANTIGEN)
Envelope	2	RLCAYCCNIVNV	1.1612 (Probable ANTIGEN)
Envelope	3	TLAILTALRL	1.1138 (Probable ANTIGEN)
Envelope	4	FLLVTLAIL	0.9645 (Probable ANTIGEN)
Envelope	5	FLLVTLAIL	0.9645 (Probable ANTIGEN)
Envelope	6	VTLAILTALR	0.8404 (Probable ANTIGEN)
Envelope	7	LAFVVFLLVTLAILT	0.8229 (Probable ANTIGEN)

Table 3. List of best MHC-II binding epitopes of N, S, M and E proteins which were ranked by VaxiJen antigenicity probability.

Antigen	Rank	Sequence	VaxiJen antigenicity probability
Nucleocapsid	1	YFYYLGTGPEAGLPY	0.9657 (Probable ANTIGEN)
Nucleocapsid	2	PRWYFYYLGTGPEAG	0.8083 (Probable ANTIGEN)
Nucleocapsid	3	RWYFYYLGTGPEAGL	0.7505 (Probable ANTIGEN)
Nucleocapsid	4	KKQQTVTLLPAADLD	0.7325 (Probable ANTIGEN)
Nucleocapsid	5	KQQTVTLLPAADLDD	0.6264 (Probable ANTIGEN)
Spike	1	IPFAMQMAYRFNGIG	1.2828 (Probable ANTIGEN)
Spike	2	QIPFAMQMAYRFNGI	1.1851 (Probable ANTIGEN)
Spike	3	PTNFTISVTTEILPV	1.1349 (Probable ANTIGEN)
Spike	4	NFNFNGLTGTGVLTE	1.0516 (Probable ANTIGEN)
Membrane	1	LACFVLAAVYRINWI	1.2905 (Probable ANTIGEN)
Membrane	2	INWITGGIAIAMACL	1.1352 (Probable ANTIGEN)
Membrane	3	TWICLLQFAYANRNR	0.8808 (Probable ANTIGEN)
Membrane	4	TSRTLSYYKLGASQR	0.7675 (Probable ANTIGEN)
Envelope	1	VKPSFYVYSRVKNLN	1.2319 (Probable ANTIGEN)
Envelope	2	VTLAILTALRLCAYC	0.8599 (Probable ANTIGEN)
Envelope	3	LAFVVFLLVTLAILT	0.8229 (Probable ANTIGEN)
Envelope	4	FLLVTLAILTALRLC	0.6311 (Probable ANTIGEN)
Envelope	5	AFVVFLLVTLAILTA	0.6299 (Probable ANTIGEN)

According to the size of the binding site in MHC I and MHC II, epitopes with lengths of 9-12 and 15 amino acid residues were selected, respectively. The prediction of specific epitopes that would stimulate B-cells was similarly successful. As shown in Table 4, the most robust epitopes that were found in the results of the most servers were sorted based on the highest antigenicity probability which was reported by Vaxigen server.

Table 4. List of best B-cell binding epitopes of N, S, M and E proteins which were ranked by VaxiJen antigenicity probability.

Antigen	Rank	Sequence	VaxiJen antigenicity probability
Nucleocapsid	1	TRRIRGGDGKMKDLSP	1.1467 (Probable ANTIGEN)
Nucleocapsid	2	ASSRSSSRSRNSSRNS	0.9557 (Probable ANTIGEN)
Nucleocapsid	3	ESKMSGKGQQQQGQTVTK	0.8021 (Probable ANTIGEN)
Nucleocapsid	4	NSSPDDQIGYYRRATRRIRGGDGKMKDLSP	0.7116 (Probable ANTIGEN)
Nucleocapsid	5	ADETQALPQRQKKQQT	0.6949 (Probable ANTIGEN)
Nucleocapsid	6	RRGPEQTQGNFGDQELIRQG	0.6789 (Probable ANTIGEN)
Nucleocapsid	7	TGSNQNGERSGARSKQ	0.6333 (Probable ANTIGEN)
Nucleocapsid	8	NQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTAT	0.6198 (Probable ANTIGEN)
Spike	1	NSASFSTFKCYGVSPTKLNDLCFT	1.4194 (Probable ANTIGEN)
Spike	2	MDLEGKQGNFKNL	1.2592 (Probable ANTIGEN)
Spike	3	DEVRQIAPGQTGKIADYNYKLP	1.0899 (Probable ANTIGEN)
Spike	4	VSGTNGTKRFD	0.8493 (Probable ANTIGEN)
Spike	5	VNNSYECDIPIGA	0.8182 (Probable ANTIGEN)
Spike	6	LTPGDSSSGWTA	0.6890 (Probable ANTIGEN
Spike	7	YTMSLGAENSVAYSNN	0.6434 (Probable ANTIGEN)
Membrane	1	VTLACFVLAAVYRINW	1.3759 (Probable ANTIGEN)
Membrane	2	NLVIGFLFLTWICLLQ	1.1176 (Probable ANTIGEN)
Membrane	3	GDSGFAAYSRYRIGNY	0.8980 (Probable ANTIGEN)
Membrane	4	HHLGRCDIKDLP	0.7503 (Probable ANTIGEN)
Membrane	5	FRLFARTRSMWSFN	0.7155 (Probable ANTIGEN)

[DOI: 10.52547/vacres.8.1.36]

Membrane	6	PKEITVATSRTLSYYKLGAS	0.5935 (Probable ANTIGEN)
Envelope	1	CNIVNVSLVKPSFYVY	0.7373 (Probable ANTIGEN)
Envelope	2	LCAYCCNIVNVSLVKP	0.7286 (Probable ANTIGEN)
Envelope	3	VNVSLVKPSFYVYSRV	0.6685 (Probable ANTIGEN)
Envelope	4	YVYSRVKNLNSSRV	0.5405 (Probable ANTIGEN)

Subsequently, epitopes stimulating IFN-γ secretion were also identified, and one epitope with the highest score of IFNepitope server was selected for each N, S, M, and E proteins (Table 5).

Table 5. List of best IFN-γ binding epitopes of N, S, M and E proteins.

Antigen	Sequence	VaxiJen antigenicity probability
Nucleocapsid	QGTDYKHWPQIAQFA	0.6625 (Probable ANTIGEN)
Spike	VVFLHVTYVPAQEKN	1.1720 (Probable ANTIGEN)
Membrane	ITVATSRTLSYYKLG	0.8002 (Probable ANTIGEN)
Envelope	RLCAYCCNIVNVSLV	1.2823 (Probable ANTIGEN)

Designing the Multiepitope Structure and Determining its Features

The final multiepitope structure consisted of 5 main parts that included: HBHA molecule as an adjuvant which was located in the N-terminal region of the developed structure and was connected to a multiepitope section through an "EAAAK" rigid linker. This section consisted of 4 main parts that included the first rank epitopes of MHC-I, MHC-II, B- cell and IFN-γ of N, S, M, and E proteins (Table 2, 3 and 4), respectively. As shown in Fig. 1, each of these epitopes was also connected to a "KP" rigid linker so that the epitopes could be presented to the immune system in a linear manner, without movement or interaction with each other (Fig. 1).

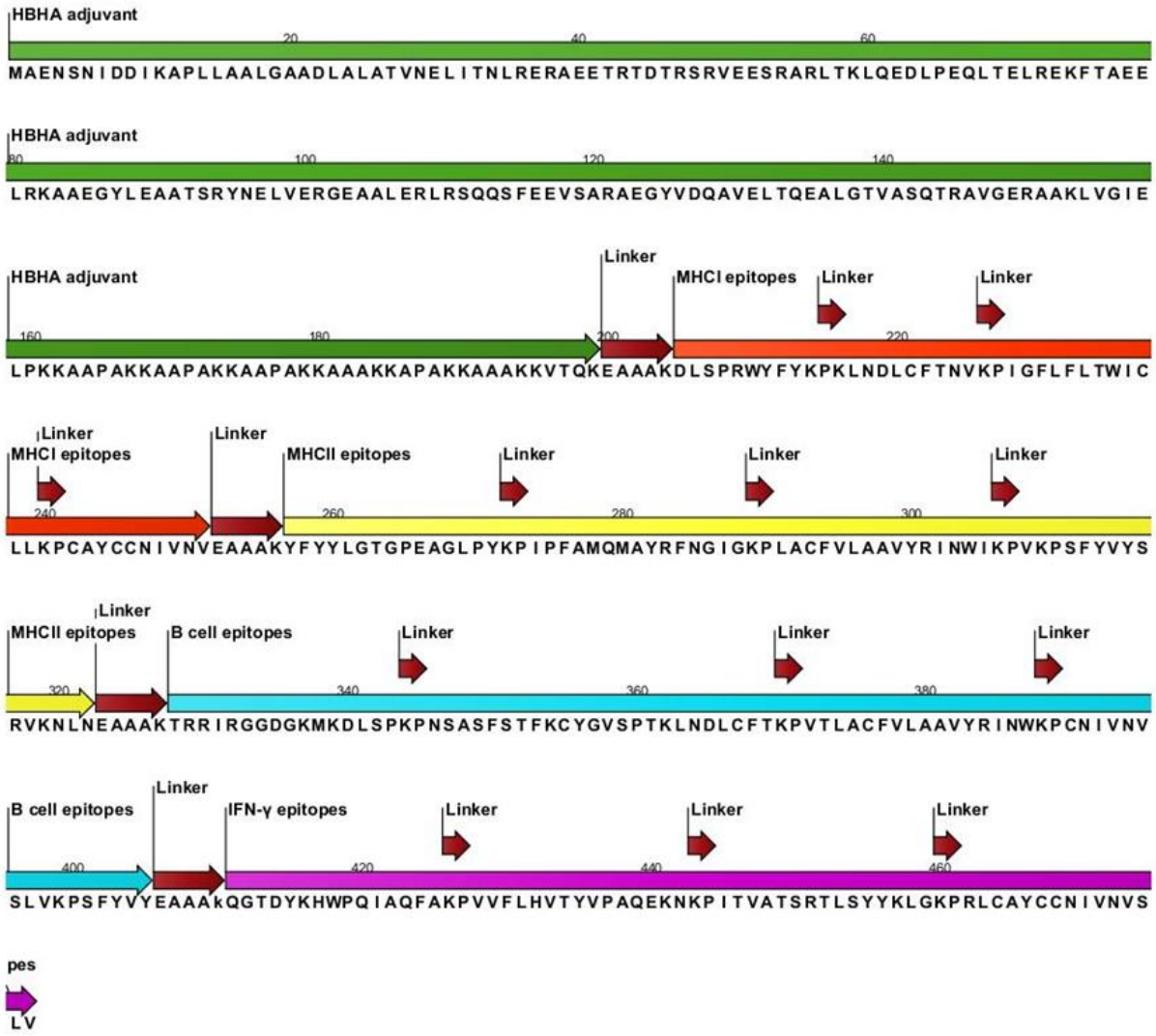


Fig. 1. Schematic representation of amino acid sequence of the designed structure (annotated using CLC Main Workebench software). The green color represents HBHA adjuvant, the red color is the MHC-I epitopes of N, S, M, and E proteins, the yellow color represents the MHC-II epitopes of N, S, M, and E proteins, the blue color shows B cell epitopes of N, S, M, and E proteins and the purple color refers to the INF-γ epitopes of N, S, M, and E proteins.

The total antigenicity probability of the multiepitope section was 0.8762 which implied this structure could be considered as a promising stimulator for inducing the immune system. Determining physicochemical properties by applying

[DOI: 10.52547/vacres.8.1.36]

ProtParam server was performed which indicated 476 amino acid residues, 52.75735 kDa Mw and 9.53 theoretical pI. Moreover, the aliphatic index was 85.40 and the grand average hydropathicity (GRAVY) was -0.210. Also, the instability index (II) was computed to be 36.39, which classifies this recombinant structure as stable. Estimation of half-life were 30 h in mammalian reticulocytes, *in-vitro*, more than 20 h in yeast *in-vivo* and more than 10 h in *Escherichia coli*, *in-vivo*. Evaluation of protein solubility revealed that this structure is soluble with probability 0.995682. Also, the allergenicity prediction by two different servers showed that the designed vaccine was known as non-allergene protein.

Determining the Characteristics of the Secondary Structure and the Best Tertiary Structure

Different states of this multiepitope structure were determined and results of SOMPA server predicted that the distribution of alpha helices, extended strands, beta turns and random coils of the construct were 50.42%, 18.07%, 3.36% and 28.15%, respectively (Fig. 2).

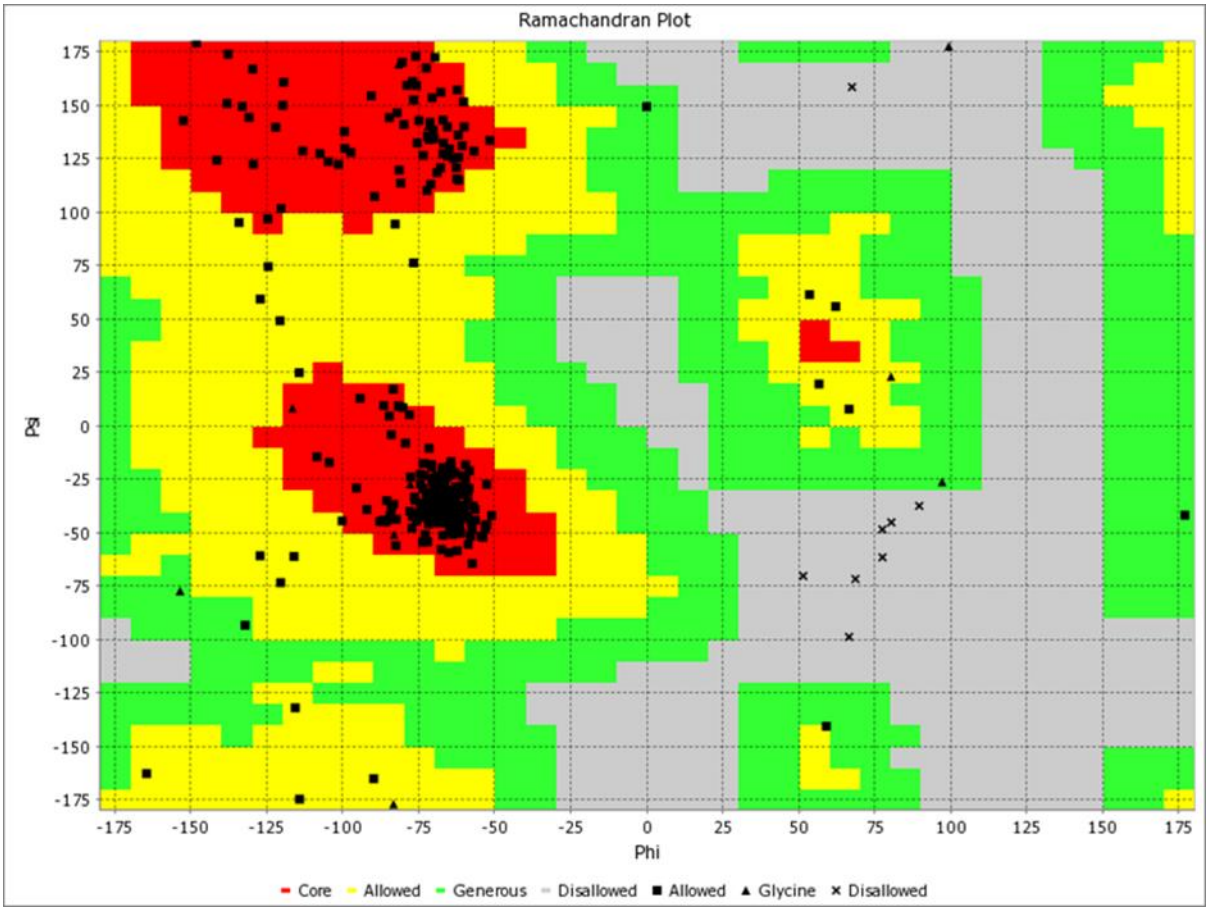
Fig. 2. Schematic secondary structure of the designed construct.

Using I-TASSER server, the selected best pdb file had a reliability coefficient score of 2.22 and estimated RMSD 12.6±4.3Aº. To improve the accuracy of the initial pdb model, GalaxyRefine server and Ramachandran plot provided the final selected 3D model while 93% of amino acid residues were located in phipsi core region and 4% of them were located in phipsi allowed region and 1% in phipsi outside region (Fig. 3).

The 3D verification done using Verify3D server indicated 52.73 score for the best refined model (Fig. 4).

[DOI: 10.52547/vacres.8.1.36]

Fig. 3. Ramachandran plot of the final 3D model. 93% of amino acid residues were located in phipsi core region and 4% of them were located in phipsi allowed region and 1% in phipsi outside.



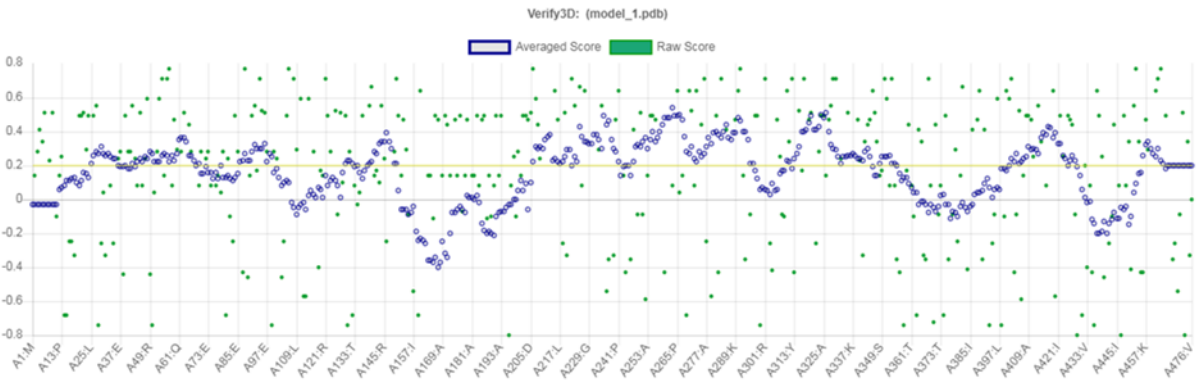


Fig. 4. Verify3D plotting validated the tertiary structure of best refined model with the score of 52.73.

[DOI: 10.52547/vacres.8.1.36]

Furthermore, analysis by ProSA server on the quality of the model presented Z-score of -5.55 (Fig. 5).

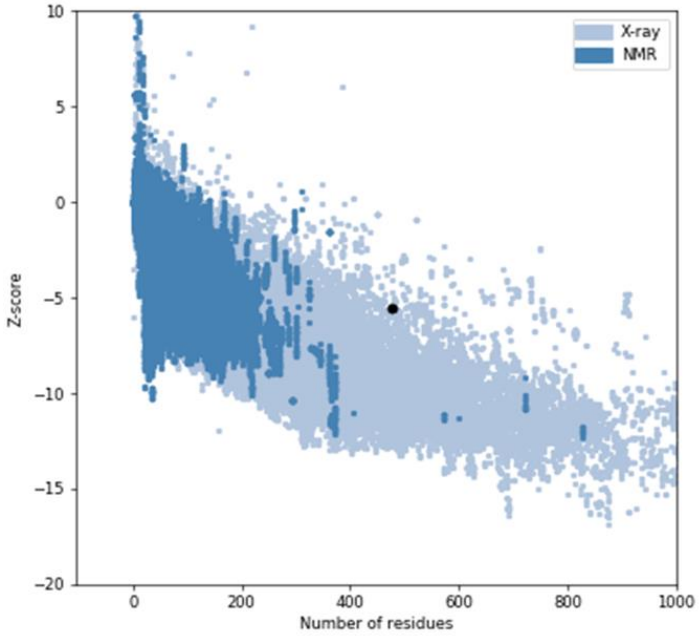


Fig. 5. Z-score plot. Overall quality of the refined model was calculated to be -5.55.

Molecular docking

Interaction of HBHA molecule fused to the multiepitope fragment with TLR4/MD2 receptor proteins by ClusPro server showed that the HBHA molecule attached to the multiepitope fragment was able to bind well to its TLR4/MD2 receptor. The score of the cluster center was -965.3 with 63 cluster members. The lowest energy of this binding was calculated to be -1310.6 kcal/mol (Fig. 6).

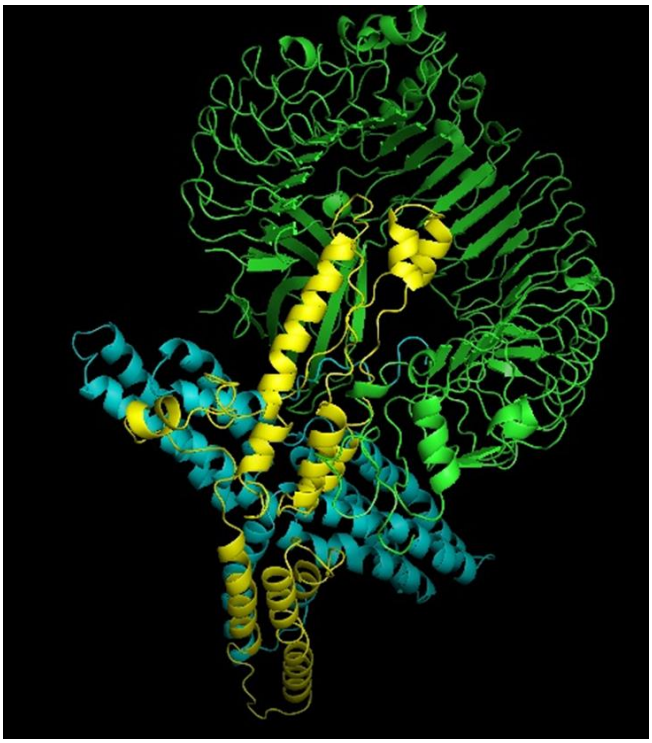


Fig. 6. Docking of HBHA molecule fused to the multiepitope fragment with TLR4/MD2 receptor proteins. The green color indicates TLR4/MD2 receptor. The blue color shows HBHA molecule. The yellow color is part of the designed structure. This picture has been visualized by PyMOL software.

[DOI: 10.52547/vacres.8.1.36]

DISCUSSION

The outbreak of COVID-19 pandemic is one of the biggest public health challenges of recent years. The transmission rate of this disease among individuals and the possibility of death is very high, since a significant proportion of the infected people are asymptomatic or have mild symptoms and do not seek medical attention [29, 30]. Therefore, the importance of vaccination is well understood because it homogenizes all members of the society, whether people at high risk or those with stronger immune systems. Other benefits of vaccinating against pandemic diseases include stabilizing the public mental health and economic growth of the afflicted countries [31].

Interest in developing new vaccination strategies against the threat of infectious diseases has resulted in development of a new generation of safe vaccines which only use the immunogenic components of the pathogens [32]. The formulation and the profile of these vaccines are under investigations by the researchers and the pharmaceutical manufacturers. In addition, Gram-positive and Gram-negative bacteria can be used in the production of these vaccines as a platform for production and expression of immunological proteins [33, 34]. Recombinant DNA-based techniques are the third generation of vaccines which along the second-generation (consisted of the effective immunogenic components of the microorganisms) [35] were added to the first generation of attenuated and inactivated vaccines [36].

Today, the recombinant protein technology has allowed producing protein antigens instead of live attenuated or inactivated bacterial and viral vaccines. In addition, large-scale production of recombinant proteins at a lower cost is one of the major advantages of this technology [37]. However, these small peptides are weaker in stimulating the immune system than the whole virus or bacterium which are found in conventional vaccines. Moreover, they have a shorter half-life due to their low molecular weight and are rapidly eliminated from the renal filtration system. Therefore, to stimulate the immune system more effectively, and have a longer half-life of the protein-based vaccines, these peptides must be combined with heavier molecules, known as adjuvants [38, 32].

HBHA as an agonist for TLR4/MD2 receptor and is considered as a known molecular adjuvant in multiepitope vaccine [26, 39]. This complex is able to initiate the initial production of anti-inflammatory cytokines and the production of chemokines [40]. In this study, using *in-silico* strategies, a multiepitope structure against COVID-19 was designed. In this regard S, M, N and E proteins of SARS-CoV-2 virus were analyzed through bioinformatics tools to identify the most effective epitopes for engagement in MHC-I, MHC-II, B-cells and IFN-γ categories to induce effective immune responses.

The resulting immunodominant epitope-base structure was connected through rigid peptide linkers to provide epitopes linearity for better presentation of the epitopes to the immune system. Repetition in the linkers was based on the regulation of physicochemical properties of the structure. GRAVY and instability index (II) were computed to be -0.210 and 36.39 which indicated a stable putative protein since a negative GRAVY indicates that the target protein is non-polar and instability index more than 40 indicates unstable proteins [41]. As it is shown in Fig. 2, in the C-terminal of the designed structure which is included with a multiepitope fragment, random coil (Cc) and extended strand (Ee) are the dominant states. These states create exposed and hydrophilic regions that are prominent features of effective epitopes [42, 43].

The *in-silico* docking procedure was used to investigate the potential side-effects of the immunodominant multiepitope structure on HBHA orientation which might disturb the protein-protein docking of HBHA molecule and TLR4/MD2 receptor. Fortunately, the results demonstrated successful binding of the designed structure and TLR4/MD2 receptor with the lowest energy. Correct docking is perceived to promote longer half-life and the consequent longer stimulation of the immune system.
The most important immunopathological event resulting from SARS-CoV-2 infections is related to acute respiratory distress syndrome (ARDS).

Cytokine storm and uncontrolled systemic inflammatory responses due to SARS-CoV-2 diffusion cause the release of large amounts of inflammatory cytokines such as IFN-α, IFN-γ, IL-1β, IL-6, IL-12, 18, IL- 33, TGFβ and chemokines such as CCL2, CCL3, CCL5, CXCL8, CCL2, CCL3, CCL5 and CXCL8 [4]. Cytokine storms cause severe immune system attacks, ARDS, and multiple organ failure which eventually lead to death from severe SARS-CoV-2 infection. According to recent reports, humoral and cellular immune systems are stimulated following the presentation of antigens of SARS-CoV-2. Therefore, in COVID-19 patients, nucleocapsid protein specific antibodies and IgM peak at day 9 after being infected and IgG peak after week 2 have been observed [44]. In another study, researchers observed specific antibodies against the spike protein on the third day after the onset of the symptoms [45].

IgG and IgM responses were observed in patients with COVID-19, especially specific antibodies against the receptor-binding domain of the spike and the nucleocapsid proteins. Detected IgG against these two proteins was mainly IgG1 isotype. In the case of stimulated cellular immunity against this virus, IFN-γ ELISpot analysis of infected individuals revealed a 2-fold increase in the numbers of IFN-γ-secreting T cells rather than the healthy people [46]. In 36 cases of patients with COVID-19, specific CD4+ and CD8+ T cells against nucleocapsid proteins were detected which subsequently resulted in the release of IFN-γ [47]. Researchers declared that SARS-CoV-2 elicits strong, extensive, and highly effective T cell responses, confirming that exposure to infection may prevent the recurrence of COVID-19 [48].

Therefore, designing vaccines based on specific MHC-I, MHC-II, B-cell and IFN-γ epitopes that produce memory cells in all four groups such as the construct studied here can be an efficient strategy. Further investigations consisting of production of this protein-based vaccine and its application in animal models are required to test the efficiency of this multiepitope construct *in-vivo*.

ACKNOWLEDGEMENT

The authors would like to thank the Department of Pathobiology, Faculty of Veterinary Medicine, Lorestan University, for providing computer facilities for this study.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

[DOI: 10.52547/vacres.8.1.36]

REFERENCES

1. Su S, Wong G, Shi W, Liu J, Lai AC, Zhou J et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends in microbiology. 2016;24(6):490-502.
2. Perlman S, Netland J. Coronaviruses post-SARS: update on replication and pathogenesis. Nature reviews microbiology. 2009;7(6):439-50.
3. Tazehkand M, Hajipour O. Evaluating the vaccine potential of a tetravalent fusion protein against coronavirus (COVID-19). J Vaccines Vaccin. 2020;11:411.
4. Li X, Geng M, Peng Y, Meng L, Lu S. Molecular immune pathogenesis and diagnosis of COVID-19. Journal of Pharmaceutical Analysis. 2020.
5. de Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. Nature Reviews Microbiology. 2016;14(8):523.
6. Kuo L, Masters PS. Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus. Journal of virology. 2002;76(10):4987-99.
7. Hiscox JA, Wurm T, Wilson L, Britton P, Cavanagh D, Brooks G. The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus. Journal of Virology. 2001;75(1):506-12.
8. Tung FY, Abraham S, Sethna M, Hung S-L, Sethna P, Hogue BG et al. The 9-kDa hydrophobic protein encoded at the 3′ end of the porcine transmissible gastroenteritis coronavirus genome is membrane-associated. Virology. 1992;186(2):676-83.
9. Vennema H, Godeke GJ, Rossen J, Voorhout WF, Horzinek MC, Opstelten D et al. Nucleocapsid□ independent assembly of coronavirus□ like particles by co□ expression of viral envelope protein genes. The EMBO journal. 1996;15(8):2020-8.
10. Kohama H, Umemura M, Okamoto Y, Yahagi A, Goga H, Harakuni T et al. Mucosal immunization with recombinant heparin-binding haemagglutinin adhesin suppresses extrapulmonary dissemination of Mycobacterium bovis bacillus Calmette-Guerin (BCG) in infected mice. Vaccine. 2008;26(7):924-32.
11. Rashidian E, Forouharmehr A, Nazifi N, Jaydari A, Shams N. Computer-Aided Design of a Novel Poly-Epitope Protein in Fusion with an Adjuvant as a Vaccine Candidate Against Leptospirosis. Current Proteomics. 2021;18(2):113-23.
12. Nezafat N, Ghasemi Y, Javadi G, Khoshnoud MJ, Omidinia E. A novel multi-epitope peptide vaccine against cancer: an in silico approach. Journal of theoretical biology. 2014;349:121-34.
13. Gale M, Tan S-L, Katze MG. Translational control of viral gene expression in eukaryotes. Microbiology and molecular biology reviews. 2000;64(2):239-80.
14. Nazifi N, Tahmoorespur M, Sekhavati MH, Haghparast A, Behroozikhah MA. Assessment of signal peptides to optimize interleukin 2 (IL-2) folding and expression. Current Proteomics. 2019;16(3):188-98.
15. Zhang Q, Wang P, Kim Y, Haste-Andersen P, Beaver J, Bourne PE et al. Immune epitope database analysis resource (IEDB-AR). Nucleic acids research. 2008;36(suppl_2):W513-W8.
16. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S et al. Reliable prediction of T□ cell epitopes using neural networks with novel sequence representations. Protein Science. 2003;12(5):1007-17.
17. Larsen MV, Lundegaard C, Lamberth K, Buus S, Brunak S, Lund O et al. An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions. European journal of immunology. 2005;35(8):2295-303.
18. Sturniolo T, Bono E, Ding J, Raddrizzani L, Tuereci O, Sahin U et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. Nature biotechnology. 1999;17(6):555.
19. Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. Bioinformatics. 2001;17(12):1236-7.
20. Chou PY, Fasman GD. Prediction of the secondary structure of proteins from their amino acid sequence. Advances in enzymology and related areas of molecular biology. 1978;47:45-148.
21. Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. Journal of virology. 1985;55(3):836-9.
22. Saha S, Raghava GPS, editors. BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties2004; Berlin, Heidelberg: Springer Berlin Heidelberg.
23. Saha S, Raghava G. Prediction of continuous B□ cell epitopes in an antigen using recurrent neural network. Proteins: Structure, Function, and Bioinformatics. 2006;65(1):40-8.
24. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic acids research. 2017;45(W1):W24-W9.
25. Jaydari A, Forouharmehr A, Nazifi N. Determination of immunodominant scaffolds of Com1 and OmpH antigens of Coxiella burnetii. Microbial pathogenesis. 2019;126:298-309.
26. Rashidian E, Gandabeh ZS, Forouharmehr A, Nazifi N, Shams N, Jaydari A. Immunoinformatics Approach to Engineer a Potent Poly-epitope Fusion Protein Vaccine Against Coxiella burnetii. International Journal of Peptide Research and Therapeutics. 2020:1-11.
27. Ehsan R, Ali F, Narges N, Amin J, Nemat S. Computer-Aided Design of a Novel Poly-Epitope Protein in Fusion with an Adjuvant as a Vaccine Candidate Against Leptospirosis. Current Proteomics. 2020;17:1-11. doi:http://dx.doi.org/10.2174/1570164617666200319144331.
28. Shams N, Gandabeh ZS, Nazifi N, Forouharmehr A, Jaydari A, Rashidian E. Computational Design of Different Epitope-Based Vaccines Against Salmonella typhi. International Journal of Peptide Research and Therapeutics. 2019:1-13.
29. Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung S-m, Hayashi K et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). International journal of infectious diseases. 2020;94:154.
30. Kim G-u, Kim M-J, Ra SH, Lee J, Bae S, Jung J et al. Clinical characteristics of asymptomatic and symptomatic patients with mild COVID-19. Clinical Microbiology and Infection. 2020.
31. Committee NVA. Protecting the public's health: critical functions of the Section 317 Immunization Program—a report of the National Vaccine Advisory Committee. SAGE Publications Sage CA: Los Angeles, CA; 2013.
32. Jaydari A, Nazifi N, Forouharmehr A. Computational design of a novel multi-epitope vaccine against Coxiella burnetii. Human Immunology. 2020.
33. Wilson RL, Hruby DE. Commensal bacteria as a novel delivery system for subunit vaccines directed against agents of bioterrorism. Advanced drug delivery reviews. 2005;57(9):1392-402.
34. Jorge S, Dellagostin OA. The development of veterinary vaccines: a review of traditional methods and modern biotechnology approaches. Biotechnology Research and Innovation. 2017;1(1):6-13.
35. Delany I, Rappuoli R, Seib KL. Vaccines, reverse vaccinology, and bacterial pathogenesis. Cold Spring Harbor perspectives in medicine. 2013;3(5):a012476.
36. Tahamtan A, Charostad J, Hoseini Shokouh SJ, Barati M. An Overview of History, Evolution, and Manufacturing of Various Generations of Vaccines. J Arch Mil Med. 2017;5(3):e12315. doi:10.5812/jamm.12315.
37. Andersen DC, Krummen L. Recombinant protein expression for therapeutic applications. Current opinion in biotechnology. 2002;13(2):117-23.
38. Werle M, Bernkop-Schnürch A. Strategies to improve plasma half life time of peptide and protein drugs. Amino acids. 2006;30(4):351-67.
39. Kardani K, Bolhassani A, Namvar A. An overview of in silico vaccine design against different pathogens and cancer. Expert Review of Vaccines. 2020:1-28.
40. Kanzler H, Barrat FJ, Hessel EM, Coffman RL. Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. Nature medicine. 2007;13(5):552-9.
41. Forouharmehr A, Nassiri M, Ghovvati S, Javadmanesh A. Evaluation of different signal peptides for secretory production of recombinant bovine pancreatic ribonuclease A in gram negative bacterial system: an in silico study. Current Proteomics. 2018;15(1):24-33.
42. 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. Iranian Journal of Applied Animal Science. 2017;7(3):429-36.
43. Li Y, Liu X, Zhu Y, Zhou X, Cao C, Hu X et al. Bioinformatic prediction of epitopes in the Emy162 antigen of Echinococcus multilocularis. Experimental and therapeutic medicine. 2013;6(2):335-40.
44. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. nature. 2020;579(7798):270-3.
45. Amanat F, Stadlbauer D, Strohmeier S, Nguyen T, Chromikova V, McMahon M et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. medRxiv. 2020:2020.03.17.20037713. doi:10.1101/2020.03.17.20037713.
46. Ni L, Ye F, Cheng M-L, Feng Y, Deng Y-Q, Zhao H et al. Detection of SARS-CoV-2-specific humoral and cellular immunity in COVID-19 convalescent individuals. Immunity. 2020.
47. Le Bert N, Tan AT, Kunasegaran K, Tham CY, Hafezi M, Chia A et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature. 2020;584(7821):457-62.

[DOI: 10.52547/vacres.8.1.36]

48. Sekine T, Perez-Potti A, Rivera-Ballesteros O, Strålin K, Gorin J-B, Olsson A et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. Cell. 2020. doi:https://doi.org/10.1016/j.cell.2020.08.017.

[DOI: 10.52547/vacres.8.1.36]