Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2021.033

RESEARCH ARTICLE

Designing of Multiepitope-based Subunit Vaccine (MESV) against Prevalent Serotype of Foot and Mouth Disease Virus (FMDV) using Immunoinformatics Approach

Rimsha Riaz[1], Muhammad Sarwar Khan[1*], Faiz Ahmad Joyia[1] and Muhammad Anjum Zia[2]

[1]Centre of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture Faisalabad, Pakistan
[2]Department of Biochemistry, University of Agriculture Faisalabad, Pakistan
*Corresponding author: sarwarkhan_40@hotmail.com

ARTICLE HISTORY (20-638)

Received: December 29, 2020
Revised: February 22, 2021
Accepted: February 25, 2021
Published online: March 13, 2021

Key words:
B-cell epitopes
CTL epitopes
FMD virus serotype O
Multiepitope-based subunit vaccine (MESV)
PanAsia-II
Reverse vaccinology

ABSTRACT

Foot-and-Mouth Disease (FMD) is causing 100% morbidity in adult animals with a high mortality rate in young animals due to cardiomyopathy. The disease-causing virus (FMDV) has been classified into seven serotypes and several topotypes. Amongst serotypes, the serotype O is predominantly present in Pakistan. A vaccine that should be safer and appears as an alternative to conventional vaccines that have failed to control the disease is required. Keeping in view, a multiepitope-based subunit vaccine (MESV) is designed using computational biology-based approaches. The sequence and structural characterization of capsid proteins and $3C^{pro}$ of O PanAsia-II sub-lineage are carried out for screening the antigenic proteins. Further, the B- and T-cell epitopes are selected based on toxicity, antigenicity, allergenicity, conservancy, and topology. The Cytotoxic T lymphocyte (CTL) epitopes are confirmed through molecular interactions with selected bovine alleles. The designed MESV is found stable, hydrophilic, antigenic, non-allergen, exposed extracellularly, non-toxic, and soluble. It will stimulate the protective immune response against the viral challenge as it has shown a strong binding affinity with the TLR4. These results suggest that the developed MESV will prove to be a potential candidate vaccine in controlling the disease against serotype O. For expression and accumulation of stable protein the sequence that encodes the multiepitope-based subunit vaccine is codon-optimized following the codon preference of *E. coli* and is cloned in pET-28a (+).

To Cite This Article: Riaz R, Khan MS, Joyia FA and Zia MA, 2021. Designing of Multiepitope-based Subunit Vaccine (MESV) against prevalent serotype of Foot and Mouth Disease Virus (FMDV) using immunoinformatics approach. Pak Vet J, 41(3): 400-408. http://dx.doi.org/10.29261/pakvetj/2021.033

INTRODUCTION

Livestock is a highly-regarded component of the Agriculture sector as it contributes 11.69% to the Gross Domestic Product (GDP) of Pakistan. More than eight million families are dependent on this sector for their livelihood and earn 35-40% of their income by livestock farming (Azam and Shafique, 2017). Animal population and their production are heavily affected due to diseases like Peste des Petits Ruminants (PPR), bovine mastitis, Foot and Mouth Disease (FMD), Caprine arthritis encephalitis, Hemorrhagic Septicemia (HS), etc. Amongst them, FMD is a highly contagious and life-threatening disease caused by Foot-and-Mouth Disease Virus (FMDV) - the first known animal virus. The disease causes an economic loss of more than 10 billion US dollars annually in Pakistan and 6.5-21 billion US dollars globally (Knight-Jones and Rushton, 2013). FMD leads to infertility, trade restrictions, and a reduction in milk and meat production. Also, it causes mortality in young animals (due to cardiomyopathy) and morbidity among adults (Pattnaik *et al.*, 2012; Truong *et al.*, 2018).

FMDV is a non-enveloped RNA virus of the *Picornaviridae* family and a sole pathogenic member of the genus *Aphthovirus*. It has a positive-sense single-stranded genome of approximately 8.4 kilobases encased in a protein capsid that encompasses 60 copies of structural proteins (VP1, VP2, VP3, and VP4), of which VP1-3 are localized on the surface of the capsid and play a significant role in attributing immunogenicity. The VP4 is found to be entirely buried inside the capsid. Among these structural proteins, VP1 is the mainstay in eliciting viral-specific immune response due to both T- and B-cell epitopes, provoking humoral and cell-mediated immunity (Mahdy *et al.*, 2019). FMDV is classified into seven serologically distinct serotypes based on VP1 protein i.e., FMDV

serotype O, C, Asia-1, A, SAT1-3, and above 60 subtypes (Abubakar and Kanwal, 2012). Of these serotypes, Asia-1, A, and O are found in Pakistan. However, Serotype O is the most prevalent in the country contributing to more than 60% of the total cases (Waheed *et al.*, 2011). Due to high antigenic variation, serotype O has been further classified into 11 topotypes. The molecular studies have disclosed that the O Pan Asia-II sub-lineage is stemmed from Pakistan and is spread all around the world (Brito *et al.*, 2013). Thus, this serotype is selected for the current study to design a potent vaccine as a part of Integrated Disease Management (IDM) in Pakistan.

Stringent precautionary measures and vaccination are the most effective and valuable strategies of IDM to prevent epidemics outbreaks of FMD. Although conventional vaccines are helping a lot in controlling the disease but failed to control virus shedding. Moreover, there are some limitations associated with these vaccines, like heat lability, limited shelf-life, impotency to immunize against multiple serotypes, massive virus culturing, incomplete virus inactivation, and a chance of escaping virulent virus and causing disease. To address these challenges, there is a dire need for the next-generation FMD vaccines i.e., peptide vaccine, subunit vaccine, and DNA vaccine (Mignaqui *et al.*, 2019). With the advent of immuno-informatics, the use of CTL and B-cells epitopes in designing the Multiepitope-based Subunit Vaccine (MESV) is a novel strategy in eliciting an efficient adaptive immune response. Their immunogenicity could be enhanced by engineering suitable adjuvant with them. Hence, the current study aims to predict the effective B- and T-cell epitopes that could be used as promising vaccine candidates for the construction of multiepitope-based subunit vaccine against FMDV serotype O.

MATERIALS AND METHODS

Immunoinformatic Analyses of Antigenic Proteins

Sequence and structural characterization of proteins: The amino acid sequences of capsid proteins, VP4, VP2, VP3, VP1, 2A, and $3C^{pro}$ were retrieved from NCBI in FASTA format. The physicochemical parameters of target proteins were investigated using the ProtParam tool (https://web.expasy.org/protparam) in the ExPASy server. The secondary structure analysis was assessed using SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html). The 3D structures of the proteins are accessible in the RCSB PDB.

Immunogenic Profiling: Further, the immunogenic assessment (antigenicity and allergenicity) of all the proteins was examined. AllerTop v.2.0 tool (https://www.ddg-pharmfac.net/AllerTOP/) estimated the allergenicity while their antigenicity was calculated by the VaxiJen v.2.0 tool (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) using 0.4 as a threshold value. After comparison, the desired proteins were further subjected to various analyses to design an effective vaccine.

Prediction and evaluation of B-cell and T-cell epitopes: Prediction of 16-mer long B cell epitopes was done using ABCPred (http://crdd.osdd.net/raghava/abcpred/ABC_submission.html) and IEDB-AR (http://tools.iedb.org/main/bcell/). Determination of B cell epitopes was based on flexibility, antigenicity, hydrophilicity, surface accessibility, and presence of beta turns. T-cell epitopes were predicted via an online available tool, NetMHC pan 4.1 (http://www.cbs.dtu.dk/services/NetMHCpan/). The epitopes of 9 residues length were selected against different BoLA alleles.

The predicted epitopes were assessed for antigenicity, allergenicity, and toxicity through online accessible servers VaxiJen v2.0, AllerTOP and ToxinPred (http://crdd.osdd.net/raghava/toxinpred/) respectively. Additionally, the topology and epitope conservancy were checked by TMHMM v2.0 server (http://www.cbs.dtu.dk/services/TMHMM/) and epitope conservancy analysis tool of IEDB-AR.

Structure modeling and molecular interaction: The 3D structures of CTL epitopes were predicted via CABS-fold server using de-novo structure prediction mode (http://biocomp.chem.uw.edu.pl/CABSfold/). The protein sequences of BoLA alleles were retrieved from the IPD-MHC database and submitted to SWISS-MODEL (https://swissmodel.expasy.org/interactive) for structure modeling that was assessed using the Ramachandran plot. To analyze the vaccine potency for inducing an effective immune response, blind docking was carried out between CTL epitopes (ligand) and BoLA alleles (protein) by an online docking tool PATCHDOCK (https://bioinfo3d.cs.tau.ac.il/PatchDock/patchdock.html). The docked models were further modified by FIREDOCK (http://bioinfo3d.cs.tau.ac.il/FireDock/php.php) LigPlus and Discovery Studio were used to visualize 3D protein-ligand interaction.

Multiepitope vaccine designing and its immunoprofiling: All the B-cell and T-cell epitopes were joined by the GPGPG linker to maintain their immunogenic activities. An adjuvant was engineered to boost up the immunity using the EAAAK linker. An efficient and safer vaccine should be soluble, non-allergen, non-toxic, and antigenic. Toxicity and allergenicity were evaluated by ToxinPred and AllerTOP v2.0. VaxiJen v2.0 was used to calculate the antigenicity of the MESV construct.

Structural characterization: The physical and chemical properties of the MESV construct were accessed by ProtParam. The solubility of the vaccine construct was predicted using SOLpro (http://scratch.proteomics.ics.uci.edu). The sequence was submitted to SOPMA to determine its secondary structure. It was ascertained that the template for modeling the 3D structure of the vaccine construct is not available, therefore a CABS-fold server was used for the prediction of its tertiary structure (http://biocomp.chem.uw.edu.pl/CABSfold/). Refinement was done by Modrefiner and further verified the structure through PROCHECK, and verify3D.

Molecular docking of Toll-Like Receptors (TLRs) with MESV: The molecular interaction between the host receptor (TLR4) and the chimeric vaccine was evaluated using ClusPro 2.0 (https://cluspro.bu.edu/login.php). The best-interacted complex was selected, and their interfaced amino acids were visualized using UCSF Chimera.

Codon optimization and *in silico* cloning: With an aim of cloning and expression of final MESV in an expression vector, codon optimization of MESV nucleotide sequence was executed through JCAT server (http://www.jcat.de/) using *Escherichia coli* (strain K12) as a host organism. To facilitate the cloning of the MESV construct in pET-28a(+), the recognition sequences for *Bam*HI and *Hind*III were engineered at the flanks of the construct, and *in silico* cloning was done using SnapGene tool.

RESULTS

Sequence analysis of curated proteins: The O PanAsia-II sublineage of FMDV was selected for the retrieval of amino acid sequences of capsid proteins, VP1, VP2, VP3, and VP4, 2A and $3C^{pro}$ using an accession number GU384683.1. The physical and chemical properties of all target proteins determined via ProtParam (Gasteiger *et al.*, 2013) categorized them as stable proteins with the instability indexes of 24.36 (VP1), 31.73 (VP2), 21.95 (VP3), 26.18 (VP4), 7.62 (2A), and 21.27 ($3C^{pro}$). Further, the results showed high aliphatic indices ((VP1) 82.32, (VP2) 78.58, (VP3) 69.32, (VP4) 44.94, (2A) 130.00, ($3C^{pro}$) 81.03) and low GRAVY values ((VP1) -0.380, (VP2) -0.245, (VP3) -0.129, (VP4) -0.698, (2A) 0.178, ($3C^{pro}$) -0.045) of all the proteins indicating them as thermostable and hydrophilic, respectively. The molecular weight of VP1, VP2, VP3, and $3C^{pro}$ was almost equal, i.e., 23.28kD, 24.34kD, 23.91kD, and 23.00kD, respectively. The molecular weight of 2A is 1.91kD and VP4 is 8.77kD.

Structural characterization of curated proteins: Secondary structures of selected proteins were explored by the SOPMA server. It predicts the structure using the homology method by aligning the related sequences of the same family. Table 1 describes the percentage frequency of beta turns, alpha helices, random coils, and extended sheets predicted in all the proteins. The transmembrane (TM) topology was also verified and found that all the proteins are exposed to the membrane and would likely be involved in the immune response during infection. During PSI-BLAST analysis, it was found that structures of all the target proteins are already available in Protein Data Bank (PDB), hence they were retrieved using the following PDB IDs: 1FMD (VP4), 1BBT (VP2), 5DDJ (VP3), 5NE4 (VP1) and 2WV4 ($3C^{pro}$).

Estimation of allergenicity and antigenicity of proteins: The allergenicity of the retrieved proteins was evaluated using AllerTop v2.0 (Dimitrov *et al.*, 2014) and observed VP4 and 2A as allergens while other proteins were characterized as non-allergens. The antigenic score of the proteins was calculated through VaxiJen using the threshold value of 0.4 (Meza *et al.*, 2017). Except for 2A (0.1943), all the proteins were revealed to stimulate protective immunity in the host due to high VaxiJen scores: VP1(0.5891), VP2(0.4428), VP3(0.4941), VP4 (0.4255), and $3C^{pro}$ (0.4584). Comparing both parameters, VP2, VP3, VP1 and $3C^{pro}$ were selected for the immunogenic analyses, fulfilling the criteria of being non-allergen and antigenic.

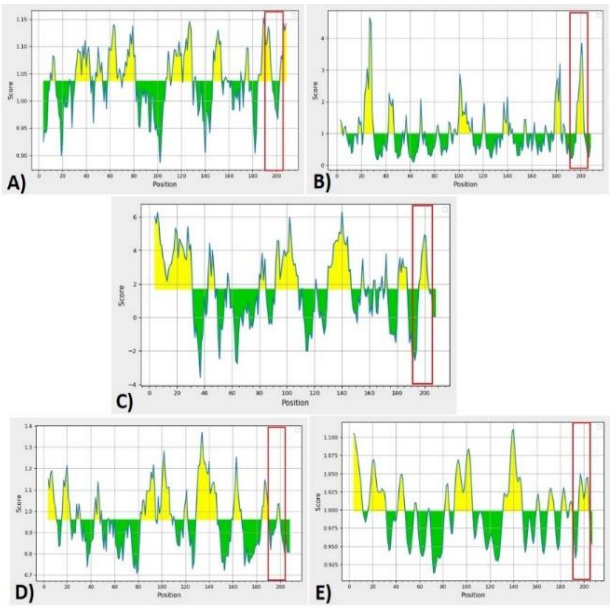

Fig. 1: Prediction of B-cell epitopes of VP1, representing the highly potent region as yellow peaks in graphs (A) Kolaskar and Tongaonkar antigenicity prediction (B) Emini surface accessibility prediction (C) Parker hydrophilicity prediction (D) Chou and Fasman beta-turn prediction (E) Karplus and Schulz flexibility prediction.

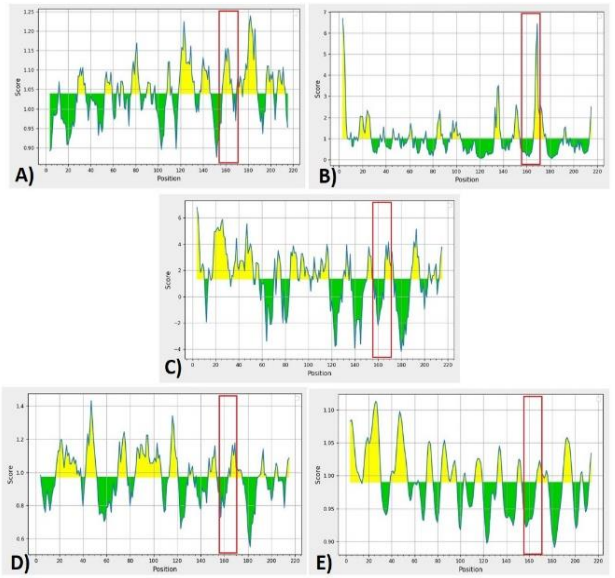

Fig. 2: Prediction of B-cell epitopes of VP2, representing the highly potent region as yellow peaks in graphs (A) Kolaskar and Tongaonkar antigenicity prediction (B) Emini surface accessibility prediction (C) Parker hydrophilicity prediction (D) Chou and Fasman beta-turn prediction (E) Karplus and Schulz flexibility prediction.

Prediction, assessment and selection of B-cell epitopes: The antigenic proteins were subjected to ABCPred for the detection of linear B-cell epitopes setting threshold 0.51 and 16-mer epitope length (Dhanda *et al.*, 2019). Its predictions are based on the recurrent neural network using fixed epitopes size. Total 19, 21, 23, and 21 B-cell epitopes were predicted from VP1, VP2, VP3, and $3C^{pro}$. Only 2 epitopes from VP2, 2 from VP1, and 5 from VP3 were short-listed, excluding all the non-antigenic, toxic, allergen, and intracellular epitopes. Furthermore, their conservancy was analyzed that is 100% for all the epitopes with the minimum identity of 87.50%.

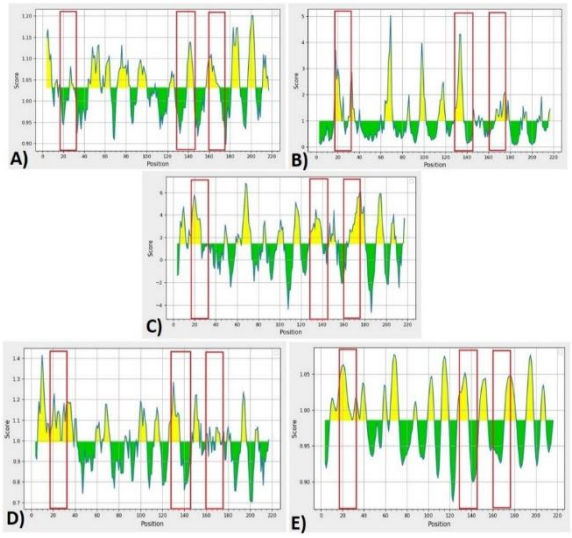

Fig. 3: Prediction of B-cell epitopes of VP3, representing the highly potent region as yellow peaks in graphs (A) Kolaskar and Tongaonkar antigenicity prediction (B) Emini surface accessibility prediction (C) Parker hydrophilicity prediction (D) Chou and Fasman beta-turn prediction (E) Karplus and Schulz flexibility prediction.

Table 1: Details of secondary structures of VP4, VP2, VP3, VP1, 2A, and $3C^{pro}$ determined by SOPMA server

Proteins	Alpha helix	Extended strand	Beta turn	Random coils
VP1	27.49%	20.38%	6.16%	45.97%
VP2	11.93%	30.28%	5.05%	52.75%
VP3	15.00%	29.55%	5.91%	49.55%
VP4	38.82%	4.71%	0.00%	56.47%
2A	38.89%	11.11%	22.22%	27.78%
$3C^{pro}$	17.37%	32.39%	10.80%	39.44%

B-cell epitopes prediction is not limited to informatics tools, but it must be evaluated through different eminent features as well. It includes beta-turn prediction, surface accessibility, flexibility, antigenicity, and hydrophilicity. The selected epitopes were further analyzed for these characteristics by IEDB Analysis Resource. Kolaskar and Tongaonkar antigenicity scale is used to estimate the antigenic tendency of proteins through the physicochemical properties and their relativeness in known epitopes. It calculated the antigenic score of VP1 as 1.037 (average), 0.887 (minimum) and 1.153 (maximum) (Fig. 1A), of VP2 as 1.039 (average), 0.876 (minimum) and 1.239 (maximum) (Fig. 2A) and of VP3 as 1.031 (average), 0.898 (minimum) and 1.201 (maximum) (Fig. 3A). The threshold values for all the antigenic proteins are equal to their average values. The higher antigenic score depicted that they could efficiently start the immune response against infection. For eliciting a better defensive response against the infection, the epitopes must be present in hydrophilic regions that are mostly found extracellularly. Therefore, Emini surface accessibility tool with the threshold value of 1.00 for all the proteins and Parker hydrophilicity prediction with a threshold value of 1.681 (VP1), 1.438 (VP3), and 1.331 (VP2) were utilized for the assessment of surface accessibility and hydrophilicity of B-cell epitopes, respectively. The calculated values for surface accessibility of VP1 were 1.000 (average), 0.086 (minimum) and 4.656 (maximum) (Fig. 1B), of VP2 were 1.000 (average), 0.057 (minimum) and 6.6.93 (maximum) (Fig. 2B) and of VP3 were 1.000 (average), 0.070 (minimum) and 5.023 (maximum) (Fig. 3B) while the calculated values for hydrophilicity of VP1 were 1.681 (average), -3.586 (minimum) and 6.271 (maximum) (Fig. 1C), of VP2 were 1.438 (average), -4.671 (minimum) and 6.829 (maximum) (Fig. 2C) and of VP3 were 1.331 (average), -4.186 (minimum) and 6.771 (maximum) (Fig. 3C). It is known that β-turns are hydrophilic and exposed on the cellular surface, so the predicted B-cell epitopes are reconfirmed and their structures are analyzed for the presence of beta turns by Chou and Fasman β-turns prediction. The output represented average (0.960) with minimum (0.707) and maximum (1.369) for VP1 (Fig. 1D), average (0.971) with minimum (0.550) and maximum (1.431) for VP2 (Fig. 2D), and average (0.994) with minimum (0.704) and maximum (1.414) for VP3 (Fig. 3D). Epitopes are significant only when they will be flexible enough to interact freely with the alleles or antibodies to stimulate an immune response. To estimate the elasticity of the epitopes, Karplus and Schulz flexibility analysis is performed and got 0.998 as both average and threshold with the minimum (0.913) and maximum (1.111) for VP1 (Fig. 1E), 0.990 as both average and threshold with the minimum (0.891) and maximum (1.113) for VP2 (Fig. 2E) and 0.985 as both average and threshold with the minimum (0.872) and maximum (1.077) for VP3 (Fig. 3E). Finally, the potential linear B-cell epitopes after screening by all the parameters were found to be PLLAIHPSEARHKQKI from VP1, TAHITVPFVGVNRYDQ from VP2 and TDPKTADPAY GKVFNP, GMEPPRTPEAAAHCIH and PYLSAADY AYTASDTA from VP3 (Table 2) and highlighted on the graphical representation of Chou and Fasman beta turns prediction, Kolaskar and Tongaonkar antigenicity prediction, Emini surface accessibility prediction, Parker hydrophilicity prediction, and Karplus and Schulz flexibility prediction of VP1 (Fig. 1), VP2 (Fig. 2) and VP3 (Fig. 3). Moreover, these linear epitopes were also confirmed by mapping them on their 3D structures and visualized by chimera (Fig. 4).

Table 2: Selection of linear B-cell epitopes based on epitope conservancy, allergenicity, antigenicity, transmembrane (TM) topology, and toxicity

Proteins	Epitope Sequence	ABCPred Score	Epitope Conservancy	Minimum Identity	Allergenicity	VaxiJen Score	TM Topology	Toxicity
VP2	TAHITVPFVGVNRYDQ	0.84	100.00% (100/100)	87.50%	Non-Allergen	0.5826 (Antigenic)	Outside	Non-Toxic
VP1	PLLAIHPSEARHKQKI	0.87	100.00% (100/100)	87.50%	Non-Allergen	0.8448 (Antigenic)	Outside	Non-Toxic
VP3	TDPKTADPAYGKVFNP	0.93	100.00% (100/100)	93.75%	Non-Allergen	0.5955 (Antigenic)	Outside	Non-Toxic
	PYLSAADYAYTASDTA	0.87	100.00% (100/100)	87.50%	Non-Allergen	0.7889 (Antigenic)	Outside	Non-Toxic
	GMEPPRTPEAAAHCIH	0.84	100.00% (100/100)	87.50%	Non-Allergen	0.5345 (Antigenic)	Outside	Non-Toxic

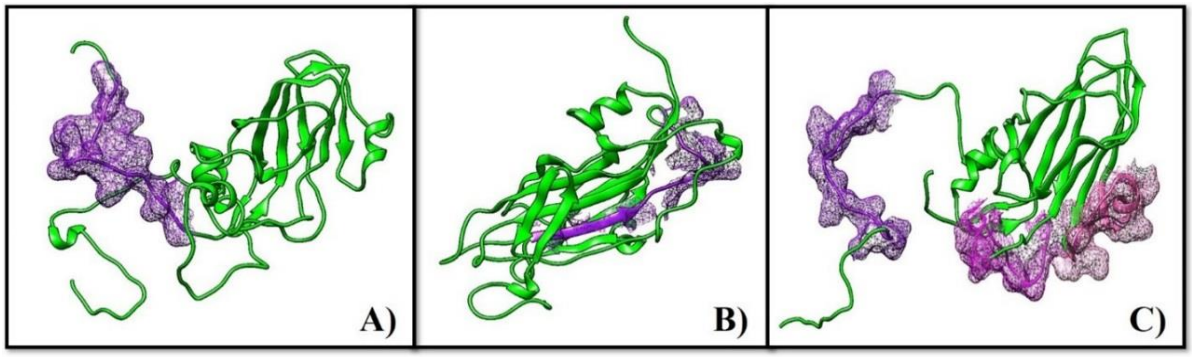

Fig. 4: Mapping of B-cell epitopes on 3D structures of (A) VP1, (B) VP2 and (C) VP3.

Prediction, assessment, and selection of T-cell epitopes: The CTL epitopes play a decisive role in eliciting long-lasting immunity by eliminating both the infected cells and the circulating virus. NetMHC pan 4.1 servers were utilized to predict the MHC class-I epitopes against the BoLA alleles of haplotypes A10, A13, A14, A18, A19, A20, and A31 (Reynisson *et al*., 2020). The 9-residue long peptide was selected for the prediction keeping the threshold value for strong binders as 0.5%. Total 23, 17, 17, and 18 T-cell epitopes from VP1, VP2, VP3, and $3C^{pro}$ were predicted, respectively. Only four epitopes were finalized that fulfilled the criteria of non-toxicity, high antigenicity, conservancy, non-allergenicity, and localized extracellularly. The epitopes are HKPWTLVVM from VP2, GESADPVTA from VP1, and YLSAADYAY and AGKDFELRL from VP3 (Table 3).

Structural modeling of CTL epitopes and BoLA alleles: The T-cell epitopes were submitted to the CABS-fold server to predict their structures under the de-novo modeling approach. The densest cluster with a low RMSD value among the resulted trajectories was chosen for further analyses. The stick model presentation of all the CTL peptides is shown in Fig. 5A. IPD-MHC database has complete information about BoLA alleles (Maccari *et al.*, 2018), so the protein sequences of BoLA-1:02101, BoLA-3:02701, BoLA-1:02301, and BoLA-2:01201 were retrieved from here and submitted to SWISS-MODEL for homology-based modeling. For modeling of BoLA-1:02101, the template with PDB ID 6at5.1.A having 74.37% sequence identity was selected. The model was assessed by a Ramachandran plot that showed 98.90% of residues in the favored region. The template used for structural modeling of BoLA-3:02701 was 6avg.2.D with a sequence identity of 72.93%. It showed 95.24% residues in the Ramachandran plot lied in a favorable region. The structure predicted for BoLA-1:02301 used 6avf.1.E as a template and Ramachandran plot analysis indicated 97.07% residues in the favorable region while 0.37% lied in Ramachandran outliers and 1.28% in rotamers outliers. The PDB ID 6avg.1.D was selected as a template for the modeling of the BoLA-2:01201 structure. The structure was confirmed as 96.69% amino acids fall in Ramachandran favorable region while only 0.74% fall in Ramachandran outliers and 0.43% in rotamer outliers. All the predicted structures were visualized by the UCSF chimera. The structures of BoLA alleles and their Ramachandran plots are shown in Fig. 5B.

Molecular interaction of T-cell epitopes and BoLA allele: Molecular docking is important during computer-aided drug design because it helps to predict the molecular interaction between different proteins and ligands before conducting *in vitro* and *in vivo* experimentations. Therefore, the blind docking between CTL epitopes and BoLA alleles was done using PATCHDOCK and the docked complexes were improved by FIREDOCK. The best-docked complex having the least global energy was chosen for interaction (Schneidman-Duhovny *et al.*, 2005). The complexes were visualized for the interaction by Discovery Studio and LigPlus (Fig. 6). The pictorial illustration of the complexes showed the strong binding affinity between the protein and the ligand due to the formation of both hydrogen bonding and hydrophobic interactions. The bond length of the hydrogen bond determines the binding strength between the protein and the ligand. To find the details of the interfaced residues LigPlot was used to visualize the interaction. It showed four hydrogen bonds Trp48-Asp5, Tyr48-Thr8, Thr70-Gly1, and Gly260-Glu2 with a bond length of 2.41Å, 1.95 Å, 2.44 Å, and 2.87 Å, respectively in GESADPVTA+ BoLA-1:02101 docked complex. BoLA-3:02701+ HKPWTLVVM complex formed two hydrogen bonds: between Glu227 and Trp4 at 2.7Å and between Glu228 and Trp4 at 2.93Å. The seven hydrogen bonds were observed during the molecular interaction of BoLA-1:02301 and YLSAADYAY: Tyr30 with Asp6 at 3.9Å, Arg83 with Tyr9 at 3.12Å, Arg83 with Tyr9 at 2.83Å, Glu84 with Tyr9 at 3.11Å, Tyr120 with Tyr9 at 3.11Å, Glu171 with Tyr1 at 2.19Å and Ser176 with Tyr1 at 3.01Å. Likewise, the seven hydrogen bonds were identified within the BoLA-2:01201+AGKDFELRL complex. The bonds lengths between Glu84 and Arg8 were 2.95Å, Arg86 and Glu6 were 2.44Å, Thr91 and Lys3 were 2.85Å, Trp168 and Gly2 were 2.68, Glu173 and Gly2 were 2.49, Arg176 and Asp4 were 2.63Å and between Thr184 and Leu7 were 2.93. These bonding showed the higher affinity of the latter two epitopes for BoLA alleles as compared to the former ones.

Designing and analyses of MESV construct: The five B-cell epitopes and four cytotoxic T-cell epitopes were linked through the GPGPG linker to design the MESV against FMD virus serotype O. To improve the immune response, CTB was engineered at the N-terminus of the construct using EAAAK linker. The MESV construct is 285 residues long as shown in Fig. 7A. The final vaccine construct was determined for its toxicity, allergenicity, and antigenicity.

Table 3: Selection of T cell epitopes based on allergenicity, antigenicity, toxicity, transmembrane (TM) topology, and epitope conservancy

Proteins	Epitope Sequence	MHC Class-I Alleles	VaxiJen Score	Allergenicity	Toxicity	TM Topology	Epitope Conservancy	Minimum Identity
VP2	HKPWTLVVM	BoLA-3:02701	0.8742 (Antigenic)	Non-Allergen	Non-Toxic	Outside	100.00% (100/100)	88.89%
VP1	GESADPVTA	BoLA-6:01401 BoLA-1:02101	0.5132 (Antigenic)	Non-Allergen	Non-Toxic	Outside	100.00% (100/100)	88.89%
VP3	YLSAADYAY	BoLA-1:02301 BoLA-2:02501 BoLA-2:02601 BoLA-3:02701 BoLA-2:02201	1.0049 (Antigenic)	Non-Allergen	Non-Toxic	Outside	100.00% (100/100)	88.89%
	AGKDFELRL	BoLA-2:01201	0.7581 (Antigenic)	Non-Allergen	Non-Toxic	Outside	100.00% (100/100)	88.89%

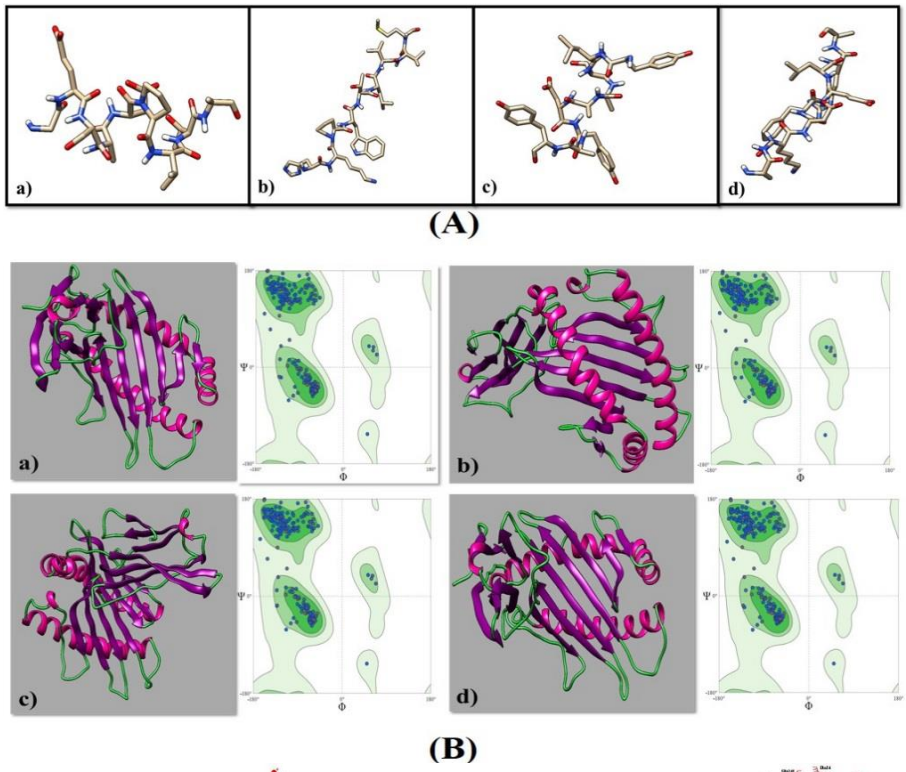

Fig. 5: Structures of T-cell epitopes and BoLA alleles (A) Stick representation of T-cell epitopes models (a) GESADPVTA (b) HKPWTLVVM (c) YLSAADYAY (d) AGKDFELRL. (B) Tertiary structures of BoLA alleles and their validation by Ramachandran plot (a) BoLA-1:02101 (b) BoLA-3:02701 (c) BoLA-1-02301 (d) BoLA-2:01201.

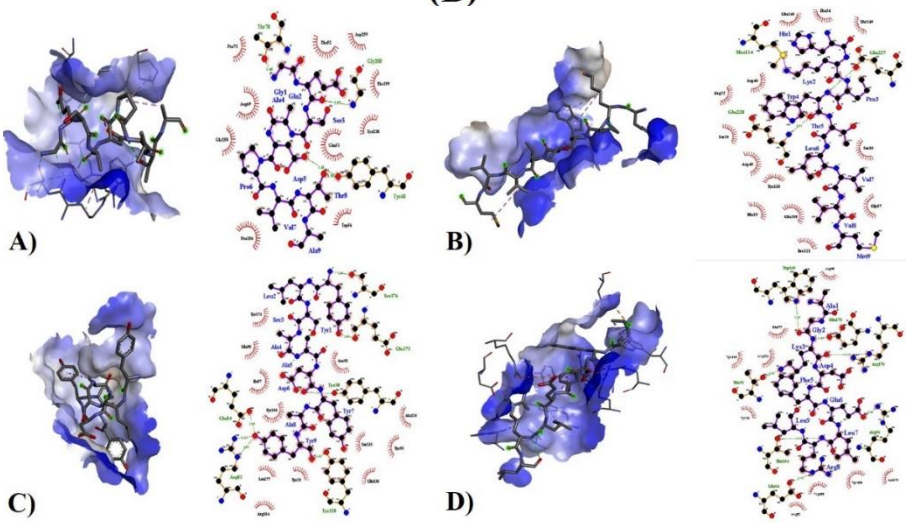

Fig. 6: Docked complexes visualized by Discovery Studio and LigPlus (A) BoLA-1:02101+ GESADPVTA (B) BoLA-3:02701+ HKPWTLVVM (C) BoLA-1-02301+ YLSAADYAY (D) BoLA-2:01201+AGKDFELRL. The ligand bonds are in purple, green dotted lines show hydrogen bonds and the sporked arcs depict the protein residues interaction with ligands through hydrophobic interactions.

The antigenicity score of the construct was calculated by VaxiJen and classified it as antigenic with a 0.5691 score. Similarly, AllerTOP and ToxinPred designated it as non-allergen and non-toxic, respectively. SOLpro categorized the MESV as soluble with a score of 0.762745. Then the physical and chemical properties of the MESV were analyzed. Its molecular weight and aliphatic index were calculated as 29.94kD and 67.58, respectively. The protein was predicted to be stable due to the 28.57 instability index. The average hydropathicity of MESV was found to be -0.333; its negative value presents that it is hydrophilic. Moreover, its half-life was found to be 30h in mammalian cells (*in vivo*) while more than 20h in yeast and above 10 hours in *E. coli* was observed during *in vitro*.

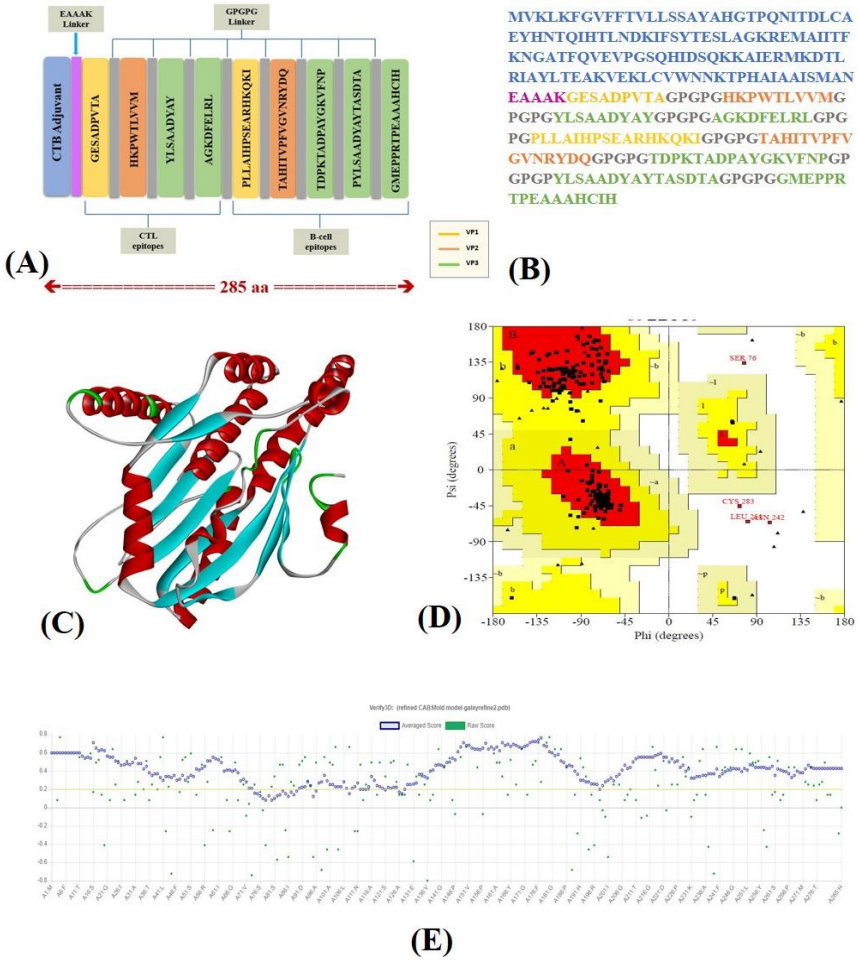

Fig. 7: Designing and 3D structure analysis of MESV construct against FMD virus serotype O. (A) Diagrammatic illustration of MESV construct. CTB adjuvant linked with the first CTL epitope by EAAAK linker (purple). CTL and B-cell epitopes linked with each other by GPGPG linkers (grey). (B) The protein sequence of MESV. (C) The 3D structure of the MESV construct. The quality of the model is evaluated by (D) Ramachandran plot and (E) Verify3D.

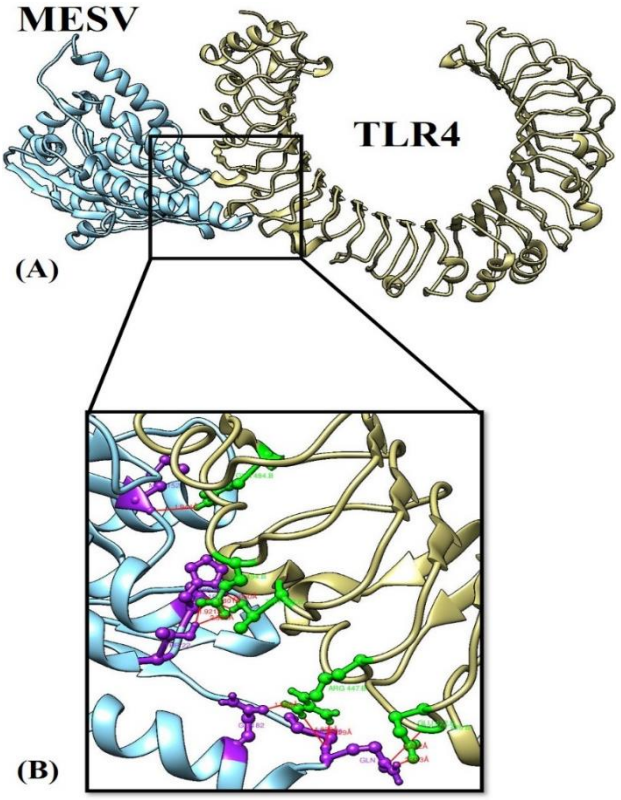

Fig. 8: Molecular docking of MESV and TLR4. (A) The docked complex of MESV+TLR4 where TLR4 colored in dark khaki and MESV colored in sky blue. (B) The interfaced amino acids of MESV (colored in purple) and TLR4 (colored in green) are involved in hydrogen bonding. The red line shows the hydrogen bond.

The secondary structure of the final vaccine construct was predicted using SOPMA, and it was found that 127 residues (44.56%) have participated in developing the random coils, while 12 (4.21%) in beta turns, 87 (30.53%) in alpha-helices, and 59 (20.70%) in extended strand formation. The tertiary structure of the MESV was formed through a CABS-fold server, refined by ModRefiner (Fig. 7B). The refined model had 0.758 RMSD with a TM score of 0.9864 revealing that the overall structure is good. The results of the Ramachandran plot developed by PROCHECK also validated the structure portraying 88.1% amino acid residues lied in its most favorable region, whereas 7.3, 3.7, and 0.9% residues lied in the additional allowed section, disallowed section, and generously allowed section, respectively (Fig. 7C). The structure was also examined by verify 3D during quality analysis. The results indicated 88.07% of residues have averaged $\geq$ 0.20 3D-1D score and its graphical plot is shown in Fig. 7D. The overall evaluation of the tertiary structure had shown that the quality of the refined model was good.

Molecular docking of MESV with bovine TLR4: It is imperative to investigate whether the designed vaccine will interact with the host immune receptors or not, the final vaccine was docked with the bovine TLR4 through ClusPro 2.0 (Kozakov *et al*., 2017). The docked complex having the maximum clustering members (46) and the least energy (-744.7kCal) was selected. UCSF Chimera was then used to visualize the interaction and 10 hydrogen bonds were observed in the complex. The interfaced amino acids of TLR4 and MESV during hydrogen bonding and their bond lengths are shown in Fig. 8.

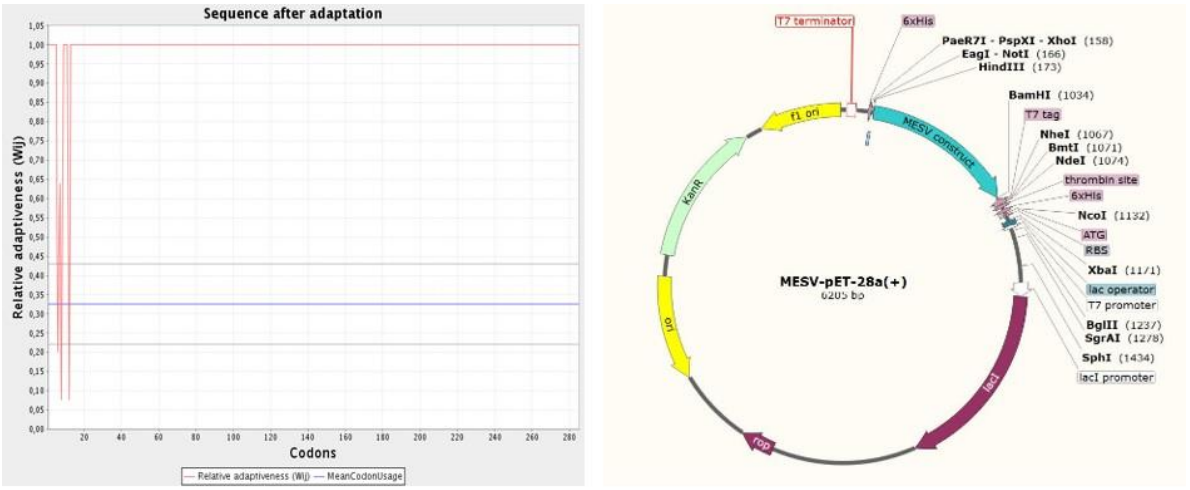

Fig. 9: (A) Graphical representation of the codon optimization of the final vaccine construct. (B) *In silico* cloning of optimized MESV construct in pET-28a(+) at *Bam*HI+*Hind*III.

***In silico* cloning in *E. coli*:** The expression of MESV against FMD virus serotype O was confirmed by performing *in silico* cloning in *E. coli* (strain K12) cells. The transgene expression in a selective host organism is related to its codon usage, it should be the same in both species. Hence, codons of MESV were optimized according to *E. coli* using the JCAT server (Fig. 8A) (Grote *et al.*, 2005). The improved MESV construct had 855 nucleotides with an ideal GC content of 55.90%, and a 0.97 CAI value. It depicted the probability of a high expression level of the target protein. Then for restriction-ligation cloning, the recognition sites of *Hind*III (AAGCTT) and *Bam*HI(GGATCC) were engineered at the N-terminal and C-terminal of the sequence, respectively. Finally, the refined MESV construct was cloned *in vitro* in pET-28a(+) at *Bam*HI+*Hind*III and got the recombinant vector of 6.205 kb (Fig. 9B).

DISCUSSION

FMD poses a continuously increasing global burden on the livestock industry despite eliminated from several developed countries. Although there are several vaccines available in the market, they fail to completely eradicate the disease. The major bottlenecks behind the failure of the current vaccines are the absence of cross-immunity among serotypes, mismatching between field and vaccine strain, cold chain requirement during storage and transportation, limited shelf-life, the need for boosters, and lack of DIVA capability (Lyons *et al*., 2016). Hence, to design a vaccine eliciting a broad-spectrum immunity is a highly demanding scientific challenge. Due to advancements in reverse vaccinology and bioinformatics tools, MESVs will be promising against the viral diseases in inducing strong defensive response as it comprises different immunodominant regions of the pathogen (Goumari *et al*., 2020). Therefore, the forecasting of highly antigenic epitopes is mandatory (Nain *et al*., 2020).

FMD virus serotype O is selected for the current study as it has been found responsible for more than 60% of all outbreaks in Pakistan. Its capsid encoding polyprotein (P12A) was selected for the current study as it contains several antigenic determinants that help in accelerating both humoral and cellular immunity through neutralizing antibodies (Liu *et al*., 2017). Also, $3C^{pro}$ was chosen due to its requirement for P1 processing into individual capsid proteins. For ideal vaccine development, the immunogenic profiling of candidate proteins is very important. After evaluating the allergenicity and antigenicity of the proteins using bioinformatics tools, only non-allergen and antigenic proteins among the curated proteins are subjected to immunoinformatics analyses to design a multiepitope-based vaccine against FMDV. An ideal MESV should have both CTL and B-cell epitopes to confer an efficient adaptive immune response (Ul Qamar *et al*., 2020). Thus, both types of epitopes were predicted and screened after immunogenic profiling. Finally, the four CTL epitopes and five B-cell epitopes were preferred for designing the vaccine. The selected B-cell epitopes were PLLAIHPSEARHKQKI from VP1, TAHITVPFVGVNR YDQ from VP2 and TDPKTADPA YGKVFNP, GMEPP RTPEAAAHCIH, and PYLSAADY AYTASDTA from VP3, whereas the CTL epitopes are HKPWTLVVM from VP2, GESADPVTA from VP1 and YLSAADYAY, and AGKDFELRL from VP3. The vaccine will be potent only if it stably interacts with the immune receptors of the host organism. Hence to evaluate the potency of CTL epitopes in our vaccine, the molecular interaction was studied between the selected CTL epitopes (ligands) and respective BoLA alleles (proteins). It depicted the distances of H-bonds in the range of 2 to 3Å that indicated the strong and stable binding affinities between the epitopes and the receptors (Fu *et al*., 2018) and will also help in stimulating an efficient immune response against the disease.

The use of the GPGPG linker for joining the epitopes in MESV maintained the immunogenic behavior of epitopes due to beta-turns formed by “G” and “P” residues. CTB was engineered in MESV as it is a strong anti-microbial agent and evokes mucosal immunity by binding with GM1 ganglioside receptors on the mucosal epithelium. The final vaccine was found non-allergen, non-toxic, and immunogenic. The 285 amino acids long chimeric vaccine was found to be highly stable due to an instability index less than 40 i.e., 28.57, and hydrophilic

with a GRAVY value of -0.333. The secondary structure indicated the presence of 30.53% alpha-helices, 20.70% extended strands, 4.21% beta-turns, and 44.56% coils in the multiepitope vaccine. The TLRs are the highly conserved proteins that detect the pathogen-associated molecular patterns (PAMPs) on the pathogen and play a major role in inducing an innate immune response. The molecular interaction between MESV and TLR4 has shown 10 hydrogen bonds in the docked complex, indicating that the vaccine will be efficiently recognized by TLR4 and will provoke a strong immune response in the host organism against FMD. Mostly the bacterial system is used for protein expression analysis because it is easy to culture and yield a high level of recombinant protein. For attaining high expression, the codon usage of foreign genes and host must be compatible. That is why the codon-optimized MESV coding nucleotide sequence was cloned *in vitro* in pET-28a(+) by restriction-ligation cloning. This vector is used due to the His-tag so that the vaccine could be easily purified.

Conclusions: FMD has been a devastating disease for decades and is now considered a global burden for the livestock industry. Despite utilizing inactivated and attenuated vaccines, the disease is still endemic in various countries due to high antigenic diversity among its serotypes and even among the subtypes of a single serotype. Thus, we constructed a chimeric subunit vaccine against FMD virus serotype O utilizing immunogenic epitopes from the antigenic proteins through bioinformatics tools and reverse vaccinology. The vaccine proposed here encourages further vaccine development against other serotypes of the FMD virus. Hence, it needs both *in vitro* and *in vivo* experimental confirmations using model animals to validate the efficiency of MESV.

Author contributions: MSK conceived and designed the project and the research work is conducted by RR under the supervision of MSK. The manuscript is prepared by RR and critically reviewed by MSK, FAJ, and MAZ.

Acknowledgments: The authors are grateful to PARB (Pakistan Agricultural Research Board) and Higher Education Commission (HEC) for providing the funds to MSK.

REFERENCES

Abubakar M, Kanwal S and Saeed A, 2012. Persistence, Emergence and Distribution of Foot and Mouth Disease Virus (FMDV); Global and Pakistan Perspectives. Pak J Life Soc Sci 10:84-90.

Azam A and Shafique M, 2017. Agriculture in Pakistan and its impact on economy-A Review. Int J Adv Sci Technol 103:47-60.

Brito BP, Perez AM, Jamal SM, *et al*., 2013. Foot-and-mouth disease virus serotype O phylodynamics: genetic variability associated with epidemiological factors in Pakistan. Transbound Emerg Dis 60:516-24.

Dhanda SK, Mahajan S, Paul S, *et al*., 2019. IEDB-AR: immune epitope database-analysis resource in 2019. Nucleic Acids Res 47:W502-6.

Dimitrov I, Bangov I, Flower DR, *et al*., 2014. AllerTOP v. 2-a server for in silico prediction of allergens. J Mol Model 20:2278-83.

Fu Y, Zhao J and Chen Z, 2018. Insights into the molecular mechanisms of protein-ligand interactions by molecular docking and molecular dynamics simulation: a case of oligopeptide binding protein. Comput Math Methods Med 2018:1-12.

Gasteiger E, Gattiker A, Hoogland C, *et al*., 2013. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784-8.

Goumari MM, Farhani I, Nezafat N, *et al*., 2020. Multi-Epitope Vaccines (MEVs), as a novel strategy against infectious diseases. Curr Proteomics 17:354-64.

Grote A, Hiller K, Scheer M, *et al*., 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 33:W526-31.

Knight-Jones TJD and Rushton J, 2013. The economic impacts of foot and mouth disease – What are they, how big are they and where do they occur? Prev Vet Med 112:161-73.

Kozakov D, Hall DR, Xia B, *et al*., 2017. The ClusPro web server for protein–protein docking. Nat Protoc 12:255-78.

Liu X, Lv J, Fang Y, *et al*., 2017. Expression and immunogenicity of two recombinant fusion proteins comprising foot-and-mouth disease virus structural protein VP1 and DC-SIGN-Binding glycoproteins. Biomed Res Int 2017:1-9.

Lyons NA, Lyoo YS, King DP, *et al*., 2016. Challenges of generating and maintaining protective vaccine-induced immune responses for foot-and-mouth disease virus in Pigs. Front Vet Sci 3:102-13.

Maccari G, Robinson J, Bontrop RE, *et al*., 2018. IPD-MHC: nomenclature requirements for the non-human major histocompatibility complex in the next-generation sequencing era. Immunogenetics 70:619-23.

Mahdy SE, Sijing L, Lin S, *et al*., 2019. Development of a recombinant vaccine against foot and mouth disease utilizing mutant attenuated *Listeria ivanovii* strain as a live vector J Virol Methods 273:113722-30.

Meza B, Ascencio F, Sierra-Beltran AP, *et al*., 2017. A novel design of a multi-antigenic, multistage and multi-epitope vaccine against Helicobacter pylori: An *in silico* approach. Infect Genet Evol 49:309-17.

Mignaqui AC, Ruiz V, Durocher Y, *et al*., 2019. Advances in novel vaccines for foot and mouth disease: focus on recombinant empty capsids. Crit Rev Biotechnol 39:306-20.

Nain Z, Karim MM, Sen MK, *et al*., 2020. Structural basis and designing of peptide vaccine using PE-PGRS family protein of *Mycobacterium ulcerans*-an integrated vaccinomics approach. Mol Immunol 120:146-63.

Pattnaik B, Subramaniam S, Sanyal A, *et al*., 2012. Foot-and-mouth disease: global status and future road map for control and prevention in India. Agric Res 1:132-47.

Schneidman-Duhovny D, Inbar Y, Nussinov R, *et al*., 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 33:W363-7.

Truong DB, Goutard FL, Bertagnoli S, *et al*., 2018. Benefit-Cost analysis of foot-and-mouth disease vaccination at the farm-level in South Vietnam. Front Vet Sci 5:26-36.

Ul Qamar MT, Shahid F, Aslam S, *et al*., 2020. Reverse vaccinology assisted designing of multiepitope-based subunit vaccine against SARS-CoV-2. Infect Dis Pover 9:132-45.

Waheed U, Parida S, Khan QM, *et al*., 2011. Molecular characterisation of foot-and-mouth disease viruses from Pakistan, 2005–2008. Transbound Emerg Dis 58:166-72.