**Journal of Animal & Plant Sciences, 35(4): 2025, Page: 1020-1036**
**ISSN (print): 1018-7081; ISSN (online): 2309-8694**
https://doi.org/10.36899/JAPS.2025.4.0087

# IMMUNOINFORMATICS-DRIVEN MULTI-EPITOPE VACCINE DESIGN FOR FOOT-AND-MOUTH DISEASE VIRUS

A. Alamri[1#], S. Imran[2#], M. A. Abdel-Maksoud[3], S. Almutairi[4], H. Ebaid[5], B. H. Kiani[6], Q. Akram[7], T. Hussain[2] and M. A. Naeem[8*]

[1] Biochemistry Department, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
[2]Department of Biotechnology, Virtual University, Islamabad, Pakistan
[3] Botany and Microbiology department, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
[4]Botany and Microbiology department, College of Science, King Saud University, Saudi Arabia
[5]Department of Zoology, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia.
[6]Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, Massachuesetts, 01609, USA
[7]Department of Pathobiology (Microbiology), University of Veterinary and Animal Sciences, Lahore (Narowal Campus) Narowal 51600, Pakistan
[*8]Department of Basic Sciences (Pharmacology), University of Veterinary and Animal Sciences, Lahore (Narowal Campus) Narowal 51600, Pakistan
[#]Equal Contributor
*Corresponding Author's email:muhammadahsannaeem@ymail.com

## ABSTRACT

Foot-and-mouth disease (FMD) is the most destructive for livestock with reservoirs in cattle, buffalo, sheep, goats, and pigs and foot-and-mouth disease virus (FMDV) with seven serotypes. Current vaccination strategies suffer from difficulties with antigenic variability coupled with high costs. This paper aims to design a multi-epitope subunit vaccine against FMDV by the use of immunoinformatics approach in order to improve effectiveness. We used immunoinformatics to design a subunit vaccine that included two T-cell epitopes linked with AAV and three B-cell epitopes linked with KK. T- and B-cell epitopes were joined by a GPSL linker. A Pan HLA-DR binding epitope, PADRE, was attached at both ends using EAAAK linkers. Physicochemical properties, allergenicity, and antigenicity of the vaccine were evaluated, along with secondary and tertiary structure predictions and molecular docking studies with the Toll-like receptor 9 (TLR-9). The vaccine had a predicted to be non-allergic and with high antigenic property (0.73). Physicochemical analysis showed to be 135 amino acids, stable (21.29 stability index), and basic (pI of 10.51). The overall 3D structure showed robust binding affinity against the cattle TLR-9 receptor. It was confirmed, in silico cloning, effective transformation into prokaryotic expression vector pET-28a (+).The subunit vaccine developed based on immunoinformatics has great promise in the form of a high antigenicity level, stabilized physicochemical properties, and interactions well-favored with TLR-9. This indicates potential further experimental validation of effectiveness as a candidate vaccine against FMD.

**Keywords:** FMDV,subunit vaccine, immunoinformatics, non-structural protein

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Published first online June 10, 2025 Published final July 29, 2025

## INTRODUCTION

Foot and Mouth Disease (FMD) is a highly economically impactful transboundary ailment that affects cloven-hoofed animals, including but not limited to cattle, buffalo, sheep, goats, and pigs. The virus causing this disease is referred to as the FMD virus (FMDV). This virus has been classified within the genus Aphthovirus within the family Picornaviridae (Raoof, 2024). FMD is characterized by the development of vesicles or blisters at any part of the animals' bodies, including the hooves, mouth, nose, and feet (Raoof, 2024). These vesicles often cause erosion of their skin, which results in weakness and inability to feed (Prasad *et al*., 2024). FMDV is a small, non-enveloped, positive, and single-stranded RNA virus that is covered by an icosahedral capsid composed of four structural proteins: VP1, VP2, VP3, and VP4 (Islam, 2024). The virus has 7 serotypes based on immunological characteristics, which include O, A, C, Asia 1, and Southern African Territories (SAT 1, SAT 2, and SAT 3), of which each of them has different genetically and epidemiologically variable

subtypes (Ali *et al*., 2024). These serotypes are spread all over the world, though some serotypes are predominantly observed in certain regions. For instance, only SAT1, SAT2, and SAT3 serotypes are seen in Africa, while the Asia1 serotype is predominant in Asia, sometimes including the Near East and parts of Eurasia (Chestley*et al*., 2022). Immunization against or exposure to one serotype does not result in immunity to other serotypes (Kitching*et al*., 2007). This leads to a problem in disease control, particularly in areas where the serotypes overlap. FMD pandemics in extensive regions of Asia, Africa, the Middle East, and South America. Serotypes O and A are the most common (Prempeh*et al*., 2001). The disease prevails in more than 100 countries, and hence, there is a high risk of spread to FMD-free countries. The economic impact of FMD is very significant, with estimated losses ranging between US$ 6.5 and US$ 21 billion globally (Layessa*et al*., 2024). These losses arise from reduced milk production, poorer livestock growth, deaths of young stock, and a decline in traction power in regions that depend on draft animals, among others (Arzt*et al*., 2011). Another outcome of FMD outbreaks can be abortion of livestock, and this eventually adds up to losses for the farmer economically (Admassu*et al*., 2015). Control of FMD efforts is mainly effective quarantine, immunization, cleanliness, and sanitation (Ndeevelo, 2023). Most FMD vaccines are commercially available and are usually inactivated with chemicals after production to kill the virus (Schneider *et al*., 2014). These are available in aluminum-based and oil-emulsion-based preparations with monovalent, bivalent, and multivalent structures (Kamel*et al*., 2019). Inactivated vaccines have been reasonably good in limiting FMD epidemics around the world and delivering immunity for at least a period of one year (Badr*et al*., 2024). The traditional vaccines, however, suffer from several shortcomings: the preparation needs to be refrigerated, booster doses are periodic, and they deliver immunity against only the homologous type (Lombard, 2012). This has led to further investigation into alternative vaccines in recombinant viral vaccines, DNA vaccines, peptide vaccines, and subunit vaccines. Recent advances in bioinformatics, genetic research, and recombinant DNA technology have provided an opportunity to develop broad-spectrum vaccines that overcome these deficits (Choudhury *et al*., 2021). For example, DNA-based vaccines can increase the expression levels of antigens as well as elicit both cellular and humoral immune responses (Medina & Diaz San Segundo, 2024). The concept of building new vaccines through the use of immunoinformatics is a new area that presents more approaches towards developing vaccines against FMDV strains, which commercialized vaccines fail to immunize against (Gunasekera*et al*., 2017). These vaccines employ conserved epitopes, thereby prohibiting the continuous production of new FMDV strains (Ribeiro *et al*., 2024). After which, the antigenicity, allergenicity, and immunological characteristics of said vaccines are computed (Manvani*et al*., 2024). To give more impetus to the development, reverse translation and codon optimization are added on to the nucleotide sequence of the vaccine protein to be generated in an in vitro environment for efficacy tests (Akter*et al*., 2024). The gains concerning control of FMD have primarily been made through the contributions of traditional vaccines. However, alternative vaccines currently in development will likely introduce an improved level of more effective control for the disease (Medina and Diaz San Segundo, 2024). This would not be possible without further research into bioinformatics and genetics combined with the methodology of recombinant DNA techniques (Waseem *et al*., 2024). This then forms a basis for further research into the in vitro expression of this protein and then efficacy studies in intact animals.

## MATERIALS AND METHODS

**Collection of B and T cell epitopes:** Using a systematic search by the Picornaviridae database (www.viprbrc.org), B and T cell epitopes from structural and non-structural proteins of FMDV were identified. Data filtered through "Cow" hosts and MHC classes I/II are presented as: IEDB ID, Epitope Sequence, Protein name, Assay results, MHC alleles; this helps identify immunogenic epitopes for vaccines (Table 1).

**Construction of a multiple-epitope vaccine sequence:** A systematic immunoinformatics approach combined several computational tools towards the design of a multi-epitope vaccine sequence by seeking to enhance its immunogenicity and potency. Five epitopes were chosen based on their predicted antigenicity and immunogenicity. For the augmentation of the construct of the vaccine, the PADRE sequence was incorporated both at the N-terminal and C-terminal ends of the sequence. The B-cell epitopes were joined through a KK linker, while T-cell epitopes were connected through the AAY linkers. Moreover, to facilitate the interplay of B and T-cell epitopes, the GPSL linker was used. The EAAAK linkers were further added on both sides to stabilize the N-terminal PADRE sequence and enhance its immunogenic response, while a GGGS linker was added at the C-terminal of the PADRE sequence followed by a sequence to bring about flexibility in the construct. This design strategy promotes both effective immunogenic responses to epitopes and proper space between epitopes that should enhance the individual immunogenic properties of the epitopes. With the vaccine construct assembled, it then underwent further rounds of bioinformatics analysis to assess its physicochemical properties, allergenicity, and potential toxicity; this

means that the final vaccine candidate ensures safety and efficacy for validation studies.

**Table 1:Immunogenic B and T Cells'Epitopes of FMDV Structural and Non-Structural Proteins.**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **For Structural Proteins** | | | | | | | |
| **IEDB ID** | **Epitope Sequence** | **Protein Name** | **Assay type category** | **Assay result** | **MHC allele name and class** | **Method** | **Measurement** |
| 14516 | ETQIQRRQHTDV SFIMDRFV | VP1 | T Cell | +ve high | BoLA class II | 3H-Thymidine | Proliferation |
| 42475 | MRKTKLAPTVA H GVF | VP4 | T Cell | +ve | BoLA class II | 3H-Thymidine | Proliferation |
| | | | B Cell | N/A | N/A | ELISA | Qualitative Binding |
| **For Non-Structural Proteins** | | | | | | | |
| 2B | O1K | PFFFSDVRSNFSKLV | | | B cell | Hohlich BJ *et al*., 2003 | |
| 2C | O1K | LKARDINDIFAILKN | | | B cell | -do- | |
| 3B | O1K | PVKKPVALKVKAKN | | | B cell | -do- | |

**Physiochemical Characteristics:** The bioinformatics tool ProtParam (http://web.expasy.org/protparam/) was applied to analyze the physicochemical properties of a protein. These parameters included the molecular mass, theoretical isoelectric point, amino acid composition, aliphatic index indicating thermal stability, GRAVY score indicating hydrophobicity, and instability index predicting the stability of a protein and its half-life in vivo/in vitro (Gasteiger*et al*., 2005).

**Prediction of allergenicity of vaccine:** The predictability of allergenicity of the multi-epitope vaccine was done through the combination of web tools AlgPred (http://www.imtech.res.in/raghava/algpred/) with the website AllerCatPro (https://allercatpro.bii.a-star.edu.sg), which are generally recognized predictors in bioinformatics (Maurer-Stroh *et al*., 2019). A hybrid algorithm was used for the analysis, based on AllerCatPro version 1.7 and AllergenFP v.1.0 (http://www.ddg-pharmfac.net/AllergenFP/), and designed to predict the potential allergenicity of protein sequences. The accuracy reached by the prediction process was 85% at a threshold of 0.4, considered to be a high level of precision for the prediction of allergenic proteins. This hybrid approach applied here consisted of six different well-established prediction methods in the field, renowned for their reliability and known to detect allergenic characteristics within sequences. Additionally, the addition of these tools has enabled full assessment and re-validation of all allergenicity predictions, ensuring that safer multi-epitope vaccines are produced.

**Prediction of Antigenicity Vaccine:** ANTIGENpro(http://scratch.proteomics.ics.uci.edu/) is a computational tool that predicts protein antigenicity using a sequence-based, alignment-free method that depends on neither the pathogen involved. The protocol to predict antigen works in two stages and utilizes multiple representations in combination with 5 variants of machine learning algorithms derived from primary protein sequences. First and foremost, it achieves two types of datasets related to known protective antigens and non-antigens for training the predictive model. It generates multiple sequence-based features for a set of protein sequences and could include physicochemical properties and structural information derived from high-throughput protein microarray data, making its ground strong for predictions that should reflect significant accuracy. Then, the application extracts sequence-based features into features using the machine learning framework as input through classification techniques against distinguishing between antigenic and non-antigenic proteins. The second iteration refines this further using highly complex algorithms, which ultimately translate to improving the generalization capabilities of the model onto various datasets and improving predictive performance on data not yet seen. Cross-validation method evaluation of ANTIGENpro is done with highly robust cross-validation methods that assess the classification accuracy of protective antigens against a backdrop of negative examples. This cross-validation process shows that ANTIGENpro is able to reach a cross-validation accuracy of about 76% on external pathogen proteomes, thus making it potentially useful for vaccine development and immunogenicity studies. It is part of the SCRATCH suite of predictors, accessible to researchers as a tool to perform ease protein antigenicity analyses.

**Prediction of secondary protein structure:** The methodology used PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) to predict secondary structures of proteins. The PSIPRED 3.2 results achieved

a Q3 score accuracy of 81.6%. PSIPRED also identified the homologous sequences, which were then used to build a PSSM.

**Prediction of tertiary protein structure:** RaptorX (http://raptorx.uchicago.edu/documentation/#goto2), a web-based protein structure prediction server, processed the secondary/tertiary structures, solvent exposure, disordered regions, and binding sites of the multi-epitope vaccine. Probabilistic graphical models were used to present GDT scores, un-normalized GDT scores, and P-values for model quality. Processing took about ~35 minutes for sequences <200 amino acids. 3D models were visualized by DiscoveryStudio. Briefly, 3D protein structure model (4uz0A-583378_1) refinement protocol, originally obtained from RaptorX, employed GalaxyRefine to generate five best-refined models. GalaxyRefine employs a side-chain repacking and molecular dynamics simulation refinement protocol that has been demonstrated to perform in CASP experiments. The target of this protocol is to improve the quality of both local and global structure by correcting predicted model mistakes. To validate the best-refined model, different metrics were used: GDT-HA (Global Distance Test - High Accuracy), RMSD (Root Mean Square Deviation), MolProbity Score, Clash Score, Poor Rotamers, and the percentage of residues in favored regions of the Ramachandran plot (Ramachandran Favored). To select the best-refined model from GalaxyRefine, different key metrics were evaluated: GDT-HA (Global Distance Test - High Accuracy), RMSD (Root Mean Square Deviation), MolProbity Score, Clash Score, Poor Rotamers, and the percentage of residues in favored regions of the Ramachandran plot.

**Vaccine's molecular docking with the receptor (TLR9):** The TLR9 receptor (PDB ID: 5Y3M) had its predicted binding pockets via CASTp http://sts.bioe.uic.edu/castp/ using a 1.4 Å probe radius with the theory of alpha-shapes. HDOCK executed the protein-protein docking protocol through atomic form description, surface fix coordination, and scoring.To evaluate the best model from HDOCK docking results, we considered factors such as docking score, confidence score, ligand RMSD, and interface residues. The methodology involved prioritizing models based on these metrics to identify a balance between binding strength and structural stability. This approach allowed us to systematically compare each model's performance.

**Reverse translation, optimization of the codon, and *in-silico* replication of the final vaccine construct:** Using a web-based tool (http://www.bioinformatics.org/sms2/rev_trans.html) for eukaryotic expression, the protein sequence was reverse-translated into a nucleotide sequence. For expression in Escherichia coli, codon optimization using JCAT (http://www.jcat.de) (Grote *et al*., 2005) was done in addition to incorporating a prokaryotic ribosome binding site, restriction enzyme cleavage sites, and rho-independent transcription termination. The optimized sequence was in silico cloned into pET-28a(+) using SnapGene with the addition of XhoI and NdeI restriction sites at the C- and N-terminals, respectively, for compatibility with the vector for efficient cloning and protein expression.

## RESULTS

**Epitopes collection and subunit vaccine designing:** The epitopes corresponding to both T and B cells for structural and non-structural Foot-and-Mouth Disease (FMD) proteins were compiled from IEDB tool (Table 1). Among these, 1 epitope were identified for the VP1 structural protein, and 1 for VP4. In contrast, 3 epitopes were associated with non-structural proteins, comprising 1 for 2B, 1 for 2C, and 1 for 3B. These epitopes were then amalgamated using an immunoinformatics approach to formulate a subunit vaccine construct(Fig. 1).

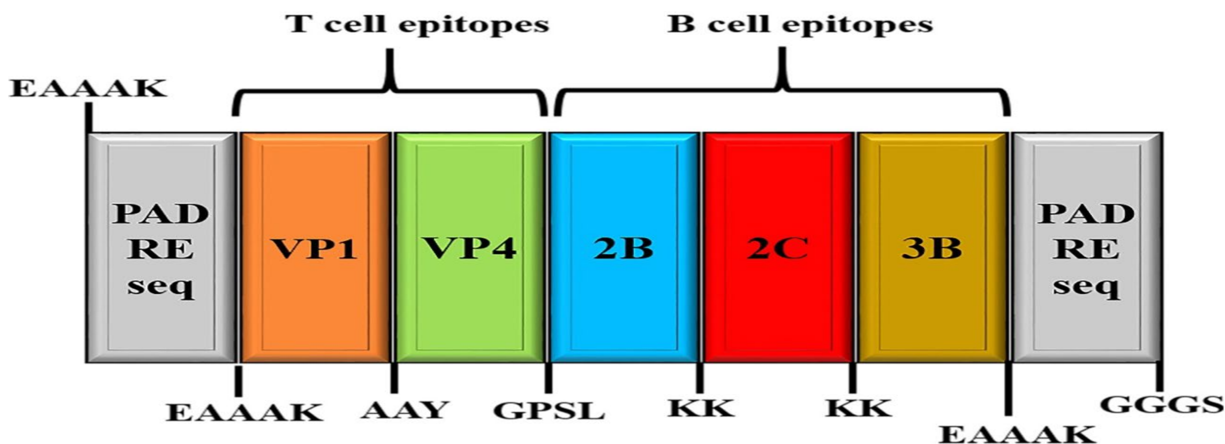


**Fig. 1: A representation of the completed subunit vaccine design**

Cytotoxic T-lymphocytes (CTLs) represent a CD8+ subset of T-cell responses responsible for eliminating intracellularly infected target cells by various pathogens such as viruses, bacteria, or protozoa (Kaufmann, 1988). Upon encountering the Major Histocompatibility Complex class I (MHC-I) presented antigen specific to their receptor during infection, CTLs undergo cell cycle activation, undergo multiple mitosis divisions, and eventually differentiate into effector cells (Pandey *et al*., 2018). Helper T-lymphocytes (HTLs) constitute a key component of both humoral and cell-mediated immune responses (Millan *et al*., 1998). Hence, it is probable that epitopes specific to HTL receptors will hold significant importance in the context of preventive and immunotherapeutic vaccination (Pandey *et al*., 2018). A vital element of humoral immunity is the B-cell response. Following antibody development, the epitope corresponding to the B-cell receptor is pivotal for vaccine design (Xu and Kulp, 2019). The vaccine's core sequence comprises several epitopes spanning 135 amino acid residues; among them, the initial 06-18 and last 119-131 residues are denoted as PADRE sequences. After the first PADRE sequence, T-cell epitopes (24-61 amino acids) and B-cell epitopes (66-113 amino acids) are positioned. The PADRE sequence and B-cell epitope are connected by the EAAAK linker, while B-cell epitopes are linked by the KK linker. The final B-cell epitope and initial T-cell epitope are linked by the GPSL linker, and the T-cell epitopes are connected by AAY linkers.

**Allergenicity Prediction of Vaccine:** The prediction of allergenicity by AlgPred shows that our SV does not have any property that is related to being an allergen, because the score we obtained was -0.69, which is far below the threshold value required to qualify the protein as being an allergen; it is more than below -0.4. The score is even so low that there's a tendency for having an allergy from this protein due to the algorithm used by AlgPred, which combines multiple approaches, like SVM and motif-based algorithms, in order to predict the allergenic potential. This dataset, consisting of 10,075 allergens and an equal count of 10,075 non-allergens, offers better robustness to the model due to proper training and validation. The characteristics of our SV features as deduced from various features extracted from AlgPred (Table 2).

**Table 2. Comprehensive Allergenicity Assessment of the Vaccine Candidate via AlgPred.**

| Mapping of IgE epitope | Contain No experimentally proven IgE epitope |
|---|---|
| MAST result | Non-Allergen |
| SVM method based on amino acid composition | Allergen<br>(-0.24802341) Minimal value= -0.4<br>+ve prediction value= 47.13 %<br>-ve Prediction value= 89.71 % |
| SVM method based on dipeptide composition | Allergen (0.094) Threshold=-0.2<br>Positive predictive value= 63.1%<br>Negative Predictive value= 85.56% |
| BLAST result of ARPs | Non-Allergen |
| Hybrid Approach | Non-Allergen |

These features include amino acid composition, dipeptide frequencies, and motifs specific to allergenicity. These features were integrated quite successfully into the predictive models that gained high accuracy; for example, the area under the receiver operating characteristic curve obtained on its validation set was 0.98 while the Matthew's correlation coefficient achieved by AlgPred 2.0 was 0.85. These metrics highlight the good promise of bioinformatics tools for leading to effective identification of proteins that are unlikely to provoke allergic reactions based on structural and compositional properties. Overall, the import of our AlgPred predictions is that our SV is non-allergenic, and its quantification could prove critical in its applications in the realm of food safety and therapeutic developments.

**Antigenicity Prediction of Vaccine:** In doing the analysis on ANTIGENpro, the score obtained for the SV protein is 0.729693, which implies a high possibility that the antigen may be present in this protein. Indeed, such a score far exceeds the threshold value established at 0.4; therefore, it is very likely that this protein carries antigens and the possibility of developing an immune response to the administration of this protein, thus making it a good candidate for further exploration in immunology studies. Immunogenic potential The vaccine scored high antigenicity with a score of 0.73, above the threshold of 0.4, meaning it has a good chance of inducing an immune response.Predictive metrics that corroborate the conclusion reached with the score from ANTIGENpro are also represented in the table: hydrophilicity, surface accessibility, and epitope prediction. These bioinformatics analyses generally support the characterization of the SV protein as an antigen, which is further justified for study in vaccine development or therapeutic applications.

**Physiochemical Properties and Domain Identification:** The use of the ProtParam tool yields many important physico-chemical properties in the outcome of the protein analysis (Table 3).

**Table 3. ProtParam Computed Physico-chemical Characteristics of the Subunit Vaccine.**

| # of Amino acids | 135 |
|---|---|
| Molecular Wt. | 146.5 kDa |
| Theoretical PI | 10.51 (Basic) |
| Total negatively charged residues (Asp + Glu) | 9 |
| Total positively charged residues (Arg + Lys) | 27 |
| Formula | $C_{673}H_{1087}N_{187}O_{174}S_2$ |
| Extinction coefficient (280nm) $M^{-1}cm^{-1}$ (relative to H2O) | 12490 |
| Abs 0.1% (=1 g/l), assuming cysteine is formed from all cys-residues | 0.853 |
| Estimated half-life (mammalian reticulocytes) | 1 hour (in-vitro) |
| Yeasts | >30 min (in-vivo) |
| E. coli | > 10-hour (in-living system) |
| Instability index (II) | 21.29, Stable |
| Aliphatic index | 82.67 |
| Grand average of hydropathicity (GRAVY) | -0.076 |

It contains 135 amino acids and has a molecular weight of 146.5 kDa, thus showing that the size of the protein was relatively large. Its theoretical isoelectric point pI was calculated to be 10.51, and it was therefore classified as a basic protein. From the charge distribution, 9 total negatively charged residues are aspartic acid and glutamic acid, while 27 total positively charged residues are arginine and lysine, so the total charge will be strong positive. The chemical formula deduced from the amino acid composition is C673H1087N187O174S2, since the protein has a complex structure. The extinction coefficient at 280 nm is listed as 12490 $M^{-1}cm^{-1}$ and will be useful for estimating protein concentration in solution. The absorbance at 0.1% (1 g/l) concentration, assuming all cysteine residues are reduced, was determined to be 0.853. Regarding stability, the half-life of this protein has been estimated at very different values in various organisms: about 1 hour in mammalian reticulocytes, over 30 minutes in yeasts, and more than 10 hours in E. coli, indicating that the protein is better stabilized in the prokaryotes than in the eukaryotes. The II is 21.29 and, therefore, indicates a stable protein. The aliphatic index reads at 82.67, which indicates that the stability of the protein is due to a very large degree of hydrophobicity in any environment. The last is the grand average of hydropathicity (GRAVY) score, which is -0.076, meaning that the protein has a generally slightly hydrophilic nature as well and should influence such behavior toward interactions within cellular environments.

**Secondary Structure Prediction:** The online tool for bioinformatics, PSIPRED, was used to make a prediction of the secondary structures of the SV protein (Fig. 2). The results of this tool were categorized into three structural elements: alpha helices, beta strands, and coils, each element in different colors for proper visualization.

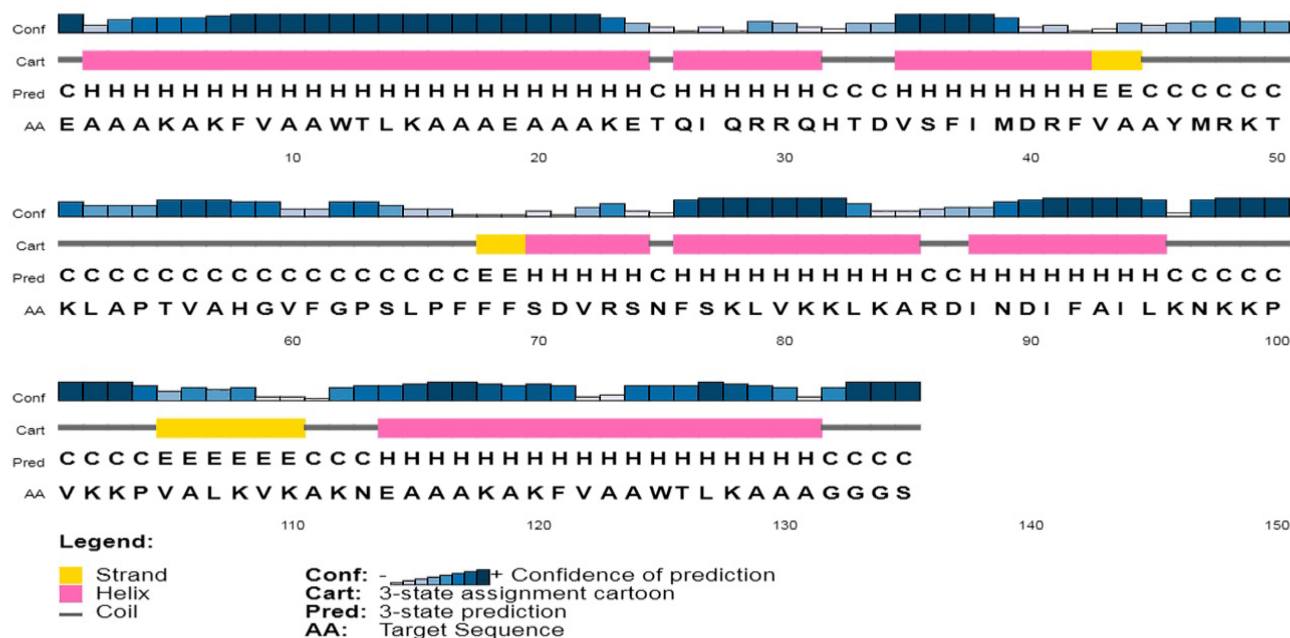


**Fig. 2: Secondary structure obtained for the SV is represented graphically**

Prediction uses a complex algorithm that involves artificial neural networks to analyze the primary sequence of the protein and then outputs the structural states based on its evolutionary relationship with known proteins. More precisely, the position-specific scoring matrix generated by PSI-BLAST is used to predict the secondary structural elements. It is possible to measure the accuracy of predictions quantitatively using the Q3 score, a fraction of correctly predicted residues across all three secondary structure types. That score is the score of maximum importance because indeed it measures the predictive ability of the algorithm to distinguish among helices, strands, and coils. Output The output not only provides a probabilistic estimation of the structural state of each amino acid but also opens up further analyses such as predicting the tertiary structure and functional annotation based on the delineated structural framework. Color coding in the results is helpful in enhancing interpretability, where researchers can rapidly assess how these secondary structure elements distribute and arrange themselves within the SV protein.

**Tertiary structure prediction and validation:**Tertiary structure of the designed subunit vaccine as predicted by RaptorX (Fig. 3a). It has two distinct domains.

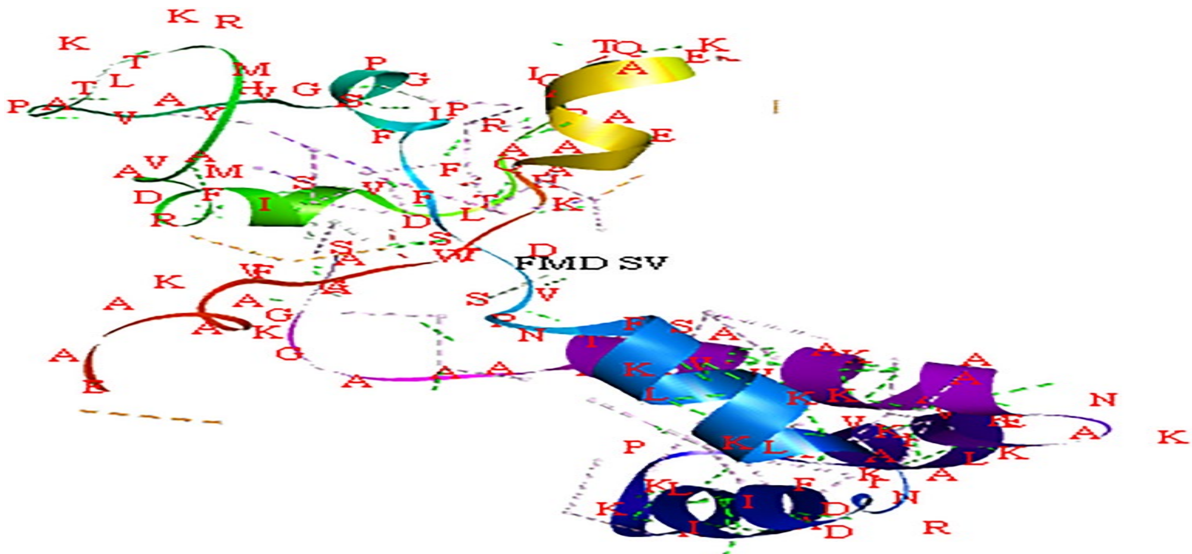


**Fig. 3a: Structure of the final SV model and its validation. The figure depicts the multi-epitope vaccine's final 3D model, which was created through homology modeling and refining.**

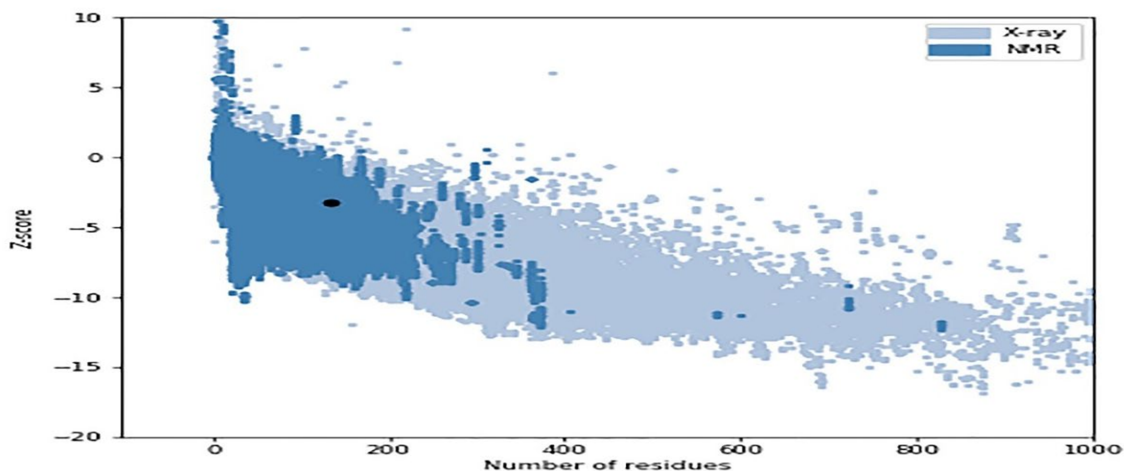


**Fig. 3b: The 3D structure's Z-score according to PROSA validation is (-3.23)**

The best template was PDB ID 4uz0A. Statististic analysis gave a P-value of 0.0157, meaning that there is significant quality in the alignment. Overall, the uGDT score for the SV was calculated to be 66, showing moderate structural accuracy. Total design and simulation included all the 135 residues of the SV protein, which incidentally also proved that 34% of these residues were disordered. This will directly impact the protein's stability and functionality. Quality assurance with the ProSA-web model resulted in a Z-score of -3.23 (Fig. 3b).

The score indicates that proteins of its size have acceptable quality, thus giving more weight to the feasibility of the predicted structure. To further validate the construct, a Ramachandran plot was built: dihedral angles along the structure of the protein. It seemed that 86 out of 135 residues were within the most favored regions (A, B, L). This constitutes 69.9%. Moreover, 32 residues fell into additional allowed regions: a, b, l, and p. These residues occur in generously allowed regions in only 3 cases (2.4%); in 126 total cases, residues were found in disallowed regions in only 1, and those accounted for a total of just 2 residues, or 1.6% (Fig. 3c).

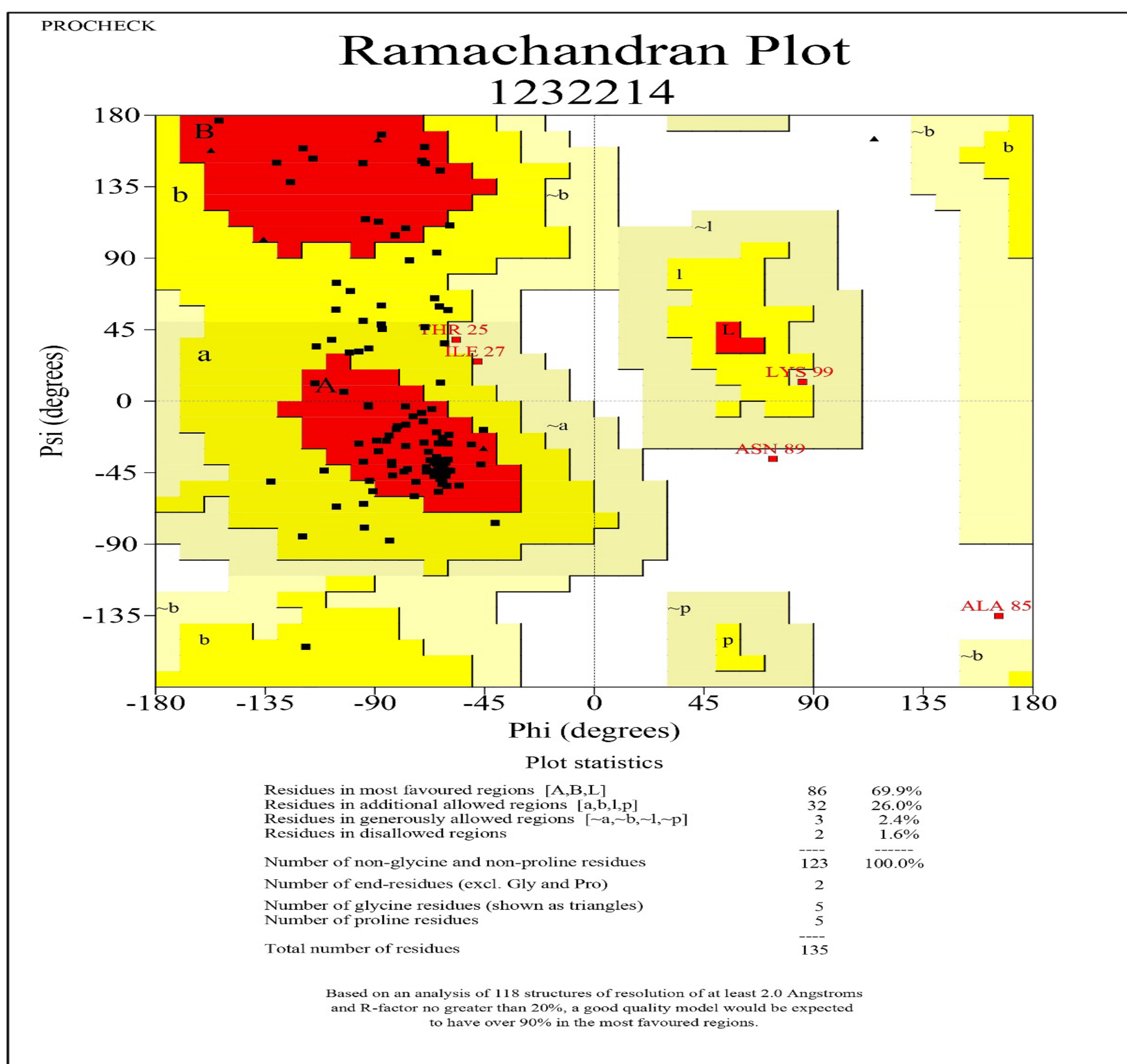


**Fig. 3c: Ramachandran plot analysis of the improved model reveals that 69.9%, 26%, 2.4%, and 1.6% of the residues, respectively, are in the preferred, additional allowed, generously allowed, and forbidden regions.**

Also, on the curve, five prolines and five glycines were shown, which influence the protein folding and stability due to specific properties of their conformation. As 70% of the residues are within favored regions and nearly a small percentage is in disallowed areas, the SV is assumed to have an average-quality structure for further experimental verification and possible therapeutic utility.

**Vaccine's molecular docking with the receptor (TLR-9):**The PatchDock server made a docking analysis of the SV protein with bovine TLR-9, using the adjuvant provided by TLR-2. Results suggested that the interaction between the SV protein and TLR-9 was significant and was characterized by favorable affinities and conformations of the docking (Fig. 4a, b,&c).In the initial set of models, 4uz0A-583378_1 was the best with a P-value of 1.6e-02 and score 44. It had structural similarities in terms of uGDT/GDT as 25/39 and sequence identity as 7/11 with respect to its template 4uz0A. 4tqlA-583378_1 came next with a P-value of 1.6e-02 but slightly lower score of 43; it had structural similarities of 21/33 for GDT and sequence identity of 11/17 with respect to its template 4tqlA. In another experiment involving the use of different models (modl1-583378_2, modl2-583378_2, etc.), all five highest-ranked models had the same statistical characteristics: each had a P-value of close to zero (0.0e+00) indicating high confidence in their predictions based on scoring systems that provided scores of around six for all highest-ranked entries (modl1-modl5). These models, however, did not represent any significant sequence or structural alignments because both their unnormalized GDT values varied from around about approximately forty-one to forty-two out of sixty possible positions for GDT measurements but had no measurable sequence identities with respect to known templates (modl1-modl5). Among models in consideration, Model 1 was the best since it had the highest GDT-HA score of 0.9727, which it tied with Model 5 but was slightly superior overall due to other considerations. Model 1 also had an excellent MolProbity score of 1.885, which reflected good structural quality and tied this score with Model 3. It also reflected a Clash Score of 8.6 and no bad rotamers (0%), further recommending it as the best-optimized structure. Model 3 was essentially the same as Model 1 in MolProbity score and Ramachandran favored residues but was slightly disadvantaged in GDT-HA compared to Model 1. Model 4 had the lowest Clash Score of 4.8 but was disadvantaged with lower GDT-HA score and higher MolProbity than Models 1 and 3. Finally, although Model 5 tied Model's GDT-HA score at a high level (0.9727), it had an enormously higher MolProbity score than Models' best scores. Overall, when all of these parameters were taken together in combination—namely its high structural similarity uncovered by GDT-HA combined with excellent structural quality uncovered by low MolProbity—Model' balance among these significant parameters made it the clear best-balanced and most refined structure among those generated by GalaxyRefine's refinement process(Supplementary Data 2).

The results showed that Model 1 had the lowest docking score of -240.36 but a high RMSD of 84.73 Å. Model 3 offered a balanced performance with a docking score of -226.20 and an RMSD of 38.22 Å, making it the most stable choice among those with strong binding predictions. Ultimately, Model 3 was selected as the best option due to its favorable combination of docking strength and structural stability metrics (confidence score: 0.8211)(Supplementary Data 3).

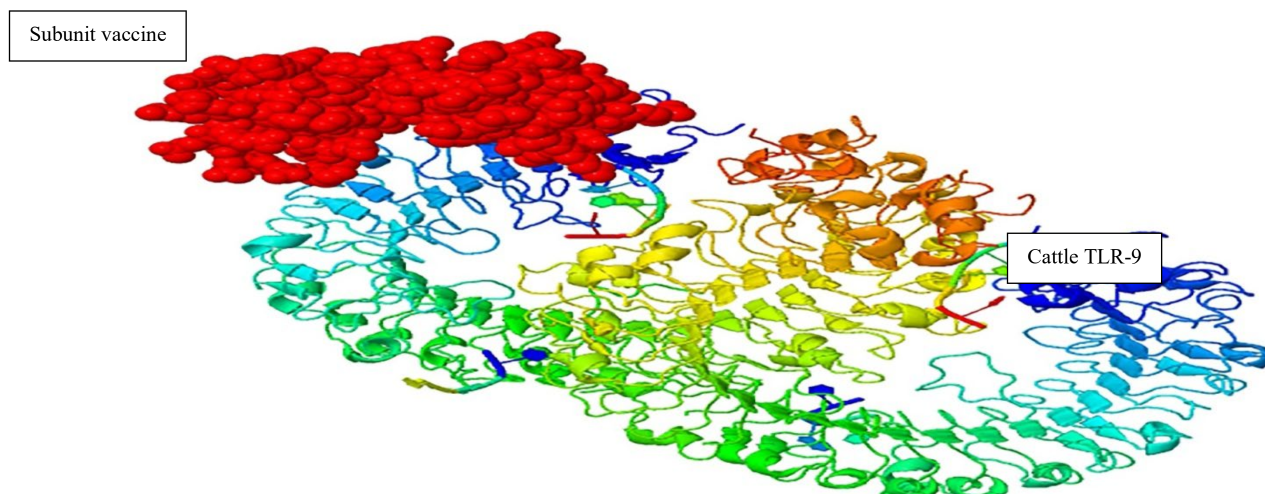


**Fig. 4a: SV constructs docked with a bovine TLR-9 complex (PDB ID: 5Y3M). In the docked complex downloaded from the HDOCK website, the receptor TLR-9 has been depicted as multi-colored ribbon-like bands, whereas red-colored beads represent the SV as a ligand.**

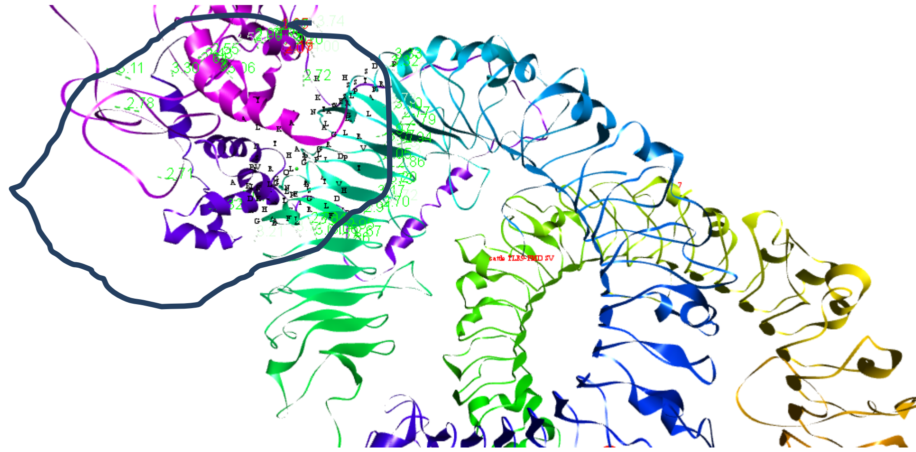

**Fig. 4b: Representation of intermolecular bonds and amino acids involved in docking between FMD-SV and cattle TLR-9**

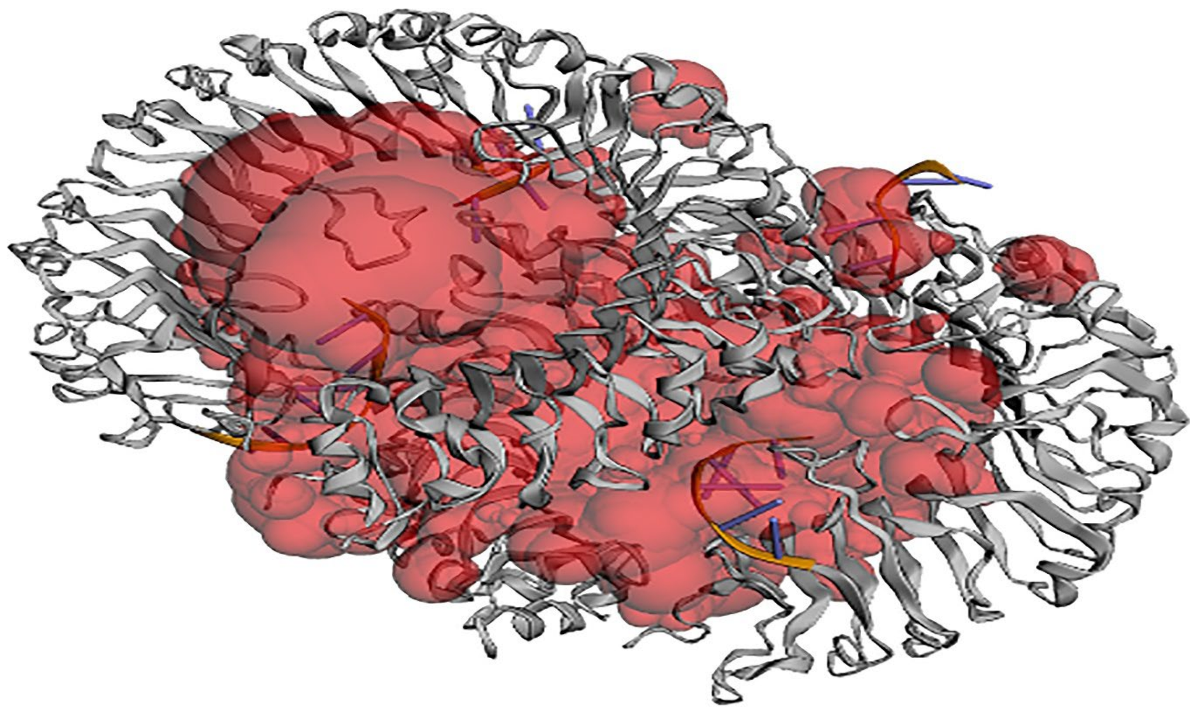

**Fig. 4c: CASTp server for determining bovine TLR-9 binding pockets. Grey ribbon structures represent TLR-9 and red balloons represents FMD-SV.**

The docking score was checked quantitatively and showed that indeed there was a rather strong interaction, whereas this interaction is important to determine the effects it may have on the immune system. Further support for the study was achieved through the use of bioinformatics tools that calculated the binding energy and modes of interaction, indicating that the SV protein may trigger TLR-9-mediated signaling pathways. The docking results, as illustrated in Fig. 4a, include a representation of complex formation with some crucial residues for the interaction highlighted: The structural integrity of the docked complex was then inspected using PyMOL, suggesting that critical amino acid residues present within the SV protein and TLR-9 themselves are involved in major contributions to interaction stability. Further validation may be done by performing energy minimization and molecular dynamics simulations to test whether predicted docking poses are both energetically favorable and biologically relevant. These results thus emphasis the possibility of using SV proteins as an adjuvant in TLR9-targeting vaccines that open ways toward understanding their role in modulating immune responses through the activation of TLR signaling pathways. The incorporation of quantitative data with

bioinformatics analyses helped us further advance our understanding of protein-protein interactions within the immune signaling frameworks and pave the way toward their future therapeutic applications.

**Reverse translation and codon optimization:** Reverse translation of the sequence of the SV protein produced a sequence that was 405 base pairs in length. It was then optimized for expression in the K12 strain of *Escherichiacoli*, one of the major hosts in the production of recombinant proteins (Supplementary data 4). Optimization was carried out to maximize the efficiency of translation and also to change the codon usage for it to favor *E. coli*. The GC content of the optimized sequence was found to be 47.16%, within a favorable range for stability and expression in bacterial systems. Additionally, the value of the CAI for the optimized sequence was determined to be 0.96, implying that the optimized sequence should have high adaptation for using the translational machinery of *E. coli*. This would imply that the optimized sequence is going to be efficiently expressed in this bacterial host, which makes it even further applicable to applications in molecular biology and biotechnology. A graphical representation of the optimized or modified sequence is shown in Fig. 5, which indicates changes during the optimization process.

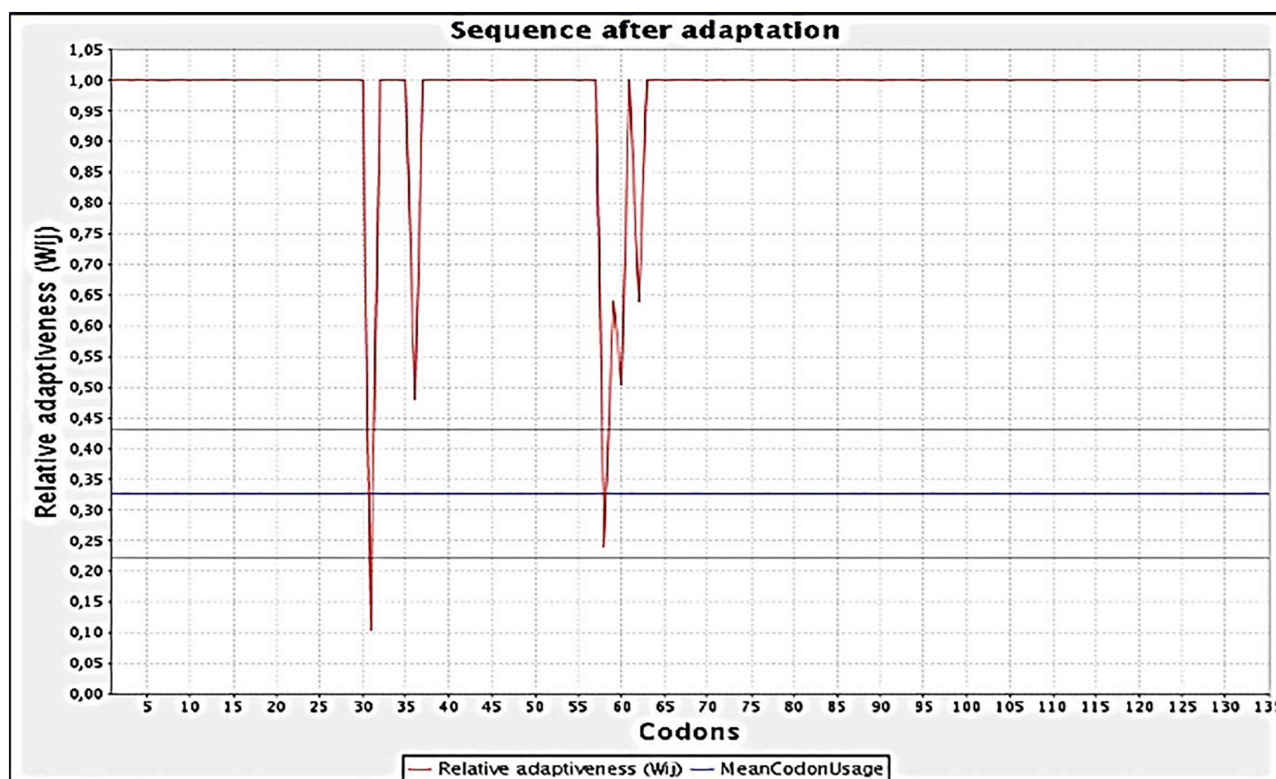


**Fig. 5: Graphical display of codon-optimized sequence. Red peaks show that practically all codons have undergone optimization. The relative adaptiveness of each codon is represented by the Y-axis.**

***In-silico* optimizing and colonization of vaccine protein:** The gene sequence was then cloned into the prokaryotic expression vector pET28a(+) using a set of techniques focusing on ensuring optimal expression of the vaccine. The gene was engineered to contain a set of restriction sites, XbaI at the N-terminus and XhoI at the C-terminus, to ensure the gene was inserted into the vector. A start codon (ATG) was added at the C-terminus just before the XhoI site in order to ensure proper initiation of translation, whereas a 6xHis tag was added after the XbaI site with this purpose to purify and detect proteins. Using the SnapGene tool, the optimized gene sequence was cloned into the pET28a(+) vector. This software thus made possible very careful design and visualization of the genetic construct, thus ensuring that all elements required for effective expression would be present in this construct. The cloning was made very efficient, and not only that, but downstream applications like the analysis of protein expression in E. coli as well. Computational validation by SnapGene showed a high probability of the expression of the targeted protein, compatible with other groups' results that show good immune responses from similar constructs. For the quantitative evaluation, yield and expression levels were forecasted, and predictions were a good expression in bacterial hosts. Bioinformatics analysis based on

antigenicity and the stability of the construct had been computed using several algorithms. In general, this systematic approach provides evidence of the integration of molecular cloning techniques with bioinformatics in vaccine development to achieve better efficacy in combating pathogens such as SARS-CoV-2 (Fig. 6).

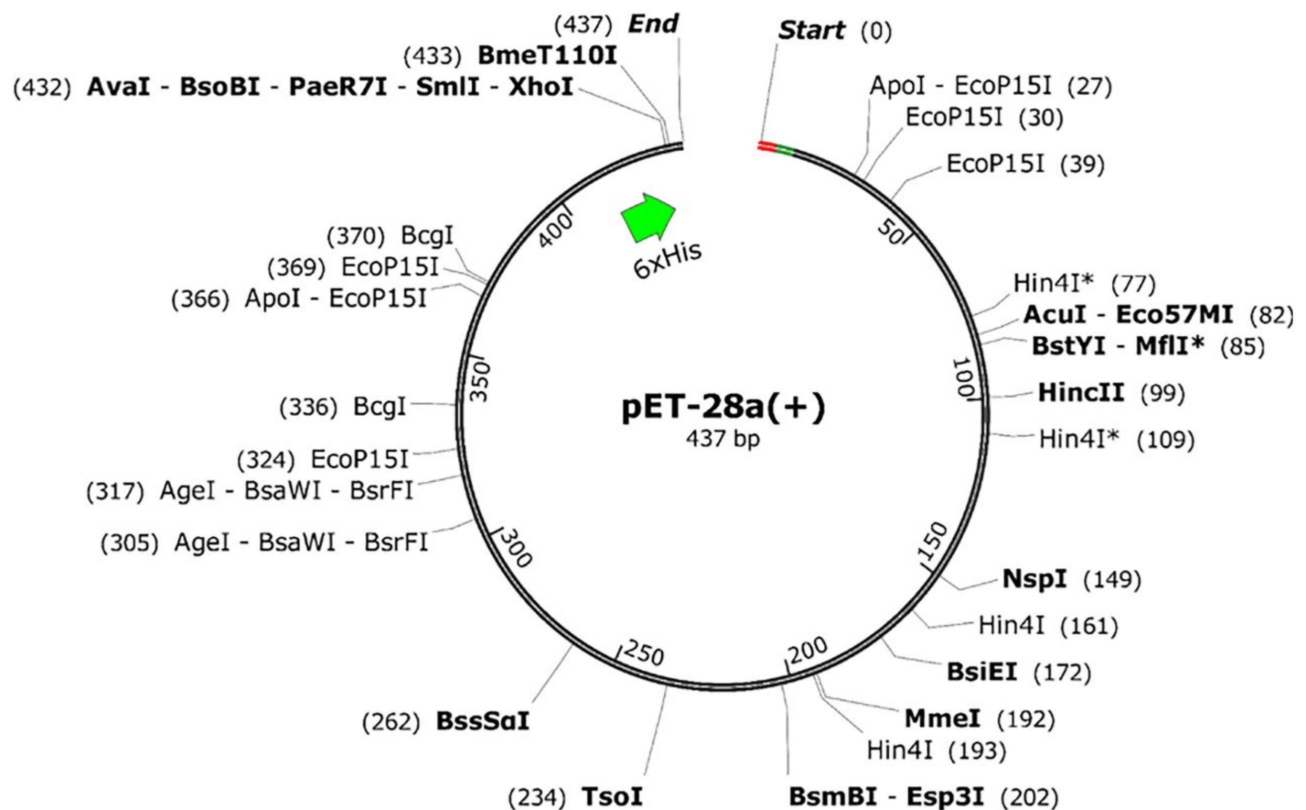


**Fig. 6: The final vaccine design was in silico restricted cloned into the pET 28a (+) prokaryotic expression vector, with the empty circle representing the "subunit vaccine." the vector is displayed by the black circle.**

## DISCUSSION

One of the most significant livestock illnesses, Foot-and-Mouth Disease (FMD), primarily affects large and small ruminants. FMD poses a serious hazard to cattle in more than 100 countries, not only harming struggling farmers but also impacting the national economy. Due to the disease's global trade constraints, controlling it is a priority (Gongal*et al*., 2022). FMD control techniques may vary from country to country based on the disease's distribution and the nation's economy. Vaccination has proven to be effective in controlling FMD when used judiciously. However, the effectiveness and application of current FMD vaccines are constrained by several issues, including high costs, the duration of immunity, maintaining the cold chain, unsatisfactory antigenic matching studies, changing antigenic load, antigen instability, and the need for boosters (Haynie, 2023).

The development of an epitope vaccine is made possible by our improved understanding of how antigens are recognized at the molecular level. Identification of B-cell and T-cell epitopes is the basis of the strategy used to create an epitope vaccination. These epitopes can trigger a particular immune response and are immunodominant (Dermime*et al*., 2004). To make the molecule immunogenic, the B-cell and T-cell epitopes can be joined together. The field of immunological study has greatly benefited from the development of computational immunology, or immunoinformatics, a subfield of bioinformatics.Immunoinformatics research greatly facilitates the investigation of possible epitopes of B and T-cells for pathogens. Immunologists can quickly assess possible sequences and their binding locations using a variety of bioinformatics technologies. This results in the creation of new, powerful vaccinations. In the current study, we developed an efficient multi-epitope vaccination using a variety of bioinformatics tools and algorithms. FMDV has seven serotypes and several strains, as is well known. Moreover, even within a strain, antigenic characteristics are evolving (Mahapatra and Parida, 2018).

The FMDV capsid proteins consist of the four constitutional proteins VP1, VP2, VP3, and VP4.These

proteins include numerous epitopes that can rouse immunological cell and humoral responses as well as the antibodies formation (Cox and Coulter, 2024). Our objective was to create an all-purpose different epitopes of FMDV vaccination. For this reason, we performed various combinations of sequences using Clustal Omega for 28 FMD viral strains that were available on NCBI (Smith, 2015). The protein sequences were positioned to underline the common residues across all serotypes. The multiple homologs of the VP1-4 proteins highlighted the most preserved regions in seven aphthovirus serotypes. Using seeded guide trees and the HMM profile-profile technique, the multiple sequence alignment tool Clustal Omega aligns protein sequences. Afterward, the epitopes of B and T cells of 28 FMDV strains were determined by consensual sequencing. Genius software verified the consensus sequence with the default settings for the parameters. Desktop software with extensions is called Geneious. It is a device for compiling and examining data sequences. To benefit from system software compatibility, Java is the primary programming language used in Geneious. The bulk of software modification is built utilizing one-step bioinformatics toolkits and is performed by non-computer specialists. 1-Tasser predicted the protein structure from the consensual sequence. Rampage was utilized to determine the general form quality that 1-Tasser had predicted. Before refinement, the quality was initially 88, and it was afterward raised to 93. Also, this functional examination of proteins highlights the three conserved domains in protein structure. These domains contained the superfamily Rhv, VP4-2, and Rhv-like. The viral protein VP4 component is found in domain VP4-2 of all three domains. This domain comes from the Piconaviridae family. MYRISTOYLATED is unique due to its prolonged structure and tiny size (VP4). Together with the N-terminus of VP1, externalization of VP4 is important in producing the infection.

The protein known as picornavirus capsid protein contains the Rhv domain. In other words, Rhv is a member of the same superfamily domain as the capsid of the picornavirus domain. To identify the epitopes of B-cell, the repetitive regions were checked for surface accessibility, epitope immunogenicity, epitope conservation, and wettability. Moreover, repetitive areas were examined, and epitopes of T-cells were identified. The conserved area epitopes demonstrated the characteristics that were mentioned before. To forecast the locations of antigenic sites, the immune epitope database (IEDB) was employed (http://tools.immuneepitope.org/bcell/). Using a default threshold value of 1, Kolaskar and Tongankar immunogenicity techniques were used to estimate antigenic locations (Kolaskar and Tongaonkar, 1990). The consensus sequence technique was used to determine the B-cell epitopes. To anticipate cell epitopes, cow alleles for the haplotypes A18, A19, A14, A20, A11, A13, and A10 were used. After eliminating the redundant epitopes, epitopes were chosen. The length of the epitopes used for B and T cells was chosen. In this way, 5 B-cell and 18 T-cell numbers were anticipated. The epitopes of B-cell were estimated following IEDB along with the star and end locations as well as the length of the epitope. The peptide "QNSMDTQLGDNAISGGSNEGSTDTTSTHTNNTQNND" also displays the longest chain of 36 amino acids in addition to the anticipated peptides.

The surface accessibility epitope was predicted using the Emini surface accessibility prediction tool, which has a default threshold value of 1.000, from the IEDB database (http://tools.immuneepitope.org/tools/bcell/iedb input) (Fleri*et al*., 2017). Epitope conservancy is projected using the IEDB epitope conservancy decision-making model (http://tools.immuneepitope.org/tools/conservancy/iedbinput). This approach was used to show how peptide sequence and epitopes based on sequence identity are both conserved. The minimal value of 1.897 was used with the Parker hydrophilicity tool (http://tools.immuneepitope.org/bcell/) to forecast hydrophilicity. In the IEDB study, the B cell epitopes were used to generate a polypeptide list of reactive sites using cattle B cells. The five most important Ag determinants were first selected and then they were coupled with GPGPG linkers. The GPGPG linkers play vital roles in epitope vaccines. The GPGPG linker, which is best to attach to beta turns, increased space between G and P and decreased binding ability in the core binding sites. It's possible that the secondary and tertiary protein structures were unaffected. By displaying the chosen epitopes through HLA-II, this linker not only makes it easier for the immune system to process antigens but also stops junctional epitopes from developing. The GPGPG spacer was used to induce HTL responses using polypeptides and DNA vaccination. I-TASSER was used to predict the structure of the B and T cell epitopes immediately following their combination with the GPGPG spacer. The projected models were improved using Rampage and the finer. For the accuracy factor, the structure of B-cell epitopes was evaluated both before and after refining. The structure's maximum quality was found to be 90% before improvement and 96% following improvement. Two distinct domains in the protein were also discovered by the systematic study of B-cell. Their motif names were G3DSA and PF08935. Surprisingly, the viral protein VP4 component was shown to depend on the motif PF08935. Both before and after being refined to account for the accuracy factor, the T-cell epitope structure's quality was assessed. The overall structural grade was 87% before refinement; 91% after refinement. The protein's pattern is revealed by the systematic study of the T cell antigenic protein. The protein that makes up

the B cell epitope contains the same pattern, known as G3DSA. A strategy for creating the "Universal FMD Vaccine" has been established in this work. These many viral epitopes may help protect the animals against FMDV strains that are heterologous.

**Conclusion:** A multi-epitope subunit vaccine against Foot-and-Mouth Disease Virus was designed in this study, driven by an immunoinformatics approach targeting the B- and T-cell epitopes from the structural and non-structural proteins. The antigenicity, stability, and the non-allergenic properties of the vaccine construct were high. The binding affinity to the TLR9 receptor was strong. Molecular docking and structural analyses supported the findings. Its ability to be expressed efficiently in *Escherichia coli* was confirmed by codon optimization and in silico cloning. The computational results are promising; however, it is crucial that the vaccine's efficacy, safety, and immunogenicity are evaluated in vivo through animal trials. This research lays a strong foundation for a universal FMD vaccine that addresses the challenges created by the virus's antigenic diversity and evolution of strains.

**Acknowledgements:** The authors extend their appreciation to the ongoing Research Funding Program (ORF-2025-552) King Saud University, Riyadh, Saudi Arabia.

## REFERENCES

Admassu, B., K. Getnet, A. Shite and S. Mohammed (2015). Review on foot and mouth disease: distribution and economic significance. Acad. J. Anim. Dis. 4(3): 160-169.DOI: 10.5829/idosi.ajad.2015.4.3.95165.

Akter, S., M.S. Rahman, M.R. Islam, M. Akther, H. Anjume, M.Marjia and M. Sultana (2024). Development of recombinant proteins for vaccine candidates against serotypes O and A of Foot-and-Mouth Disease virus in Bangladesh. Access Microbiol. 6(6): 000713-v4. https://doi.org/10.1099/acmi.0.000713.v3.

Ali, M.Z., D. Dewan, M.M.R. Khan and M.M. Meher (2024). Prevalence, genetic characteristics and economic losses of Foot-and-Mouth Disease Virus (FMD) in global and Bangladesh context: A Review. Ukr. J. Vet. Agric. Sci. 7(2): 8-19. https://doi.org/10.32718/ujvas7-2.02.

Arzt, J., B. Baxt, M. J. Grubman, T. Jackson, N. Juleff and J. Rhyan (2011). The pathogenesis of foot-and-mouth disease II: viral pathways in swine, small ruminants, and wildlife; myotropism, chronic syndromes, and molecular virus-host interactions. Transbound. Emerg. Dis. 58(4): 305-326. doi:10.1111/j.1865-1682.2011.01236.x.

Badr, A., A.A. Zaghawa, M.A. Nayel, A.A. Salama, A.M. Elsify and W.S. Mousa (2024). Foot-and-Mouth Disease (FMD) current and novel vaccines: progress and limitations. J. Curr. Vet. Res. 6(1): 16-30. Doi:10.21608/JCVR.2024.352699

Chestley, T., P. Sroga, M. Nebroski, K. Hole, H. Ularamu, O. Lung and C. Nfon (2022). Development of reverse-transcriptase, real-time PCR assays to distinguish the Southern African Territories (SAT) serotypes 1 and 3 and topotype VII of SAT2 of Foot-and-Mouth Disease Virus. Front. Vet. Sci. 9: 977761.https://doi.org/10.3389/fvets.2022.977761

Choudhury, S.M., X. Ma, W. Dang, Y. Li and H. Zheng (2021). Recent development of ruminant vaccine against viral diseases. Front. Vet. Sci. 8: 697194. https://doi.org/10.3389/fvets.2021.697194

Cox, J.C. and A.R. Coulter (2024). Advances in adjuvant technology and application. In Animal parasite control utilizing biotechnology (pp. 49-112). CRC Press.

Dermime, S., D.E. Gilham, D.M. Shaw, E.J. Davidson, E.K.Meziane, A. Armstrong and P.L. Stern (2004). Vaccine and antibody-directed T cell tumour immunotherapy. Biochim. Biophys. Acta (BBA)-Rev. Cancer. 1704(1): 11-35. https://doi.org/10.1016/j.bbcan.2004.03.002.

Fleri, W., S. Paul, S.K. Dhanda, S. Mahajan, X. Xu, B. Peters and A. Sette (2017). The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. Front. Immunol. 8: 278. https://doi.org/10.3389/fimmu.2017.00278

Gasteiger, E., C. Hoogland, A. Gattiker, M.R. Wilkins, R.D. Appel and A. Bairoch (2005). Protein identification and analysis tools on the ExPASy server. In The Proteomics Protocols Handbook, pp. 571-607. Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Gongal, G., H. Rahman, K.C. Thakuri and K. Vijayalakshmy (2022). An overview of transboundary animal diseases of viral origin in South Asia: what needs to be done? Vet. Sci. 9(11): 586. https://doi.org/10.3390/vetsci9110586.

Gunasekera A.H. and D.H. Subasinghe (2011). Dairy industry in Sri Lanka. Sri Lanka Vet. J. 58: 29-34.

Grote, A., K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D.C. Hempel and D. Jahn (2005). JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 33(suppl_2): 526-531. https://doi.org/10.1093/nar/gki376

Haynie, T. (2023). Recombinant DNA vaccine design as a potential strategy against bovine foot and mouth disease virus (FMDV). M.Sc. thesis (unpublished). Louisiana State University and Agricultural and Mechanical College. Doi: 10.31390/gradschool_theses.5772.

Hohlich, B.J., K.H. Wiesmuller, T. Schlapp, B. Haas, E. Pfaff and A. Saalmuller (2003). Identification of foot-and-mouth disease virus-specific linear B-cell epitopes to differentiate between infected and vaccinated cattle. J. Virol. 77: 8633-8639. https://doi.org/10.1128/jvi.77.16.8633-8639.2003

Islam, Z. (2024). Functional characterization of coxsackievirus A9-eGFP vector. Master's Thesis (Unpublished).Department of Biology, Faculty of Science, University of Turku, Finland.

Kamel, M., A. El-Sayed and H. Castañeda Vazquez (2019). Foot-and-mouth disease vaccines: recent updates and future perspectives. Arch. Virol. 164: 1501-1513. https://doi.org/10.1007/s00705-019-04216-x

Kaufmann, S.H.E. (1988). CD8+ T lymphocytes in intracellular microbial infections. Immunol. Today. 9(6): 168-174. https://doi.org/10.1016/0167-5699(88)91292-3.

Kitching, P., J. Hammond, M. Jeggo, B. Charleston, D. Paton, L. Rodriguez and R. Heckert (2007). Global FMD control—is it an option? Vaccine. 25(30): 5660-5664. https://doi.org/10.1016/j.vaccine.2006.10.052.

Kolaskar, A.S. and P.C. Tongaonkar (1990). A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 276(1-2): 172-174. https://doi.org/10.1016/0014-5793(90)80535-Q.

Layessa, M.L., E.Z. Gebremedhin, E.J. Sarba and W.M. Bune (2024). Assessing farmers' willingness to pay for FMD vaccines and factors influencing payment: a contingent valuation study in central Oromia, Ethiopia. BMC Vet. Res. 20(1): 313.https://doi.org/10.1186/s12917-024-04169-7.

Lombard, M.F. (2012). Vaccine use for foot and mouth disease control. In The FAO/OIE Global Conference on Foot and Mouth Disease Control, p. 103.

Mahapatra, M. and S. Parida (2018). Foot and mouth disease vaccine strain selection: current approaches and future perspectives. Expert Rev. Vaccines. 17(7): 577-591. https://doi.org/10.1080/14760584.2018.1492378

Manvani, R., H. Purohit, C.R. Sahoo, M. Rajput and S. Shah (2024). Immunoinformatics: an interdisciplinary technique for designing and engineering vaccine antigen. In Reverse Vaccinology, pp. 87-99. Academic Press. https://doi.org/10.1016/B978-0-443-13395-4.00012-5.

Maurer-Stroh, S., N.L. Krutz, P.S.Kern, V. Gunalan, M.N. Nguyen, V. Limviphuvadh, F. Eisenhaber and G.F. Gerberick (2019). AllerCatPro-prediction of protein allergenicity potential from the protein sequence. Bioinformatics. 35: 3020-3027. doi:10.1093/bioinformatics/btz029. https://doi.org/10.1093/bioinformatics/btz029

Medina, G.N. and F. Diaz San Segundo (2024). Virulence and immune evasion strategies of FMDV: implications for vaccine design. Vaccines. 12(9): 1071. https://doi.org/10.3390/vaccines12091071.

Millan, C.L. B., R. Weeratna, A.M. Krieg, C.A. Siegrist and H.L. Davis (1998). CpG DNA can induce strong Th1 humoral and cell-mediated immune responses against hepatitis B surface antigen in young mice. Proc. Natl. Acad. Sci. 95(26): 15553-15558.https://doi.org/10.1073/pnas.95.26.15553

Ndeevelo, P. (2023). Mathematical analysis of foot and mouth disease with optimal control: a case study of FMD in Namibia. Doctoral dissertation (unpublished). Namibia University of Science and Technology. http://ir.nust.na:8080/jspui/handle/10628/1014.

Pandey, R.K., T.K. Bhatt and V.K. Prajapati (2018). Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. Sci. Rep. 8(1): 1-11. https://doi.org/10.1038/s41598-018-19456-1

Prasad, R.D., K.D. Sonawane, R.S. Prasad, S.R. Prasad, N. Prasad, R.K. Shrivastavand Z. Guo (2024). A review on livestock viral disease and their management. ES General. 5: 1227. http://dx.doi.org/10.30919/esg1227

Prempeh, H., R. Smith and B. Müller (2001). Foot and mouth disease: the human consequences. The health consequences are slight, the economic ones huge. BMJ. 322(7286): 565-566.doi: https://doi.org/10.1136/bmj.322.7286.565

Raoof, H.S. (2024). FOOT AND MOUTH DISEASE (FMD). Assiut Vet. Med. J. 70(182): 1-8. https://doi.org/10.21608/avmj.2024.274826.1232

Ribeiro, B.F.S., D.L.N. Rodrigues, J.C. Ariute, M.V. Dias, W.G. Nogueira, R. T. Ramos and F.F. Aburjaile (2024). Bioinformatics approach to design peptide vaccines for viruses. In Reverse

Vaccinology, pp. 265-287. Academic Press. https://doi.org/10.1016/B978-0-443-13395-4.00024-1

Schneider, W.M., M.D. Chevillotte and C.M. Rice (2014). Interferon-stimulated genes: a complex web of host defenses. Annu. Rev. Immunol. 32: 513-545. https://doi.org/10.1146/annurev-immunol-032713-120231

Smith, D.R. (2015). Buying in to bioinformatics: an introduction to commercial sequence analysis software. Brief. Bioinform. 16(4): 700-709.https://doi.org/10.1093/bib/bbu030

Waseem, M., H. Qureshi, K. Maheen, M. Kanwal, A. Basheer, M. Faheem and S.B. Jamal (2024). Development of subunit vaccine: a step forward toward cost-effective technology of vaccine candidate discovery. In Reverse Vaccinology, pp. 41-60. Academic Press. https://doi.org/10.1016/B978-0-443-13395-4.00017-4

Xu, Z. and D.W. Kulp (2019). Protein engineering and particulate display of B-cell epitopes to facilitate development of novel vaccines. Curr. Opin. Immunol. 59: 49-56.https://doi.org/10.1016/j.coi.2019.03.003

## Supplementary Data

**Supplementary Data 1**
**Protein sequence of Subunit Vaccine**
**EAAAK**AKFVAAWTLKAAA**EAAAK**ETQIQRRQHTDVSFIMDRFV**AAY**MRKTKLAPTVAHGVF**GPSL**PFFFSDV RSNFSKLV**KK**LKARDINDIFAILKN**KK**PVKKPVALKVKAKN**EAAAK**AKFVAAWTLKAAA**GGGS**

**Supplementary Data 2: 3D model 4uz0A-583378_1 obtained from RaptorX and refined by Galaxy to get 5 more models**

| Model | GDT-HA | RMSD | MolProbity | Clash score | Poor rotamers | Rama favored |
|---|---|---|---|---|---|---|
| Initial | 1.0000 | 0.000 | 3.150 | 79.5 | 2.1 | 90.3 |
| MODEL 1 | 0.9727 | 0.364 | 1.885 | 8.6 | 0.0 | 93.5 |
| MODEL 2 | 0.9688 | 0.368 | 1.950 | 8.6 | 0.0 | 91.9 |
| MODEL 3 | 0.9688 | 0.359 | 1.885 | 8.6 | 0.0 | 93.5 |
| MODEL 4 | 0.9609 | 0.392 | 1.911 | 4.8 | 2.1 | 93.5 |
| MODEL 5 | 0.9727 | 0.392 | 2.127 | 8.6 | 2.1 | 93.5 |

**Supplementary Data 3: Summary of the Top 10 Models obtained from HDOCK server**

| Rank | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Docking Score** | -240.36 | -230.07 | -226.20 | -222.33 | -222.01 | -220.89 | -220.73 | -218.08 | -216.12 | -215.69 |
| **Confidence Score** | 0.8590 | 0.8322 | 0.8211 | 0.8095 | 0.8085 | 0.8050 | 0.8045 | 0.7960 | 0.7896 | 0.7881 |
| **Ligand RMSD (Å)** | 84.73 | 80.85 | 38.22 | 45.16 | 90.38 | 84.79 | 82.14 | 76.84 | 91.57 | 79.65 |
| **Interface Residues** | model_1 | model_2 | model_3 | model_4 | model_5 | model_6 | model_7 | model_8 | model_9 | model_10 |

**Supplementary Data 4**

**Reverse translation of Untitled to a 405 base sequence of most likely codons:**
gaagcggcggcgaaagcgaaatttgtggcggcgtggaccctgaaagcggcggcggaagcggcggcgaaagaaacccagattcagcgccgccagcataccgatgtgagc tttattatggatcgctttgtggcggcgtatatgcgcaaaaccaaactggcgccgaccgtggcgcatggcgtgtttggcccgagcctgccgtttttttttagcgatgtgcgcagcaac tttagcaaactggtgaaaaaactgaaagcgcgcgatattaacgatatttttgcgattctgaaaaacaaaaaaccggtgaaaaaaccggtggcgctgaaagtgaaagcgaaaaac gaagcggcggcgaaagcgaaatttgtggcggcgtggaccctgaaagcggcggcgggcggcggcagc
**Codon optimize sequence for E. coli K12 strain:**
gaagctgctgctaaagctaaattcgttgctgcttggaccctgaaagctgctgctgaagctgctgctaaagaaacccagatccagcgtcgtcaacacaccgacgttagcttcatcat ggaccgtttcgttgctgcttacatgcgtaaaaccaaactggctccgaccgttgctcatggcgtattcggcccgtctctgccgttcttcttctctgacgttcgttctaacttctctaaactg gttaaaaaactgaaagctcgtgacatcaacgacatcttcgctatcctgaaaaacaaaaaaccggttaaaaaaccggttgctctgaaagttaaagctaaaaacgaagctgctgcta aagctaaattcgttgctgcttggaccctgaaagctgctgctggtggtggttct