*Innov Biosyst Bioeng*, 2024, vol. 8, no. 3, 3–17
doi: 10.20535/ibb.2024.8.3.296662

Original article

# DESIGNING A MULTI-EPITOPE VACCINE CANDIDATE TO MERS-CoV: AN *IN SILICO* APPROACH

**M.N. Majeed[1,2], A. Iqbal[1], N. Murtaza[2], L.D. Herrera-Zúñiga[3], S. Siddique[4], M. Raza[1], M. Hussain[1], M. Sajid[1*]**

[1]University of Okara, Punjab, Pakistan
[2]University of Central Punjab, Lahore, Pakistan
[3]Metropolitan Autonomous University, Mexico City, Mexico
[4]National Yunlin University of Science and Technology, Douliu City, Taiwan

*Corresponding author: infobiotec@uo.edu.pk

Received 14 January 2024; Accepted 10 July 2024

**Background.** Middle East Respiratory Syndrome Coronavirus (MERS-CoV), associated with severe respiratory illness, originates from the Middle East region. The virus is transmitted from animals to humans, with the dromedary camel serving as a significant reservoir. The virus's high fatality rate has spurred research into vaccine development and therapeutics.
**Objective.** This study aimed to employ an *in silico* approach to design a potential vaccine candidate against MERS-CoV, focusing on the M protein as an antigen.
**Methods.** The FASTA sequence of M protein was used to predict B cell and major histocompatibility complex class I and class II epitopes. The best epitopes were selected from these predicted epitopes. The vaccine candidate's construct consisted of epitopes, linkers, and a tag. The sequence of the vaccine candidate's construct, consisting of 390 amino acids, was back-translated, optimized, and then inserted into a plasmid for cloning and expression using SnapGene. The 3D structure of the vaccine candidate is docked with TLR-4 receptor. Molecular dynamics simulation was run for this docked complex using GROMACS gmx, version 2021.4.
**Results.** Through computational modeling and analysis, we developed a novel vaccine candidate with promising structural and functional properties. Our results suggest that the designed vaccine candidate has the potential to induce a robust immune response.
**Conclusions.** This *in silico* approach presents a promising MERS-CoV vaccine candidate designed to trigger both humoral and cellular immune responses. This candidate holds the potential to provide broad-spectrum protection against MERS-CoV.

**Keywords:** MERS-CoV; vaccine candidate's; molecular docking; molecular dynamics simulation; bioinformatics approaches.

## Introduction

MERS-CoV is an infectious virus that was first reported in Jeddah, Saudi Arabia, in June 2012 [1]. MERS-CoV is a zoonotic viral pathogen. It likely uses bats as its natural host [2] camels may acquire MERS-CoV from bats and it also transmit from camels to camels [3] and by direct contact to these camels (dromedary camels) it can be transmitted to Humans [4]. MERS-CoV enters the human body via the lower respiratory tract, leading to the development of severe respiratory symptoms that may ultimately result in respiratory failure or affect other organs [5]. As of July 4th, 2023, health authorities globally have recorded a cumulative total of 2,613 instances of MERS-CoV, with 945 fatalities [6]. Most of MERS vaccines currently being developed are based on the MERS-CoV S protein [7]. The genome of MERS-CoV contains 30,119 nucleotides [8] and encodes four structural proteins, containing spike (S), membrane (M), nucleocapsid (N), envelope (E), it also have some accessory proteins for example 3, 4a, 4b, 5 and 8 [9]. In this study, we used modern-day computational methods to design a novel multi-epitope vaccine candidate against MERS-CoV using its M protein as an antigen. The main protein in the viral envelope is M protein. By networking with all the other structural proteins, they play a crucial part in viral assembly [10]. The M protein is the prevalent protein of the envelope. Its length could be 217-230 amino acid, but it could go up to 270 residues in some variants [11]. To use a novel approach consisting in utilizing advanced computational methods and immunoinformatics tools for the design of an *in silico* vaccine candidate against MERS-CoV. This approach involves the identification of potential antigenic peptides, structural modeling of vaccine candidate, and validation through molecular docking and MD simulations. Previous studies based on developing *in silico* vaccine candidates against MERS-CoV lacks the *in silico* cloning technique. Our study also aims to demonstrate the

feasibility and efficacy of *in silico* cloning and optimization in the development of a MERS-CoV vaccine candidate. The benefits include accelerated vaccine development timelines, reduced need for extensive laboratory experimentation, and the ability to rapidly respond to emerging viral threats. In this study we aimed to design a multi epitope based *in silico* vaccine candidate against MERS-CoV using its Beta coronavirus England 1 strain. The M protein of this strain consist of 219 amino acids. Constructing an in-silico Vaccine candidate against the MERS-CoV M antigen involves several steps. Initially, the viral antigen is used to predict potential epitopes which can elicit immune response. Then these epitopes are linked together using linkers. Three-dimensional structure of these linked epitopes is then predicted, and that structure is then used to dock with receptor and MD simulation is then run on the complex formed by vaccine candidate and receptor dock.

## Materials and methods

### *Antigen*

M antigen of MERS-CoV [Beta coronavirus England 1] is selected as an antigen, because of its important role in virus cell cycle. The M antigen is essential for the assembly and maturation of viruses. It has value of 0.5504 for Overall Prediction for the Protective Antigen predicted by VaxiJen 2.0 [12]. Protein sequence of MERS-CoV M protein [Beta coronavirus England 1] was downloaded from NCBI [13] with accession number >AFY13313.1, this protein has 219 amino acids.

### *Epitopes prediction and evaluation*

B cell and T cell epitopes for the M antigen were predicted by IEDB [14]. As T cell epitopes we predicted was of two types of major histocompatibility complex (MHC) class I and MHC class II epitopes. IEDB tool predicted linear epitopes for B cell and used ANN 4.0 and NN-align 2.3 (NETMHCII 2.3) method to predict MHC class I and MHC class II Epitopes respectively. Then Firstly, epitopes having ic50 value equal or more than 100 was deleted, then VaxiJen v2.0 server was used to predict vaccigenicity of remaining epitopes. Epitopes with Vaxijen value lower than 0.6 was deleted, but in case of B cell epitopes threshold value in Vaxijen 2.0 server was 0.4. epitopes having good Vaxijen value are than used in AllerTop server [15] to predict there allergenicity, the allergic epitopes are deleted. At last, ToxinPred tool [16] was used to check toxicity of epitopes, all the epitopes were nontoxic. So, in the end we only had 21 potential epitopes, which are candidates for a vaccine construct with 9 MHC class II, 11 MHC class II, and 1 B cell epitopes.

### *Vaccine candidate's construction*

All the epitopes were joined together using linkers and then to provide more stability and solubility to vaccine candidate an adjuvant is also connected to vaccine candidate's N-terminal with the aid of a linker. Then a tag is added to the C terminal to have an identification of our vaccine candidate in a complex or mixture of proteins. Then physiochemical properties were examined using the ProtParam online server [17]. PsiPred tool [18] was used to predict secondary structure of our vaccine candidate's construct and I-Tasser is used to predict tertiary structure. I-TASSER [19], which stands for Iterative Threading ASSEmbly Refinement, is a hierarchical method used for predicting protein structure and annotating structure-based functions. The process begins by identifying structural templates from the Protein Data Bank (PDB) using a multiple threading approach called LOMETS. These templates are then used to construct full-length atomic models through iterative simulations based on template fragments. This allows for the derivation of information regarding the protein's functions. The finest-modeled structure underwent refinement using the Galaxy Refine server [20]. The final 3D model was subject to analysis using ProSAweb [21] and Ramachandran plot. The Ramachandran plot, obtained from PROCHECK [22], evaluates the torsional angles of each residue in the protein and categorizes them as allowed, favored, or outliers. ProSA web was utilized to identify errors in the generated 3D models by analyzing the atomic coordinates of the model. ProSA web provides a Z-score indicating the overall quality of the model and plots residue energies of the protein.

### *Cloning*

EMBOSS bactranseq [23] is used to backtranslate sequence of vaccine candidate into nucleotide sequence and then it is refined using Jcat [24] and then inserted into a vector by inducing cut in vector using SnapGene [25].

### *Docking and Molecular dynamic simulation*

Structure of TLR-4 was retrieved from RCSB in pdb format using PDB id 3FXI [26], all the ligands and water molecules are than deleted from

this structure. This structure is then used as receptor. Cluspro 2.0 server [27] is used for docking. Refined structure of vaccine candidate is than docked with this TLR-4 receptor. The best scored model obtained from Cluspro is than subjected towards MD simulation. MD simulation is performed using GROMACS [28] gmx, version 2021.4 CHARMm36 force field [29] at 100. The docked complex was positioned at the center of a triclinic box with a space of 1.5 nm from all edges and solvated with the TIP3P water model. Sodium-Chloride (NaCl) was added at physiological concentrations. After that, 3 independent runs were simulated, and the systems were minimized with different shutdown criteria using the steep of the steepest descent method to ensure independent simulations. Next, for each system, we carried out two equilibration steps: a 1 ns NVT equilibration followed by a 10 ns NPT equilibration. After the equilibrations, two systems were run for 50 ns, and the last one for 100 ns of a production run were performed. Generated graphs, data and files were visualized using qtGrace [30] and Microsoft excel 2019 and VMD [31], respectively. The methodology of *in silico* vaccine candidate development is illustrated in Fig. 1.

## Results

### *Antigen selection*

The M protein of MERS-CoV is selected as an antigen and candidate to construct an in-silico vaccine candidate against it has multiple functions. It helps shape the virus and assemble new viral particles, allowing them to be released from infected cells. The M protein also assists in moving viral components within the cell and helps the vi-

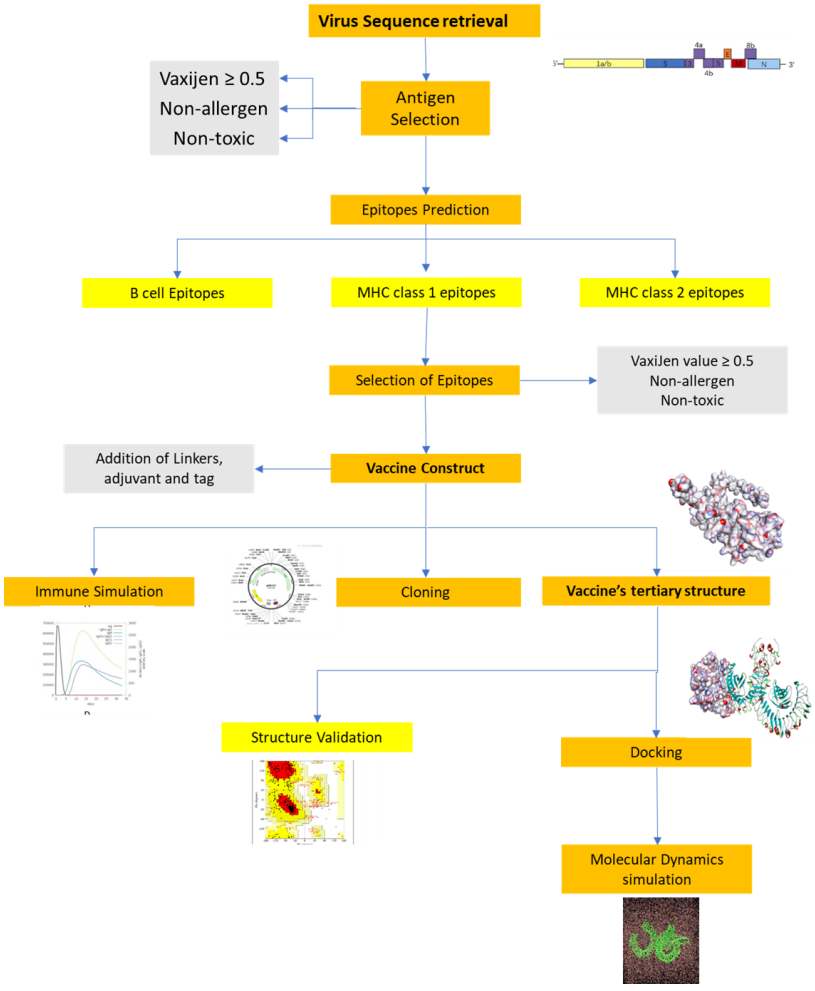



**Figure 1:** Schematic illustration of complete methodology of *in silico* epitopes based vaccine candidate development

rus evade the host's immune system. Additionally, it may contribute to the development of symptoms and damage caused by the virus. FASTA sequence of M protein is downloaded from NCBI with accession number AFY13313.1, It have 219 amino acids.

>AFY13313.1 M protein [Beta coronavirus England 1]

MSNMTQLTEAQIIAIIKDWNFAWSLIFLL
ITIVLQYGYPSRSMTVYVFKMFVLWLLWP
SSMALSIFSAVYPIDLASQIISGIVAAVSAM
MWISYFVQSIRLFMRTGSWWSFNPETNC
LLNVPFGGTTVVRPLVEDSTSVTAVVTNG
HLKMAGMHFGACDYDRLPNEVTVAKP
NVLIALKMVKRQSYGTNSGVAIYHRYKA
GNYRSPPITADIELALLRA.

Antigenicity of this sequence is predicted by VaxiJen and have 0.5504 VaxiJen value. It is non allergen and nontoxic as predicted by AllerTop and ToxinPred respectively. Secondary structure predicted by PSIPred of M protein suggests that 50 amino acids participate in formation of beta-strands and 99 amino acids are participated in formation of alpha-helix, remaining 70 amino acids makes coil as shown in Fig. 2. Tertiary structure is predicted by I-TASSER.

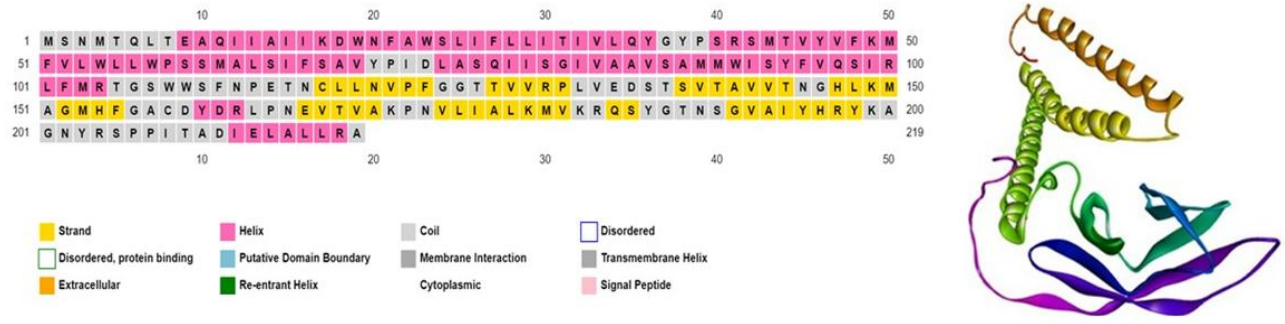
*Epitopes prediction and evaluation*

B cell and T cell epitopes were predicted using IEDB tool and only epitopes which are non-allergic, nontoxic, having ic50 value below 100 and had VaxiJen value 0.6 or above (threshold value in case of B-cell is 0.4) was selected, predicted by Aller Top, ToxinPred, IEDB and VaxiJen server 2.0 respectively. So, we had 21 potential epitopes, which are candidates for a vaccine with 9 MHC class II, 11 MHC class II, and 1 B cell epitope. All the epitopes and there vaxijen values for MHC class I and MHC clas II are mentioned in Table 1 and Table 2 respectively.

Only 1 b cell epitope is selected as a candidate for a vaccine KMVKRQSYGTN**,** its vaxijen value is 0.4474 and it is also nontoxic and non-allergen. Population coverage for MHC class I epitopes and MHC class-II epitopes separately and combined is assessed against whole world and for different world regions using IEDB server. The population coverage for combined MHC class I and MHC class II epitopes was estimated to be 89.78%. MHC class I and MHC class II epitopes as combined showed highest percentage of population coverage (94.34%) in North America followed by Europe (93.01%), North Africa (85.50%), West Indies (83.95%), South Asia (82.40%), East Africa (81.80%), East Asia (81.43%), West Africa (80.30%), Central Africa (75.54%), Northeast Asia (70.91%), Southwest Asia (65.28%), South Africa (59.39%), Oceania (55.02%) as shown in Fig. 3.

*Vaccine candidate construction and evaluation*

All the individual epitopes were linked together using different linkers, depending upon there types. GPGPG linker is used to link MHC class I epitopes and AAY linker is used to link all the epitopes of MHC class II epitopes. B cell epitope is also linked to MHC class I epitopes with the help of GPGPG linker. An adjuvant is added at the N terminal of vaccine candidate's construct, to provide the vaccine candidate's more stability and improve its solubility in the expression system. This adjuvant is linked with B cell epitope using EAAAK linker. The adjuvant we used is heat-liable enterotoxin MUTANT S63K which can be retrieved from RCSB.org using PDB ID 1LT4 [32].




a b

**Figure 2:** (a) Secondary structure of M antigen of MERS-CoV and (b) tertiary structure of M antigen predicted by I-TASSER

**Table 1:** Potential MHC class I epitopes candidates for vaccine and their toxicity vaxijen value and allergenicity

| Epitope no. | Epitopes | VaxiJen Value | Allerginicity | Toxicity |
|---|---|---|---|---|
| 1 | AIIKDWNFA | 1.286 | Non-allergen | Non-toxic |
| 2 | FMRTGSWWSF | 0.7623 | Non-allergen | Non-toxic |
| 3 | ITADIELALL | 0.945 | Non-allergen | Non-toxic |
| 4 | LITIVLQYGY | 0.8467 | Non-allergen | Non-toxic |
| 5 | LLITIVLQY | 0.8719 | Non-allergen | Non-toxic |
| 6 | LWPSSMALSI | 0.6861 | Non-allergen | Non-toxic |
| 7 | NVLIALKMVK | 1.0571 | Non-allergen | Non-toxic |
| 8 | SLIFLLITIV | 0.9646 | Non-allergen | Non-toxic |
| 9 | TADIELALLR | 0.7197 | Non-allergen | Non-toxic |

**Table 2:** Potential MHC class II epitopes candidates for vaccine and their toxicity vaxijen value and allergenicity

| Epitope no. | Epitopes | VaxiJen Value | Allergenicity | Toxicity |
|---|---|---|---|---|
| 1 | DWNFAWSLI | 2.3238 | Non-allergen | Non-toxic |
| 2 | FAWSLIFLL | 1.0535 | Non-allergen | Non-toxic |
| 3 | IIKDWNFAW | 1.6834 | Non-allergen | Non-toxic |
| 4 | ITADIELAL | 0.9197 | Non-allergen | Non-toxic |
| 5 | IVLQYGYPS | 0.9152 | Non-allergen | Non-toxic |
| 6 | LKMAGMHFG | 0.9909 | Non-allergen | Non-toxic |
| 7 | MRTGSWWSF | 1.0057 | Non-allergen | Non-toxic |
| 8 | NFAWSLIFL | 1.6504 | Non-allergen | Non-toxic |
| 9 | PNVLIALKM | 1.1168 | Non-allergen | Non-toxic |
| 10 | VLIALKMVK | 1.28954 | Non-allergen | Non-toxic |
| 11 | WNFAWSLIF | 1.3815 | Non-allergen | Non-toxic |

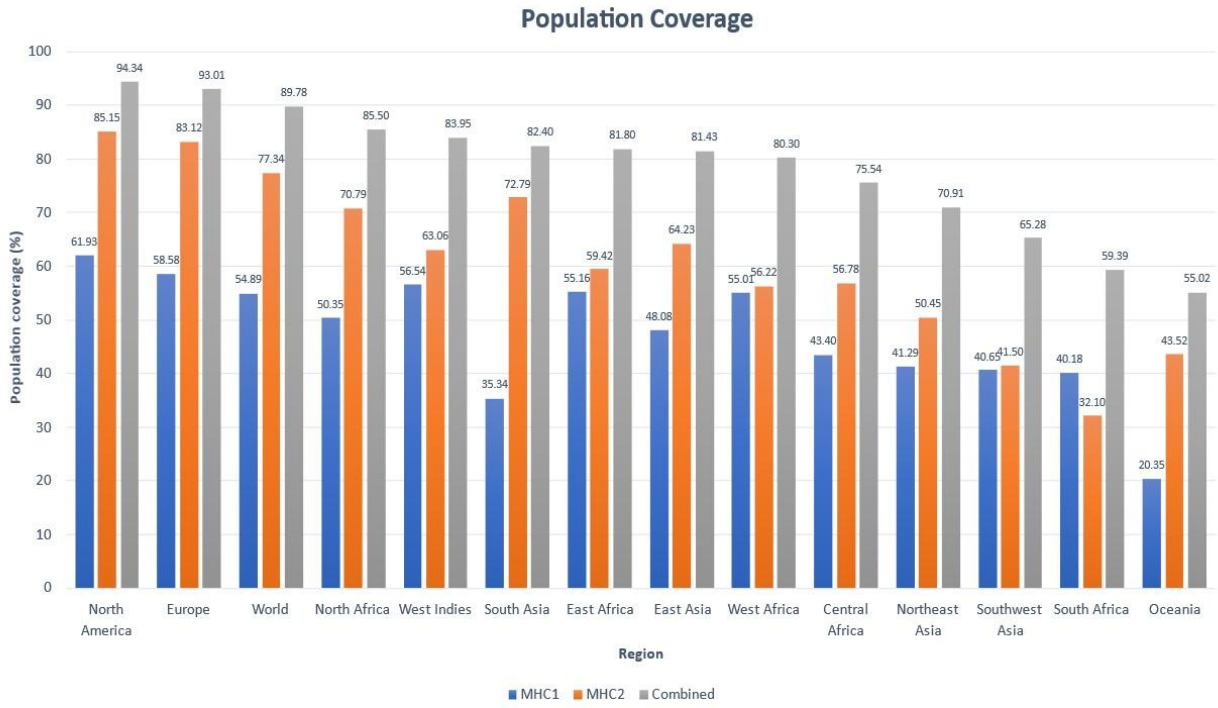



**Figure 3:** Population coverage of MHC class I and MHC class II epitopes

FASTA sequence of adjuvant is

APQTITELCSEYRNTQIYTINDKILSYTES
MAGKREMVIITFKSGETFQVEVPGSQHI
DSQKKAIERMKDTLRITYLTETKIDKLCV
WNNKTPNSIAAISMKN

At the end of the vaccine candidate a 6x Histidine tag is attached for identification of our specific protein (vaccine candidate) in a complex or mixture of proteins. The final vaccine candidate's construct has an adjuvant (heat-liable enterotoxin MUTANT S63K) attached to b cell with linker EAAAK and then GPGPG linker is used to attach B cell epitope to MHC class I epitopes and all the MHC class I epitopes. To link MHC class I and MHC class II epitopes we used AAY and all the epitopes of MHC class II are also linked by AAY linker, the arrangement of adjuvant, linkers and epitopes can be seen in Fig. 4.

| Vaccine Construct | |
|---|---|
| Adjuvant | Heat-liable enterotoxin MUTANT S63K which |
| Linker | EAAAK |
| B-cell epitope | KMVKRQSYGTN |
| Linker | GPGPG |
| MHC-I epitope | AIIKDWNFA |
| Linker | GPGPG |
| MHC-I epitope | FMRTGSWWSF |
| Linker | GPGPG |
| MHC-I epitope | ITADIELALL |
| Linker | GPGPG |
| MHC-I epitope | LITIVLQYGY |
| Linker | GPGPG |
| MHC-I epitope | LLITIVLQY |
| Linker | GPGPG |
| MHC-I epitope | LWPSSMALSI |
| Linker | GPGPG |
| MHC-I epitope | NVLIALKMVK |
| Linker | GPGPG |
| MHC-I epitope | SLIFLLITIV |
| Linker | GPGPG |
| MHC-I epitope | TADIELALLR |
| Linker | AAY |
| MHC-II epitope | DWNFAWSLI |
| Linker | AAY |
| MHC-II epitope | FAWSLIFLL |
| Linker | AAY |
| MHC-II epitope | IIKDWNFAW |
| Linker | AAY |
| MHC-II epitope | ITADIELAL |
| Linker | AAY |
| MHC-II epitope | IVLQYGYPS |
| Linker | AAY |
| MHC-II epitope | LKMAGMHFG |
| Linker | AAY |
| MHC-II epitope | MRTGSWWSF |
| Linker | AAY |
| MHC-II epitope | NFAWSLIFL |
| Linker | AAY |
| MHC-II epitope | PNVLIALKM |
| Linker | AAY |
| MHC-II epitope | VLIALKMVK |
| Linker | AAY |
| MHC-II epitope | WNFAWSLIF |
| 6X histidineTag | HHHHHH |

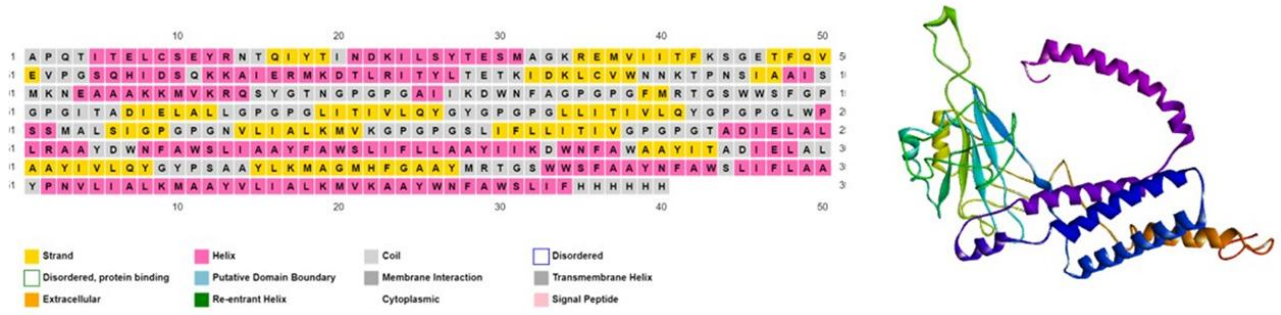
**Figure 4:** Final vaccine candidate construct with 3 types of epitopes, an adjuvant and linkers (linking epitopes and adjuvant together)

Final vaccine candidate construct

APQTITELCSEYRNTQIYTINDKILSYTES
MAGKREMVIITFKSGETFQVEVPGSQHI
DSQKKAIERMKDTLRITYLTETKIDKLCV
WNNKTPNSIAAISMKNEAAAKKMVKRQ
SYGTNGPGPGAIIKDWNFAGPGPGFMR
TGSWWSFGPGPGITADIELALLGPGPGLI
TIVLQYGYGPGPGLLITIVLQYGPGPGLW
PSSMALSIGPGPGNVLIALKMVKGPGPG
SLIFLLITIVGPGPGTADIELALLRAAYDW
NFAWSLIAAYFAWSLIFLLAAYIIKDWNF
AWAAYITADIELALAAYIVLQYGYPSAAYL
KMAGMHFGAAYMRTGSWWSFAAYNFA
WSLIFLAAYPNVLIALKMAAYVLIALKMV
KAAYWNFAWSLIFHHHHHH

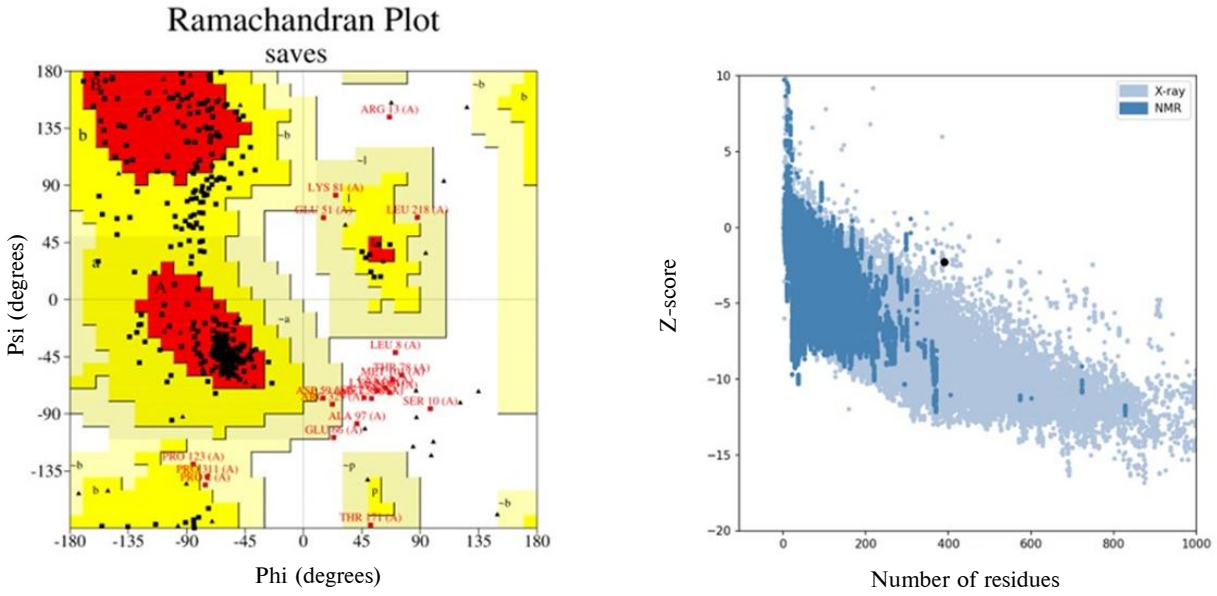
This vaccine candidate sequence is 390 amino acids long and is non allergen, nontoxic and have VaxiJen value of 0.6625 predicted by AllerTop, ToxinPred and VaxiJen server 2.0 respectively. Then Prediction of solubility of vaccine candidate's sequence for expression in Escherichia coli is predicted by SOLUPROT V1.0. solubility score here is 0.828, means it has better solubility in Escherichia coli and sequence can be subjected towards further processing's. The secondary and tertiary structure of vaccine candidate is shown in Fig. 5. PSIPRED tool is used to predict the secondary structure of the vaccine candidate's construct, demonstrating it have 153 amino acids involved in constructing alpha-helix, 99 amino acids are involved in making beta-pleated sheets and remaining 138 amino acids makes up coils. Tertiary structure is predicted using I-TASSER. This 3D structure has C-score of –1.58 and Estimated TM-score $= 0.52 \pm 0.15$, Estimated RMSD $= 10.4 \pm 4.6$ Å indicating a good 3D structure.

The predicted tertiary structure is then subjected towards GALAXYWEB server for its refinement, the refined structure is then used in furthur analysis. To evaluate the model's validity, protein structure comparisons and analyses were conducted using ProSA-web and Ramachandran plot. The Z-score for the most accurate 3D model was determined by ProSA Web. The vaccine candidate's 3D refined structure obtained a Z-score of –2.28, which falls within the typical range of scores observed for native proteins of similar size. The ERRAT Value [33] of 3D structure of vaccine candidate is 87.9581 [33]. The Ramachandran plot analysis predicted by UCLA PROCHECK reveals that the majority of residues (96.3%) are located within the favored and allowed regions out of which 67.3% are in mostfavoured region, 27.2%




a b

**Figure 5:** (a) Secondary structure of vaccine candidate predicted by PsiPred and (b) tertiary structure of vaccine candidate predicted by I-TASSER




a b

**Figure 6:** Structure validation of vaccine candidate tertiary structure using (a) UCLA Ramachandran plot and (b) ProSA web Z-score

residues in additional allowed region and 1.8% in generously allowed region with only 3.7% residing in the outlier region. There are 2 end residues, 37 glycine residues and 24 proline residues. The ramachandran plot and ProsaWeb's Z-score plot can be seen in Fig. 6. The outcomes indicate that the quality of the model designed is considered acceptable.

*Immune simulation*

Using CIMMSIM server immune simulation is predicted, using default parameters as, simulation volume is 10, random seed is 12345, number of steps are 100, number of injection is one and the vaccine injection with no LPS which indicates that this vaccine candidate will have its effect in human body for 33 days [34]. The graphs generated by CIMMSIM are shown in Fig. 7.

*Cloning*

The FASTA sequence of vaccine candidate is back translated into nucleotide sequence using EMBOSS backtranseq tool, jcat.de is used to improve this sequence. By using snap gene this sequence is converted into map, and then inserted into a vector (Pbr322) by inducing cut in it at BsaAI site. Our vaccine candidate's map is 1170 bp long and vector has 4361 base pairs, in final cloning system we had vector+ vaccine candidate map of 5531 base pairs. The maps can be visualized in Fig. 8.

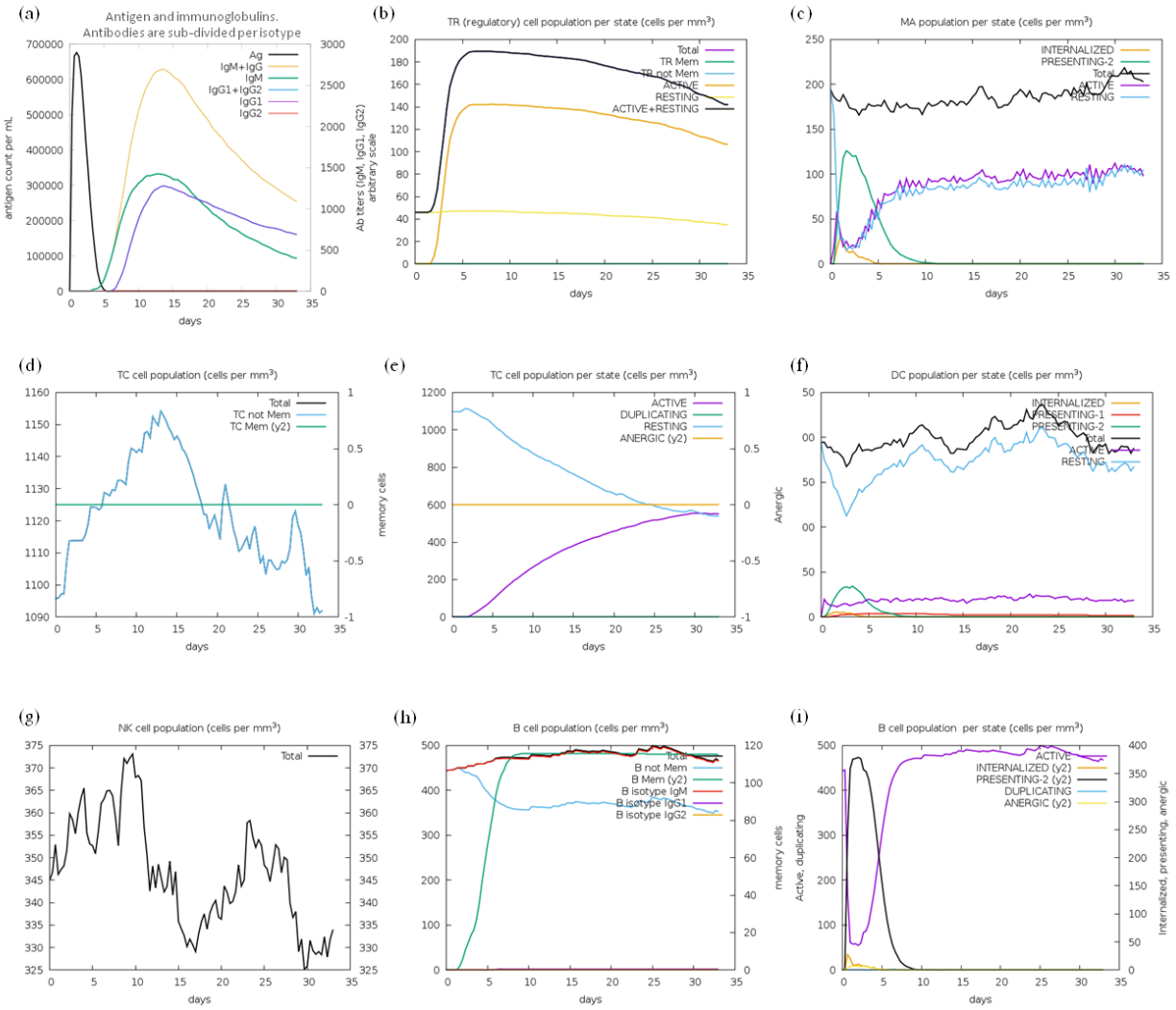



**Figure 7:** Immune Simulation: (a) amount of antigens could be more than 650000 after vaccination in first 5 days and the IgM antibodies are produced approximately after 3 days, IgG1 could be produced at day 5; (b) TR cell production: maximum of 140 active TR cells; (c) macrophages population: active macrophages starts increasing from day 1, but do not increase much for days 5−33; (d) CD8 T-cytotoxic lymphocytes. Memory cells stays constant for 33 days but number of non-memory TC cells fluctuates; (e) active CD8 T-cytotoxic lymphocytes starts increasing from day 3 and could be more than 500 after 33 days; (f) dendritic cells count: total DC could fluctuate between 160−240 in 33 days; (g) Natural Killer cells shows a higher frequency of change in numbers, the highest number of NK cells is 374 at day 10; (h) B lymphocytes memory cells increase from day 3−8 and then stay constant but non-memory cells decreases; (i) active B lymphocytes cells per mmi could be up to 500

### *Docking and Molecular dynamics simulation*

The prepared refined 3D structure of vaccine candidate is then docked with TLR-4 receptor. Cluspro 2.0 server is used for docking [35]. Cluspro will generate many models with different weighted scores. From the results 0 model have minimum energy of –1505.7.1 showing better dock complex. So, it is downloaded for further analysis.

The vaccine candidate exhibited an interaction with both chains of the TLR-4 receptor as shown in Fig. 9. For better understanding of this interaction between TLR-4 receptor and vaccine candidate we used LigPlot+ v.2.2.4 [36]. A strong interaction between TLR-4 Chain A and vaccine candidate can be observed in Fig. 10.

Vaccine candidate also interacted with the chain B of TLR-4, but only a few residues participated in this interaction as shown in Fig. 11.

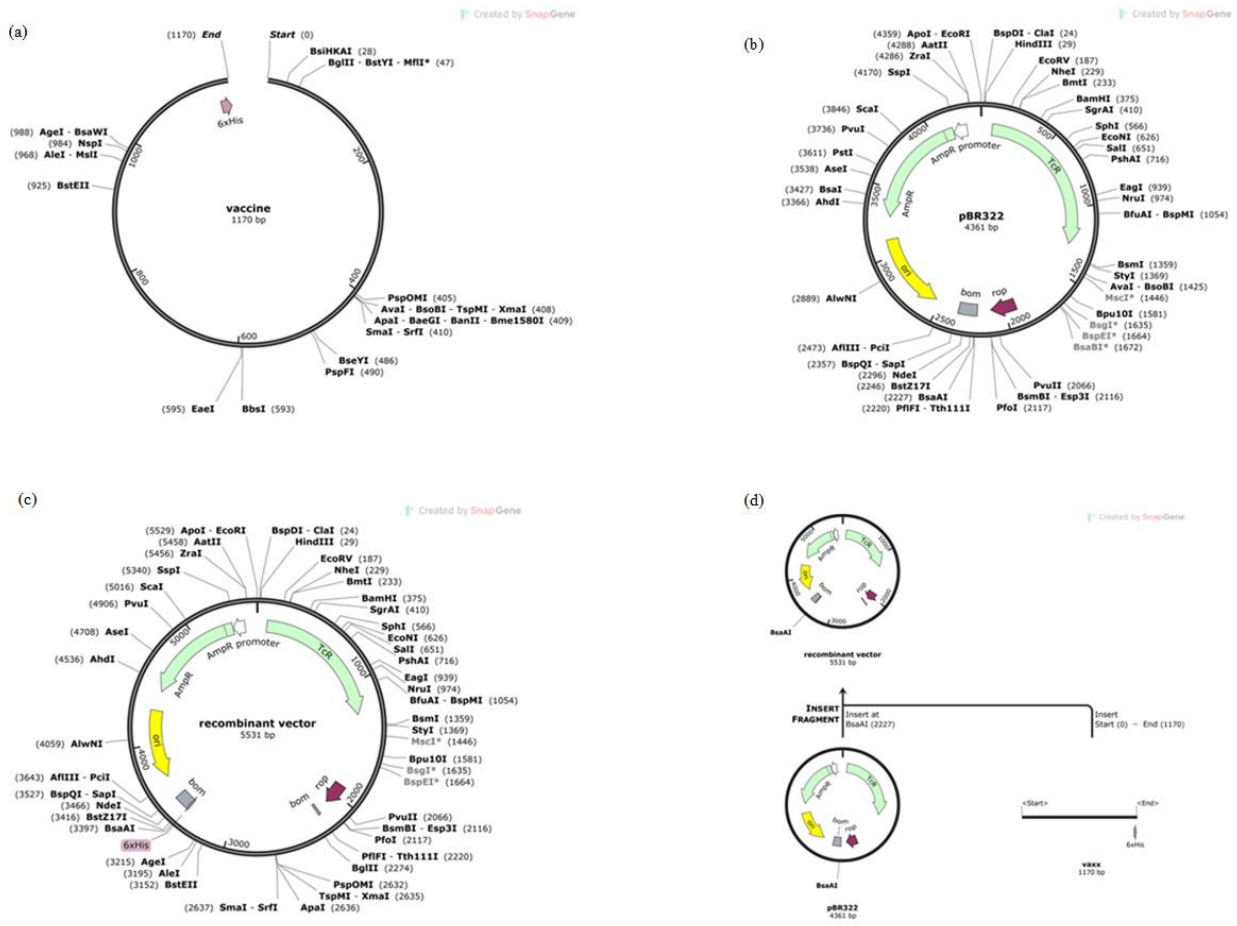



**Figure 8:** Cloning of vaccine candidate using Snapgene (a), map of vaccine candidate (b), pBR322 vector (c) recombinant plasmid having vaccine construct nucleotide sequence, (d) insertion of vaccine candidate in Pbr322 by inducing cut on BsaAI

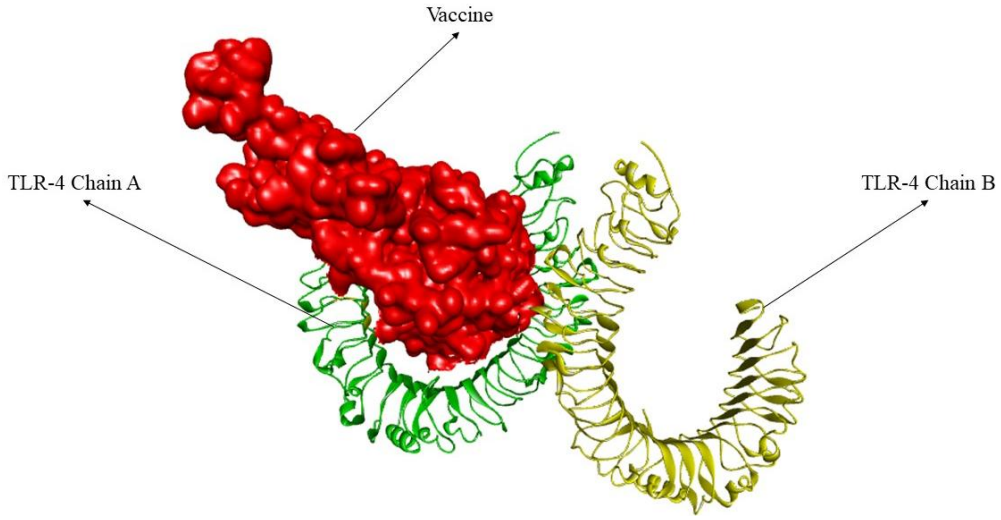



**Figure 9:** The complex formed by the docking of vaccine candidate and TLR-4 receptor

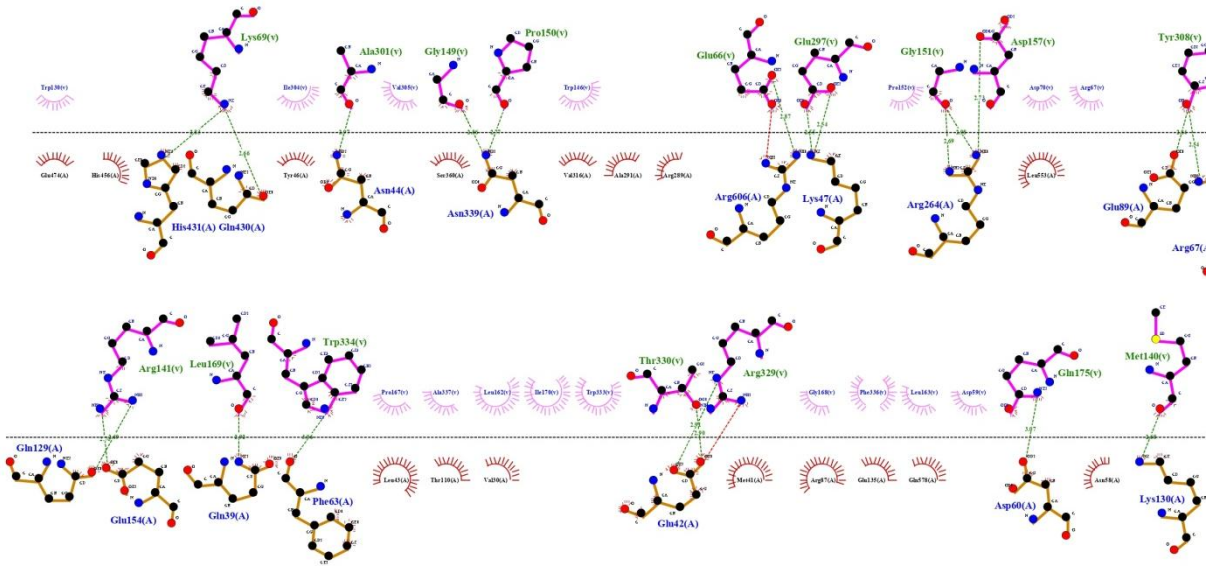



**Figure 10:** Interaction between TLR-4 receptor's chain A and vaccine candidate visualized by DIMPLOT tool of LigPlot+, green lines show hydrogen bonds

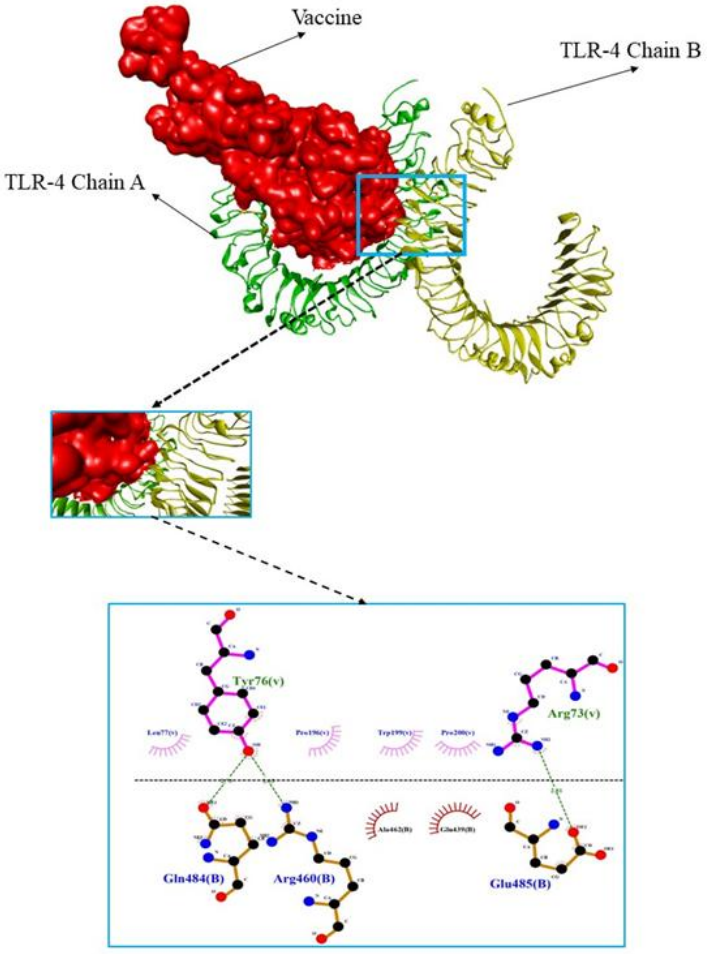



**Figure 11:** Interaction between chain B of TLR-4 and vaccine candidate

The best scored model is than subjected towards MD simulation. MD simulation is performed using GROMACS gmx, version 2021.4 and CHARMm36 force field at 100 ns. The docked complex was positioned at the center of a triclinic box with a space of 1.5 nm from all edges and solvated with the TIP3P water model. Sodium-Chloride (NaCl) was added at physiological concentrations. After that, 3 independent runs were simulated, and the systems were minimized with different shutdown criteria using the steep of the steepest descent method to ensure independent simulations. Next, for each system, we carried out two equilibration steps: a 1 ns NVT equilibration followed by a 10 ns NPT equilibration. After the equilibrations, two systems were run for 50 ns, and the last one for 100 ns of a production run were performed. Generated graphs and files were visualized using qtGrace [30], Microsoft Excel 2019 and VMD [31]. Fig. 12 shows all graphs generated after the MD simulation of complex.

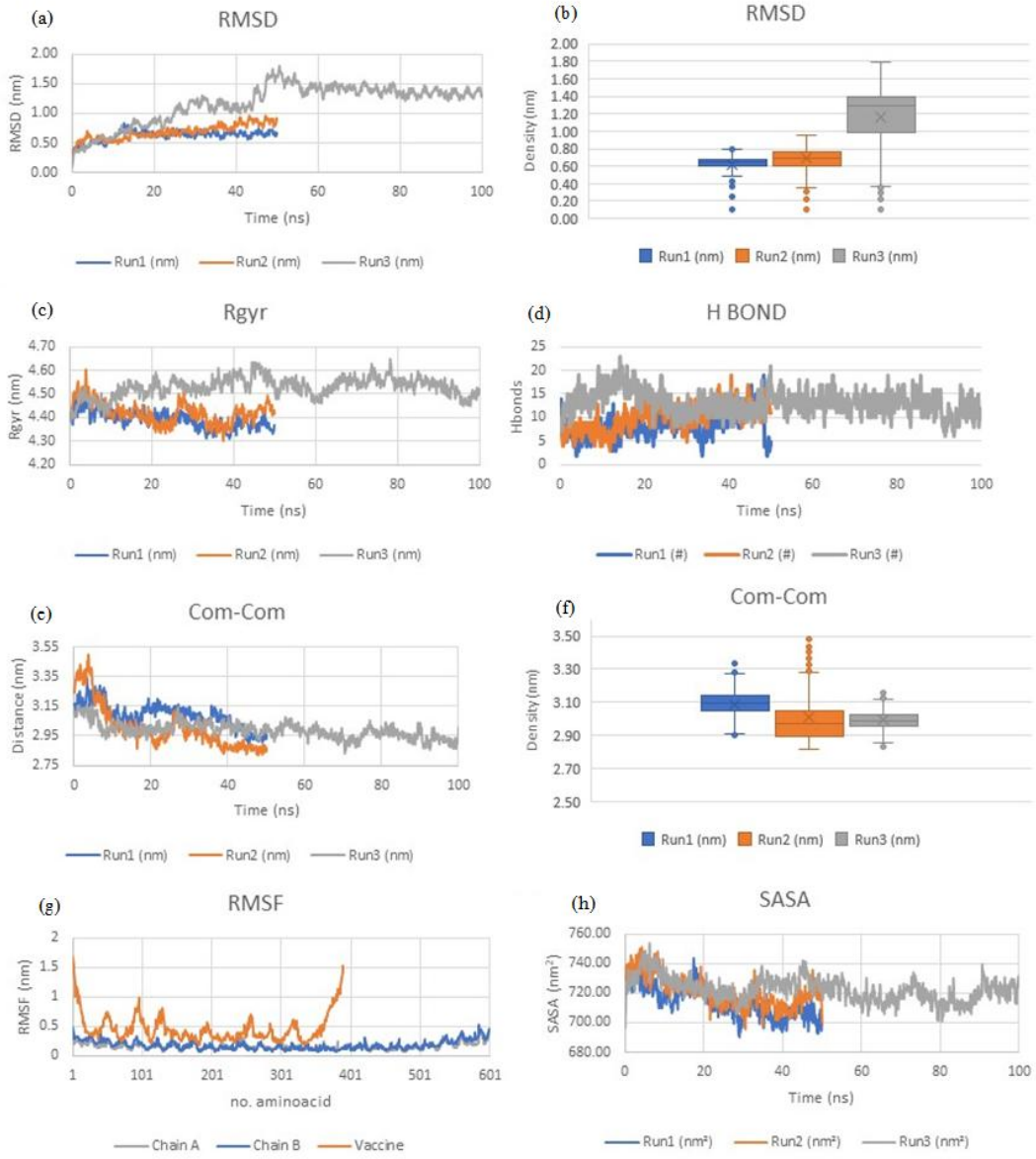



**Figure 12:** (a) RMSD at 0.1 ns all the three runs was at 0.11nm butt gradually it increased and 1st 50 ns run ended up on 0.64nm and 2nd 50 ns was at 0.88nm and at 100 ns the RMSD value is 1.31; (b) comparison of RMSD density of all three runs; (c) rate of gyration is 4.39 nm at 0.1 for all runs but fluctuate to 4.37 nm and 4.41 nm for 50 ns runs and 4.50 nm for 100ns runs; (d) number of Hydrogen Bond fluctuations; (e) center of mass gradually decreases for all three runs; (f) comparison of density of Center of mass of all three runs; (g) shows RMSF of all three chains, Chain A and Chain B are of receptor both constitutes of 601 residues and third chain is of vaccine candidate which constitutes of 390 residues; (h) SASA value is 696.58 on 0.1 ns and is 699.10 nm and 669.63 nm for 2 50 ns runs and 723.62 at 100 ns

## Discussion

Our study demonstrates the feasibility and utility of *in silico* vaccine candidate design in combating emerging viral threats such as MERS-CoV. By leveraging computational modeling and bioinformatics tools, we have expedited the identification and optimization of a potential vaccine candidate targeting the M protein of MERS-CoV. Middle east respiratory syndrome is an infrequent yet serious respiratory sickness. So, designing a vaccine against MERS-CoV is really a need of time. Researchers have employed immunoinformatics to create cutting-edge multi-epitope-driven vaccine models for viruses like MERS-CoV, SARS-CoV-2, Ebola [37, 38]. Shafi Mahmud and colleagues also prepared an *in silico* vaccine candidate against MERS-CoV but they used S glycoprotein as an antigen [39]. At present, there is neither a FDA approved drug, vaccine nor a targeted treatment available for MERS. The vaccine candidate was developed to confer immunity by utilizing multiple small antigenic peptide fragments.

We retrieved FASTA sequence of M protein of MERS-CoV and then predicted three types of epitopes B cell epitopes, MHC class I epitopes and MHC class II epitopes using IEDB. Top epitopes were chosed on the basis of their allergenicity, antigenicity and IC50 value. So, our vaccine candidate has three types of epitopes (1 B cell, 9 MHC class I, 11 MHC class II), linkers were used to link all epitopes, a 6x Histidine tag added at the end of vaccine candidate and an adjuvant was added in start of vaccine candidate's construct. Heat-liable enterotoxin MUTANT S63K was used as an adjuvant and is linked to B cell epitope using EAAAK linker. GPGPG and AAY linkers were used to link MHC class I and MHC class II epitopes respectively. After construction of vaccine candidate, we performed *in silico* cloning in SnapGene using pBR322 as vector and then we predicted tertiary structure of constructed vaccine candidate using I-Tasser. This 3D structure has a C-score of –1.58 and Estimated TM-score = $0.52 \pm 0.15$ Estimated RMSD = $10.4 \pm 4.6$ Å indicating a good 3D structure. and then we refined this tertiary structure using GalaxyRefine. This structure was docked with TLR-4 using Cluspro 2.0. the best scored docked complex was downloaded and subjected towards MD simulations. The RMSD and RMSF characteristics of the vaccine candidates remained under 2 Å throughout the simulation period. These outcomes establish the integrity of the vaccine candidate complexes and their limited mobility under the simulation conditions. As our designed vaccine candidate, molecular docking and molecular dynamics simulation is completed based on computational programs and softwares, there are chances that the predicted biological responses and interactions are not 100% accurate. The complexity of biological systems can be difficult to replicate with precision in a digital environment. although our designed vaccine candidate proved good in molecular docking and molecular dynamics simulations, but it must be validated through in vitro and in vivo experimentations. The MERS virus may undergo antigenic variation, which could affect the effectiveness of the vaccine designed against a specific M antigen sequence. Continuous monitoring and potential updates to the vaccine design may be necessary to address viral mutations. The immune simulation study suggested that our developed vaccine candidate is probable to elicit an adequate immune response upon subsequent exposure following the final injection.

## Conclusions

We used modern bioinformatics approaches to design vaccine candidate against MERS-CoV using Matrix protein as an antigen. This vaccine candidate constituents upon 391 amino acids and 1170 nucleotide if we back translate the amino acid sequence using EMBOSS Backtranseq tool. Using snapgene *in silico* cloning of vaccine candidate was performed and pBR322 was used as a vector. The 3D structure of vaccine candidate is docked with TLR-4 receptor using Cluspro 2.0 and then the best scored model is subjected towards Molecular dynamics simulation. MD simulation is performed using GROMACS gmx, version 2021.4. The *in silico* immune simulation confirmed immune cell response against antigen. The *in silico* vaccine candidates developed against MERS-CoV earlier was based on using nucleocapsid protein and spike protein as antigen, meanwhile *in silico* cloning was also not reported in those works. But we used M protein to construct a vaccine candidate against MERS-CoV, as it is the most abundant structural protein in the virus and is crucial for the assembly and shape of the virus. It interacts with other structural proteins to form the viral envelope. we used Snapgene for *in silico* cloning of our vaccine candidate, which facilitates virtual optimization and valida-

tion of our back translated vaccine candidate's construct. However, the experimental support is indispensable to certify the vaccine candidate use against MERS.

## Interests disclosure

No potential conflict of interest is reported by the authors.

## References


[1] Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med. 2012 Nov 8;367(19):1814-20. DOI: 10.1056/NEJMoa1211721

[2] Munster VJ, Adney DR, van Doremalen N, Brown VR, Miazgowicz KL, Milne-Price S, et al. Replication and shedding of MERS-CoV in Jamaican fruit bats (Artibeus jamaicensis). Sci Rep. 2016 Feb 22;6:21878. DOI: 10.1038/srep21878

[3] Ali MA, Shehata MM, Gomaa MR, Kandeil A, El-Shesheny R, Kayed AS, et al. Systematic, active surveillance for Middle East respiratory syndrome coronavirus in camels in Egypt. Emerg Microbes Infect. 2017 Jan 4;6(1):e1. DOI: 10.1038/emi.2016.130

[4] Alhakeem RF, Midgley CM, Assiri AM, Alessa M, Al Hawaj H, Saeed AB, et al. Exposures among MERS case-patients, Saudi Arabia, January-February 2016. Emerg Infect Dis. 2016 Nov;22(11):2020-2. DOI: 10.3201/eid2211.161042

[5] Assiri A, Al-Tawfiq JA, Al-Rabeeah AA, Al-Rabiah FA, Al-Hajjar S, Al-Barrak A, et al. Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study. Lancet Infect Dis. 2013 Sep;13(9):752-61. DOI: 10.1016/S1473-3099(13)70204-4

[6] European Centre for Disease Prevention and Control. MERS-CoV worldwide overview [Internet] ECDC; 2023 [cited 2024 Jun 10]. Available from: https://www.ecdc.europa.eu/en/middle-east-respiratory-syndrome-coronavirus-mers-cov-situation-update

[7] Zhou Y, Jiang S, Du L. Prospects for a MERS-CoV spike vaccine. Expert Rev Vaccines. 2018 Aug;17(8):677-86. DOI: 10.1080/14760584.2018.1506702

[8] Ramadan N, Shaib H. Middle East respiratory syndrome coronavirus (MERS-CoV): A review. Germs. 2019;9(1):35-42. DOI: 10.18683/germs.2019.1155

[9] van Boheemen S, de Graaf M, Lauber C, Bestebroer TM, Raj VS, Zaki AM, et al. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. mBio. 2012 Nov 20;3(6):e00473-12. DOI: 10.1128/mBio.00473-12

[10] Perrier A, Bonnin A, Desmarets L, Danneels A, Goffard A, Rouillé Y, et al. The C-terminal domain of the MERS coronavirus M protein contains a trans-Golgi network localization signal. J Biol Chem. 2019 Sep 27;294(39):14406-21. DOI: 10.1074/jbc.RA119.008964

[11] de Haan CA, Rottier PJ. Molecular interactions in the assembly of coronaviruses. Adv Virus Res. 2005;64:165-230. DOI: 10.1016/S0065-3527(05)64006-7

[12] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8(1):4. DOI: 10.1186/1471-2105-8-4

[13] Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2021 Jan 8;49(D1):D10-7. DOI: 10.1093/nar/gkaa892

[14] Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. 2019 Jan 8;47(D1):D339-43. DOI: 10.1093/nar/gky1006

[15] Dimitrov I, Flower DR, Doytchinova I. AllerTOP - a server for in silico prediction of allergens. BMC Bioinformatics. 2013;14 Suppl 6(Suppl 6):S4. DOI: 10.1186/1471-2105-14-S6-S4

[16] Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP. In silico approach for predicting toxicity of peptides and proteins. PLoS One. 2013 Sep 13;8(9):e73957. DOI: 10.1371/journal.pone.0073957

[17] Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press; 2005. p. 571-607. DOI: 10.1385/1-59259-890-0:571

[18] McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics. 2000 Apr;16(4):404-5. DOI: 10.1093/bioinformatics/16.4.404

[19] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015 Jan;12(1):7-8. DOI: 10.1038/nmeth.3213

[20] Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res. 2012 Jul;40(Web Server issue):W294-7. DOI: 10.1093/nar/gks493

[21] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007 Jul;35(Web Server issue):W407-10. DOI: 10.1093/nar/gkm290

[22] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Mol Biol. 1996 Dec 6;264(3):585-602. DOI: 10.1006/jmbi.1996.0663

[23] Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. 2022 Jul 5;50(W1):W276-9. DOI: 10.1093/nar/gkac240

[24] Grote A, Hiller K, Scheer M, Мьnch R, Nörtemann B, Hempel DC, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 2005 Jul 1;33(Web Server issue):W526-31. DOI: 10.1093/nar/gki376

[25] Dey J, Mahapatra SR, Patnaik S, Lata S, Kushwaha GS, Panda RK, et al. Molecular characterization and designing of a novel multiepitope vaccine construct against Pseudomonas aeruginosa. Int J Pept Res Ther. 2022;28(2):49. DOI: 10.1007/s10989-021-10356-z

[26] Park BS, Song DH, Kim HM, Choi BS, Lee H, Lee JO. The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. Nature. 2009 Apr 30;458(7242):1191-5. DOI: 10.1038/nature07830

[27] Desta IT, Porter KA, Xia B, Kozakov D, Vajda S. Performance and its limits in rigid body protein-protein docking. Structure. 2020 Sep 1;28(9):1071-81.e3. DOI: 10.1016/j.str.2020.06.006

[28] Bekker H, Berendsen H, Dijkstra E, Achterop S, Vondrumen Rv, Vanderspoel D, et al. Gromacs-a parallel computer for molecular-dynamics simulations. In: DeGroot RA, Nadrchal J, editors. Physics Computing '92. Singapore: World Scientific Publishing; 1993. p. 252-6.

[29] Bjelkmar P, Larsson P, Cuendet MA, Hess B, Lindahl E. Implementation of the CHARMM force field in GROMACS: Analysis of protein stability effects from correction maps, virtual interaction sites, and water models. J Chem Theory Comput. 2010 Feb 9;6(2):459-66. DOI: 10.1021/ct900549r

[30] Turner PJCfC, Land-Margin Research OGIoS, Technology B, OR. XMGRACE, Version 5.1. 19. 2. 2005.

[31] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996 Feb;14(1):33-8. DOI: 10.1016/0263-7855(96)00018-5

[32] van den Akker F, Pizza M, Rappuoli R, Hol WG. Crystal structure of a non-toxic mutant of heat-labile enterotoxin, which is a potent mucosal adjuvant. Protein Sci. 1997 Dec;6(12):2650-4. DOI: 10.1002/pro.5560061220

[33] Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993 Sep;2(9):1511-9. DOI: 10.1002/pro.5560020916

[34] Castiglione F, Bernaschi M. C-ImmSim: playing with the immune response. In: proceedings of 16h International Symposium on Mathematical Theory of Networks and Systems; Jul 5—9, 2004; Leuven.

[35] Porter KA, Xia B, Beglov D, Bohnuud T, Alam N, Schueler-Furman O, et al. ClusPro PeptiDock: efficient global docking of peptide recognition motifs using FFT. Bioinformatics. 2017 Oct 15;33(20):3299-301. DOI: 10.1093/bioinformatics/btx216

[36] Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model. 2011 Oct 24;51(10):2778-86. DOI: 10.1021/ci200227u

[37] Tahir Ul Qamar M, Rehman A, Tusleem K, Ashfaq UA, Qasim M, Zhu X, et al. Designing of a next generation multiepitope based vaccine (MEV) against SARS-COV-2: Immunoinformatics and in silico approaches. PLoS One. 2020 Dec 22;15(12):e0244176. DOI: 10.1371/journal.pone.0244176

[38] Shi J, Zhang J, Li S, Sun J, Teng Y, Wu M, et al. Epitope-based vaccine target screening against highly pathogenic MERS-CoV: An in silico approach applied to emerging infectious diseases. PLoS One. 2015;10(12):e0144475. DOI: 10.1371/journal.pone.0144475

[39] Mahmud S, Rafi MO, Paul GK, Promi MM, Shimu MSS, Biswas S, et al. Designing a multi-epitope vaccine candidate to combat MERS-CoV by employing an immunoinformatics approach. Sci Rep. 2021;11(1):15431. DOI: 10.1038/s41598-021-92176-1

М.Н. Маджид[1,2], А. Ікбал[1], Н. Муртаза[2], Л.Д. Еррера-Зуньїга[3], С. Сіддіке[4], М. Раза[1], М. Хуссейн[1], М. Саджид[1]

[1]Університет Окара, Пенджаб, Пакистан
[2]Університет Центрального Пенджабу, Лахор, Пакистан
[3]Столичний автономний університет, Мехіко, Мексика
[4]Національний університет науки і технологій Юньлінь, Доулю, Тайвань

**РОЗРОБЛЕННЯ МУЛЬТИЕПІТОПНОЇ ВАКЦИНИ-КАНДИДАТА ПРОТИ MERS-CoV: *IN SILICO* ПІДХІД**

**Проблематика.** Коронавірус Близькосхідного респіраторного синдрому (MERS-CoV), асоційований із важкими респіраторними захворюваннями, походить із регіону Близького Сходу. Вірус передається від тварин до людини, при цьому основним резервуарним хазяїном є одногорбий верблюд. Високий рівень смертності від вірусу спонукав до розроблення вакцин і терапевтичних засобів.
**Мета.** Із застосуванням *in silico* підходу розробити потенційну вакцину проти MERS-CoV, зосереджуючись на білку M як антигені.
**Методика реалізації.** Послідовність білка M у форматі FASTA використовувалася для прогнозування епітопів B-клітин і епітопів головного комплексу гістосумісності I і II класу. З цих передбачуваних епітопів були обрані найкращі. До складу вакцини-кандидата входять епітопи, лінкери та мітки. Послідовність конструкції вакцини-кандидата, що складається з 390 амінокислот, було зворотно транскрибовано, оптимізовано і вставлено в плазміду з метою клонування та експресії за допомогою SnapGene.

Тривимірну структуру вакцини-кандидата було доковано з рецептором TLR-4. Моделювання молекулярної динаміки докованого комплексу було виконано за допомогою GROMACS gmx, версія 2021.4.
**Результати.** За допомогою комп'ютерного моделювання й аналізу розроблено нову вакцину-кандидата із перспективними структурними та функціональними властивостями. Наші результати свідчать про те, що розроблена вакцина-кандидат має потенціал викликати сильну імунну відповідь.
**Висновки.** Застосований *in silico* підхід пропонує до розгляду перспективну вакцину-кандидата проти MERS-CoV, яка здатна викликати як гуморальну, так і клітинну імунну відповідь. Цей кандидат має потенціал забезпечити захист широкого спектру від MERS-CoV.

**Ключові слова:** MERS-CoV; вакцина-кандидат; молекулярний докінг; моделювання молекулярної динаміки; біоінформатичні підходи.