JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
https://doi.org/10.1080/07391102.2022.2141887

Design of a novel multiple epitope-based vaccine: an immunoinformatics approach to combat monkeypox

Chandni Hayat[a], Muhammad Shahab[b], Salman Ali Khan[d], Chaoqun Liang[b], Xiuyuan Duan[b], Haleema Khan[c], Guojun Zheng[b] and Zaheer Ul-Haq[d,e]

[a]Department of Biochemistry, Computational Medicinal Chemistry Laboratory, UCSS, Abdul Wali Khan University, Mardan, Pakistan; [b]State Key Laboratories Of Chemical Resources Engineering, Beijing University of Chemical Technology, Beijing, China; [c]Department of Chemistry, UCSS, Abdul Wali Khan University, Mardan, Pakistan; [d]Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan; [e]Third World Center for Science and Technology, H.E.J. Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan

Communicated by Ramaswamy H. Sarma

ABSTRACT

Monkeypox virus is an infectious agent that causes fever, Pneumonitis encephalitis, rash, lymphadenopathy and bacterial infection. The current outbreak of monkeypox has reawakened the global health concern. In the current situation of increasing viral infection, no vaccine or drug is available for monkeypox. Thus, there is an urgent need for viable vaccine development to prevent viral transmission by boosting human immunity. Herein, using immunoinformatics approaches, a multi-epitope vaccine was constructed for the Monkeypox virus. In this connection, B-Cell and T-cell epitopes were identified and joined with the help of adjutants and linkers. The vaccine construct was selected based on promising vaccine candidates and immunogenic potential. Further epitopes were selected based on antigenicity score, non-allergenicity and good immunological properties. Molecular docking reveals strong interactions between TLR-9 and the predicted vaccine construct. Finally, molecular dynamics simulations were performed to evaluate the stability and compactness of the constructed vaccine. The MD simulation results demonstrated the significant stability of the polypeptide vaccine construct. The predicted vaccine represented good stability, expression, immunostimulatory capabilities and significant solubility. Design vaccine was verified as efficient in different computer-based immune response investigations. Additionally, the constructed vaccine also represents a good population coverage in computer base analysis.

ARTICLE HISTORY
Received 1 August 2022
Accepted 24 October 2022

KEYWORDS
Monkeypox; epitope; vaccine; B-cell; T-cell; immunoinformatics

Introduction

Human monkeypox (MPX) is a rare viral infection caused by smallpox-like orthopoxvirus. In April 2022, hundreds of MPXV-positive cases were reported from more than 30 countries. At the end of July 2022, World Health Organization (WHO) declared the recent monkeypox (MPX) outbreak a global health emergency. MPXV is a close family member of the variola virus, the causative agent of smallpox, which killed 300 million people worldwide in the twentieth century (Kenneth, 1993). MPXV has a 200-kilobyte double-stranded DNA genome (Shchelkunov et al., 2001). Although MPXV is not a natural host for the virus, it was named when the virus was discovered in a laboratory for the first time in 1958. Monkeypox is a zoonotic disease, and the infection is frequently severe in children (Magnus et al., 2009). MPXV is transmitted mostly by contact with infected individuals and animals, along with infected material. The diseases can be prevented by avoiding close contact with infected individuals and animals, and also contaminated products.

Monkeypox is a neglected disease prevalent in Central and Western Africa. However, it has recently gained international attention as a result of more than 100 confirmed and suspected cases (by 21 May 2022), affecting more than ten countries especially Australia, North America and Europe. It is predicted that the number of instances would increase worldwide (Potter et al., 2007). Bunge et al. analyze the trend of suspected and confirmed cases of MPXV over the last 50 years. The number of cases has increased over time: only 47 cases were reported during 1970–1979, 356 cases in 1980–1989, 520 cases in 1990–1999, 10,166 cases in 2000–2009 and over 19 thousand cases in 2010–2019 (Bunge et al., 2022). In 2020 and 2021, the WHO Bulletin reported about 6200 and 9400 confirmed and suspected cases, respectively (Impouma et al., 2018). Smallpox immunizations protect against monkeypox, and TPOXX, a smallpox medication, is promising and effective against monkeypox. Two US FDA-approved vaccines that can protect against monkeypox are JYNNEOS and ACAM2000. Unfortunately, there is still a need to develop a safe and effective multi-epitope vaccine against monkeypox.

CONTACT Zaheer Ul-Haq zaheer.qasmi@iccs.edu Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan; Guojun Zheng zhenggj@mail.buct.edu.cn State Key Laboratories Of Chemical Resources Engineering, Beijing University of Chemical Technology, Beijing, China

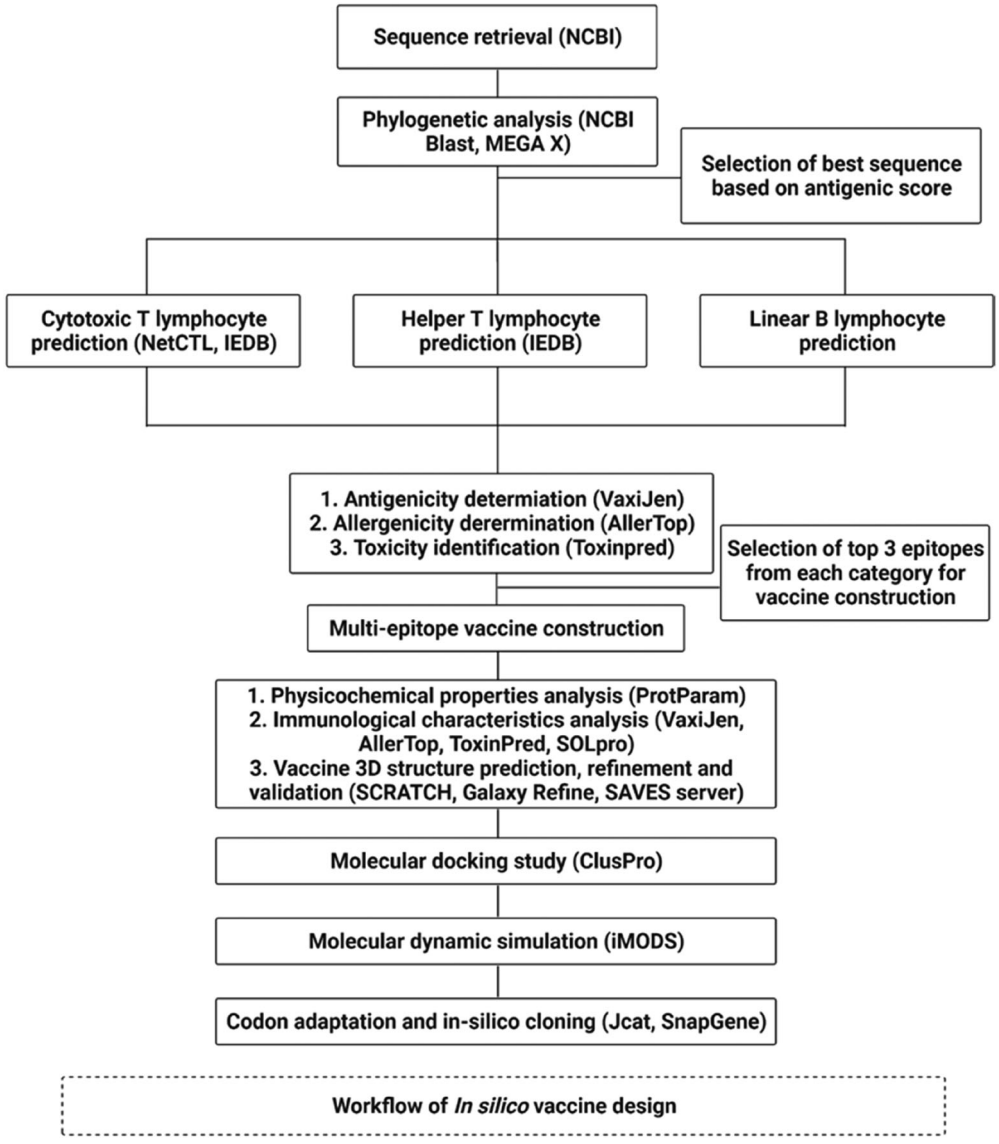

Figure 1. Workflow and tools used in this study.

Resistance to antiviral medications is prevalent, therefore novel treatments that target viral sites are critical for treating monkeypox (Fine et al., 1988). To prevent uncontrollable spread, more intensive surveillance, and research on MPXV biology, history, transmission mode, host interactions, drug targets and vaccine development are extremely crucial. Governments, academics, industries and healthcare systems work together to achieve this goal.

The conventional approach to vaccine development is risky and time-consuming. However, the immunoinformatics-based multi-epitope vaccine approach is an attractive alternative. Different vaccines are developed by using immunoinformatics approaches against human infections caused by H. pylori (Naz et al., 2015), Ebola virus (Bazhan et al., 2019), Marburg virus (Hasan et al., 2019), Mokola Rabies virus (Aa et al., 2017) and Crimean-Congo hemorrhagic virus (Nosrati et al., 2019). In this study, we designed an effective, safe and thermodynamically stable epitope-based vaccine for monkeypox, to trigger innate and adaptive immune responses. Here we employed an immunoinformatics approach to design non-allergic, and top-most antigenic epitopes for vaccine development, and the workflow is presented in Figure 1. The findings provide the way for the construction of the Monkeypox vaccine; however further experimental validation is needed to reduce the Monkeypox disease.

Methodology

Sequence retrieval

The reference sequence of the monkeypox (accession no: AAQ09810.1) was first retrieved from NCBI (https://www.ncbi.nlm.nih.gov/protein/ AAQ09810.1), and then subjected to protein-protein blast (Blastp) against the non-redundant database and the resulting top 6 sequences (including the reference sequence) were retrieved as FASTA format. Multiple sequence alignment was conducted by the MUSCLE v3.6 program. Phylogenetic analysis of all 6 sequences was performed using Mega X. Different immunoinformatics tools and software were used for the development of the multi-epitope vaccine against monkeypox. First, the sequence was submitted to online tools Vaxijen 2.0 (Doytchinova & Flower,

2007a) and AllerTop (Dimitrov et al., 2014) to check antigenicity and allergenicity. Further, the Soluprot (https://loschmidt.chemi.muni.cz/soluprot/) (Hon et al., 2021) tools were employed to check the soluble protein expression in *Escherichia coli*. Physical and chemical characteristics were predicted via the online server Expasy (https://web.expasy.org/protparam/) (Wilkins et al., 2005). Similarly, the secondary structure was predicted via the online server, Psipred (http://bioinf.cs.ucl.ac.uk/psipred) (Jones, 1999).

Prediction of B-cell and T-cell epitopes

Selecting effective B-cell epitopes requires surface accessibility. Thus, the Emini surface accessibility tool was used to evaluate surface accessibility. IEDB Resource (http://tool.iedb.org/main/) (Nielsen et al., 2003) was used to predict the B-cell and T-cell epitopes and the binding scores were calculated for T-cell epitopes with MHC-I and MHC-II. The threshold value was kept at 0.5 and IC_{50} scores were assigned by the online server. A threshold score of more than 0.5 was taken as a good epitope candidate. IC_{50} and binding affinity are inversely proportional to each other. Mean if IC_{50} value is small, epitope binding affinity to MHC-II will be high. IC_{50} values <10 nM $<$ 100 nM <1000 nM means high, intermediate and low binding affinity with MHC-II.

Protein construct allergenicity prediction

To check the antigenic and allergic behavior of B-cells and T-cells epitopes, online servers VaxiJen v2.0 (Doytchinova & Flower, 2007b) and AllerTop (Dimitrov et al., 2014) were used. The epitopes with a higher value than the reference and non-allergic were chosen for further research.

Population coverage

Using the online IEDB Analysis Resource, MHC-I and MHC-II epitopes were used to predict population coverage (tools.iedb.org/population/). A vaccine construct was designed using B and T cell epitopes with high binding affinity, non-allergenicity and antigenicity. Based on earlier works (Bhattacharya et al., 2020; Shams et al., 2020), AAY and GPGPG, linkers were used to connect all MHC-I and MHC-II epitopes, respectively. Further, the EAAAK linker was used to bind the adjuvant to the N-terminal of the vaccine construct (Arai et al., 2001).

Secondary and tertiary structure prediction and validation

PSIPRED 3.3 online server (http://bioinf.cs.ucl.ac.uk/psipred/) and TrRosetta web (Du et al., 2021) server were used to predict the secondary and tertiary structures of the vaccine construct. Galaxy Refine (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) was used to validate the vaccine. For further verification, ProSA-Web (Wiederstein & Sippl, 2007) Ramachandran Plot and PROCHECK (https://saves.mbi.ucla.edu/) (Laskowski et al., 2006) servers were used.

Molecular docking of vaccine construct

The three-dimensional structure of TLR-9 (PDB ID: 3WPF) was downloaded from PDB. The structure was subjected to geometry correction and missing residues were added. All the water molecules were removed. The MOE software was used to correct the protonation states. Then the structure was saved in TLR-9.pdb format. The vaccine construct was docked with the human Toll-like receptor 9 using Cluspro (https://cluspro.bu.edu) (Vajda et al., 2017), the most extensively used docking server, which is based on six energy functions. Cluspro predicted the 10 models in about three hours, based on densely packed low-energy clusters for each docking parameter. This procedure illustrates how to use various choices such as creating extra restrictions files, selecting various energy parameters and analyzing the outcomes. PDBsum (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html) was used to evaluate the vaccine and TLR-9 interactions.

Codon optimization and MD simulation

The JCat tool was used to optimize codons and reverse the vaccination sequence (Grote et al., 2005). The JCat program is also used to guarantee that the vaccination sequence is expressed in a vector with a high level of expression. Three extra parameters were chosen in this tool, including Rho-independent transcription termination, restriction enzyme cleavage sites and bacterial ribosome binding sites. JCat also determines the CAI score and GC content of the vaccine sequence.

MD simulation

iMODS, an open-source modeling server for molecular dynamics simulation and image processing provides a user-friendly interface for internal normal mode analysis (NMA) of deformability, stability, and mobility of the target protein. Users can perform NMA or simulate possible trajectories between two conformations and view the results in 3D, even for several biomolecular complexes. Herein, iMODS was utilized to evaluate the dynamic stability of the TLR-9-vaccine complex.

Results

Sequence retrieval, phylogenetic analysis and sequence prioritization

Protein information of the top 6 protein sequences, obtained by BlastP against the nr-database is depicted in Table 1, along with their properties including antigenicity, allergenicity and toxicity. MUSCLE v3.6 was used for multiple sequence alignment. A phylogenetic tree showing the phylogenetic relatedness among the sequences was constructed using the MEGA X program by neighbor-joining method with a bootstrap replication of 1000, shown in Figure 2. Among the 6 sequences, the protein sequence with accession number NP_536428.1 was found to be the most potent antigenic

Table 1. Protein information of the 6 protein sequences obtained from NCBI by BlastP search.

Accession No	Protein name	Sequence	VaxiJen score	Antigenicity
NP_536428.1	Monkeypox virus Zaire-96-I-16	MKQYIVLACMCLVAAAMPTSLQQSSSSCTEEENKHHMGIDVIIKVTKQDQTPTNDKICQSVTEVTETED DEVSEEVVKGDPTTYYTIVGAGLNMNFGFTKCPKISSISESSDGNTVNTRLSSVSPGQGKDSPAI TREEALAMIKDCEMSIDIRCSEEEKDSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVKNLEFSV RIGDMCEESSELEVKDGFKYVDGSASEGATDDTSLIDSTKLKACV	0.7703	ANTIGEN
AAQ09810	secreted chemokine binding protein	MKQYIVLACMCLVAAAMPTSLQQSSSSCTEEENKHHMGIDVIIKVTKQDQTPTNDKICQSVTEVTETEDD EVSEDDEVSEEVVKGDPTTYYTIVGAGLNMNFGFTKCPKISSISESSDGNTVNTRLSSVSPGQGK DSPAITREEALAMIKDCEMSIDIRCSEEEKDSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVKN LEFSVRIGDMCEESSELEVKDGFKYVDGSASEGATDDTSLIDSTKLKACV	0.7574	ANTIGEN
AAM76335.1	secreted chemokine binding protein	MKQYIVLACMCLVAAAMPTSLQQSSSSCTEEENKHHMGIDVIIKVTKQDQTPTNDKICQSVTEVTETEDD EVSEEVVKGDPTTYYNIVGAGLNMNFGFTKCPKISSISESSDGNTVNTRLSSVSPGQGKDSPAITR EEALAMIKDCEMSIDIRCSEEEKDSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVKNLEFSVRIGD MCEESSELEVKDGFKYVDGSASEGATDDTSLIDSTKLKACVKDGDTWYYLEASGAMKASQWFK VSDKWYYVNGLGALAVNTTVDGYKVNANGEWV	0.7003	ANTIGEN
AAM76334.1	secreted chemokine binding protein [Monkeypox virus] [Streptococcus pneumoniae]	MKQYIVLACMCLVAAAMPTSLQQSSSSCTEEENKHHMGIDVIIKVTKQDQTPTNDKICQSVTEVTETEDDEV SEDDEVSEEVVKGDPTTYYNIVGAGLNMNFGFTKCPKISSISESSDGNTVNTRLSSVSPGQGKDSP AITREEALAMIKDCEMSIDIRCSEEEKDSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVKNLEFSV RIGDMCEESSELEVKDGFKYVDGSASEGATDDTSLIDSTKLKACV	0.7563	ANTIGEN
URF91554.1	hypothetical protein MPXV-SI-2022V52144_00002 [Monkeypox virus]	MKQYIVLACMCLVAAAMPTSLQQSSSSCTEEENKHHMGIDVIIKVTKQDQTPTNDKICQSVTEVTETEDDEV SEEVVKGDPTTYYTIVGAGLNMNFGFTKCPKILSISESSDGNTVNTRLSSVSPGQGKDSPAITREEAL AMIKDCEMSIDIRCSEEEKDSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVKNLEFSVRIGDMCEE SSELEVKDGFKYVDGSASEGATDDTSLIDSTKLKACV	0.7634	ANTIGEN
USS79525.1	chemokine binding protein [Monkeypox virus]	MKQYIVLACMCLVAAAMPTSLXXXXSSCTEEENKHHMGIDVIIKVTKQDQTPTNDKICQSVTEVTETEDDEV SEEVVKGDPTTYYTIVGAGLNMNFGFTKCPKILSISESSDGNTVNTRLSSVSPGQGKDSPAITREEALA MIKDCEMSIDIRCSEEEKDSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVKNLEFSVRIGDMCEESSE LEVKDGFKYVDGSASEGATDDTSLIDSTKLKACV	0.7521	ANTIGEN

protein with a VaxiJen score of 0.7703 and selected to further design a multi-epitope based vaccine for monkeypox.

Physiochemical properties

The sequence of monkeypox (NP_536428.1) was submitted to Protparam (https://web.expasy.org/protparam/), an online tool to calculate the physical and chemical properties of the vaccine. The results reveal that the vaccine construct has a total number of 246 amino acid residues, and a molecular weight of 33547.32 kDa, and the theoretical isoelectric point (PI) value of 7.60. The total number of negatively charged residues (Asp + Glu) was 36 and the total number of positively charged was 37 which indicates that the protein is positively charged, as is the case with isoelectric points over 7.0. An instability index (II) of 58.15 determined by Protparam classified our protein to be stable. The aliphatic index was 98.99, indicating that it is thermo-stable over a wide temperature range. At 0.076, the grand average of hydropathcity (GRAVY) was computed using the chemical formula C1449H2412N394O450S31.

Secondary structure analysis

The monkeypox has 11.4% β-strands, 48.9% α-helices and 39.7% coil structures, according to the secondary structure prediction using PSIPRED and 3D structure prediction using trRosetta (https://yanglab.nankai.edu.cn/trRosetta/) (Figure 3) (Zheng et al., 2021). TMHMM, an online tool, was used to predict transmembrane topology. Surface exposed residues were detected at positions 150–160, 165–175, 185–190 and 195–200, whereas residues from 10–20, 50–80, 85–105 and 115–120 were found inside the transmembrane region. The core area of the chemokine-binding protein of monkeypox was determined to have residues from positions 200–206, 210–225 and 228–235.

B-cell epitope prediction

The FASTA sequence of the monkeypox was subjected to the IEDB online server with default parameters to identify potential B-cell epitopes. Residues with a higher value than the given threshold of 0.5 were designated as B-cell epitopes and are graphically presented in Figure 4. The IEDB server predicted a total number of 6 B-cell epitopes. All six epitopes were deposited to online tools VaxiJen 2.0 and AllerTop to calculate their antigenicity and allergenicity. Among the six epitopes, 3 epitopes were finalized based on good antigenic scores and non-allergic behavior which is presented in Table 2 along with their amino acid sequence, length and position. The antigenicity analysis showed that the minimum antigenicity value was 0.5181 and the maximum value was 0.847. However, the average value of 0.62 was observed.

T-cell epitope prediction

MHC-1 epitopes

Stabilized matrix method was utilized in order to prioritize the selected epitopes using the IC_{50} threshold of 100 nM. IC_{50} value less than 100, representing strong and higher binding affinity of epitopes that bind with MHC-1. A lesser IC_{50} suggests a greater affinity for MHC-I molecules. To maximize affinity for MHC-1 alleles, the total number of epitopes was designed to be fewer than 100. Based on MHC-1 allele interactions and IC_{50} value, 89 epitopes were chosen. A total of 10 epitopes were selected showing good antigenicity, allergenicity and non-toxicity. The antigenic scores of toxic

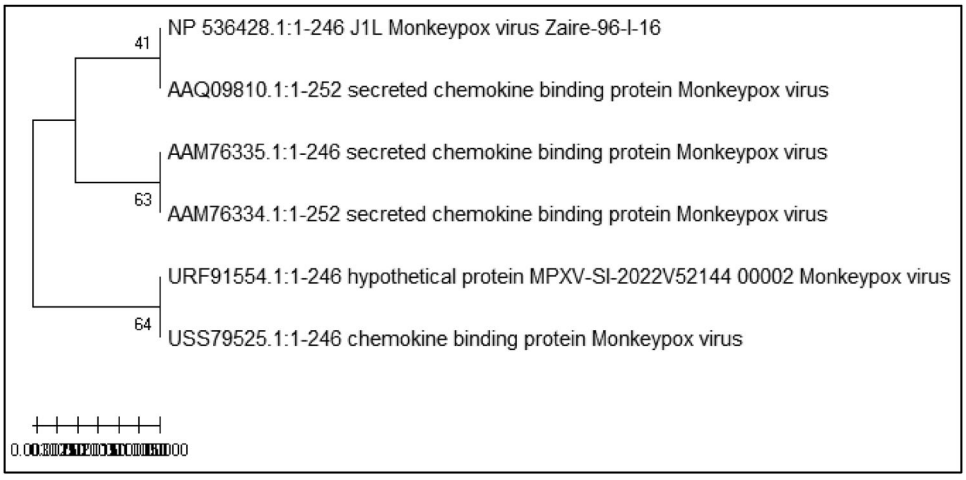

Figure 2. Phylogenetic relationship among the studied protein (AAQ09810.1), reference protein and other proteins obtained from non-redundant database using BlastP. The evolutionary distances were computed using the Poisson correction method (Magnus et al., 2009) and are in the units of the number of amino acid substitutions per site.

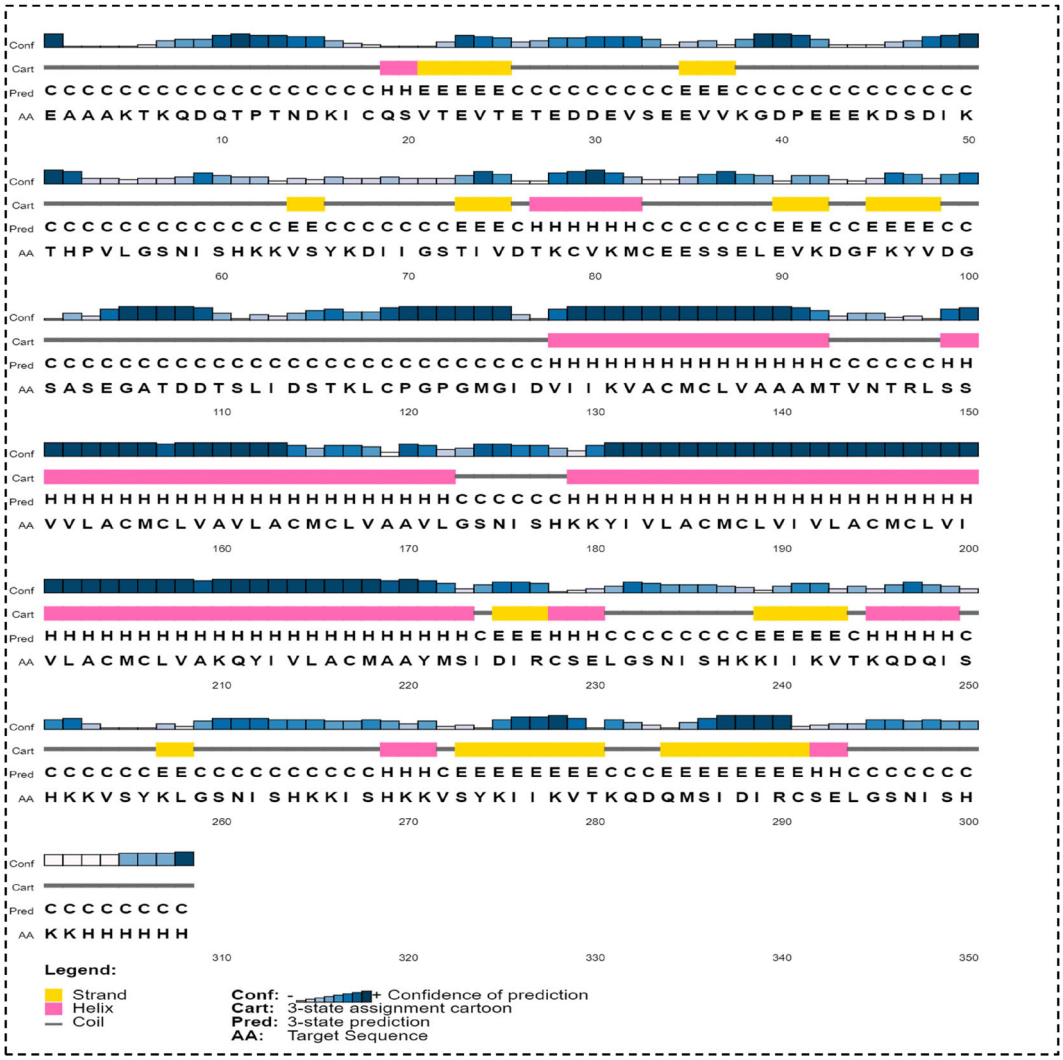

Figure 3. Graphical representation of secondary structure prediction of target protein. H; helix, E; strands and C; coils.

and allergic epitopes were ruled out. The MHC-1 epitopes were finalized that bind to alleles HLA-A*02:06, HLA-B*15:01, HLA-A*68:02 and HLA-A*03:01. NISHKKVSYK has an antigenic score of 1.3013 (Table 3).

MHC-II epitopes

MHC-II alleles interacted with 550 conserved predicted epitopes with IC_{50} less than 60 nM. 30 epitopes were selected among 550

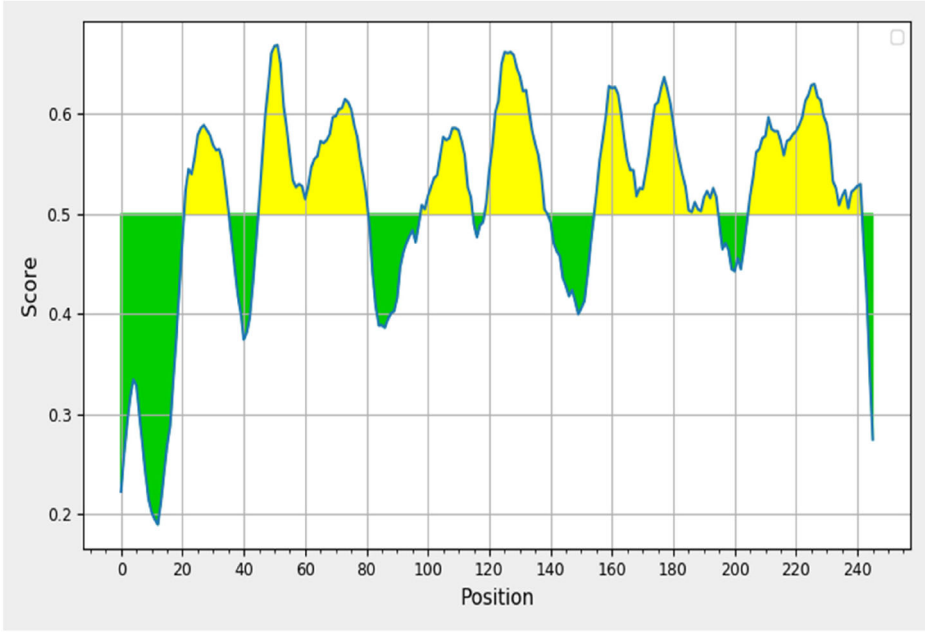

Figure 4. Graphical representation of B-cell epitope (0.5 shows the threshold value of the yellow color epitopic portion while the green color represents the non-epitopic portion).

Table 2. IEDB analysis resource predicts a list of Bepipred linear epitopes.

Start	End	Peptide	Length
46	81	TKQDQTPTNDKICQSVTEVTETEDDEVSEEVVKGDP	36
156	195	EEEKDSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVK	40
206	242	MCEESSELEVKDGFKYVDGSASEGATDDTSLIDSTKL	37

Table 3. IEDB analysis resource predicts a list of Bepipred linear epitopes.

Start	End	Peptide antigenic score	Length
5	14	MGIDVIIKV 0.5798	10
8	17	ACMCLVAAAM 0.9166	10
45	53	TVNTRLSSV 1.1279	9
6	14	VLACMCLVA 0.7627	9
6	15	VLACMCLVAA 0.8605	10
28	37	VLGSNISHKK 0.9404	10
4	13	YIVLACMCLV 1.1068	10
5	13	IVLACMCLV 1.1304	9
5	10	IVLACMCLVA 1.0097	10
2	10	KQYIVLACM 0.7219	9

that interacted with five MHC-II alleles. 13 epitopes were selected for further study based on their allergenicity, toxicity and antigenicity. The epitopes MSIDIRCSE, LGSNISHKK, IIKVTKQDQ, ISHKKVSYK and LGSNISHKK were considered the top binder, with alleles HLA-DRB1*03:01, HLA-DRB5*01:01, HLA-DRB1*07:01, HLA-DRB5*01:01 HLA-DRB4*01:01 and HLA-DRB1*03:01 (Table 4).

Construction of vaccine

Vaccine ensembles were performed by joining a total number of 3 B-cell epitopes and 19 T-cell epitopes (10 MHC-I and 9 MHC-II). 50S ribosomal protein is used as an adjuvant for the construction of vaccines. Through different linkers, adjuvant were combined with B-cell epitope to make a specific immune response. At the C-terminus of the vaccine sequence, a 6× His tag was inserted in order to facilitate the protein identification and purification processes. The final vaccine construct was deposited to the online tools VaxiJen2.0 and AllerTop to evaluate the antigenicity and

Table 4. Antigenicity prediction using the Kolaskar and Tongaonkar technique.

Start	End	Peptide antigenic score	Length
5	13	MSIDIRCSE 2.4214	9
22	30	LGSNISHKK 1.0759	9
40	49	IIKVTKQDQ 1.2531	9
30	38	ISHKKVSYK 1.3341	9
26	34	LGSNISHKK 1.0759	9
15	23	ISHKKVSYK 1.3341	9
39	47	IIKVTKQDQ 1.2531	9
3	11	MSIDIRCSE 1.1304	9
8	16	LGSNISHKK 1.0759	9

allergenicity. Prediction of antigenicity and allergenicity found that the construct was a probable antigen (VaxiJen score 0.8780) and non-allergen. The generated vaccine construct is shown below.

>vaccine protein

EAAAKTKQDQTPTNDKICQSVTEVTETEDDEVSEEVVKGDPEEEK-DSDIKTHPVLGSNISHKKVSYKDIIGSTIVDTKCVKMCEESSELEVKDG-FKYVDGSASEGATDDTSLIDSTKLCPGPGMGIDVIIKVACMCLVAAA-MTVNTRLSSVVLACMCLVAVLACMCLVAAVLGSNISHKKYIVLACM-CLVIVLACMCLVIVLACMCLVAKQYIVLACMAAYMSIDIRCSELGSNIS-HKKIIKVTKQDQISHKKVSYKLGSNISHKKISHKKVSYKIIKVTKQDQMSI-DIRCSELGSNISHKKHHHHHH

Population coverage

The MH Class-I allele was found to be present in 84.06% of the world's population, followed by East Asia (74.06%), South Asia (78.07%), Hong Kong (66.57%), Europe (75.1%), North-East Asia (74.39%), Southwest Asia (69.03%), South America (80.72%), Central America (75.32%) and South Africa (80.56%) as depicted in Figure 5. The lowest population was found in Hong Kong. South America and South Africa have the highest percentage of MH Class-I and MH Class-II alleles in the population followed by East Asia. Three epitopes in MH

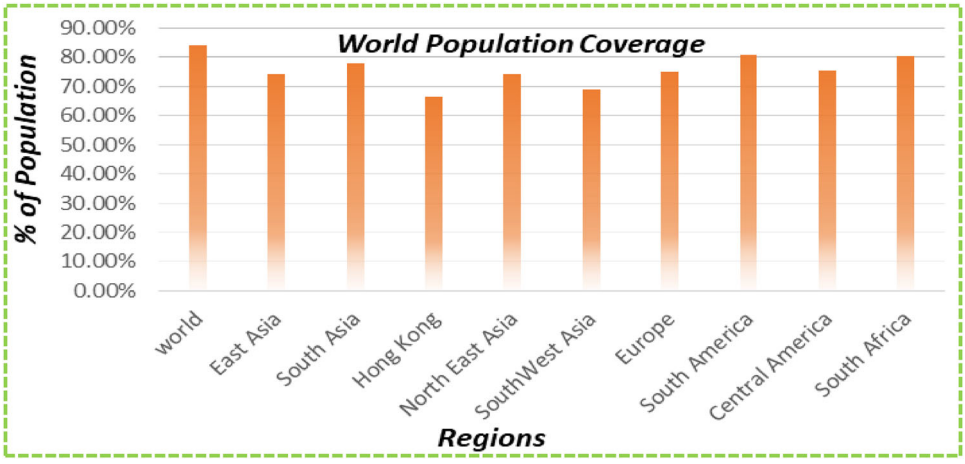

Figure 5. Bar-graph represents the population coverage of final 10 epitopes in different regions of the world.

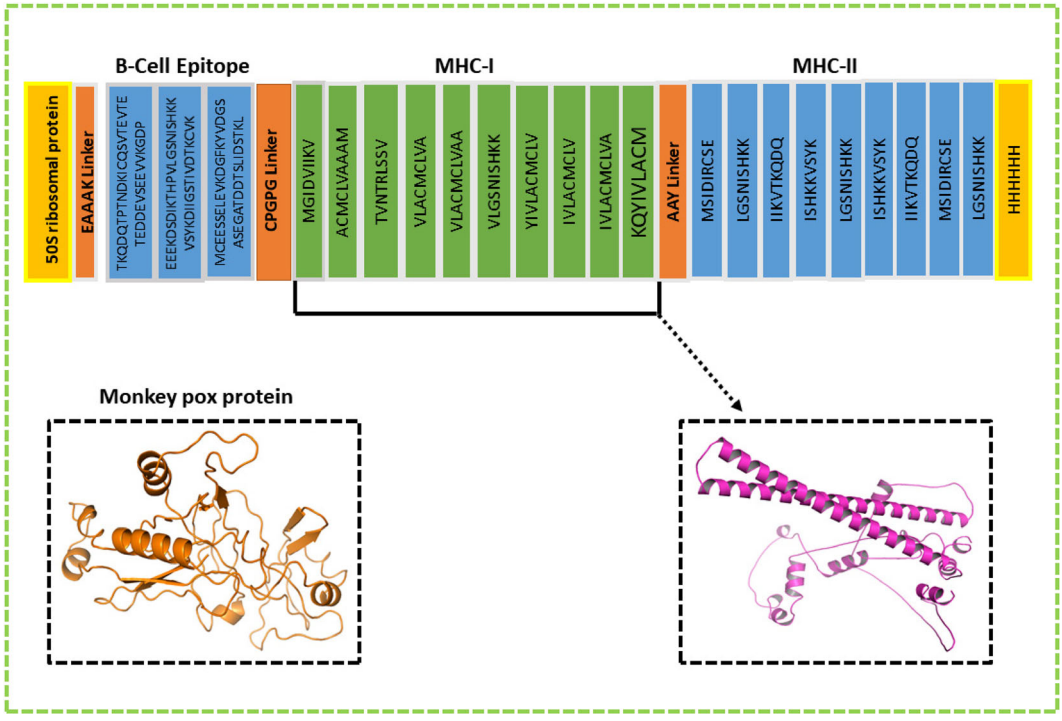

Figure 6. Graphical and 3D view of vaccine construction. 50 s ribosomal protein, EAAAL Linker, CPGPG Linker and AAY Linker are shown.

Class-I (TVNTRLSSV, YIVLACMCLV and IVLACMCLVA) are crucial for the majority of interactions. Five Epitope MH Class-II alleles (LGSNISHKK, MSIDIRCSE, MSIDIRCSE, ISHKKVSYK and LGSNISHKK); describe a considerable coverage in contrast to the whole world population. For IVLACMCLVA, the proportion of concentrated population coverage in the world was anticipated at 70.57%. The population coverage data for the abundant binders to MH Class-I and MH Class-II alleles reveal 88.06 and 75.06% coverage, respectively.

3D structure prediction and validation

The PROSA 3D server was used to predict the 3D structure of the multi-epitope vaccine sequence, as a result, ten structures were predicted for a given query sequence. Model five was selected for further investigation (Figure 6). Using the ERRAT, ProSA-web and PROCHECK servers, the structure was validated and any potential errors in the projected tertiary structure were corrected. The ERRAT server projected the overall quality of the vaccine 3D structure, with an estimated quality score of 95.0%. The Z-score was calculated to determine whether the input structure was within the range of natural proteins of similar size. The computed *Z*-score for the input structure was −6.67, indicating that it was outside the usual range for natural proteins of the same size, as shown in Figure 7. For Ramachandran analysis, the PROCHECK server shows the total number of 308 amino acid residues, 96.9% of the residues in the most favored areas,

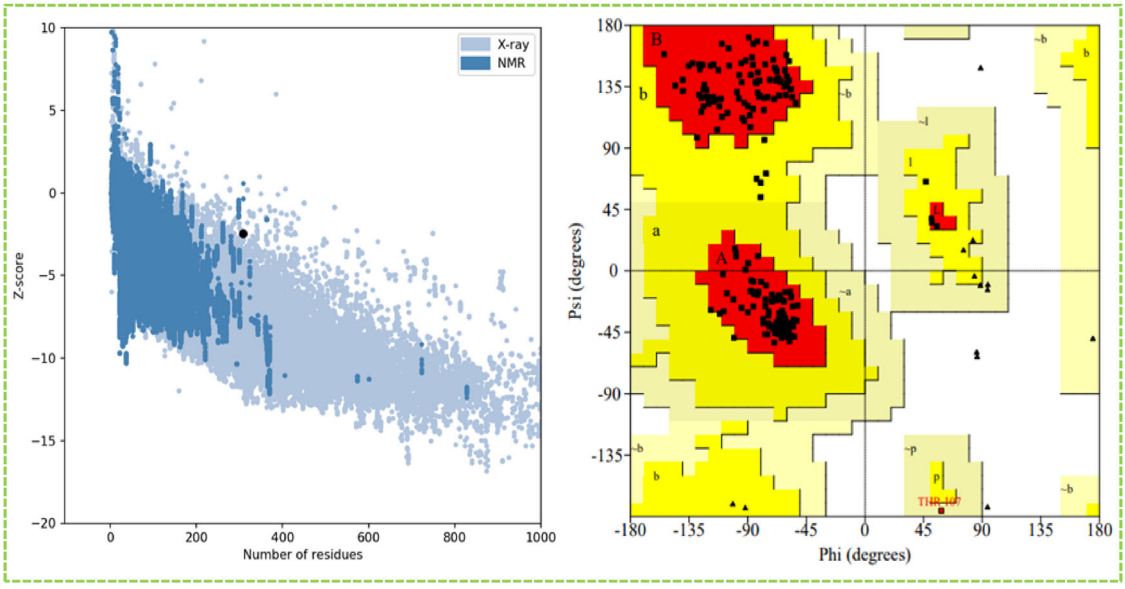

Figure 7. Validation of the final vaccine 3D models. PROSA 3D structure validation showing corresponding *Z*-score of −6.67. Ramachandran plot shows most favored (96.99%), allowed (2.8%), generously allowed (0.3%) and disallowed regions (0.0%), respectively.

2.8% in extra permitted regions, 0.3% in generously allowed regions, and 0.0% in the forbidden region, with 0.0% residues in disallowed regions.

Molecular docking analysis

To analyze the interaction and forecast the final 3D complex, TLR-9 and the vaccine construct were docked using the Clustpro docking server. A total of ten models were obtained. Using the Pymol program, all ten docking models were visually examined and analyzed. Among the 10 models, the first model showed a good docking result with a total of 11 H-bond interactions with a binding score of −1202.4. For a graphical illustration of the residual interaction between the vaccine construct and TLR-9, the PDBsum online database was used. To evaluate the hydrogen bonding between the vaccine construct and TLR-9 complex, a graphical picture was generated (Figure 8). 11 hydrogen bond contacts between TLR-9 and the vaccine construct were observed. Analysis of the vaccine-TLR complex revealed that hydrogen bonds were formed between HIS203-ASN145, THR202-ASN145, GLU223-ASN145, HIS203-SER149, HIS506-LYS179, HIS531-LYS179, ARG482-LYS180, ASN372-CYS188, THR395-CYS195, THR395-CYS195 and ARG337-VAL199 at a distance of 3.32, 2.86, 2.88, 3.27, 2.83, 2.80, 2.53, 3.03, 3.05, 3.05 and 2.81 Å, respectively.

Molecular dynamic simulation

MD simulation was performed by using iModS. iModS analyze structure by adjusting complex force field with various time interval. In the heat map, high co-related area and low RMSD indicated good interactions of individual residues. Figure 9 represents a more detailed explanation of IModS results. Figure 9A represents the NMA mobility in protein structure, Figure 9B represents the deformability and Figure 9C represents the B-factor. Figure 9D shows Eigen value, Figure 9E showed variance and Figure 9F shows the complex elastic network and covariance.

Immune simulations

The vaccine's potential to induce a robust immunological response if delivered globally was evaluated using the C-ImmSim program. After the primary reaction, both secondary and 3D immune response was high. Different antibodies IgG and IgM were detected. Correspondingly, IFN-γ and IL-2 were also observed (Figure 10).

Codon optimization

The Jcat server was used to identify reverse translation and codon optimization in *E. coli* in order to locate expression in the vaccine. The vaccine sequence consisted of nucleotides, whereas the CAI was 0.53 and the GC content was 61.25, indicating a high level of expression. Two main restriction sites, BanI and Eco241 were added. The restriction sites and vaccines were cloned with Snap gene software. The vaccines along with another site in the cloning vector are depicted in Figure 11.

Discussion

As the world recovers from the pandemic of COVID-19, a new infectious disease, monkeypox widespread and has clusters across Europe, America and Australia. There is currently no antiviral treatment that has been licensed by the Food and Drug Administration (FDA) of the United States particularly against MPXV. Alternative preventative measures include smallpox vaccinations such as JYNNEOS and ACAM2000. Despite these preventive measures, the widespread transmission of monkeypox necessitates the development of novel therapeutics against MPXV.

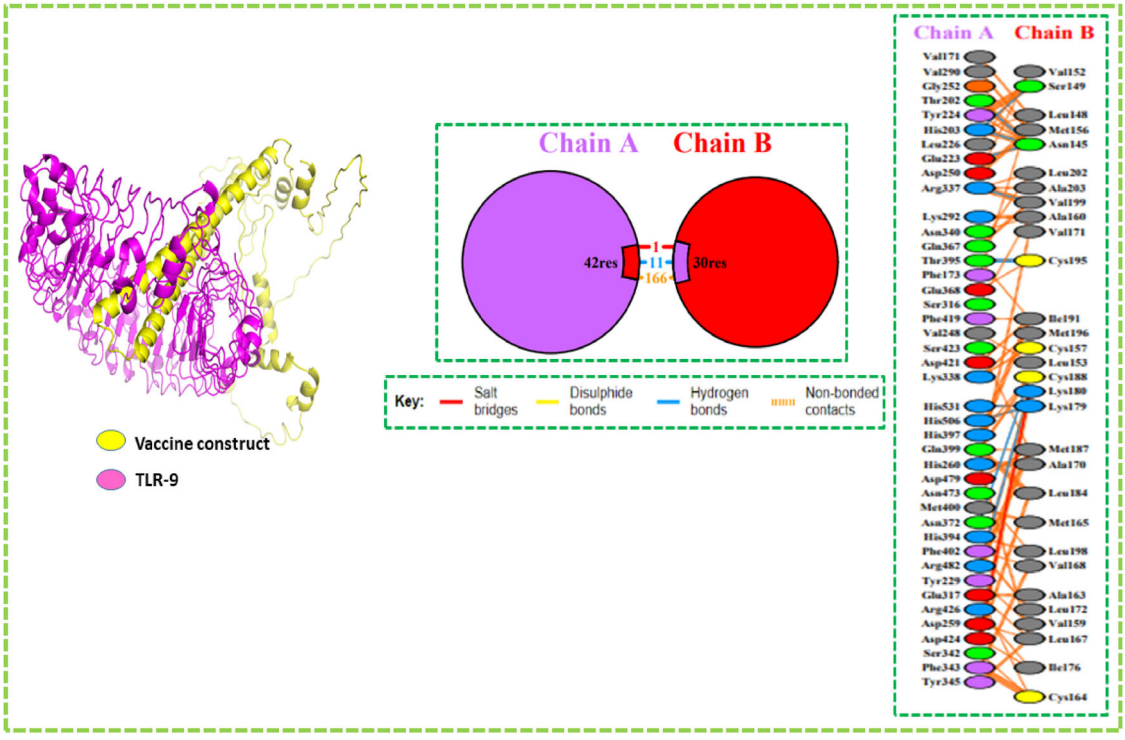

Figure 8. TLR-9 (PDB ID: 3WPF)-vaccine docked complex. The TLR-9 is shown in magentas, while the yellow color represents the multi-epitope subunit vaccine. Left panel graphically represents the hydrogen bonds interaction between TLR-9 and vaccine complex.

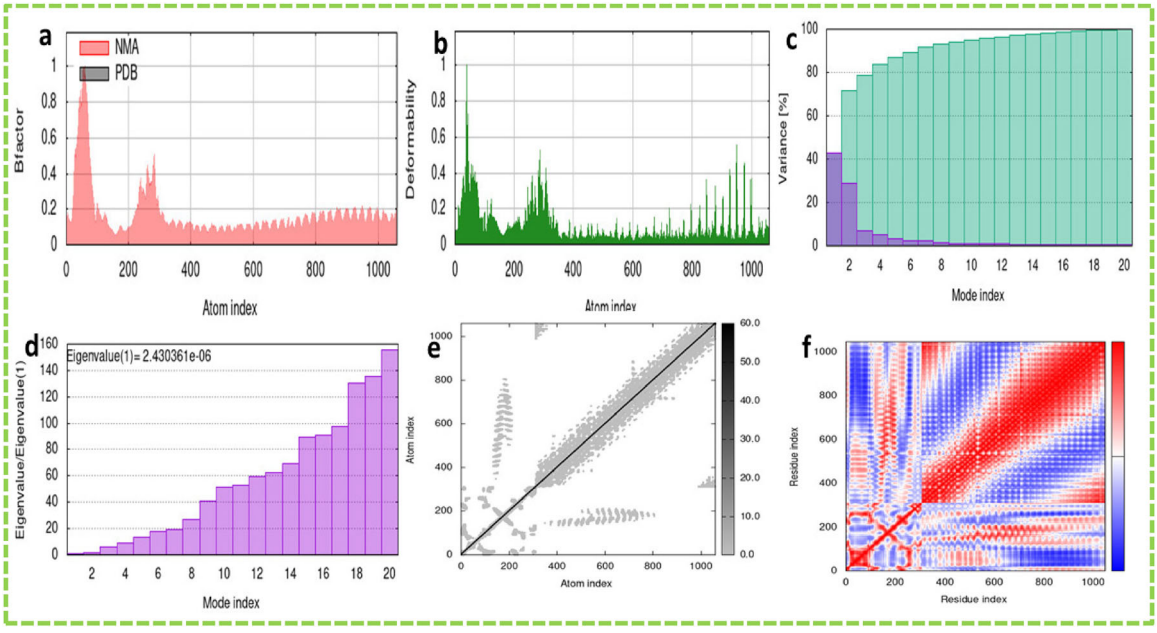

Figure 9. MD simulation of multi-epitope vaccine complexed with the TLR-9.

The implementation of advanced bioinformatics tools is more beneficial than conventional methods (Oli et al., 2020). Bio-computational research, especially reverse vaccinology, is an attractive alternative to developing an epitopic vaccination for monkeypox in the current situation of rapidly spreading disease. Therefore, using immunoinformatic approaches, our research aims to design a vaccine against monkeypox. Using a similar method, Bazhan *et al*. designed a T-cell multi-epitope vaccine model that was highly immunogenic in mice against the Ebola virus (Bazhan et al., 2019).

The FASTA sequence of monkeypox was obtained from the NCBI database. The antigenicity, non-toxicity and non-allergenic behavior of the proteins were investigated. MHC-I and MHC-II epitopes on B-cells and T-cells were predicted. By using linkers CPGPG and AAY, selected epitopes joined with each other. The effectiveness of the predicted vaccine has demonstrated that our design vaccine is both non-allergic and has a high antigenic score (0.8780), along with a good solubility expression inside *E. coli* (0.5075). Constructed vaccine further proceeded for physicochemical properties analysis, where molecular weight

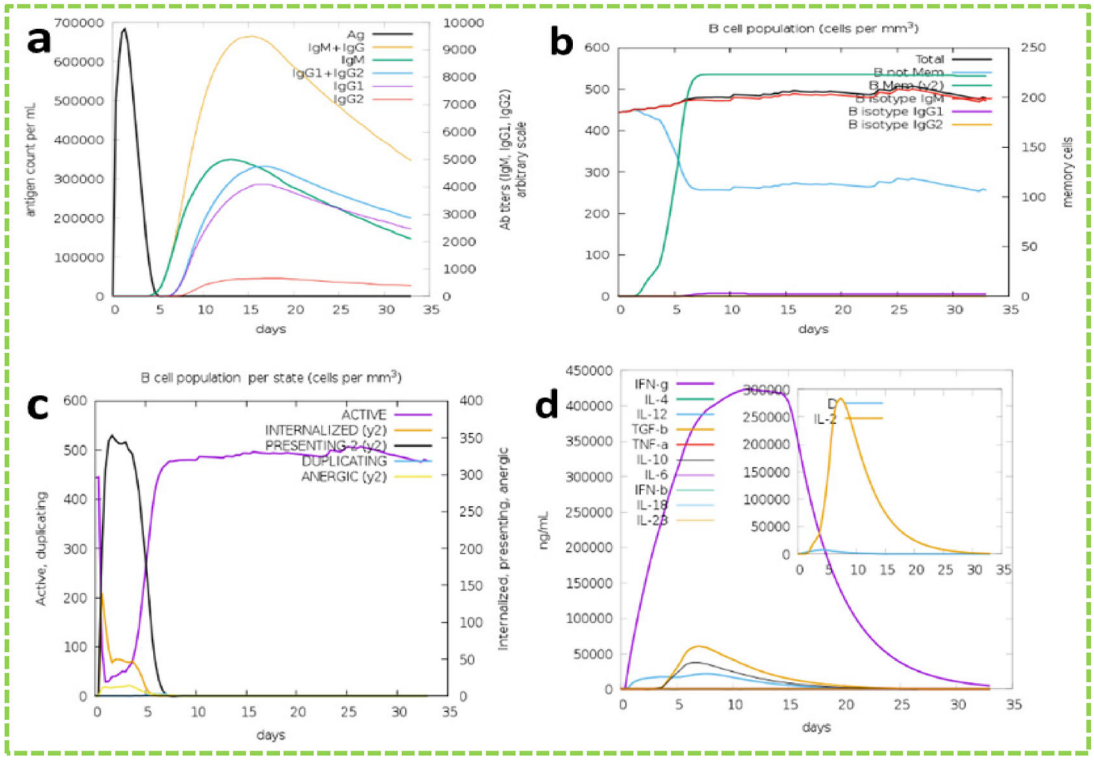

Figure 10. Vaccine immune simulation through C-ImmSim server. (A) The production of antibodies. (B) The population of B-cell. (C) Cytokines production and (D) T cell-population.

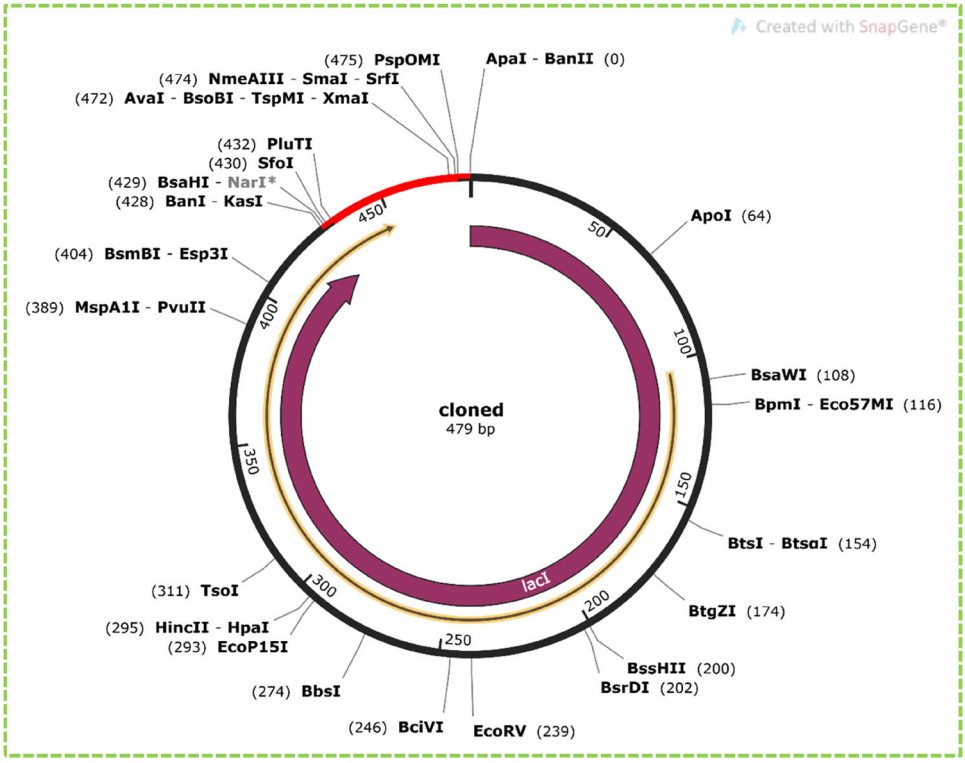

Figure 11. The cloning of the final vaccine where the black color shows the vector and the red color shows the vaccine insert.

was 33547.32, and GRAVY was 0.076, which represents our vaccine is hydrophobic. Similar in silico methodologies were also used by Foroutan *et al.* against *Toxoplasma gondii* to analyse the antigenicity and physiochemical properties of their model vaccine, in addition to laboratory validation (Foroutan, 2020). It was demonstrated that this strategy for the design of vaccines was successful in eliciting an immunological response in mice.

The vaccine's stability was demonstrated by its low instability score of 58.15 and its high aliphatic index of 98.99. For secondary structure, PSIPRED V3.3 online tool was used, which was

found to be 11.4% β-strands, 48.9% α-helices and 39.7% coil structures while the 3D structure of constructed vaccine was modeled by online server trRosetta. 3D model validation was performed by using ERRAT and PROCHECK. Ramachandran plot represents the suitable stereochemical statics, essential for the 3D structure. Results verified that most of the residues were inside the allowed regions. The *Z* score predicted by ProSa-Web was −6.67, although the score was outside the usual range for natural proteins of the same size; however with *Z*-scores of −5.26 and −9.5 kcal/mol, Droppa-Almeida *et al.* and Rekik *et al.* concluded that the predicted 3D structures were accurate and of high quality (Droppa-Almeida et al., 2018; Rekik et al., 2015). Taken together the stereochemical quality of the predicted model is acceptable to be used further.

Previous studies have suggested the importance of toll-like receptors in immune response stimulation (Iwasaki & Medzhitov, 2004). Thus, the vaccine construct was subjected to molecular docking studies with TLR-9. By Cluspro server docking was performed and the results reveal that the vaccine construct exhibits strong interactions with the binding site residues of TLR-9 which indicated the ability of the vaccine to elicit immunogenic responses. Finally, MD simulation was used to validate the stability of the vaccine construct and TLR-9 docked complex. The results indicated the stable molecular interactions between the vaccination and immunological receptor, assuring the molecular stability of the multi-epitope vaccine complex in a cellular environment. Immune simulation results predicted cellular immune responses that were similar to those observed in nature. High levels of T-cytotoxic, memory cells and Ig production were observed, along with an increase in IFN-γ and IL-2 levels. Jcat software is used to predict the best protein expression in *E. coli* K12 strain for codon optimization to enhance transcription and translation efficacy. Similar to the results of previous immunoinformatics vaccine designed studies, the designed vaccines may offer protection against MPXV (Ismail et al., 2020; Tahir Ul Qamar et al., 2020).

This study suggested an alternate vaccine method based on the multi-epitope assembly of MPXV genome protein components to deal with antigenic complexity. The predicted vaccine is believed to be immunogenic based on immunoinformatics techniques and it could contribute to eradicating the disease. However, *in vitro* immunological assays are needed to validate the potency of the vaccine.

Disclosure statement

The authors declare that they have no conflicts of interest.

Funding

This research was funded by the National Key R&D Program of China No. 2021YFC2102900 and Beijing Natural Science Foundation No. L212001.

Author's contributions

C.H. & M.S.: Contribute Equally, Conceptualization, Methodology, Visualization, Data Curation. C.L. & X.D.: Investigation, S.A.K.: Writing - review & editing, Data Curation. H.K.: Methodology write-up. G.Z. & Z.U.-H.: Supervision, Project administration, Writing - review & editing, Software, Funding acquisition.

Funding

This research was funded by the National Key R&D Program of China No. 2021YFC2102900 and Beijing Natural Science Foundation No. L212001.

ORCID

Zaheer Ul-Haq http://orcid.org/0000-0002-8530-8711

Data availability statement

NCBI database in FASTA format with ID (AAQ09810.1).

References

Aa, M., O, H., Kaa, E., A, H., & Ma, H. (2017). Epitope-based peptide vaccine design against Mokola rabies virus glycoprotein G utilizing in silico approaches. *Immunome Research*, *13*(3), 2. https://doi.org/10.4172/1745-7580.1000144

Arai, R., Ueda, H., Kitayama, A., Kamiya, N., & Nagamune, T. (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering*, *14*(8), 529–532. https://doi.org/10.1093/protein/14.8.529

Bazhan, S. I., Antonets, D. V., Karpenko, L. I., Oreshkova, S. F., Kaplina, O. N., Starostina, E. V., Dudko, S. G., Fedotova, S. A., & Ilyichev, A. A. (2019). In silico designed ebola virus T-cell multi-epitope DNA vaccine constructions are immunogenic in mice. *Vaccines*, *7*(2), 34. https://doi.org/10.3390/vaccines7020034

Bhattacharya, M., Sharma, A. R., Patra, P., Ghosh, P., Sharma, G., Patra, B. C., Lee, S.-S., & Chakraborty, C. (2020). Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. *Journal of Medical Virology*, *92*(6), 618–631. https://doi.org/10.1002/jmv.25736

Bunge, E. M., Hoet, B., Chen, L., Lienert, F., Weidenthaler, H., Baer, L. R., & Steffen, R. (2022). The changing epidemiology of human monkeypox—A potential threat? A systematic review. *PLoS Neglected Tropical Diseases*, *16*(2), e0010141. https://doi.org/10.1371/journal.pntd.0010141

Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014). AllerTOP v. 2—a server for in silico prediction of allergens. *Journal of Molecular Modeling*, *20*(6), 2278–2276. https://doi.org/10.1007/s00894-014-2278-5

Doytchinova, I. A., & Flower, D. R. (2007a). Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine*, *25*(5), 856–866. https://doi.org/10.1016/j.vaccine.2006.09.032

Doytchinova, I. A., & Flower, D. R. (2007b). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*(1), 1–7. https://doi.org/10.1186/1471-2105-8-4

Droppa-Almeida, D., Franceschi, E., & Padilha, F. F. (2018). Immune-informatic analysis and design of peptide vaccine from multi-epitopes against Corynebacterium pseudotuberculosis. *Bioinformatics and Biology Insights*, *12*, 1177932218755337.

Du, Z., Su, H., Wang, W., Ye, L., Wei, H., Peng, Z., Anishchenko, I., Baker, D., & Yang, J. (2021). The trRosetta server for fast and accurate protein structure prediction. *Nature Protocols*, *16*(12), 5634–5651. https://doi.org/10.1038/s41596-021-00628-9

Fine, P. E., Jezek, Z., Grab, B., & Dixon, H. (1988). The transmission potential of monkeypox virus in human populations. *International Journal of Epidemiology*, *17*(3), 643–650. https://doi.org/10.1093/ije/17.3.643

Foroutan, M. (2020). Vaccination with a novel multi-epitope ROP8 DNA vaccine against acute Toxoplasma gondii infection induces strong B and T cell responses in mice. *Comparative Immunology, Microbiology and Infectious Diseases*, *65*, 101413.

Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C., & Jahn, D. (2005). JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Research*, ***33***(Web Server issue), W526–W531. https://doi.org/10.1093/nar/gki376

Hasan, M., Azim, K. F., Begum, A., Khan, N. A., Shammi, T. S., Imran, A. S., Chowdhury, I. M., & Urme, S. R. A. (2019). Vaccinomics strategy for developing a unique multi-epitope monovalent vaccine against Marburg marburgvirus. *Infection, Genetics and Evolution*, *70*, 140–157. https://doi.org/10.1016/j.meegid.2019.03.003

Hon, J., Marusiak, M., Martinek, T., Kunka, A., Zendulka, J., Bednar, D., & Damborsky, J. (2021). SoluProt: Prediction of soluble protein expression in Escherichia coli. *Bioinformatics*, *37*(1), 23–28. https://doi.org/10.1093/bioinformatics/btaa1102

Impouma, B., Archer, B. N., Lukoya, O. C., Hamblion, E. L., & Fall, I. S. (2018). World Health Organization Regional Office for Africa weekly bulletin on outbreaks and other emergencies. *Emerging Infectious Diseases*, *24*(7), 1394–1395. https://doi.org/10.3201/eid2407.180573

Ismail, S., Ahmad, S., & Azam, S. S. (2020). Vaccinomics to design a novel single chimeric subunit vaccine for broad-spectrum immunological applications targeting nosocomial Enterobacteriaceae pathogens. *European Journal of Pharmaceutical Sciences*, *146*, 105258.

Iwasaki, A., & Medzhitov, R. (2004). Toll-like receptor control of the adaptive immune responses. *Nature Immunology*, *5*(10), 987–995. https://doi.org/10.1038/ni1112

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, *292*(2), 195–202. https://doi.org/10.1006/jmbi.1999.3091

Kenneth, F. K. (1993). *The Cambridge world history of human disease*. Cambridge University Press.

Laskowski, R., MacArthur, M., & Thornton, J. (2006). *PROCHECK: Validation of protein-structure coordinates.*

Magnus, P. v., Andersen, E. K., Petersen, K. B., & Birch-Andersen, A. (2009). A pox-like disease in cynomolgus monkeys. *Acta Pathologica Microbiologica Scandinavica*, *46*(2), 156–176. https://doi.org/10.1111/j.1699-0463.1959.tb00328.x

Naz, A., Awan, F. M., Obaid, A., Muhammad, S. A., Paracha, R. Z., Ahmad, J., & Ali, A. (2015). Identification of putative vaccine candidates against Helicobacter pylori exploiting exoproteome and secretome: A reverse vaccinology based approach. *Infection, Genetics and Evolution*, *32*, 280–291. https://doi.org/10.1016/j.meegid.2015.03.027

Nielsen, M., Lundegaard, C., Worning, P., Lauemøller, S. L., Lamberth, K., Buus, S., Brunak, S., & Lund, O. (2003). Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Science*, *12*(5), 1007–1017. https://doi.org/10.1110/ps.0239403

Nosrati, M., Behbahani, M., & Mohabatkar, H. (2019). Towards the first multi-epitope recombinant vaccine against Crimean-Congo hemorrhagic fever virus: A computer-aided vaccine design approach. *Journal of Biomedical Informatics*, *93*, 103160.

Oli, A. N., Obialor, W. O., Ifeanyichukwu, M. O., Odimegwu, D. C., Okoyeh, J. N., Emechebe, G. O., Adejumo, S. A., & Ibeanu, G. C. (2020). *Immunoinformatics and vaccine development: an overview. ImmunoTargets and Therapy*, ***9***, 13–30.

Potter, M. A., Sweeney, P., Iuliano, A. D., & Allswede, M. P. (2007). Performance indicators for response to selected infectious disease outbreaks: A review of the published record. *Journal of Public Health Management and Practice : JPHMP*, *13*(5), 510–518. https://doi.org/10.1097/01.PHH.0000285205.40964.28

Rekik, I., Chaabene, Z., Grubb, C. D., Drira, N., Cheour, F., & Elleuch, A. (2015). In silico characterization and Molecular modeling of double-strand break repair protein MRE11 from Phoenix dactylifera v deglet nour. *Theoretical Biology and Medical Modelling*, *12*(1), 1–14. https://doi.org/10.1186/s12976-015-0013-2

Shams, N., Shakarami Gandabeh, Z., Nazifi, N., Forouharmehr, A., Jaydari, A., & Rashidian, E. (2020). Computational design of different epitope-based vaccines against Salmonella typhi. *International Journal of Peptide Research and Therapeutics*, *26*(3), 1527–1539. https://doi.org/10.1007/s10989-019-09959-4

Shchelkunov, S. N., Totmenin, A. V., Babkin, I. V., Safronov, P. F., Ryazankina, O. I., Petrov, N. A., Gutorov, V. V., Uvarova, E. A., Mikheev, M. V., Sisler, J. R., Esposito, J. J., Jahrling, P. B., Moss, B., & Sandakhchiev, L. S. (2001). Human monkeypox and smallpox viruses: Genomic comparison. *FEBS Letters*, *509*(1), 66–70. https://doi.org/10.1016/S0014-5793(01)03144-1

Tahir Ul Qamar, M., Shokat, Z., Muneer, I., Ashfaq, U. A., Javed, H., Anwar, F., Bari, A., Zahid, B., & Saari, N. (2020). Multiepitope-based subunit vaccine design and evaluation against respiratory syncytial virus using reverse vaccinology approach. *Vaccines*, *8*(2), 288. https://doi.org/10.3390/vaccines8020288

Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., & Kozakov, D. (2017). New additions to the C lus P ro server motivated by CAPRI. *Proteins*, *85*(3), 435–444. https://doi.org/10.1002/prot.25219

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(Web Server issue), W407–W410. https://doi.org/10.1093/nar/gkm290

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server. *Methods in Molecular Biology (Clifton, N.J.)*, *112*, 531–552. https://doi.org/10.1385/1-59259-584-7:531

Zheng, W., Zhang, C., Li, Y., Pearce, R., Bell, E. W., & Zhang, Y. (2021). Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Reports Methods*, *1*(3), 100014. https://doi.org/10.1016/j.crmeth.2021.100014