Received: 14 March 2024 | Revised: 19 April 2024 | Accepted: 2 May 2024

DOI: 10.1002/prot.26703

RESEARCH ARTICLE

PROTEINS WILEY

In silico design and evaluation of a multiepitope vaccine targeting the nucleoprotein of *Puumala orthohantavirus*

Kunal Bhattacharya[1,2] | **Nongmaithem Randhoni Chanu**[1,3] | **Saurav Kumar Jha**[4] | **Pukar Khanal**[5] | **Keshav Raj Paudel**[6]

[1]Pratiksha Institute of Pharmaceutical Sciences, Guwahati, Assam, India

[2]Royal School of Pharmacy, The Assam Royal Global University, Guwahati, Assam, India

[3]Faculty of Pharmaceutical Science, Assam Downtown University, Guwahati, Assam, India

[4]Department of Biological Sciences and Bioengineering (BSBE), Indian Institute of Technology, Kanpur, Uttar Pradesh, India

[5]Department of Pharmacology and Toxicology, KLE College of Pharmacy, Belagavi, KLE Academy of Higher Education and Research (KAHER), Belagavi, India

[6]Centre for Inflammation, Centenary Institute and University of Technology Sydney, Faculty of Science, School of Life Sciences, Sydney, New South Wales, Australia

Correspondence
Kunal Bhattacharya, Pratiksha Institute of Pharmaceutical Sciences, Guwahati, Assam 781026, India.
Email: kunal22101994@gmail.com

Keshav Raj Paudel, Centre for Inflammation, Centenary Institute and University of Technology Sydney, Faculty of Science, School of Life Sciences, Sydney, NSW 2007, Australia.
Email: keshavraj.paudel@uts.edu.au

Abstract

The Puumala orthohantavirus is present in the body of the bank vole (*Myodes glareolus*). Humans infected with this virus may develop hemorrhagic fever accompanying renal syndrome. In addition, the infection may further lead to the failure of an immune system completely. The present study aimed to propose a possible vaccine by employing bioinformatics techniques to identify B and T-cell antigens. The best multi-epitope of potential immunogenicity was generated by combining epitopes. Additionally, the linkers EAAAK, AAY, and GPGPG were utilized in order to link the epitopes successfully. Further, C-ImmSim was used to perform in silico immunological simulations upon the vaccine. For the purpose of conducting expression tests in *Escherichia coli*, the chimeric protein construct was cloned using Snapgene into the pET-9c vector. The designed vaccine showed adequate results, evidenced by the global population coverage and favorable immune response. The developed vaccine was found to be highly effective and to have excellent population coverage in a number of computer-based assessments. This work is fully dependent on the development of nucleoprotein-based vaccines, which would constitute a significant step forward if our findings were used in developing a global vaccination to combat the Puumala virus.

KEYWORDS

B and T-cells, immunoinformatics, molecular docking, multiepitope vaccine, pET-9c vector, Puumala virus

1 | INTRODUCTION

Hantaviruses, part of the Bunyaviridae family, are zoonotic pathogens primarily carried by rodents. They can be transmitted to humans through various routes, including inhalation of virus-laden aerosols and contact with infected rodents' feces or urine.[1] The Puumala virus (PUUV) is endemic in regions such as Sweden, Finland, Norway, Russia, and parts of central Europe, with bank voles (*Myodes glareolus*) acting as chronic carriers and vectors for its transmission.[2] PUUV infection in humans can lead to nephropathia epidemica (NE), a moderate form of hemorrhagic fever characterized by kidney impairment.[3] NE cases are notably concentrated in the northernmost counties of Sweden, contributing to around 90% of reported cases.[4] While most NE cases exhibit mild hemorrhagic symptoms, a subset experiences acute renal failure characterized by oliguria, requiring dialysis in some instances, with approximately 0.5% of NE cases proving fatal.

All the authors contributed equally to this study.

Understanding the genetic makeup of PUUV is essential in developing effective countermeasures. The virus possesses a segmented, single-stranded RNA genome with negative polarity, encoding key proteins for replication and pathogenicity, including RNA polymerase in the large (L) segment, the nucleocapsid protein in the short (S) segment, and two surface glycoproteins in the medium (M) segment.[5] The genetic characteristics of Puumala virus from Finland,[6–9] Sweden,[10–12] Norway,[12] Denmark,[13] Russia,[13–16] Belgium,[17,18] Austria,[19,20] and Germany[21,22] are also determined. Over a hundred sequences of the Puumala virus, in part or whole, have been deposited in GenBank so far. Additionally, they represent thirty-eight complete S segment sequences, nine complete M segment sequences, and two full L segment sequences. Also, the evolutionary connections between several hantaviruses and the rodents they infect have been mapped out, showing they are nearly similar to one another, supporting the concept that the virus and its hosts have coevolved.[23] Various vaccine approaches are being explored, including nano-based and homology-based immunization. Notably, advances in vaccine technologies such as live-attenuated, protein subunit, inactivated, and viral vector vaccines offer hope for progress in the field.

The focus of this study is to design a novel Puumala virus multiepitope vaccine using computational tools and an immunoinformatics approach with a particular emphasis on the nucleoprotein—a critical component facilitating viral immune evasion and morphogenesis. The vaccine is computationally designed, and the natural immune response was simulated on a computer, demonstrating promising results. For efficient production, *Escherichia coli* expression is chosen as the preferred method while performing simulation using snapgene, facilitated by cloning the peptide into a pET-9c vector. However, subsequent in vitro and in vivo testing are required to determine the vaccine's potential as a viable treatment option. A multiepitope vaccine against Puumala orthohantavirus would have numerous advantages. First, by focusing on several viral epitopes, this type of vaccine can stimulate a strong immune response and increase the possibility of protecting against multiple strains of the virus.

Further, in areas where numerous orthohantaviruses co-circulate, a multiepitope vaccine may provide cross-protection against related orthohantaviruses, expanding the vaccine's usefulness. However, concerns about the cost-effectiveness of testing and production must be addressed. Higher initial costs and complexity in developing and producing multiepitope vaccines can be reduced by economies of scale and advances in technology in manufacturing. Hemorrhagic fever and renal syndrome are two of the potentially fatal outcomes of a Puumala orthohantavirus infection, and their prevention may be worth more than the cost of manufacturing. To maximize the vaccine's influence on public health while successfully managing costs, it would be advisable to target risk groups such as those living in endemic regions, healthcare personnel, and those with occupational exposure to rodents. Overall, the potential benefits of a multiepitope vaccine for Puumala orthohantavirus make it a valuable avenue for research and development, with careful consideration of cost–benefit analyses and target audience prioritization.

2 | MATERIALS AND METHODS

2.1 | Nucleoprotein selection and evaluation of its physiochemical characteristics

The selection of the nucleoprotein of the Puumala virus was based on its exceptional antigenic properties, cellular infiltration capabilities, and capacity to assemble the viral genome. The potential of nucleoprotein as a vaccine candidate is supported by its limited number of transmembrane helices and adhesive properties. Moreover, it does not exhibit any distinct amino acid sequences that are seen in proteins identified inside the *Homo sapiens* genome. It was found that its cellular localization was suitable for vaccine designing. The Puumala virus sotkamo/v-2969/81 genome was retrieved from the NCBI (https://www.ncbi.nlm.nih.gov/) database, which is part of the parent taxonomy (ID 1980486). Uniprot database (KBID: P27313; NCAP_PUUMS) and NCBI database (accession number ALJ30169) were queried for the nucleoprotein, which was shown to be one of the most antigenic proteins based on reports from the past. The ExPASyProt pram programme was utilized to assess the physicochemical parameters of the protein under investigation.[24] The PSIPRED online programme was employed to assess the secondary structure of the protein.[25] The determination of disulfide bonding was conducted using the DIANNA v1.1 programme, which utilizes a neural system foundation that has been trained.[26] The antigenic protein was identified using Vaxijen version 2.0.[27] Consideration was given to focusing research on the protein with the highest antigenicity. Therefore, we assessed the allergenicity of the samples using the AllerTOPver2.0 online server.[28] In addition, the PsiPred tool (which uses the PSI-blast method to forecast the secondary structure of a particular nucleoprotein) was used to predict the secondary structure. We employed Iterative Threading ASSEmbly Refinement (I-TASSER) as a computational tool to predict the tertiary structure of the nucleoprotein. The structure with the highest C-score was identified as the most reliable and authentic structure.[29]

2.2 | B-cell epitopes prediction

In our study, we employed the Immune Epitope Database (IEDB) Linear Epitope Prediction version 2.0 (available at http://www.cbs.dtu.dk/services/BepiPred/) to predict B-cell epitopes. This computational approach is based on the analysis of protein structures of antigen–antibody complexes, providing insights into epitope regions that are likely to be recognized by B-cells.[30]

BepiPred 2.0 utilizes a machine-learning algorithm that integrates a random forest classifier trained on both epitopes and non-epitopes derived from experimentally determined B-cell epitopes. The algorithm considers several parameters, including the propensity of amino acid residues to be part of linear epitopes, structural accessibility, and flexibility.

2.3 | T-cell epitopes prediction

We utilized the IEDB MHC-I Binding Predictions Tool (http://tools.iedb.org/mhci) to generate predictions for MHC-I epitopes.[31] The selection of *H. sapiens* as the host and the provision of the sequence in FASTA format, with the prediction technique being set to the SMM approach, were undertaken. Furthermore, the output format was configured to adhere to the XHTML table format, but all other options and parameters were retained in their default state. Additionally, we utilized the IEDB MHC-II binding predictions tool (available at http://tools.iedb.org/mhcii) to predict the MHC-II epitopes.[32] The FASTA sequence was supplied and the prediction technique employed was SMM. The HLA-DR, HLA-DQ, and HLA-DP species and loci of human beings were carefully chosen, and all alleles were selected based on default length values specific to each species and locus. The output format chosen was the XHTML table format, but all other settings remained at their default.

2.4 | Eminent features profiling of selected T and B-cell epitopes

In our study, we performed a comprehensive profiling of the eminent features of selected T and B-cell epitopes using several predictive tools to evaluate their suitability and safety as vaccine components.

2.4.1 | Antigenicity prediction

We utilized the VaxiJen v2.0 server (available at http://www.ddg-pharmfac.net/vaxijen) to predict the antigenicity of the epitopes based on their physicochemical properties. VaxiJen operates on the principle of auto cross-covariance (ACC) transformation of protein sequences into uniform vectors of principal physicochemical properties. This server is designed to discriminate between antigens and non-antigens independent of species and the immune system. A threshold of 0.4 was selected to ensure a high confidence level in antigenicity prediction, aiming to identify epitopes with robust immune eliciting potential.[33]

2.4.2 | Allergenicity assessment

To evaluate the potential allergenicity of both T and B-cell epitopes, we employed the AllerTOP v2.0 tool (https://www.ddg-pharmfac.net/AllerTOP). This server uses an alignment-free approach, which incorporates various physicochemical properties of amino acids to predict allergenicity. Identifying non-allergenic epitopes is crucial for vaccine safety, reducing the risk of adverse allergic responses upon administration.[34]

2.4.3 | Toxicity analysis

The ToxinPred server (https://webs.iiitd.edu.in/raghava/toxinpred/) was used to assess the toxicity of the epitopes. This server employs a machine learning approach, specifically Support Vector Machine (SVM) models, which are trained on a dataset of known toxic and non-toxic peptides. By predicting the toxicity of epitopes, we aimed to ensure that the selected epitopes do not exhibit harmful effects, thereby enhancing the safety profile of the proposed vaccine.

2.5 | Population coverage calculation

We conducted a comprehensive assessment of the population coverage for the selected epitopes using the IEDB's Population Coverage Analysis tool (available at http://tools.iedb.org/population). This tool is specifically designed to estimate the percentage of individuals likely to respond to a set of epitopes, based on the frequency of HLA (Human Leukocyte Antigen) alleles in different populations. This analysis is crucial for determining the potential effectiveness of the vaccine across diverse genetic backgrounds and geographical locations.[35]

2.6 | Vaccine construction

The adjuvant sequence was acquired from the UniProt database, accessible at https://www.uniprot.org/. and used to create the vaccine by exploiting all the possible epitopes to build the multiepitope vaccine. The adjuvant 50S ribosomal protein L7/L12 (UniProt ID: P9WHE3) was employed to bind to the N terminal of the vaccination and act as an adjuvant by using three main types of linkers, that is, EAAAK, GPGPG, and AAY.

2.7 | Multiepitope sequence assemblage

The adjuvant sequence, EAAAK, AAY, GPGPG linkers, B-cell epitopes, MHC-I, and II binding epitopes were manually assembled to form a candidate vaccine sequence.[36] The assembly of these components was performed manually to ensure optimal spacing and orientation, maximizing the immunogenic potential of the construct. Each element was carefully positioned to preserve its functional integrity while promoting an efficient immune response. The order and arrangement of the epitopes and linkers were designed to mimic natural immunogenic sequences as closely as possible, thereby enhancing recognition and processing by the immune system.

2.8 | Evaluation of the vaccine protein's antigenicity and allergenicity

The antigenicity of the constructed vaccine was analyzed using Vaxigen ver 2.0 (http://www.ddg-pharmfac.net/vaxijen),[33,37] and the aller-

genicity of the multiepitope vaccine was predicted with AllerTOP (https://www.ddg-pharmfac.net/AllerTOP).[34] In addition, the toxicity profile was evaluated using ToxinPred (https://webs.iiitd.edu.in/raghava/txinpred/).[38]

2.9 | Evaluation of solubility and physicochemical characteristics

ExPASy-Protparam (https://web.expasy.org/protparam/protparam-doc.html) was used to analyze the physicochemical properties, and SOLpro (http://scratch.proteomics.ics.uci.edu/) was used to analyze the solubility, of the designed multiepitope vaccine.

2.10 | Extrapolation of secondary and tertiary structure and validation and refinement of the tertiary structural design

The PsiPred tool (http://bioinf.cs.ucl.ac.uk/psipred/) was employed, along with the I-TASSER software (https://zhanggroup.org/I-TASSER/) to derive the secondary and tertiary structures of the multiepitope vaccine.[25,29,39] This was done to gain a broad understanding of the predominant helices, sheets, and loops present in the protein structure. GalaxyRefine, server accessible at https://galaxy.seoklab.org/, is widely recognized as a highly reliable tool for the refinement of tertiary structures. The process of side chain rebuilding and repacking was initially conducted as an integral component of the refinement methodology. The validation and refinement of the tertiary structure of the protein in question were conducted using the GalaxyRefine server.[40]

2.11 | 3D structure validation and molecular docking

To check the model and assure the excellent quality of the 3D structure, we employed MolProbity, a tool available at http://molprobity.biochem.duke.edu/.[41] The docking analysis was performed to investigate the interaction between the ligand-binding domain of the TLR7 receptor (PDB: 7CYN) and the proposed vaccine construct. This analysis was conducted using the Cluspro2.0 online docking server, accessible at https://cluspro.bu.edu/.[42]

2.12 | Normal mode analysis-based molecular dynamics simulations

Normal mode analysis (NMA) is a computational technique used to predict the motions of atoms and molecules within a protein structure under the assumption that the structure oscillates around an equilibrium position. When applied to the study of a multiepitope vaccine complex it offers several significant benefits. NMA provides insights into the flexibility and dynamics of a vaccine complex. The analysis can predict how the vaccine complex might change conformation under physiological conditions, which is important for its stability and efficacy. The iModS online tool, accessible at http://imods.chaconlab.org/, was employed to perform molecular dynamics simulations in order to investigate the torsional angles of the complex. The stability of the complex was assessed through the utilization of this method, which involved an examination of many factors such as the NMA mobility, deformability, B-factor, eigenvalue, variance, covariance map, and elastic network.[43]

2.13 | Codon optimization for expression analysis of vaccine peptide

The chimeric protein production process involved reverse translation of the sequence using the EMBOSS 6.0.1 backtranseq programme, available at the https://www.ebi.ac.uk/Tools/emboss/. Subsequently, codon optimization was performed using the Java Codon Adaption tool, an online web-based server accessible at http://www.jcat.de/.[44] The approach was conducted to express the synthesized peptide in *E. coli* (K12 strain), which utilizes a distinct set of codons compared to the native host. To do this, the nucleotide sequence of the construct was replicated and inserted, while specific modifications were implemented to inhibit rho-independent transcription termination, bacterial ribosome binding, and reestablishment of enzyme cleavage sites. The findings were analyzed based on the proportion of codons exhibiting a high GC content and the codon adaptation index (CAI) score of the optimized codons. These parameters assessed whether the expression would be good or bad. Generally, a CAI score between 0.8 and 1.0 is considered good, with one being considered to be the optimal score.[45,46] The expression vector used in this study was the *E. coli* pET-9c. Using the SnapGene tool (http://www.snapgene.com/), BsrBI and EagI restriction sites were inserted into the construct, allowing it to be cloned into the vector straightforwardly. In silico PCR was performed to amplify the construct by SnapGene (https://www.snapgene.com/).

2.14 | Immune simulation

The effectiveness of the vaccine construct was validated by stimulating an immunological response through the utilization of the C-ImmSim server (http://kraken.iac.rm.cnr.it/C-IMMSIM/). The server utilizes a position-specific scoring matrix (PSSM) to identify immunological epitopes and immune interactions. The simulation was conducted with all of the default parameters.[47]

3 | RESULTS

3.1 | Evaluation of antigenicity and physiochemical characteristics of nucleoprotein

The UniProt database provided the nucleoprotein sequence for Puumala virus, which was used in the vaccine designing. The viral

structural protein-nucleoprotein of the Pummala virus is essential for infection and viral replication in the host cell. Using the Vaxijen 2.0 online server, we developed an antigenicity estimate for the viral protein. A threshold value of 0.5 was applied to increase accuracy. A full-length protein antigenicity research indicated that nucleoprotein has an antigenicity of 0.5952, indicating that it is a potential antigen. Additional research confirmed the protein's usefulness. Nucleoprotein stability was improved by the addition of adjuvant 50S ribosomal protein from Mycobacterium tuberculosis, and its physiochemical properties were calculated using the ProtParam tool. The investigation showed that there were 563 different amino acids, and the computed molecular weight (MW) was 62702.03 Da. Since proteins with an isoelectric point below 7.0 are negatively charged, the calculated value of 5.24 for the theoretical isoelectric point (PI) indicates that it is negative. This protein is considered stable by the Protparam method because its computed instability index (II) is 39.85. This protein is highly stable over a wide temperature range due to its high aliphatic index of 86.66. The sum of all the atoms of carbon, hydrogen, nitrogen, oxygen, and sulfur can be found by plugging their respective numbers into the formula C2784H4470N744O848S24. Twenty different amino acids were identified through chemical analysis (Figure S1). PSIPRED and I-TASSER correctly predicted that nucleoproteins would have a secondary structure composed of 2.5% helices, 55.6% sheets, and 41.8% loops (Figure S2). Transmembrane topology predictions were made using the web-based application TMHMM. The transmembrane region was discovered to be missing residues 1–118 and 234–433. While residues 142–214 were found in the transmembrane domain, residues 119–141 and 215–233 were found to be buried deep within the nucleoprotein's core (Figure S3).

3.2 | Identification of B-cell epitopes

B-cell epitopes significantly influence the formation of a resistant immune response to viral infection. Potential B-cell epitopes have altered properties that instruct B-cells to detect and activate a wide variety of immune responses in response to a particular viral infection. These characteristics direct B-cells to recognize and activate immune responses. In particular, we anticipated using techniques based on amino acid screening for the investigation of possible B-cell epitopes in this work. We employed a technique that was based on consensus in order to anticipate probable B-cell epitopes.

Linear epitope prediction using Bepipred and a threshold score of 0.600 predicted a total of 15 linear epitopes. Linear epitopes are calculated using the Hidden Markov model-based approach, one of the best ways. Linear epitope prediction has a minimum score of 0.191 and a maximum value of 0.633. It was also observed that the average prediction score was 0.483 (Figure 1A, Table S1).

Using the Kolaskar and Tongaonkar approach, the antigenicity of experimentally discovered amino acid epitopes was investigated. Antigenicity's maximum and minimum values were 1.175 and 0.878, respectively. In general, 1.015 was the average value observed. The threshold was set at 1.055, and all values more than 1.055 were considered antigenic factors. A total of seven epitopes met the criterion and were selected for further investigation. They were discovered to have the ability to activate B-cells (Figure 1B, Table S2).

A B-cell epitope's efficacy largely depends on its surface accessibility. As a result, this forecast was made using the Emini surface accessibility tool. In total, seven epitopes passed the threshold value of 1.4. In general, we observed an average value of 1.00. The minimum value was 0.037, and the maximum value was found to be 4.285 (Figure 1C, Table S3).

Beta-turn has a hydrophilic character, is exposed to the environment, and is influential in initiating and progressing the immune system's natural defenses. Chou and Fasman's beta-turn evaluation method was applied to predict the beta-turn in nucleoproteins. There was a 1.100 threshold level for the, and the computed findings showed a minimum value of 0.559 and a maximum value of 1.416. The average value was 0.962. Beta turns in the peptide structure are more likely to occur in the region between 295 and 310 (Figure 1D). Peptide flexibility is linked to antigenicity, according to the evidence from experiments. For this reason, the Karplus and Schulz approach was developed. As demonstrated in Figure 1E, this prediction algorithm discovered that the area between 110 and 120 was more flexible. Calculated results showed a maximum value of 1.120 with a minimum of 0.895, and the tool's threshold value was set to 1.000. The average calculated value was 1.002%. Based on toxicity, allergenicity, and antigenicity, the epitopes were filtered out. Epitopes that were toxic, allergic, and non-antigenic were ruled out, while 10 epitopes were selected as B-cell epitopes that can elicit B lymphocytes response, that is, (a) DIQEDITRHEQQLVV, (b) RQTVKENKGT, (c) IRFKDDTSFEDINGIRR, (d) AQSTMKAEELTPG, (e) QIQVRNIM, (f) GWADRIREFMEK, (g) TVGTAEEKLKKKSSF, (h) LRRTQSMGIQ, (i) HLGDDMDPE, and (j) KEISNQEP.

3.3 | Identification of T-cell epitopes

3.3.1 | Prediction of MHC Class-I epitopes

By employing *H. sapiens* as the major histocompatibility complex (MHC) source and utilizing the SMM approach, an extensive exploration of human leukocyte antigen (HLA) alleles was conducted. The output interface of this application delivers the HLA-binding affinity in IC50 nM units. A low IC50 value indicates a high affinity of the epitope for the MHC Class-I molecule. A total of 273 epitopes were identified, each exhibiting IC50 values below 300. This observation indicates a strong probability of interaction with a diverse range of MHC Class-1 alleles. Out of the total of 273 epitopes, a subset of 18 epitopes was chosen for further examination in terms of their interaction with MHC Class-1 alleles, that is, (a) EEPSGQTAD, (b) KRQMADAVS, (c)VIGFAFFVK, (d) MSPVMGVIG, (e) TKPTDPTGI, (f) GDPTSPDNI, (g) QALIDQKVK, (h) RRPKHLYVS, (i) VEMDPDDVN, (j) QSMGIQLDQ, (k) YGNVLDVNA, (l) RQMADAVSR, (m) AFFVKGWAD, (n) VKPGTPTQE, (o) RIREFMEKE, (p) QIQVRNIMS, (q) IEPDDHLKE, and (r) MDPDDVNKN (Table 1).

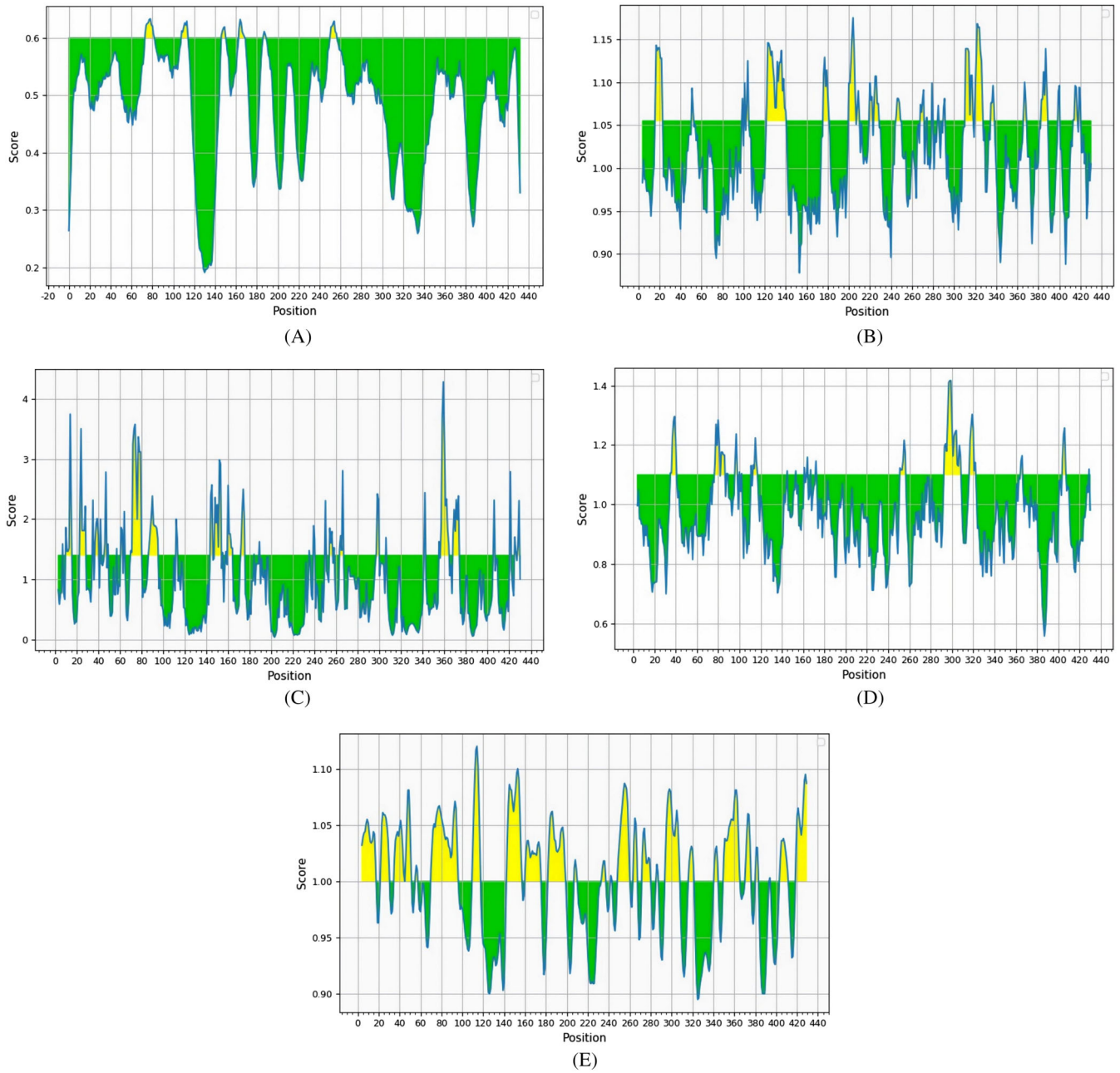

FIGURE 1 (A) Bepipred linear epitope prediction of the nucleoprotein. (B) Prediction of antigenicity of nucleoprotein by using Kolaskar and Tongaonkar antigenicity scale. (C) Emini surface accessibility prediction of nupcleorotein. (D) Prediction of beta turns in nucleoprotein by using Chou and Fasman beta turn analyzing algorithm. (E) Karplus and Schulz flexibility prediction for the flexibility of nucleoprotein.

3.3.2 | Prediction of MHC Class-II epitopes

Out of 2096 predicted epitopes, 18 highly ranked epitopes were selected for further analysis having good interaction with MHC Class-II alleles after analyzing their allergenicity, toxicity, and antigenicity, that is, (a) LKDAEKAVEMDPDDV, (b) KDAEKAVEMDPDDVN, (c) DAEKAVEMDPDDVNK, (d) AEKAVEMDPDDVNKN, (e) EKAVEMDPDDVNKNT, (f) KAVEMDPDDVNKNTL, (g) AVEMDPDDVNKNTLQ, (h) VEMDPDDVNKNTLQA, (i) RQMADAVSRKKMDTK, (j) VSRKKMDTKPTDPTG, (k) SRKKMDTKPTDPTGI, (l) RKKMDTKPTDPTGIE, (m) KKMDTKPTDPTGIEP, (n) KMDTKPTDPTGIEPD, (o) MDTKPTDPTGIEPDD, (p) DTKPTDPTGIEPDDH, (q) TKPTDPTGIEPDDHL, and (r) KPTDPTGIEPDDHLK (Table 2).

3.4 | Assembly of multiepitope subunit vaccine

A chimeric vaccine was developed by including a collection of 28 epitopes, comprising 10 B-cell epitopes, 18 MHC Class-I epitopes, and 18 MHC Class-II epitopes. The adjuvant utilized in the

TABLE 1 Most potential non-allergen, nontoxic, 18 T-cell epitopes with interacting MHC-I alleles.

Peptide	Allele	Start	End	Length	IC_{50}
EEPSGQTAD	HLA-C*12:03	110	118	9	24.672320
	HLA-C*03:03	110	118	9	283.537169
KRQMADAVS	HLA-C*12:03	62	70	9	24.786202
	HLA-C*14:02	62	70	9	85.235399
	HLA-C*07:02	62	70	9	144.600568
	HLA-C*03:03	62	70	9	230.467664
VIGFAFFVK	HLA-C*12:03	222	230	9	24.958012
	HLA-A*11:01	222	230	9	27.880464
	HLA-C*03:03	222	230	9	167.343943
MSPVMGVIG	HLA-C*12:03	216	224	9	23.999931
TKPTDPTGI	HLA-C*14:02	76	84	9	223.161874
	HLA-C*03:03	76	84	9	241.329270
GDPTSPDNI	HLA-C*12:03	295	303	9	24.222000
	HLA-C*14:02	295	303	9	160.553582
QALIDQKVK	HLA-C*03:03	415	423	9	12.549025
	HLA-C*12:03	415	423	9	24.389898
	HLA-C*15:02	415	423	9	47.799130
	HLA-C*14:02	415	423	9	118.200686
RRPKHLYVS	HLA-C*12:03	172	180	9	23.238602
	HLA-C*07:01	172	180	9	169.890854
VEMDPDDVN	HLA-C*12:03	32	40	9	23.292172
	HLA-C*03:03	32	40	9	132.010963
	HLA-C*07:02	32	40	9	269.879564
QSMGIQLDQ	HLA-C*15:02	374	382	9	16.159205
	HLA-C*12:03	374	382	9	23.399684
	HLA-C*14:02	374	382	9	105.104182
	HLA-C*03:03	374	382	9	111.073383
	HLA-C*07:01	374	382	9	127.693855
YGNVLDVNA	HLA-C*12:03	98	106	9	23.561883
	HLA-C*03:03	98	106	9	220.602281
	HLA-C*14:02	98	106	9	226.787962
RQMADAVSR	HLA-C*12:03	63	71	9	23.779898
	HLA-A*31:01	63	71	9	37.972570
	HLA-B*27:05	63	71	9	145.401872
	HLA-C*03:03	63	71	9	165.4283735
	HLA-A*02:06	63	71	9	184.735346
AFFVKGWAD	HLA-C*12:03	226	234	9	22.867042
	HLA-C*03:03	226	234	9	74.064483
	HLA-C*14:02	226	234	9	149.492704
VKPGTPTQE	HLA-C*12:03	251	259	9	22.867042
	HLA-C*14:02	251	259	9	162.412705
RIREFMEKE	HLA-C*12:03	235	243	9	22.867042
QIQVRNIMS	HLA-C*12:03	209	217	9	22.501423
	HLA-C*03:03	209	217	9	114.712303
	HLA-C*14:02	209	217	9	204.936823
IEPDDHLKE	HLA-C*12:03	84	92	9	22.553294

(Continues)

TABLE 1 (Continued)

Peptide	Allele	Start	End	Length	IC_{50}
MDPDDVNKN	HLA-C*12:03	34	42	9	22.657396
	HLA-C*14:02	34	42	9	79.913472

manufacture of the vaccine was the 50S ribosomal protein L7/L12, identified by its UniProt ID P9WHE3. The adjuvant was conjugated to the original B-cell epitope using an EAAAK linker to induce a targeted immunological response. Furthermore, GPGPG linkers were utilized to connect B-cell and MHC Class-I epitopes. AAY linkers were employed in order to facilitate the binding of MHC Class-II epitopes. To enhance protein identification and purification, the vaccine sequence was altered through the incorporation of a 6× His tag at the C-terminus. The amino acid sequence of the resulting vaccine, with a MW of 91854.21 Da, is presented in Figure 2.

3.5 | Population coverage analysis

Using the IEDB tool for population coverage analysis, it was possible to identify the most prevalent candidate epitopes for each coverage approach. This was accomplished by analyzing IEDB database information. The distribution of MHC and HLA alleles varies in different geographic regions of the world as a direct consequence of the influence of various environmental factors. It is essential to consider population coverage as part of the process in order to develop a potentially effective vaccine that can be administered. The population exposure to MHC Class-I and Class-II combined alleles was estimated to be 100% in Europe and North America followed by East Africa, South America, and West Africa with a coverage of 99.9%. South Africa has the lowest population coverage, calculated at 95.27%. The information is summarized in Figure S4.

3.6 | Vaccine protein antigenicity and allergenicity evaluation

It was projected that the antigenicity of the vaccine protein in combination with the adjuvant would be 0.5263, based on calculations made with the VaxiJen 2.0 web server. A value of 0.5389 was determined for the antigenicity of a vaccine design that did not contain an adjuvant. Based on the findings, it was established that the vaccine design is antigenic, regardless of whether it is coupled with an adjuvant or not. Following the results of AllerTOP v2, the vaccine protein was determined to be non-allergenic regardless of whether the adjuvant was included or not. By using Toxinpred with or without an adjuvant, the toxicity of the protein was determined to be nontoxic.

3.7 | Multiepitope vaccine subunit solubility and physicochemical properties analysis

The ExPASY ProtParam tool was used to predict the physicochemical properties of proteins, and the results revealed various features related to the protein's nature. The multiepitope vaccine subunit had a MW of 91854.21 Da. The protein's predicted pI was 4.67, according to the calculations. The II was 27.24, which indicates high stability. A number larger than 40 indicates that the protein is unstable. This protein was verified to be thermostable because it had an aliphatic index of 55.35 when it was tested. Our vaccine design had a good solubility rate, as measured by the SOLpro server, and received a score of 0.848676.

3.8 | Secondary structure extrapolation

The extrapolation of secondary structure was conducted by employing the PSIPRED 4.0 software, which analyzed the inherent characteristics of the protein and subsequently inferred its secondary structure. Based on the findings, it was ascertained that 23.38% of the protein structure exhibits helix conformation, 3.40% comprises beta strands, and the remaining 73.21% of the protein adopts a coil conformation.

3.9 | Characterization of the vaccine's tertiary structure

The trRosetta tool, available at the website https://yanglab.nankai.edu.cn/trRosetta, was employed to produce an optimal model of the tertiary structure for the chimeric vaccine construct. The models were projected by employing the top five threading templates, which were chosen based on their high coverage values. The decision to refine structures was determined by selecting the model with the highest coverage score.

3.10 | Tertiary structure refinement

Following refinement, the Galaxy Refine tool produced a total of five models of the vaccine chimera. Several parameters were taken into consideration during the refinement process, including GDT-HA (0.9659), RMSD (0.371), and MolProbity (2.510). The calculation of clash score was 30.4, the score of poor rotamers was 1.3, and the Ramachandran score predicted that it would be 93.1%. Model 1 was chosen for further investigation.

TABLE 2 Most potential non-allergen, nontoxic, 18 T-cell epitopes with interacting MHC-II alleles.

Peptide	Antigenic score	Allele	Start	End	Length	Rank
LKDAEKAVEMDPDDV	1.0603	HLA-DPA1*01:03	25	39	15	96
		HLA-DPA1*03:01	25	39	15	96
KDAEKAVEMDPDDVN	1.0348	HLA-DPA1*01:03	26	40	15	97
		HLA-DPA1*03:01	26	40	15	97
DAEKAVEMDPDDVNK	0.6482	HLA-DPA1*02:01	27	41	15	97
		HLA-DPA1*01:03	27	41	15	97
		HLA-DPA1*02:01	27	41	15	97
		HLA-DPA1*03:01	27	41	15	97
AEKAVEMDPDDVNKN	0.7843	HLA-DPA1*02:01	28	42	15	97
		HLA-DPA1*01:03	28	42	15	97
		HLA-DPA1*03:01	28	42	15	97
EKAVEMDPDDVNKNT	0.6607	HLA-DPA1*02:01	29	43	15	98
KAVEMDPDDVNKNTL	0.6259	HLA-DPA1*02:01	30	44	15	97
		HLA-DPA1*01:03	30	44	15	97
AVEMDPDDVNKNTLQ	0.4232	HLA-DPA1*02:01	31	45	15	98
VEMDPDDVNKNTLQA	0.5111	HLA-DPA1*02:01	32	46	15	97
		HLA-DPA1*01:03	32	46	15	97
RQMADAVSRKKMDTK	0.4714	HLA-DPA1*01:03	63	77	15	95
		HLA-DPA1*01	63	77	15	95
VSRKKMDTKPTDPTG	0.8362	HLA-DPA1*02:01	69	83	15	98
SRKKMDTKPTDPTGI	0.9144	HLA-DPA1*02:01	70	84	15	98
RKKMDTKPTDPTGIE	1.0164	HLA-DPA1*02:01	71	85	15	98
KKMDTKPTDPTGIEP	1.1317	HLA-DPA1*01:03	72	86	15	97
		HLA-DPA1*02:01	72	86	15	97
KMDTKPTDPTGIEPD	1.1837	HLA-DPA1*01:03	73	87	15	97
		HLA-DPA1*02:01	73	87	15	97
MDTKPTDPTGIEPDD	1.2460	HLA-DPA1*01:03	74	88	15	97
		HLA-DPA1*02:01	74	88	15	97
DTKPTDPTGIEPDDH	1.0652	HLA-DPA1*01:03	75	89	15	97
		HLA-DPA1*02:01	75	89	15	97
TKPTDPTGIEPDDHL	0.7161	HLA-DPA1*02:01	76	90	15	97
KPTDPTGIEPDDHLK	0.7655	HLA-DPA1*01:03	77	91	15	97
		HLA-DPA1*02:01	77	91	15	97

3.11 | 3D structure validations

MolProbity server validated the revised tertiary structure. The structure of the protein was examined, and a Ramachandran plot was created (Figure S5). Before refining, 89.4% of the plot's region was in the favored region, 97.2% structural region in the allowed regions, and 2.8% residues in the outlier region. MolProbity produced improved results after executing the refinement. A total of 93.1% of residues were in the favored region, 98.6% in the allowed region, and 1.3% in the outlier region.

3.12 | Molecular docking with TLR7

Molecular docking was carried out to predict the interaction between the refined vaccine construct and the ligand-binding domain of the immune receptor TLR7. This was done using the online protein–protein docking server Cluspro2.0, which is designed specifically for this purpose. Docking allows for the examination of several models at the same time. When all 29 docked postures were analyzed, it was determined that model number 9 had the best-docked structure, with 18 hydrogen bonds, 34 ionic interactions, 16 salt bridges, 6 carbon bonds, 3 pi–alkyl interactions, and 1 donor–donor interaction with a docking score of −650.2, as shown in Figure 3. The vaccine and receptor-interacting residues are represented in Figure S6.

3.13 | Molecular dynamics simulation using NMA

iModS conducted an in-depth analysis of the structure by modifying the force field around the complex regarding several different time

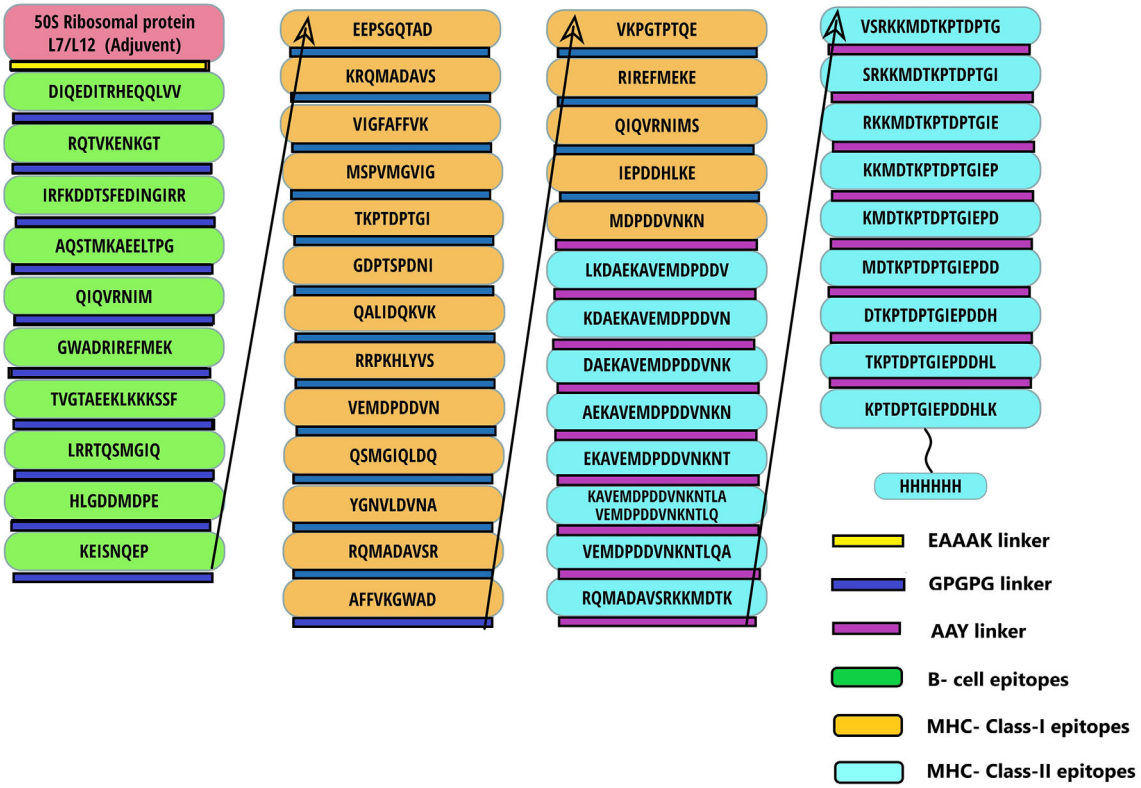

FIGURE 2 Multiepitope vaccine construct.

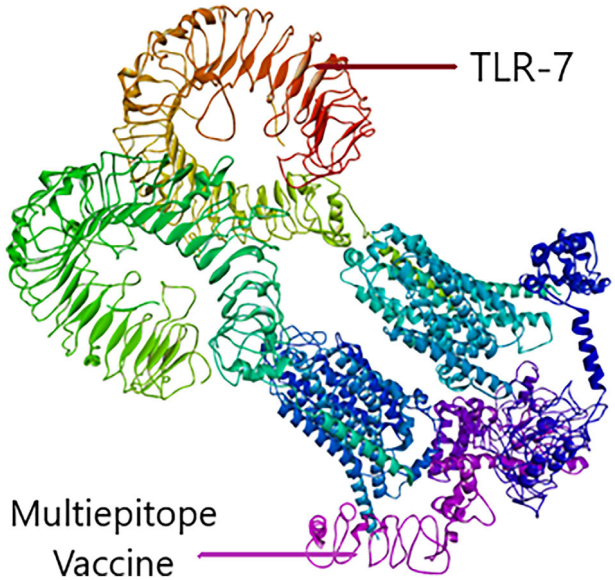

FIGURE 3 Docked complex of vaccine construct and TLR-7.

periods. The resultant model exhibits reduced distortion at the capacity level of each residue. The eigenvalue of the complex was found to be 2.184195e−06. Better relationships between individual residues were highlighted by heat maps with low RMSD and significantly co-related areas (Figure 4). The image presents a comprehensive analysis of the findings from the iModS molecular dynamics simulation. Figure 4A displays the NMA mobility in the supplied protein structure, and Figure 4B demonstrates deformability, which illustrates that there is only a small amount of deformation at each of the residues in question. Figure 4C depicts the B-factor in its entirety. The eigen values can be seen in Figure 4D, and the variance that can be accounted for is shown in Figure 4E, colored in red and green, respectively. Figure 4F,G displays the covariance and the elastic network associated with the complex.

3.14 | Codon optimization of designed vaccine peptide for expression analysis

Codons were optimized with the use of the Java codon adaptation tool, also known as JCat, so that maximum protein expression could be achieved. The optimized codon had a CAI value of 0.96 for a sequence with a GC content calculated as 72.11%. These findings suggest consistent vector expression in *E. coli*, which is supported by the GC content ranging from 30% to 75% (Figure S7). After being amplified by in silico PCR using SnapGene software, the improved sequence was cloned in a pET-9c vector to form a recombinant plasmid (Figure 5A,B).

3.15 | Immune simulation

The C-ImmSim server was used to carry out the immune simulation. This shows an immunological response that is similar to a real

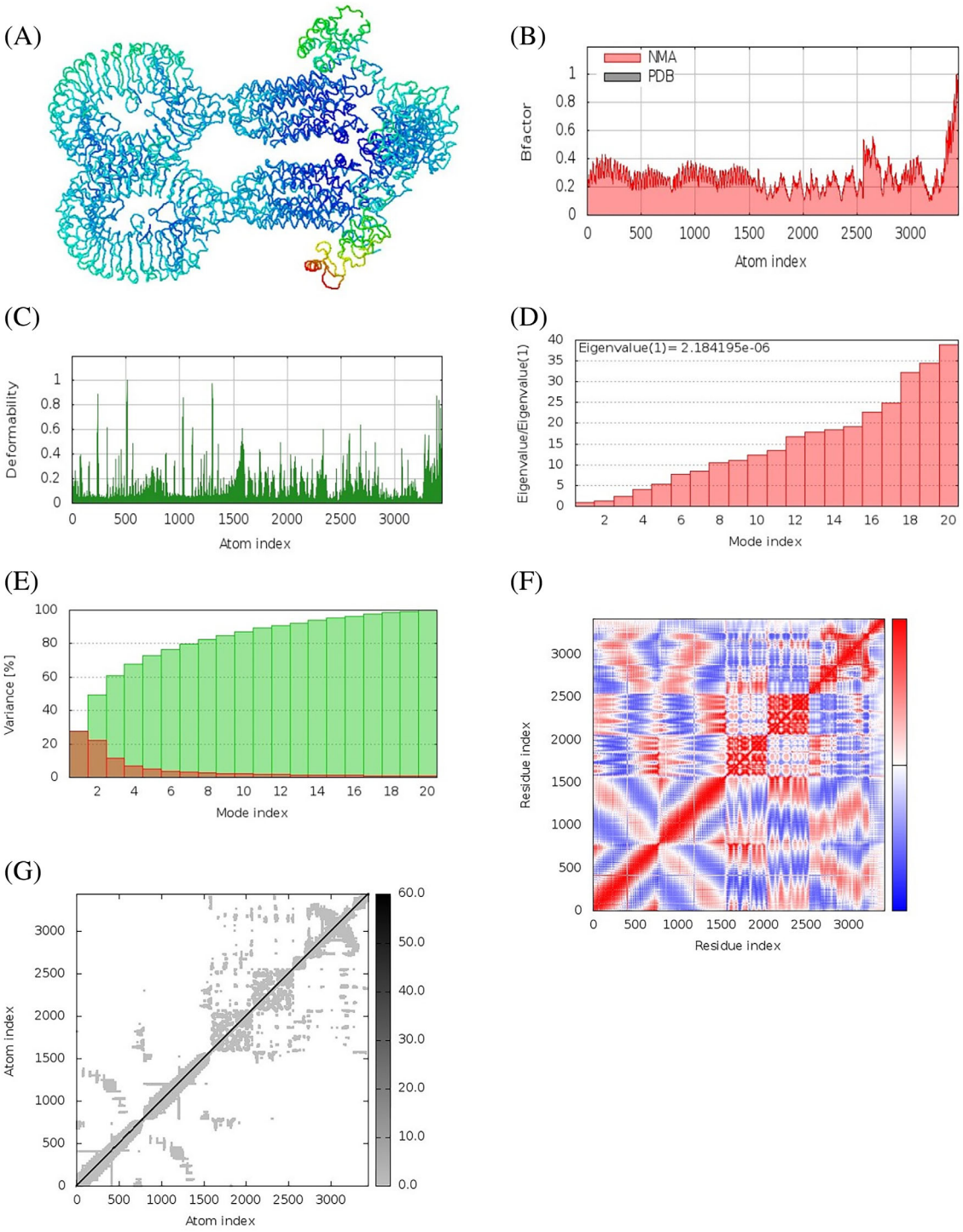

FIGURE 4 The results of molecular dynamic simulation study of vaccine construct and TLR-7 docked complex. (A) NMA mobility, (B) deformability, (C) B-factor, (D) eigenvalue, (E) variance (red color indicates individual variances and green color indicated cumulative variances), (F) covariance map (correlated [red], uncorrelated [white], or anti-correlated [blue] motions), and (G) elastic network (darker gray regions indicate more stiffer regions).

immune response. The initial response was characterized by an increase in IgM + IgG levels, followed by IgM, and IgG1 + IgG2 levels (Figure 6A). Large numbers of B cells were characteristic of the secondary and tertiary phases of the immune response (Figure 6B). The findings also revealed the development of memory cells after further exposure. There was also a rise in the concentration of helper (TH) cells (Figure 6C) and cytokines (Figure 6D).

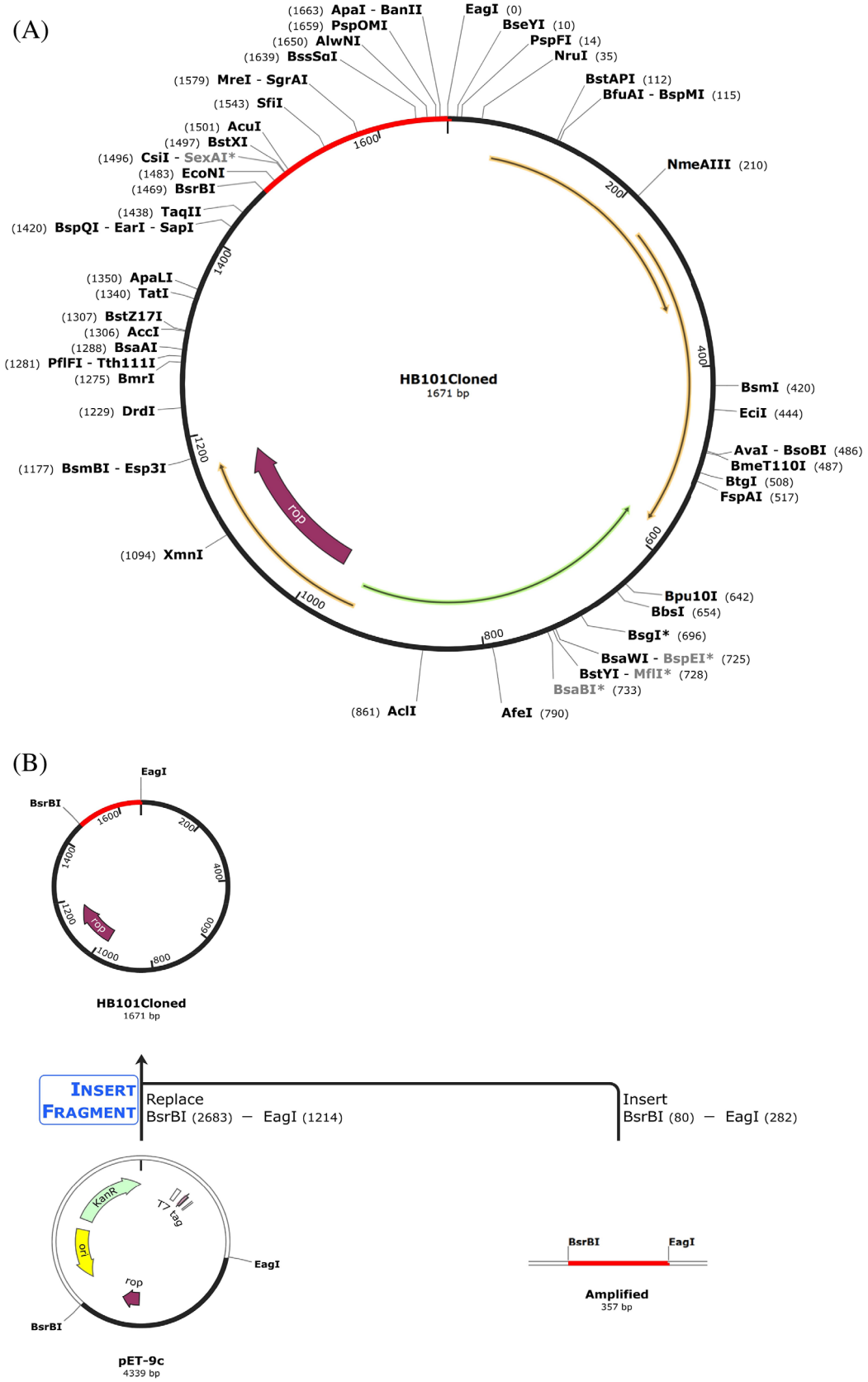

FIGURE 5 (A) In silico PCR amplification of vaccine construct followed by addition of restriction sites and cloning in pET-9c vector. (B) Recombinant plasmid obtained after cloning of peptide in vector. Vaccine construct is showed in red color and black line is representing the vector backbone.

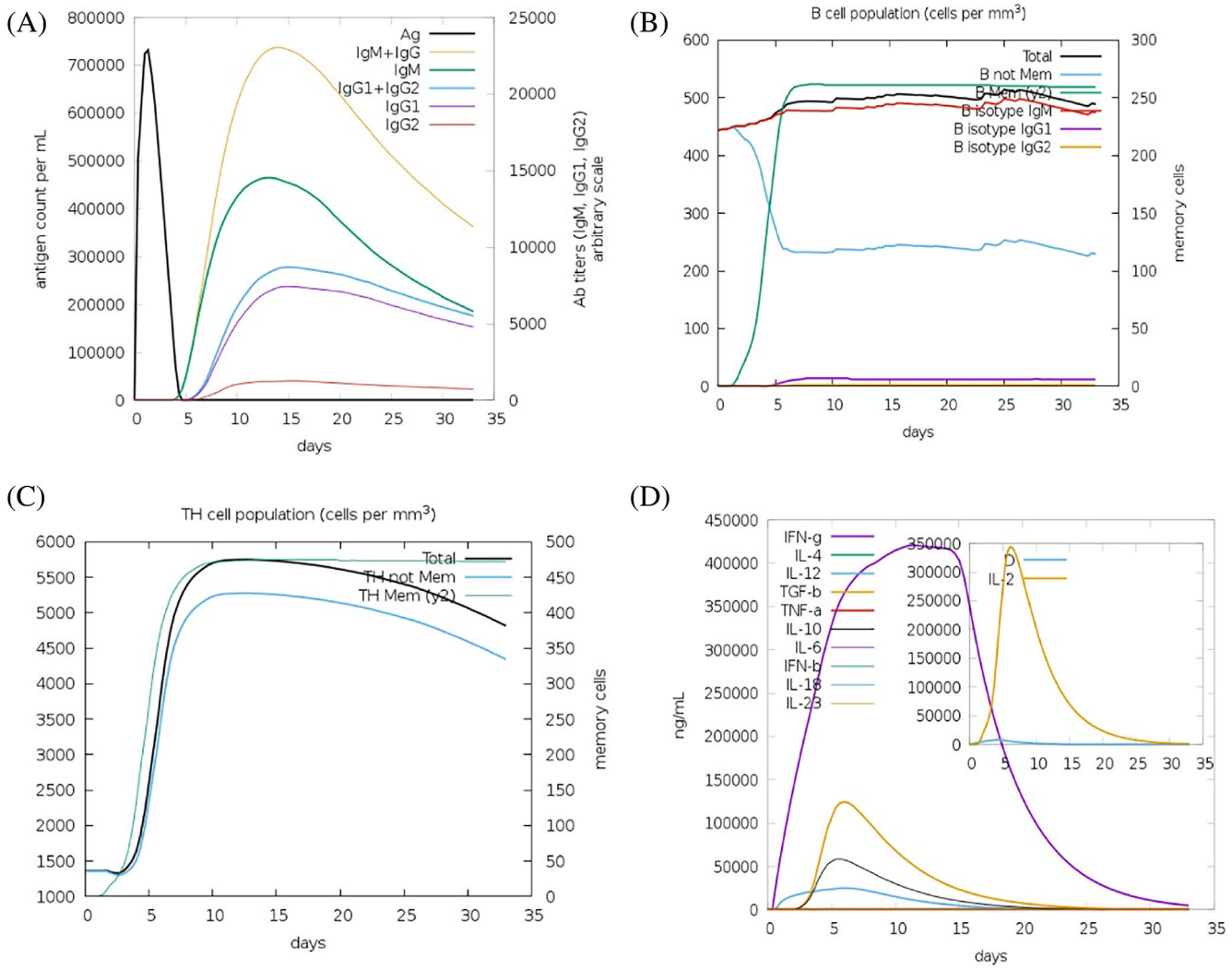

FIGURE 6 Immune simulation of vaccine constructs (A) production of immunoglobulin molecules on subsequence injection of antigens (shown by black lines); colored lines are indicating classes of immune cells. (B) Changes in B-cell population and production of memory. (C) Production of helper T-cells. (D) Elevated rates of cytokines and interleukins for effective immune response.

4 | DISCUSSION

Due to the growing number of outbreaks throughout the globe, it is urgent that we make progress in our search for a vaccine or treatment for Puumala virus. Bioinformatics methods are rising in significance for creating peptide-based vaccines, thanks to the improvement of computationally assisted sequence-based technology. Dengue virus, chikungunya virus, rhinovirus, and the SLE virus are just a few of the others that have benefited from using peptide-based vaccine design. Because of the rapid mutation rate in its genome, the RNA-based Puumala virus exhibits a wide range of resistance. As a result, the nucleoprotein in the virus is the primary focus of this investigation. Physical and chemical analysis of the target protein revealed that its antigenic helical structure made it a promising antigenic vaccine candidate. Furthermore, due to the lack of availability of the PDB structure of the associated protein in several protein data banks, bioinformatics tools were used to build the structure of the linked protein. Taken as a whole, this research demonstrates the potential of in silico methods for developing a subunit peptide vaccine based on several epitopes, which might improve humoral and cell-mediated immune responses. It used to be that B-cells were the sole option for creating a vaccine. Recent developments in computational biology have led to the emergence of a new field of clinical research: targeting MHC T-cells.[48] Antigenic drift, which occurs over time, has the potential to free the antigen from the antibody memorial response. However, even if T-cell immunity provides a long-lasting immune response, the trip from epitope to vaccine must meet certain requirements to be considered successful.[49] First, we created a database including all of the potential epitopes of the M protein. We have chosen five methods from the IEDB database to predict the antigenicity of B-cell epitopes. There is a correlation between the presence of epitopes and features of polypeptide chains such as turns, hydrophilicity, flexibility, polarity exposed surface, accessibility, and antigenic propensity. Each residue's ability to contribute to epitope formation is graphically represented in computer analysis, such as that of Kolaskar and Tongaonkar, the Emini surface, the Chou and Fasman beta-turn,

the Karplus and Schulz linear epitope, and the Bepipred linear epitope.[50] On the other hand, three of them disclose the epitope peptide sequence for additional investigation. IEDB, on the other hand, is helpful when looking at T cell binding and processing predictions.[51] We looked for putative T-cell epitopes with an IC50 value of less than 300, indicating that they are highly active against their target allele. The consensus method allows epitopes that bind to more than five MHC Class-I and MHC Class-II molecules to be sorted out for further analysis. The most important of these is the examination of the antigenic properties of both B and T cells. With the assistance of the Vaxijen v2.0 programme, the ACC value of the peptide was determined by using its physicochemical properties as the basis for the computation. Peptides have a value that is higher than the cutoff point that has been established for antigens. The antigenic epitope must be harmless in its natural environment in order to elicit a protective immune response successfully. To determine whether or not a peptide will be harmful, Toxinpred applies an SVM classifier to predict the toxicity data. The presence of allergenicity is a significant barrier to vaccine development.[52] Many current vaccines have unwanted side effects that counteract whatever protective effect they could have. AllerTop v.2.0 uses the quantitative structure–activity relationship (QSAR) to determine the presence or absence of allergens. Auto cross-covariance (ACC) indicated epitopes without allergenicity were employed to create the vaccine after being filtered out as non-allergens. However, according to the FAO/WHO methodology, a sequence is likely allergenic if it comprises at least six contiguous amino acids when compared to the amino acid database of known allergens.[28] According to AllerTop v.2.0, the FAO/WHO evaluation technique of allergenicity prediction does not agree with our chosen epitopes; hence they are not allergens. The probable antigenic epitopes were shown to be free of allergenicity and toxicity, producing immunoreactive peptides for future investigation. IEDB was used to perform a T-cell population coverage analysis due to the enormous polymorphism of MHC molecules, which can be found in thousands of different human MHC alleles (HLAs). To this end, a variety of T-cell peptides bound to various HLA molecules were examined. After meeting all criteria, 10 B-cell epitopes, 18 MHC Class-I T-cell epitopes, and 18 MHC Class-II T-cell epitopes were selected as vaccine components. 50S ribosomal protein L7/L12 is responsible for increasing the response of vaccines in pathogen recognition as well as in immune system activation and has been used as an adjuvant to increase the immunoreactive property of vaccines.[53] Suitable linkers were employed to fuse specific B-cell and T-cell epitopes in order to build the multiepitope subunit vaccine design. The spacer sequences are essential in the development process of vaccines because of their optimal effect. In light of the findings from the earlier studies, the GPGPG and AAY linkers that were present between the projected epitopes were combined to produce a prospective vaccine with the maximum humanly conceivable antigenicity. As part of the process of optimizing the sequence, an EAAAK linker was included in the sequence design in order to facilitate the connection between the adjuvant and the initial B-cell epitope. Known as a polyhistidine tag, the 6 His tag located at the C terminal of a sequence is a motif consisting of at least 6 histidine residues fused to the sequence's carboxyl (C-) terminal. It is easier for the sequence to function in a buffer state because histidine residues can attach to immobilized ions.[54]

A bioinformatic examination in conjunction with immunologic research determined that the protein sequence that had been produced possessed neither allergenic nor toxic characteristics. The designed vaccine chimera had a satisfactory antigenic score whether it was bound to adjuvant or not. The MW of the suggested vaccine protein was determined to be 91854.21, and it was then tested for solubility. The theoretical isoelectric point (pI) of the vaccine design is 4.67. The vaccine has a low II, with a score of 27.24, indicating that the developed vaccine protein is stable. According to aliphatic index estimates, the thermostability of the chimeric vaccine design was projected to be high. The secondary and tertiary structures are thought to be critical in the design of the vaccine. According to the investigation results, the secondary structure of the chimeric vaccine protein was composed of 23.38% helix, 3.40% beta strands, and 73.21% coil. GalaxyRefine was used to refine the 3D structure of a chimeric vaccine construct, resulting in the development of desirable features. The MolProbity (Ramachandran Plot) plot illustrates the traits that are required for a prospective vaccine candidate were present. The results confirmed that the majority of the residues are found in favorable areas, with only a small number of residues found in the outlier zone. It represents the acceptable desired model quality.

The next step in vaccine development is essential to achieving promising results. When it comes to predicting protein–protein interactions, immunoinformatics using computational methods like molecular docking has been proven effective. TLR7, in particular, has been implicated as a mechanism by which the immune system responds to the Puumala virus. According to one study, the action of innate immunity is highly specific to TLRs.[55] Model 9 of the vaccine-receptor complex has potential interactions proximate to the reference structure, as demonstrated by Clustpro2.0. Docked complexes that have the lowest energy values show that the interaction is stable and that there is reduced RMSD from the initial configuration. Ligands are stable inside the receptor's binding pocket because of strong electrostatic van der Waals, hydrogen bonds, and hydrophobic interactions.[56] Complex physiological movement can be captured in one shot via docking. There is a need for a more flexible environment to research protein–protein interactions. A molecular dynamics simulation was used to simulate our dynamic system's inherent behavior. Figure 4 shows the numerous factors that influenced the outcome. NMA mobility, deformation, B-factor, eigenvalue, c-variance, and elasticity of the complex are shown in the figure. The built complex is proved to be stable based on the maximum eigenvalue and shows fewer deformation changes during immunological response. The body's humoral reaction was expected to be generated as a result of vaccination. The immune simulation was used to elicit a typical immune response. Repeated exposure to antigens increases the immunological response. For several months, it was clear that memory B-cells had been established. Memory T-cells and helper T-cells simulation were also developed. After the initial exposure, there was a noticeable increase in IL-2 levels. The proposed vaccine was evaluated for immunoreactivity by expressing it in *E. coli*,[57] which was done to validate it in a simulated environment. A CAI of 0.96 and GC content of 72.11%

were achieved by optimizing the codon by the host to ensure maximal expression. Based on the above results, in vitro and in vivo research on the vaccine candidate could show it to be effective. The immunoinformatics approach used in this study for the Puumala orthohantavirus is broadly applicable to other hantaviruses and related zoonotic pathogens. Our method involves predicting and selecting B-cell and T-cell epitopes, utilizing linkers and adjuvants to enhance immunogenicity, and simulating immune responses computationally, all of which can be adapted to design vaccines against similar viruses. This approach is not only economical but also aids in identifying conserved epitopes that might promote cross-reactivity and broader vaccine coverage but also accelerates the vaccine development process by allowing for the early identification of promising vaccine constructs before experimental validation. Such adaptability underscores the potential of our methodology to contribute to global health security by facilitating rapid responses to emerging zoonotic threats.

5 | CONCLUSION

Computational methodologies were applied in the current work to design an efficient multiepitope vaccine against Puumala virus. A favorable outcome of the simulated environment encourages the illustrated method to design a promising vaccine as a time-saving and cost-effective strategy due to the low likelihood of failure. In silico vaccines with a favorable immune response and high population coverage have the potential to be tested in clinical studies. Snapgene was used to perform computational cloning in the pET-9c plasmid, which resulted in high protein expression. On the other hand, experimental validation is required to ensure the efficacy of the vaccine design against the Puumala virus. Peptide vaccines have demonstrated positive effects with a strong immune response; hence, this vaccine construct may be considered. Developing a vaccine for the Puumala virus, as a result of this research, would be extremely beneficial in eliminating the threat posed by the virus.

AUTHOR CONTRIBUTIONS

Kunal Bhattacharya: Conceptualization; methodology; software; data curation; investigation; visualization; writing – original draft; writing – review and editing; supervision; project administration. **Nongmaithem Randhoni Chanu:** Methodology; data curation; validation; writing – original draft; resources; formal analysis. **Saurav Kumar Jha:** Conceptualization; methodology; writing – original draft; writing – review and editing; formal analysis. **Pukar Khanal:** Methodology; software; investigation; visualization; writing – original draft; writing – review and editing. **Keshav Raj Paudel:** Conceptualization; methodology; investigation; writing – original draft; writing – review and editing; data curation.

ACKNOWLEDGMENT

Open access publishing facilitated by University of Technology Sydney, as part of the Wiley - University of Technology Sydney agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the Supporting Information of this article.

ORCID

Kunal Bhattacharya https://orcid.org/0000-0002-5546-1504

REFERENCES

1. Schmaljohn C, Hjelle B. Hantaviruses: A global disease problem. *Emerg Inf Dis*. 1997;3:95-104.
2. Pettersson L, Boman J, Juto P, Evander M, Ahlm C. Outbreak of Puumala virus infection, Sweden. *Emerg Inf Dis*. 2008;14:808-810.
3. Settergren B. Clinical aspects of nephropathia epidemica (Puumala virus infection) in Europe: a review. *Scandinav J Inf Dis*. 2000;32:125-132.
4. Olsson GE, Dalerum F, Hörnfeldt B, et al. Human hantavirus infections, Sweden. *Emerg Inf Dis*. 2003;9:1395-1401.
5. Sironen T, Vaheri A, Plyusnin A. Molecular evolution of Puumala hantavirus. *J Virol*. 2001;75:11803-11810.
6. Plyusnin A, Horling J, Kanerva M, et al. Puumala hantavirus genome in patients with nephropathia epidemica: correlation of PCR positivity with HLA haplotype and link to viral sequences in local rodents. *J Clin Microbiol*. 1997;35:1090-1096.
7. Plyusnin A, Mustonen J, Asikainen K, et al. Analysis of Puumala hantavirus genome in patients with nephropathia epidemica and rodent carriers from the sites of infection. *J Med Virol*. 1999;59:397-405.
8. Plyusnin A, Vapalahti O, Lehvaslaiho H, et al. Genetic variation of wild Puumala viruses within the serotype, local rodent populations and individual animal. *Virus Res*. 1995;38:25-41.
9. Vapalahti O, KallioKokko H, Salonen EM, BrummerKorvenkontio M, Vaheri A. Cloning and sequencing of Puumala virus Sotkamo strain S and M RNA segments: evidence for strain variation in hantaviruses and expression of the nucleocapsid protein. *J Gen Virol*. 1992;73:829-838.
10. Horling J, Lundkvist A, Jaarola M, et al. Distribution and genetic heterogeneity of Puumala virus in Sweden. *J Gen Virol*. 1996;77:2555-2562.
11. Horling J, Lundkvist A, Persson K, et al. Detection and subsequent sequencing of Puumala virus from human specimens by PCR. *J Clin Microbiol*. 1995;33:277-282.
12. Lundkvist A, Wiger D, Horling J, et al. Isolation and characterization of Puumala hantavirus from Norway: evidence for a distinct phylogenetic sublineage. *J Gen Virol*. 1998;79:2603-2614.
13. Davidyuk Y, Shamsutdinov A, Kabwe E, et al. Prevalence of the Puumala orthohantavirus strains in the pre-kama area of the Republic of Tatarstan, Russia. *Pathogens*. 2020;9:540.
14. Davidyuk YN, Kabwe E, Shakirova VG, et al. Characterization of the Puumala orthohantavirus strains in the northwestern region of the Republic of Tatarstan in relation to the clinical manifestations in hemorrhagic fever with renal syndrome patients. *Front Pharm*. 2019;10:447678.
15. Blinova E, Deviatkin A, Kurashova S, et al. A fatal case of haemorrhagic fever with renal syndrome in Kursk Region, Russia, caused by a novel Puumala virus clade. *Infect Genet Evol*. 2022;102:105295.
16. Xiao SY, Spik KW, Li D, Schmaljohn CS. Nucleotide and deduced amino acid sequences of the M and S genome segments of two Puumala virus isolates from Russia. *Virus Res*. 1993;30:97-103.
17. Bowen MD, Kariwa H, Rollin PE, Peters CJ, Nichol ST. Genetic characterization of a human isolate of Puumala hantavirus from France. *Virus Res*. 1995;38:279-289.

18. Escutenaire S, Chalon P, Heyman P, et al. Genetic characterization of Puumala hantavirus strains from Belgium: evidence for a distinct phylogenetic lineage. *Virus Res*. 2001;74:1-15.
19. Aberle SW, Lehner P, Ecker M, et al. Nephropathia epidemica and Puumala virus in Austria. *Eur J Clin Microbiol Infect Dis*. 1999;18:467-472.
20. Bowen MD, Gelbmann W, Ksiazek TG, Nichol ST, Nowotny N. Puumala virus and two genetic variants of Tula virus are present in Austrian rodents. *J Med Virol*. 1997;53:174-181.
21. Heiske A, Anheier B, Pilaski J, Volchkov VE, Feldmann H. A new Clethrionomys-derived hantavirus from Germany: evidence for distinct genetic sublineages of Puumala viruses in Western Europe. *Virus Res*. 1999;61:101-112.
22. Pilaski J, Feldmann H, Morzunov S, et al. Genetic identification of a new Puumala virus strain causing severe hemorrhagic fever with renal syndrome in Germany. *J Infect Dis*. 1994;170:1456-1462.
23. Plyusnin A, Vapalahti O, Ulfves K, et al. Sequences of wild Puumala virus genes show a correlation of genetic variation with geographic origin of the strains. *J Gen Virol*. 1994;75:405-409.
24. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD. Bairoch A protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook*. Springer; 2005:571-607.
25. Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT. Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res*. 2013;41:349-357.
26. Ferrè F, Clote P. DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Res*. 2006;34:182-185.
27. Doytchinova IA, Flower DR. Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines. *Open Vaccine J*. 2010;3:22-26.
28. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v.2—a server for in silico prediction of allergens. *J Mol Model*. 2014;20:2278.
29. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res*. 2015;43:W174-W181.
30. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45:W24-W29.
31. Buus S, Lauemøller SL, Worning P, et al. Sensitive quantitative predictions of peptide-MHC binding by a "Query by Committee" artificial neural network approach. *Tissue Antigens*. 2003;62:378-384.
32. Andreatta M, Karosiene E, Rasmussen M, Stryhn A, Buus S, Nielsen M. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. *Immunogenetics*. 2015;67:641-650.
33. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8:1-7.
34. Zhao S, Pan F, Cai S, Yi J, Zhou L, Liu Z. Secrets behind protein sequences: unveiling the potential reasons for varying allergenicity caused by caseins from cows, goats, camels, and mares based on bioinformatics analyses. *Int J Mol Sci*. 2023;24:2481.
35. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 2006;7:1-5.
36. Bhattacharya K, Shamkh IM, Khan MS, et al. Multiepitope vaccine design against monkeypox virus via reverse vaccinology method exploiting immunoinformatic and bioinformatic approaches. *Vaccine*. 2022;1:2010.
37. Magnan CN, Zeller M, Kayala MA, et al. High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics*. 2010;26:2936-2943.
38. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP. Peptide toxicity prediction. *Methods Mol Biol*. 2015;1268:143-157.
39. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 2008;9:1-8.
40. Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res*. 2013;41:384-388.
41. Chen VB, Arendall WB, Headd JJ. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr*. 2010;D66:12-21.
42. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein-protein docking. *Nat Protoc*. 2017;12:255-278.
43. López-Blanco JR, Aliaga JI, Quintana-Ortí ES, Chacón P. IMODS: internal coordinates normal mode analysis server. *Nucleic Acids Res*. 2014;42:271-276.
44. Grote A, Hiller K, Scheer M, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res*. 2005;33:526-531.
45. Morla S, Makhija A, Kumar S. Synonymous codon usage pattern in glycoprotein gene of rabies virus. *Gene*. 2016;584:1-6.
46. Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, Prajapati VK. Exploring dengue genome to construct a multiepitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Sci Rep*. 2017;7:1-13.
47. Shey RA, Ghogomu SM, Esoh KK, et al. In silico design of a multiepitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci Rep*. 2019;9:1-18.
48. Naveed M, Tehreem S, Arshad S, et al. Design of a novel multiple epitope-based vaccine: an immunoinformatics approach to combat SARS-CoV-2 strains. *J Infect Public Health*. 2021;14:938-946.
49. Phan AT, Goldrath AW, Glass CK. Metabolic and epigenetic coordination of T cell and macrophage immunity. *Immunity*. 2017;46:714-729.
50. Oany AR, Emran A, Jyoti TP. Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. *Drug Des Devel Ther*. 2014;8:1139-1149.
51. Khan S, Khan A, Rehman AU, et al. Immunoinformatics and structural vaccinology driven prediction of multiepitope vaccine against Mayaro virus and validation through in silico expression. *Infect Genet Evol*. 2019;73:390-400.
52. Herman RA, Song P, ThirumalaiswamySekhar A. Value of eight-amino-acid matches in predicting the allergenicity status of proteins: an empirical bioinformatic investigation. *Clin Mol Allergy*. 2009;7:1-7.
53. Qamar TU, Ahmad M, Fatima S, et al. Designing multiepitope vaccine against *Staphylococcus aureus* by employing subtractive proteomics, reverse vaccinology and immuno-informatics approaches. *Comput Biol Med*. 2021;132:104389.
54. Bornhorst JA, Falke JJ. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol*. 2000;326:245-254.
55. Duan T, Du Y, Xing C, Wang HY, Wang F. Toll-like receptor signaling and its role in cell-mediated immunity. *Front Immunol*. 2022;13:812774.
56. Patil R, Das S, Stanley A, Yadav L, Sudhakar A, Varma AK. Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of drug-designing. *PLoS One*. 2010;5:e12029.
57. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol*. 2014;5:1-17.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bhattacharya K, Chanu NR, Jha SK, Khanal P, Paudel KR. *In silico* design and evaluation of a multiepitope vaccine targeting the nucleoprotein of *Puumala orthohantavirus*. *Proteins*. 2024;1-16. doi:10.1002/prot.26703