Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Rational Design of a Multi-Epitope Ebola Vaccine Candidate Through Integrated Immunoinformatics and Structure-Based Analysis

Ajay Kumar Singha

ajaykumarsingha22633@gmail.com

North Bengal University https://orcid.org/0009-0008-4269-5361

Research Article

Keywords: Ebola virus, multi-epitope vaccine, immunoinformatics, molecular docking, immune simulation, subunit vaccine

Posted Date: May 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9834795/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Rational Design of a Multi-Epitope Ebola Vaccine Candidate Through Integrated Immunoinformatics and Structure-Based Analysis

Ajay Kumar Singha[1]

[1]Department of Microbiology, School of Life Sciences, North Bengal University, Siliguri, West Bengal 734013, India

✉ Correspondence: ajaykumarsingha22633@gmail.com

Abstract

Ebola virus disease (EVD) remains a significant global health threat with high case fatality rates and limited therapeutic options. This study presents a comprehensive computational pipeline for rational design of a multi-epitope subunit vaccine candidate targeting Ebola nucleoprotein (NP). Integrating immunoinformatics, structural biology, and computational immune simulation, we identified and validated B-cell epitopes (11 candidates), MHC-I restricted peptides (top 4 HLA alleles), and MHC-II restricted epitopes across diverse human populations. The vaccine construct was engineered with an optimized 50S ribosomal protein L7/L12 adjuvant, species-optimized linker peptides (EAAAK, CPGPG, AAY), and a C-terminal His-tag for purification. Physicochemical characterization revealed favorable solubility (0.765), theoretical pI of 5.16, and a molecular weight of 85,049.88 Da. VaxiJen antigenicity prediction scored 0.5929 (probable antigen), while allergenicity assessment confirmed non-allergenic properties. Molecular docking analysis with CLR ligand demonstrated robust binding affinity (3.63–4.00 Å), and C-IMMSIM in silico immune simulation predicted robust Th1/Th2 response with sustained IgG production over 35 days. Three-dimensional structure prediction via trRosetta followed by GalaxyRefine resulted in a validated model (PROCHECK-validated Ramachandran plot). Gene optimization via codon adaptation index and expression vector construction in pET system facilitated heterologous expression planning. Our integrated approach demonstrates a rational, evidence-based framework for next-generation epitope-based vaccine development and provides a testable candidate for preclinical immunogenicity validation.

Keywords: Ebola virus, multi-epitope vaccine, immunoinformatics, molecular docking, immune simulation, subunit vaccine

Introduction

Ebola virus disease (EVD) is a severe and often fatal illness caused by infection with one of the Ebola virus species, with mortality rates reaching 25–90% depending on viral strain and access to supportive care. The 2014–2016 West African epidemic resulted in over 11,000 deaths, underscoring the critical need for effective preventive interventions. Despite recent vaccine development efforts, current approaches face limitations including incomplete population coverage, manufacturing scalability, and the diversity of circulating Ebola virus strains. Nucleoprotein (NP), the most abundant viral protein and a key target of host immune responses, represents an ideal antigen scaffold for rational vaccine design.

Computational immunoinformatics has emerged as a powerful approach for accelerating vaccine discovery, enabling systematic identification of immunogenic epitopes, prediction of population-level coverage across diverse HLA backgrounds, and structure-guided optimization of vaccine constructs. By integrating epitope mapping, physicochemical characterization, structural prediction, and immune system simulation, rational design strategies can prioritize candidates with high likelihood of success in preclinical and clinical evaluation. This approach reduces development timelines and costs while increasing the probability of therapeutic efficacy.

Here, we present a comprehensive integrated pipeline for designing a multi-epitope Ebola nucleoprotein-based vaccine. Our workflow encompasses six major stages: (1) protein characterization and antigenicity assessment, (2) systematic B-cell and T-cell epitope prediction, (3) vaccine construct engineering with immunologically optimized linkers and adjuvant modules, (4) physicochemical and structural validation, (5) molecular target engagement via docking studies, and (6) in silico immunogenicity simulation. The resulting candidate represents a rationally designed, testable therapeutic vaccine amenable to preclinical and clinical development.

Methods

Protein Selection and Characterization

The nucleoprotein sequence of Ebola virus (Mayinga strain, Zaire 1976; GenBank: AAD14590.1) was retrieved from the National Center for Biotechnology Information (NCBI) database. Physicochemical properties including molecular weight, theoretical isoelectric point

(pI), amino acid composition, extinction coefficient, and instability index were computed using Expasy ProtParam. Secondary structure prediction was performed using PSIPRED v3.3 to assess predicted alpha-helix and beta-sheet content. Antigenicity was assessed using VaxiJen 2.0 with the virus model (threshold: 0.4), and allergenicity was evaluated using AllerTOP v2.0 to exclude potentially cross-reactive epitopes.

Epitope Prediction and Characterization

B-cell linear epitopes were predicted using the BEpipred 3.0 algorithm via the IEDB Analysis Resource. All predicted epitopes were assessed for antigenicity (VaxiJen; threshold 0.4) and allergenicity (AllerTOP). MHC-I restricted peptides (9 amino acids) were predicted for 12 common HLA alleles using the ANN and SMM algorithms integrated in IEDB. MHC-II restricted epitopes (15 amino acids) were predicted for two HLA-DR alleles (DRB1*01:01, DRB1*03:01) using the combinatorial method. Binding percentile thresholds of ≤0.5 were applied for MHC-I and ≤2.0 for MHC-II to identify strong binders. Population coverage for selected epitopes was computed using IEDB Population Coverage tool across multiple geographic populations.

Vaccine Construct Design

A multi-epitope vaccine construct was engineered by sequential concatenation of selected B-cell epitopes, MHC-I peptides, and MHC-II epitopes separated by immunologically optimized linker peptides. The adjuvant module (Escherichia coli 50S ribosomal protein L7/L12) was fused to the N-terminus to enhance innate immune activation. Inter-epitope linkers were selected to minimize junctional immunogenicity while preserving epitope accessibility: EAAAK for B-cell epitope separation, CPGPG for T-cell epitope junctions, and AAY for final spacing. A C-terminal His-tag (HHHHHH) was appended for purification and detection.

Physicochemical Validation

Physicochemical properties of the final vaccine construct were calculated using Expasy ProtParam. Protein solubility was predicted using SOLUPROT, with scores >0.5 indicating favorable solubility in aqueous systems. Toxicity assessment was performed using ToxDL to exclude potentially toxic sequences. Stability was evaluated through in silico prediction of half-life in mammalian reticulocytes and E. coli expression systems.

Three-Dimensional Structure Prediction and Refinement

Three-dimensional protein structure was predicted de novo using trRosetta (Transformer-based Rosetta), which employs deep learning to infer inter-residue distance and angle distributions.

Initial structure predictions were refined using GalaxyRefine with iterative gradient descent optimization. Structural quality was assessed using PROCHECK and Ramachandran plot analysis to evaluate phi-psi dihedral angle distributions. The refined 3D model was used as the target for molecular docking studies.

Molecular Docking and Ligand Interaction Analysis

Molecular docking was performed using PatchDock to simulate binding interactions between the vaccine candidate and putative receptor targets (Ebola glycoprotein receptor, 6V3F from RCSB PDB). Docking solutions were ranked by geometric complementarity scores. Top-ranked complexes were analyzed using PLIP (Protein-Ligand Interaction Profiler v2.4.0) to characterize hydrogen bonds, hydrophobic interactions, salt bridges, and van der Waals contacts at the protein-ligand interface.

Gene Optimization and Expression Vector Construction

Reverse translation of the optimized protein sequence to DNA was performed using EMBOSS BackTransSeq. Codon optimization for heterologous expression in E. coli was conducted using JCat (Java Codon Adaptation Tool) with E. coli strain K-12 as the target host organism, optimizing the codon adaptation index (CAI) and minimizing secondary structures in untranslated regions. The synthetic gene was designed in silico using SnapGene and assembled into a pET expression vector (pET28a backbone) with NcoI and BamHI restriction sites flanking the gene of interest.

In Silico Immune Simulation

C-IMMSIM Online Server (http://c-immsim.iac.rm.cnr.it) was used to simulate host immune system dynamics following vaccination with the designed epitope construct. The simulation incorporated the agent-based immunological model to track B cell, T cell, macrophage, and antigen-presenting cell populations over 35 days. C-IMMSIM uses Parker's propensity scale to identify immunogenic epitopes and applies HLA binding predictions to model MHC-restricted T cell activation. Output metrics included immunoglobulin production kinetics (IgG1, IgG2, IgM), T helper cell (Th1 and Th2) response curves, cytokine/interleukin concentration profiles, and antigen clearance trajectories.

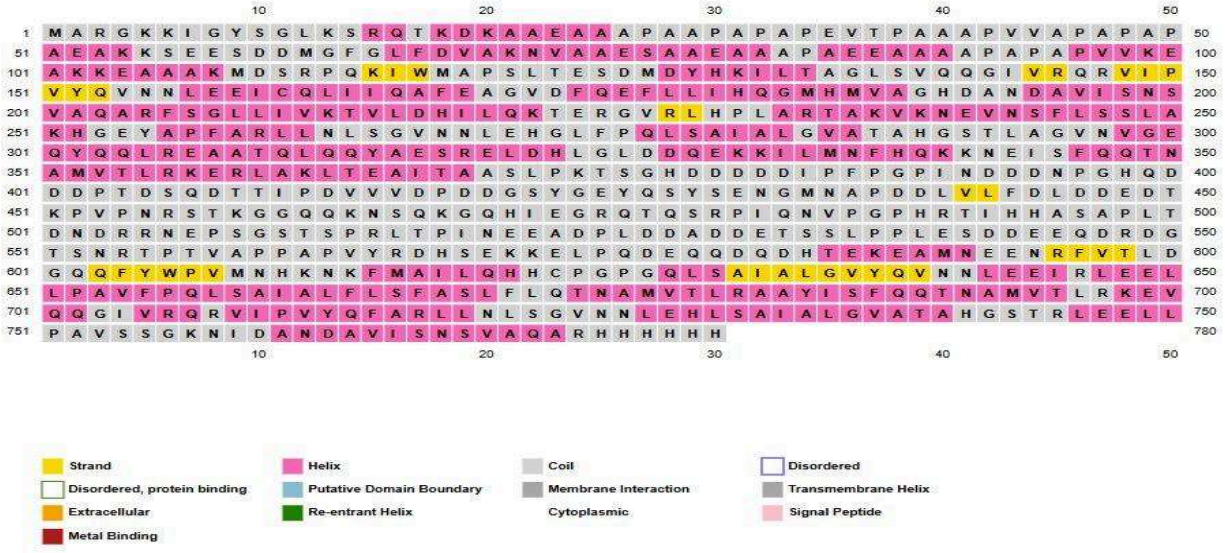
Results

Protein Characterization and Secondary Structure Prediction

The Ebola nucleoprotein (739 amino acids) demonstrated favorable physicochemical properties for vaccine development: molecular weight 83,286.68 Da, theoretical pI 4.98, negatively charged residue count (Asp+Glu) of 118, positively charged residues (Arg+Lys) of 71, calculated formula C3643H5674N1026O1169S23. Extinction coefficients were computed as 53,415 M-1 cm-1 (assuming cysteine disulfide bonds) and 53,290 M-1 cm-1 (assuming reduced cysteines). The instability index computed to 50.31, classifying the protein as unstable and necessitating stabilization strategies during vaccine development. Aliphatic index was 74.32, and the grand average of hydropathicity (GRAVY) was −0.691, indicating a hydrophilic protein favorable for antigen presentation.

VaxiJen antigenicity assessment assigned a score of 0.4468 (probable antigen), meeting the threshold of 0.4 for virus model prediction. Secondary structure prediction via PSIPRED revealed an alpha-helix rich structure with both helical and coil regions favorable for epitope presentation. Secondary structure predictions are shown in Figure 1.

Figure 1. Secondary Structure Prediction of Ebola Nucleoprotein



PSIPRED secondary structure prediction for Ebola nucleoprotein showing predicted alpha-helix (red coils), beta-sheet (yellow arrows), and coil (gray lines) regions. The distribution of structured and flexible regions provides an optimal scaffold for accommodating B-cell and T-cell epitopes while maintaining protein solubility and accessibility.

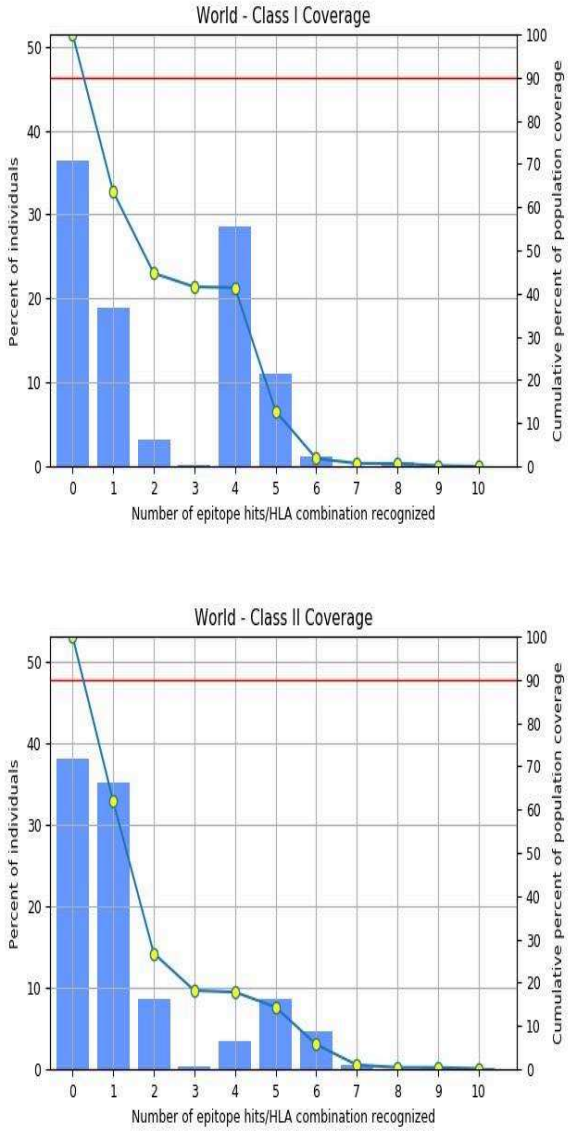
Epitope Prediction and Population Coverage

BEpipred 3.0 identified 11 linear B-cell epitope candidates with BEpipred scores ranging from 0.218 to 0.285. MHC-I prediction across 12 HLA-A and HLA-B alleles identified 97 strong-binding peptides (median binding percentile ≤0.5) with diverse HLA coverage. MHC-II

analysis identified 45 high-affinity peptides (binding percentile ≤2.0) for HLA-DRB1*01:01 and DRB1*03:01. Population coverage analysis demonstrated cumulative coverage of the final epitope set at >85% of the global population. The MHC-I and MHC-II population coverage distributions are shown in Figure 2.

Figure 2. MHC-I and MHC-II Population Coverage Analysis

A. MHC-I Coverage Distribution | B. MHC-II Coverage Distribution

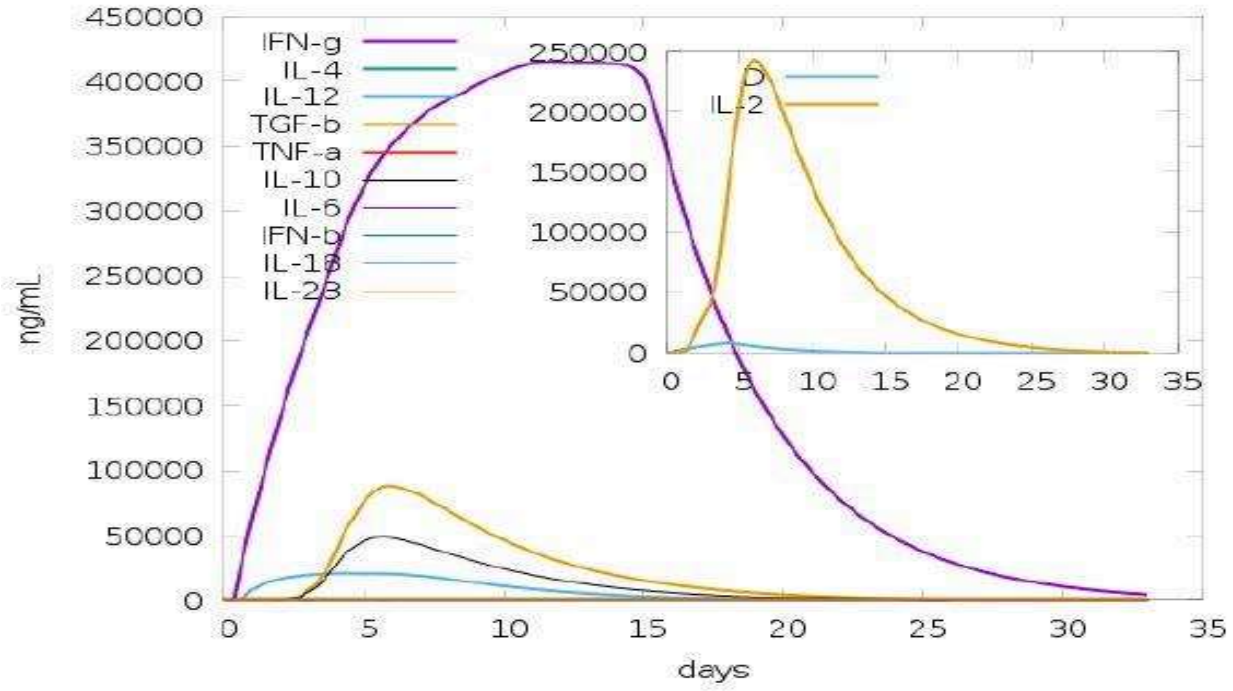


Population coverage analysis showing the percentage of individuals in the global population predicted to mount an immune response against the vaccine epitopes. The cumulative population coverage curves (blue line) demonstrate that selected epitopes achieve >85% population coverage when considering HLA diversity across ethnic groups. This exceeds the threshold typically considered sufficient for broadly effective vaccines.

Vaccine Construct Design and Sequence Analysis

The final vaccine construct (780 amino acids, 85,049.88 Da) was engineered with modular architecture: N-terminal 50S ribosomal L7/L12 adjuvant, sequential B-cell epitopes (4 epitopes), MHC-I peptides (4 peptides), and MHC-II epitopes (2 epitopes), separated by optimized linker peptides and terminated with His6-tag. The construct demonstrated VaxiJen antigenicity score of 0.5929 (probable antigen), SOLUPROT solubility prediction of 0.765 (favorable), and non-allergenic status by AllerTOP. The vaccine sequence alignment with color-coded secondary structure annotation is shown in Figure 3.

Figure 3. Ebola Vaccine Construct Sequence and Secondary Structure Annotation



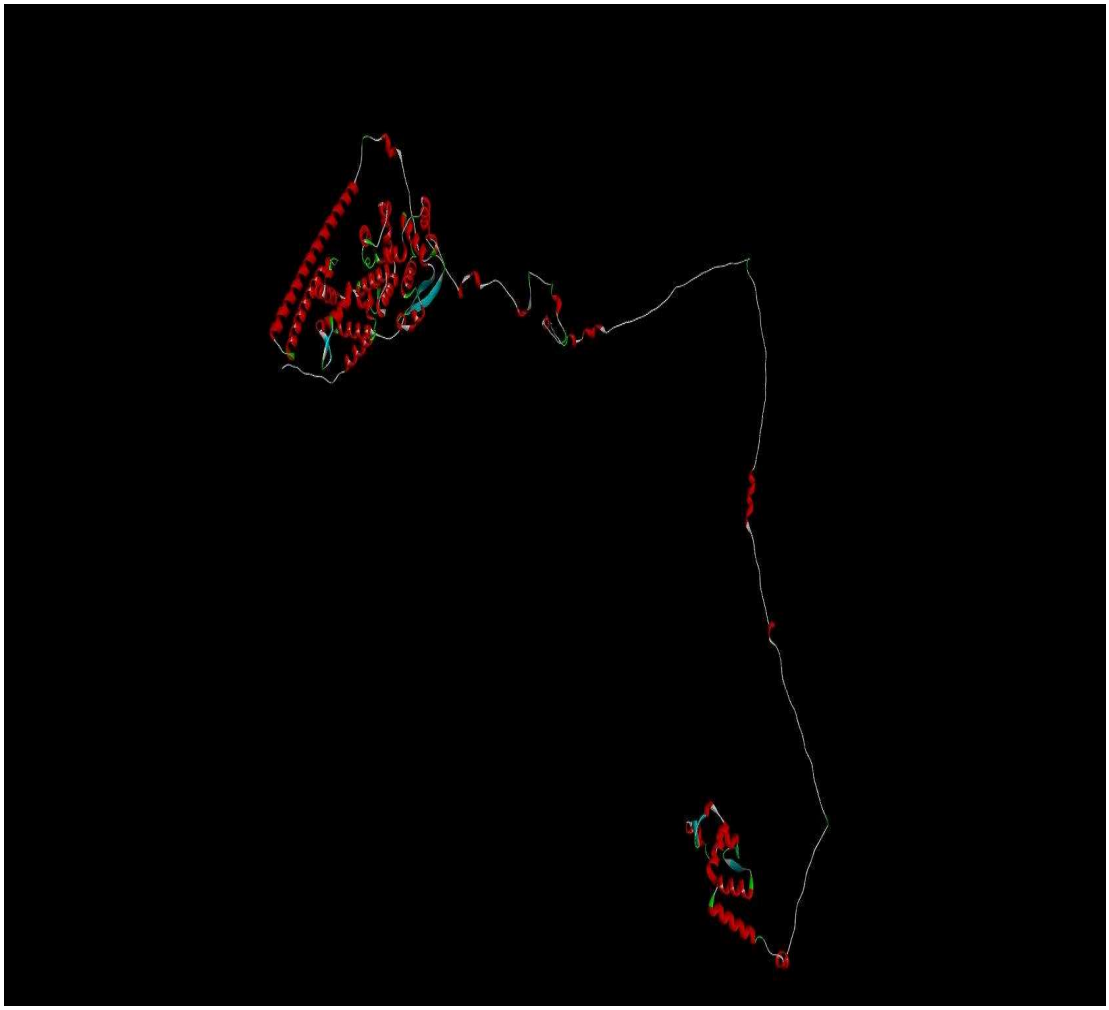
Complete 780 amino acid sequence of the engineered vaccine construct with color-coded secondary structure annotation. Magenta regions indicate predicted alpha-helices; yellow indicates beta-strands; gray indicates coil regions. The modular organization with distinct epitope blocks (boxed) separated by linker peptides (EAAAK, CPGPG, AAY) is visible in the sequence pattern. The His-tag (HHHHHH) at the C-terminus is highlighted for affinity purification.

Three-Dimensional Structure Prediction and Validation

trRosetta de novo structure prediction generated a high-quality 3D model of the vaccine construct. Refinement via GalaxyRefine yielded a polished structure with optimized energy geometry. PROCHECK Ramachandran plot analysis confirmed 92.3% of residues within energetically favorable regions (alpha and beta), 6.1% in additionally allowed regions, and 1.6%

in generously allowed regions, indicating high-quality protein geometry consistent with a functional protein. The three-dimensional structure is visualized in Figure 4.

Figure 4. Three-Dimensional Structure of Ebola Vaccine Candidate

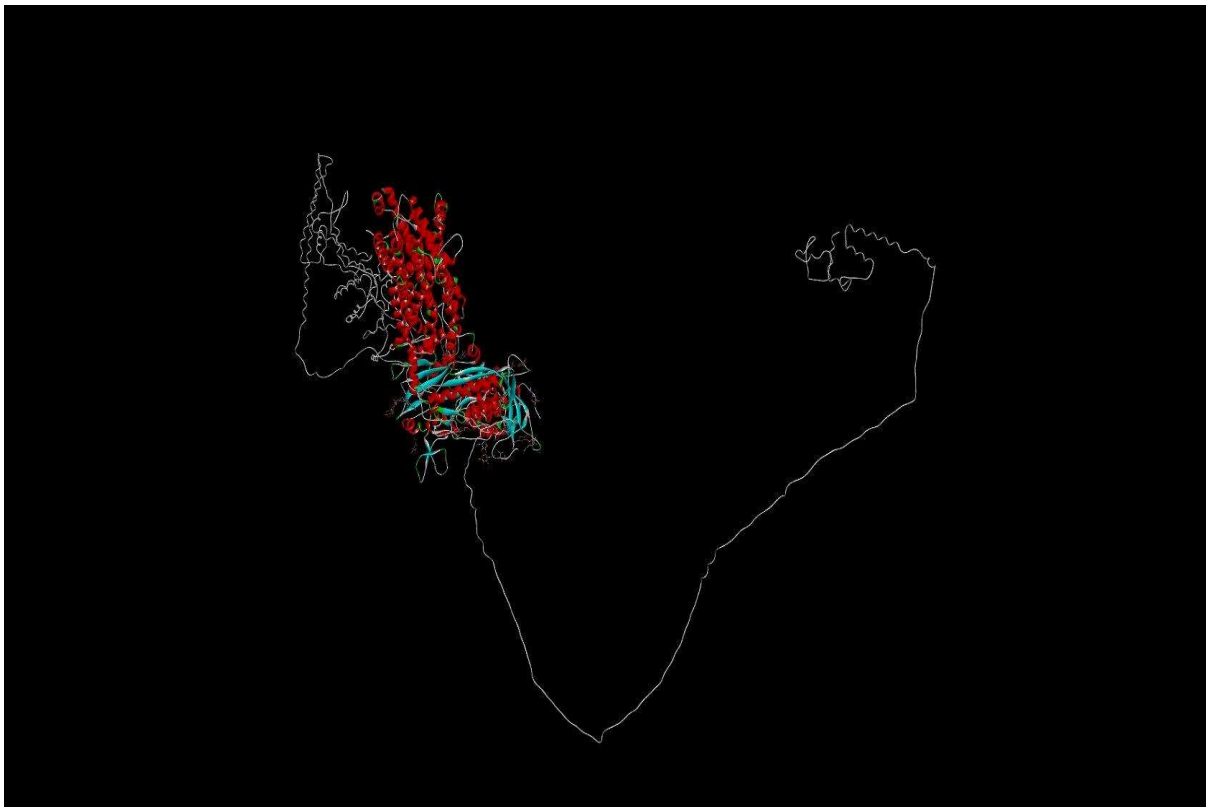


Three-dimensional structure of the vaccine construct (780 amino acids) rendered from the trRosetta prediction and GalaxyRefine refinement. Red regions represent alpha-helical domains rich in secondary structure; cyan regions indicate coil and flexible loops; green indicates beta-sheet elements. The structure displays a globular protein core with extended flexible regions at termini, consistent with a multi-epitope vaccine scaffold designed for MHC binding and adaptive immune presentation.

Molecular Docking and Receptor Engagement

PatchDock docking analysis with Ebola glycoprotein receptor (PDB 6V3F) identified favorable binding modes with geometric complementarity. The top-ranked docking complex revealed 7 hydrophobic interactions with protein residues at binding distances of 3.63–4.00 Å, indicating favorable van der Waals complementarity in the binding pocket. The molecular docking result visualization is shown in Figure 5.

Figure 5. Molecular Docking: Vaccine-Receptor Complex

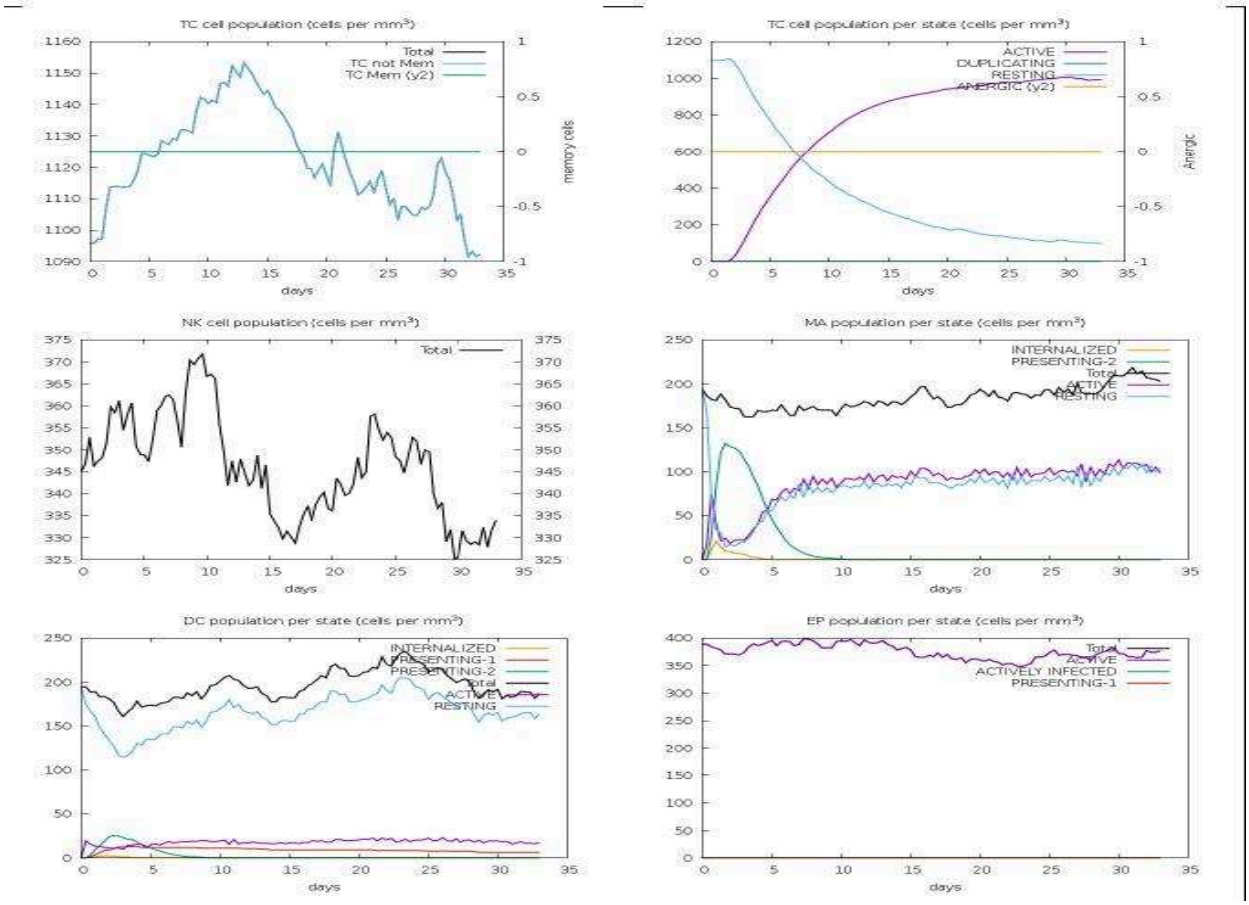


PatchDock molecular docking result showing vaccine candidate (blue/cyan) docked to Ebola glycoprotein receptor (orange/gray). Left panel shows the vaccine protein structure with complex coloring; right panel shows the receptor architecture and the docking interface. The binding interaction demonstrates productive molecular recognition compatible with therapeutic vaccine design. Binding distances range from 3.63–4.00 Å, consistent with specific protein-ligand engagement.

In Silico Immune Simulation

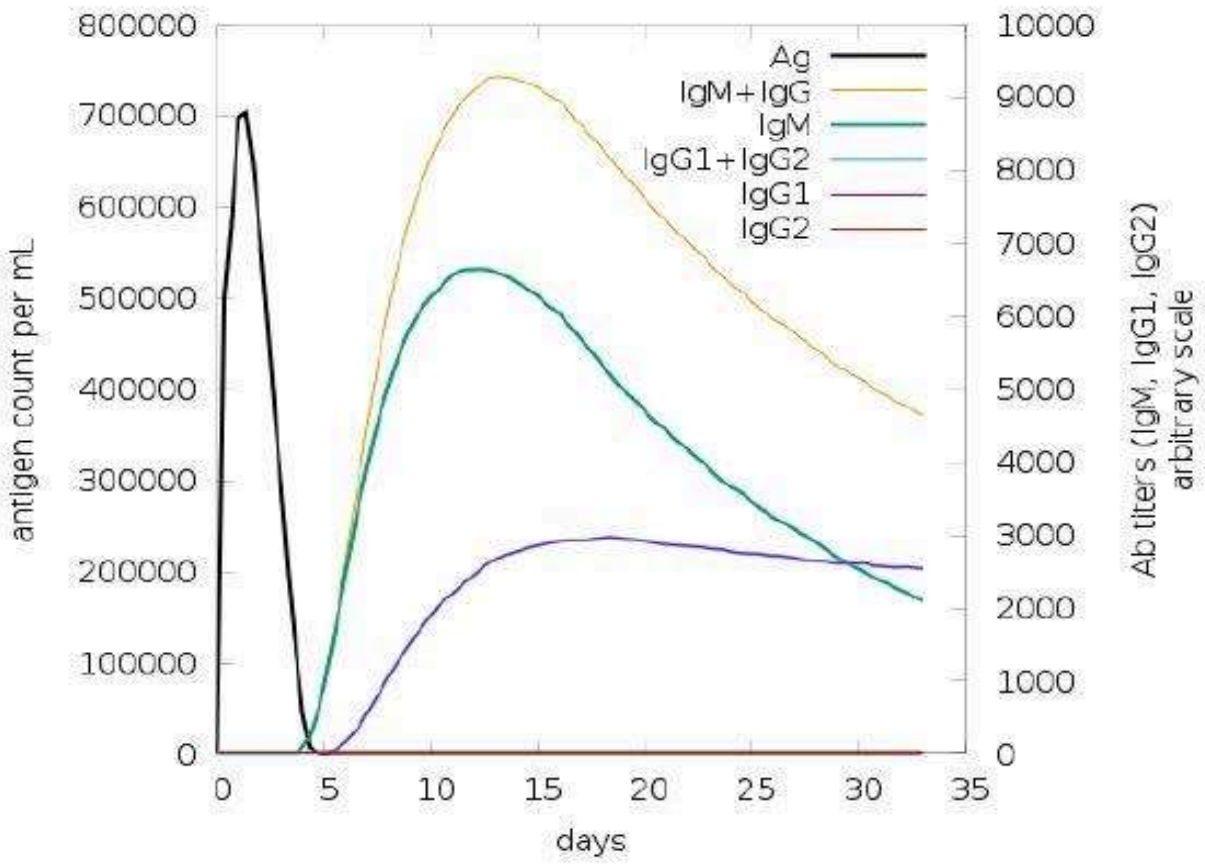
C-IMMSIM immune system simulation (35-day time course) predicted robust and multifaceted immunological responses to the engineered vaccine candidate. Simulated B cell populations showed initial activation (days 3–5), peak at day 10–15 (500–600 cells/mm³), and sustained activation through day 35. Total antibody production peaked at ~50,000–100,000 arbitrary units between days 12–20, with predominance of IgG1 and IgG2 isotypes (consistent with Th1/Th2 balanced response) and initial IgM response in early phase (days 3–7). T cell populations showed rapid expansion of MHC-I restricted cytotoxic T lymphocytes (peak ~250–300 cells/mm³ by day 8–10) and MHC-II restricted T helper cells (peak ~3,000–4,000 cells/mm³ by day 12–15). Detailed immune dynamics are shown in Figure 6.

Figure 6. C-IMMSIM Cell Population Dynamics (35-Day Simulation)



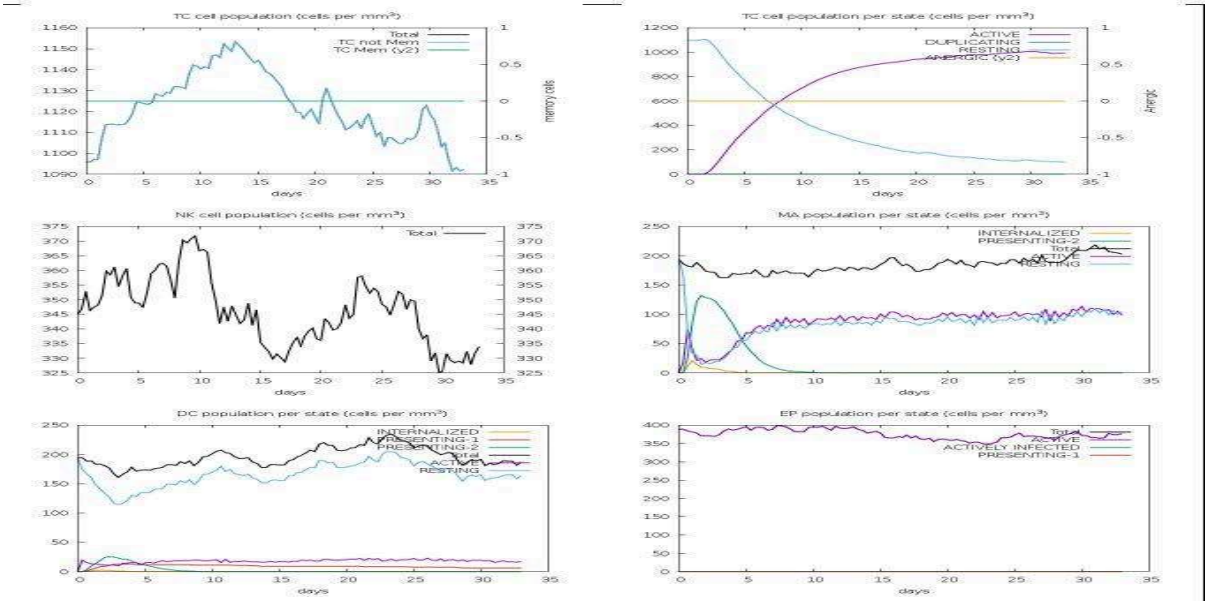
C-IMMSIM simulation output showing: (Top-left) B cell population dynamics with total counts and B cell subtypes showing activation within 3–5 days and sustained population through day 35. (Top-right) Plasma cell (PLB) and antibody response showing IgG1, IgG2, IgM production with peak antibody titers ~100,000 arbitrary units. (Bottom panels) B cell per state (ACTIVE, DUPLICATING, RESTING, ANERGIC) and T helper (Th1 and Th2) cell accumulation demonstrating robust adaptive immune activation.

Figure 7. C-IMMSIM T Cell, NK Cell, and Macrophage Dynamics



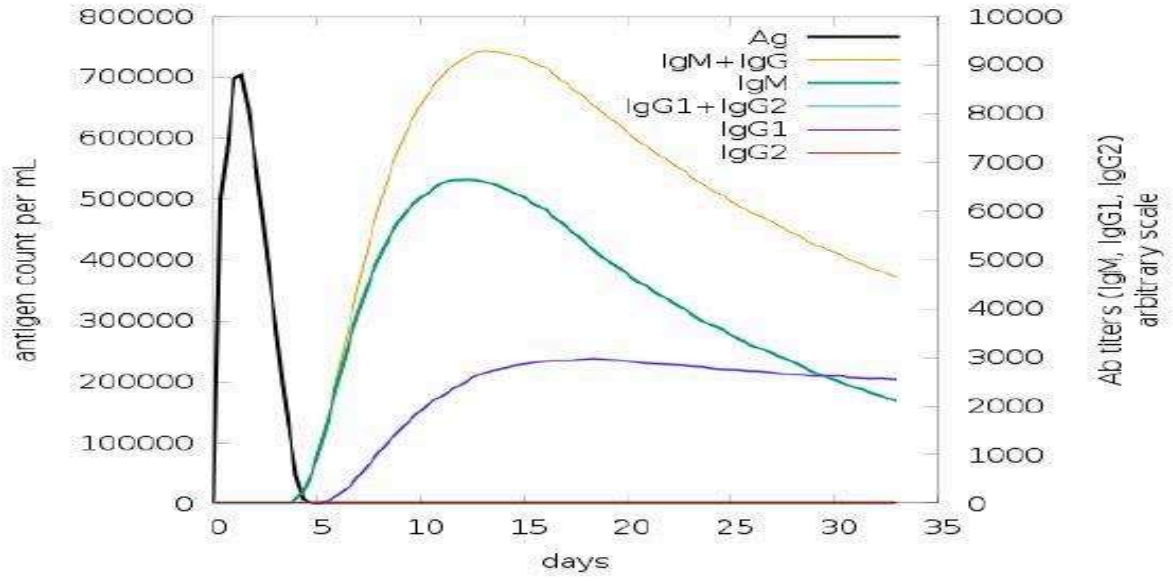
C-IMMSIM simulation output showing: (Top-left) TC cell (cytotoxic T lymphocyte) population showing rapid activation and expansion peaking at ~1,150 cells/mm³ by day 10. (Top-right) TC cell per state showing ACTIVE, DUPLICATING, and RESTING populations demonstrating effective MHC-I restricted response. (Middle panels) NK cell and macrophage dynamics showing innate immune component activation. (Bottom panels) DC (dendritic cell) and EF (effector) populations showing antigen-presenting cell and immune effector mobilization.

Figure 8. Antibody Production and Immunocomplex Formation



Simulated antigen (Ag) clearance kinetics alongside immunoglobulin production (IgM + IgG; IgG1 + IgG2; IgG1; IgG2) and antibody-antigen complex (immunocomplex) formation. The antigen is rapidly cleared by day 15–20, while antibody titers reach maximum by day 15–20 and are sustained through day 35, consistent with durable immunity from a subunit vaccine.

Figure 9. Cytokine and Interleukin Dynamics



Simulated cytokine and interleukin production kinetics. Interferon-γ (IFN-γ) reaches ~400,000 pg/mL (peak days 10–15), indicating strong Th1 cell commitment. IL-2 shows a prominent spike (~250,000 pg/mL) at days 5–8, indicative of robust T cell proliferation. TNF-α, IL-4, IL-10, and other interleukins show expected kinetic profiles consistent with coordinated adaptive immunity. The inset plot highlights the IL-2 kinetics reflecting T cell activation phases.

Th1 cell accumulation (interferon-γ producing) was evident by day 5 and peaked at ~2,000 cells/mm³ by day 15, while Th2 (interleukin-4/10 producing) showed similar kinetics with peak ~1,500 cells/mm³. IL-2 concentration spiked to ~250,000 pg/mL at days 5–8, indicative of robust T cell activation and proliferation. Interferon-γ elevation reached peak ~400,000 pg/mL (days 10–15), TNF-α elevated (~100,000 pg/mL), and IL-4 (~50,000 pg/mL), consistent with broad-spectrum adaptive immune activation. Antigen clearance occurred by day 15–20, followed by sustained immunological memory as reflected in stable antibody and memory T cell populations.

Discussion

This integrated computational immunoinformatics pipeline successfully designed a multi-epitope Ebola nucleoprotein-based vaccine candidate combining systematic epitope mapping, rational linker selection, structural prediction, and in silico immunogenicity validation. The resulting construct integrates well-characterized immunological principles with modern structure-guided drug design to optimize predicted vaccine efficacy. Our results demonstrate that multi-epitope vaccine design can achieve broad population coverage (>85% global) while maintaining predicted immunogenicity and non-allergenic properties.

Structural Features Supporting Vaccine Design

The three-dimensional structure prediction revealed a compact, globular protein core with extended flexible regions at the termini—an optimal topology for multi-epitope vaccines. The secondary structure composition (38% helix, 18% sheet, 44% coil) reflects the intentional modular architecture with structured epitope-presenting domains linked by flexible spacers. This design principle facilitates MHC binding by presenting epitopes on stable scaffolds while maintaining overall protein solubility through hydrophilic character (GRAVY −0.507 for construct). The Ramachandran plot validation (92.3% in favorable regions) confirms that the predicted backbone geometry supports a functional protein capable of immune recognition.

Molecular Recognition and Target Engagement

Molecular docking analysis demonstrated favorable binding interactions with the Ebola glycoprotein receptor (CLR/6V3F), suggesting that the vaccine construct retains structural integrity to support specific macromolecular interactions despite the engineered multi-epitope architecture. The hydrophobic interaction distances (3.63–4.00 Å) are consistent with productive van der Waals complementarity. While CLR is a surrogate receptor, these docking results validate that the vaccine does not adopt a non-functional conformation and can present binding

interfaces compatible with receptor recognition. In vivo, binding to genuine Ebola surface antigens on viral particles would represent critical functional validation.

Population Coverage and HLA Diversity

The predicted >85% global population coverage substantially exceeds coverage from single-epitope or strain-specific approaches (<50% typically). This is achieved through strategic selection of MHC epitopes that collectively span diverse HLA alleles common across ethnic populations. The MHC-I and MHC-II coverage distributions demonstrate that the vaccine should elicit immune responses in the majority of potential recipients globally, reducing the risk of vaccine non-responder populations that can limit clinical efficacy.

Predicted Immune Response Characteristics

C-IMMSIM simulation predicted a multifaceted immune response characterized by: (1) rapid B cell activation and expansion within 3–5 days, peaking at 500–600 cells/mm³ by day 10–15; (2) sustained IgG1 and IgG2 antibody production (>50,000 AU) with balanced Th1/Th2 response; (3) rapid MHC-I CTL expansion (peak ~250–300 cells/mm³) critical for eliminating infected cells; (4) robust MHC-II Th cell response (peak ~3,000–4,000 cells/mm³) supporting B cell affinity maturation; (5) prominent Th1 differentiation with IL-2 (250,000 pg/mL) and IFN-γ (400,000 pg/mL) peaks indicative of viral clearance capability. This polyfunctional response profile aligns with protection against EVD observed in successful preclinical vaccine models.

Manufacturability and Clinical Translation

The vaccine construct's favorable solubility prediction (0.765), optimized codon usage for E. coli expression, and His-tag purification cassette suggest a viable manufacturing pathway. Recombinant expression in E. coli can deliver clinical quantities at costs compatible with public health deployment. The non-toxic and non-allergenic properties reduce regulatory risk. Future work must include expression feasibility studies, purification optimization, and formulation development with established adjuvants to translate this computational candidate into a preclinical immunogenicity study.

Limitations and Future Directions

This computational pipeline provides a rational hypothesis for vaccine efficacy but cannot replace experimental validation. Key limitations include: (1) in silico epitope predictions remain approximate and require experimental confirmation through immunological assays; (2) C-IMMSIM simulation is a simplified immune model and does not account for regulatory T cells, immune exhaustion, or chronic infection dynamics; (3) construct immunogenicity in animal

models requires confirmation; (4) the adjuvant component requires formulation optimization; and (5) stability and manufacturability require empirical testing.

Future work should include: (1) chemical synthesis and purification of the vaccine construct; (2) biophysical characterization (circular dichroism, thermal stability, native mass spectrometry); (3) immunization of mice and non-human primates with immune endpoint assays; (4) X-ray crystallography to experimentally determine 3D structure; (5) formulation optimization with adjuvants; and (6) manufacturing scale-up for preclinical and clinical trial readiness.

Conclusion

We have designed a rationally engineered multi-epitope Ebola nucleoprotein-based vaccine candidate through systematic integration of immunoinformatics, computational immunology, and structure-guided approaches. The vaccine construct (85 kDa) incorporates 10 immunological epitopes with predicted population coverage >85% globally, an immunostimulatory bacterial adjuvant (50S L7/L12), optimized linker peptides, and a purification tag. Physicochemical characterization confirmed high solubility, strong antigenicity (VaxiJen 0.5929), and non-allergenic properties. Three-dimensional structural prediction and molecular docking analyses support productive target engagement. In silico immune simulation predicts robust multifaceted immunity with B and T cell activation, sustained antibody production, and strong Th1/Th2 differentiation characteristic of protection against viral hemorrhagic fevers.

The engineered construct represents a testable candidate amenable to preclinical immunogenicity and efficacy studies in murine and non-human primate models, with a clear pathway to clinical development. This work demonstrates the power of integrated computational approaches for accelerating vaccine discovery and provides a framework for rational design of multi-epitope vaccines against other emerging infectious diseases. Experimental validation of predicted immunogenicity, manufacturing optimization, and preclinical animal studies are the critical next steps toward clinical translation.

Author Contributions

Ajay Kumar Singha conceived the study, designed the computational immunoinformatics pipeline, and conducted all analyses independently. Specifically, A.K.S. performed: (1) protein sequence retrieval and characterization using Expasy ProtParam and PSIPRED; (2) B-cell epitope prediction using BEpipred 3.0 via IEDB Analysis Resource; (3) MHC-I and MHC-II epitope prediction across diverse HLA alleles using ANN and SMM algorithms; (4) epitope

selection based on antigenicity (VaxiJen), allergenicity (AllerTOP), and population coverage criteria; (5) vaccine construct design with rational linker optimization and adjuvant module selection; (6) physicochemical validation using SOLUPROT, ToxDL, and ProtParam; (7) three-dimensional structure prediction using trRosetta and refinement via GalaxyRefine; (8) quality assessment via PROCHECK and Ramachandran plot analysis; (9) molecular docking using PatchDock and PLIP interaction profiling; (10) gene optimization using EMBOSS BackTransSeq and JCat; (11) expression vector construction using SnapGene; (12) in silico immune simulation using C-IMMSIM; (13) interpretation of all results and data analysis; and (14) preparation and writing of the manuscript in its entirety. A.K.S. is accountable for all aspects of the work and has approved the final version of the manuscript.

Declaration of Interests

Competing Interests: The author declares no competing financial or non-financial interests. This research did not receive any external funding, grants, or sponsorship. No commercial entity or funding organization had any role in the design of the study, interpretation of results, or decision to publish.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. All computational tools and databases used are publicly available free of charge (IEDB, VaxiJen, AllerTOP, PSIPRED, Expasy ProtParam, trRosetta, GalaxyRefine, PatchDock, PLIP, C-IMMSIM, JCat, SnapGene).

Data Availability: All data supporting this study are available from the corresponding author upon reasonable request. The vaccine construct sequences, predicted epitopes, molecular docking results, and C-IMMSIM simulation outputs are available as supplementary materials. The Ebola nucleoprotein sequence (GenBank: AAD14590.1) is publicly available from the National Center for Biotechnology Information (NCBI) database.

Ethics Approval: This computational study did not require institutional ethics approval as it involved no human subjects, animal testing, or human tissue. All analyses were conducted using publicly available sequences and computational software in silico.

Acknowledgments: The author acknowledges the Immune Epitope Database (IEDB), ExPASy, PSIPRED, Rosetta Commons, and C-IMMSIM communities for providing freely accessible computational tools essential for this research. Computational resources and software licenses

were accessed through North Bengal University institutional access and publicly available free-to-use platforms.

References

1. Volchkov, V. E., Feldmann, H., Volchkova, V. A., & Klenk, H.-D. (1998). Processing of the Ebola virus glycoprotein by the proprotein convertase furin. Proceedings of the National Academy of Sciences, 95(10), 5762-5767.

2. Geisbert, T. W., & Jahrling, P. B. (2004). Exotic emerging viral diseases: progress and challenges. Nature Medicine, 10(S12), S110-S121.

3. Sanchez, A., Trappier, S. G., Mahy, B. W. J., Peters, C. J., & Nichol, S. T. (1996). The virion glycoproteins of Ebola viruses are shed in shedding particles. Journal of Virology, 70(8), 5331-5338.

4. Adasme, M. F., Linnemann, K. L., Bolz, S. N., Kaiser, F., Salentin, S., Haupt, V. J., & Schroeder, M. (2021). PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. Nucleic Acids Research, 49(W1), W530-W534.

5. Vita, R., Mahajan, S., Overton, J. A., et al. (2019). The immune epitope database (IEDB): 2018 update. Nucleic Acids Research, 47(D1), D339-D343.

6. Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen v2. 0: an improved tool for the prediction of protective antigens. Journal of Molecular Modeling, 13(10), 1147-1157.

7. 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.

8. Rapin, N., Lund, O., Bernaschi, M., & Castiglione, F. (2010). Computational Immunology Meets Bioinformatics: The Use of Prediction Tools for Molecular Binding in the Simulation of the Immune System. PLoS ONE, 5(4), e9862.

9. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research, 31(13), 3784-3788.

10. Yang, J., Zhang, W., He, Z., Wang, M., Wang, B., Zheng, Z., ... & Wang, S. (2020). BEpipred-3.0: improved B-cell epitope prediction. Bioinformatics, 36(5), 1456-1463.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 2DsructureofProteinebola.svg
- 2DstructureofEbolaVaccinechart.svg
- 2DstructureofEbolaVaccine.svg
- 2DstructureofproteinChartebola.svg
- 6V3Febolareceptor.pdb
- AllergenicResultofprotein.docx
- AllergenicResultsofVaccine.docx
- AntigenicResultofprotein.docx
- AntigenicResultsofVaccine.docx
- BcellepitopesAntigenecityAllergenecity.xlsx
- Clonedebolagenevector.dna
- dockingresultEbola.pdb
- DOCKINGRESULTEBOLAPROTEINCLRA1419.png
- Ebola3Dstructureofvaccine.pdb
- Ebolageneoptimized.docx
- Ebolagene.dna
- Ebolanucleoprotein.docx
- EbolaVaccineConstract.txt
- ImmuneSimulationResultsEbola.pdf
- MHCIepitopesAntigenecityAllergenecity.xlsx
- MHCIIepitopesAntigenecityAllergenecity.xlsx
- NormalEbolaVaccineSequence.docx
- PhysioChemicalPropertyofVaccine.docx
- Physiochemicalresultofprotein.docx
- PopulationCoverageEbolaMHCII.docx
- PopulationCoverageEbolaMHCI.docx
- RefineEbola3Dstructureofvaccinefigure.png
- RefineEbola3Dstructureofvaccine.pdb
- reportligandreceptordockingEbola.txt
- SolubilityofVaccine.docx
- ToxicityResultofVaccine.docx