DEVELOPMENT OF EPITOPE-BASED VACCINE TO PREVENT MARBURG VIRUS INFECTION: AN IN SILICO APPROACH

Shruti Sanjay Thorle
Department of Biotechnology and Bioinformatics[1]

Mahadev Asaram Jadhav✉
Department of Biotechnology and Bioinformatics[1]
majadhav22g@gmail.com

Dipak Pandit Chavan
Department of Biotechnology and Bioinformatics[1]

Shivani Sunil Dhadge
Department of Biotechnology and Bioinformatics[1]

[1]Deogiri College
Railway Station Road, Maharashtra, India, 431005

✉**Corresponding author**

Summary

Marburg virus (MARV) is one of the deadliest zoonotic viruses, causing severe hemorrhagic fever in humans with high mortality rates. The development of an effective vaccine is crucial to prevent potential Marburg virus outbreaks. In this study, an in silico approach was employed to design an epitope-based vaccine to prevent MARV infections. The MARV proteins nominating NP, VP24, VP35, VP30, VP40, GP & Polymerase L was analyzed for antigenicity and non-allergenicity prediction, among these proteins VP30 protein has a 0.5636 (Probable Antigen) score and it was non-allergen. For that reason, VP30 was selected for further in silico analysis. After analysis it is found that the top ranked T–cell (MHC-I) epitopes LSKPPPPPK, ESSPTNHIPR, TQLPSKPHY, SPQDCGSPSL, FEAALWQGW, T-Cell (MHC-II) epitopes IHLDKGGQF, INTMTELHM, VTPTIYHET, YTNYHPRAR, YTGIHLDKG was epitopes & B-Cell epitopes SEIGKLDET, IHLDKGGQF, MNHENLPQDQNGV, PTCNRDHDLDNLTN was found non-toxic and non-allergen. The T-Cell (MHC-I) epitope TQLPSKPHY,T-Cell (MHC-II)epitope YTNYHPRAR & B-Cell epitope SEIGKLDET was found highly antigenic, non-toxic as well as non-allergen and it was selected for molecular docking analysis. The T-Cell (MHC-I) epitope TQLPSKPHY,T-Cell (MHC-II) epitope YTNYHPRAR shows strong structural similarity and potential binding affinity with antibody. The B-Cell epitope SEIGKLDET shows poor affinity towards antibody. In silico analysis indicate that both T-Cell epitopes becomes an effective peptide vaccine to prevent MARV infection. Our findings highlight the promise of in silico vaccine design in accelerating the development of vaccines against MARV, a highly pathogenic virus with no effective cure currently available.

Keywords: epitope, MARV, in silico, allergenicity, simulation, stability, antigenicity, non-allergenicity, pathogenicity, vaccine.

DOI: 10.21303/2504-5695.2024.003411

1. Introduction

Preface Marburg contagion (MARV) complaint is a largely malign illness that induces hemorrhagic fever with a mortality rate of over 88 %. It belongs to the Filoviridae family, which is responsible for causing Ebola contagion complaints. The complaint was originally linked to two significant outbreaks that occurred coincidently in Marburg and Frankfurt, Germany, and Belgrade, Serbia, in 1967. The transmission of Marburg contagion complaint to humans primarily occurs following extended exposure to mines or grottoes harboring Rousettus club colonies. Once the contagion infects an existent, Marburg can spread through mortal-to-mortal transmission via direct contact through compromised skin or mucous membranes) with the blood, concealment, organs, or other fleshly fluids of infected individualities as well as with shells and particulars (similar to coverlets and apparel) defiled by these fluids [1].

These outbreaks were linked to laboratory conditioning involving African green monkeys (Cercopithecus aethiops) that were imported from Uganda. Also, cases of outbreaks and sporadic cases have

been proved in Angola, the Democratic Republic of the Congo, Kenya, South Africa (involving an individual with a recent trip history to Zimbabwe), and Uganda [2]. In 2008, two insulated cases of Rousettus club colonies in Uganda were proved. lately, the World Health Organization (WHO) declared MVD as an outbreak in February 2023, when Equatorial Guinea reported nine verified and 20 probable cases. Seven people out of all the verified cases failed, and all suspected cases passed away in Equatorial Guinea [3].

MARV consists of 19-KBnon-infectious single-stranded RNA genome garbling seven abecedarian structural proteins nucleoprotein (NP), viral protein 24 (VP24), VP35, VP30, VP40, Glycoprotein (GP), and Polymerase L protein (L). Presently, there are no available treatments or vaccines for MARV complaints, and several immunizations are presently witnessing clinical trials. The integration of immunogenetics and immunogenomics with bioinformatics in a new approach known as vaccinomics has been employed to enhance vaccination strategies. This methodology has been applied to the disquisition and invention of new vaccine phrasings. The application of the "Vaccinomics" strategy has proven to be essential in combatting colorful infections, similar to multiple sclerosis, malaria, and malice. Nonetheless, the methodologies employed in vaccine development end to identify mortal leukocyte antigen (HLA) ligands and T cell epitopes that are pivotal in opting for potent vaccine campaigners associated with antigen donation transporter (valve) motes [4].

Traditional vaccine development approaches are time-consuming and expensive. In silico vaccine design offers a promising volition by using bioinformatics tools to identify implicit vaccine campaigners efficiently. The World Health Organization has put MARV on a list of critical requirements to find a result for this deadly contagion, where severe outbreaks and high casualty rates have increased the significance of this call. In this study, a bioinformatics approach was applied to induce an implicit vaccine construct against MARV [5]. The viral whole proteome was first anatomized for antigenicity and stability of each protein, and also seeker proteins were uprooted for B and T cell epitope vaticination. Also, the epitopes of each seeker were named to construct a fantastic epitope implicit vaccine that was assessed computationally for its structural, immunological, and chemical characteristics to be nominated as an apparent result against MARV [4].

2. Materials and methods

The Marburg virus has a high fatality rate, so a preventive measure should be taken to prevent infection. As there is no effective therapeutic agent available against these viruses, an effective vaccine design touching all strains would be a great step for human health. The proper flowchart of the applied method is shown in **Fig. 1.**

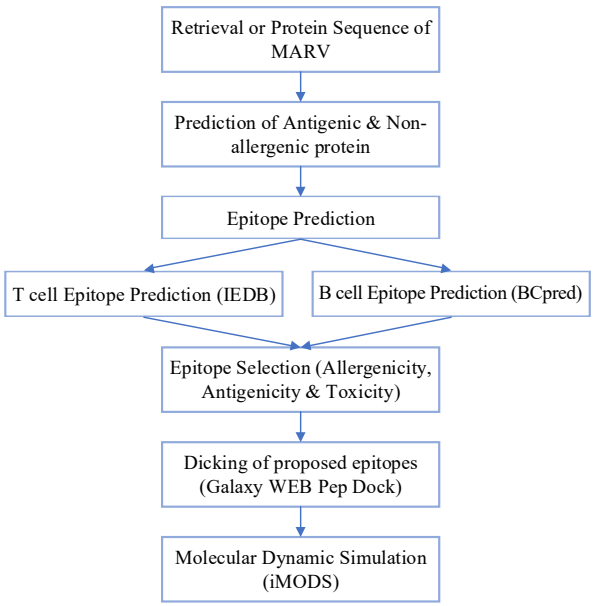


Fig. 1. Flowchart of the applied methods for in silico vaccine designing

2. 1. Retrieval of MARV protein sequences

The UniProtKB database was utilized to collect MARV and its related data like genes, proteins, organisms, amino acids, annotation scores and protein existence are collected. The primary sequences of MARV proteins: Nucleoprotein (NP), viral protein 24 (VP24), VP35, VP30, VP40, Glycoprotein (GP), and Polymerase L protein (L) were retrieved [6].

2. 2. Determination of most antigenic and non-allergenic protein

VaxiJen v2.0, an alignment-free server for identifying protective antigens and subunit vaccines, was used to assess the antigenicity of MARV proteins based on their physicochemical properties. The tool ranked the proteins according to their antigenic scores, with higher scores indicating greater antigenicity [7]. ProtParam was employed to analyze the physicochemical properties of MARV proteins, including stability-related parameters such as instability index and aliphatic index. Proteins with an instability index below 40 were considered stable, while higher values suggested instability. The aliphatic index provided insights into the stability and thermostability of the proteins based on the relative volume occupied by aliphatic side chains [8].

2. 3. Epitope prediction

Specific software tools were used to predict T-cell and B-cell epitopes within the selected MARV protein sequences. T-cell epitopes were predicted using the Immune Epitope Database (IEDB), which employs computational algorithms and experimental data to evaluate epitopes based on MHC binding affinity and immunogenicity. B-cell epitopes were predicted using BCpred, a machine learning-based approach that considers surface accessibility, flexibility, and hydrophilicity to identify potential epitopes recognized by B cells. The predicted epitopes were evaluated for their antigenicity and potential to elicit immune responses [9].

2. 4. Epitope ielection

Epitopes were selected based on a comprehensive evaluation of allergenicity, toxicity, and antigenicity using bioinformatics tools like Vaxijen, AllerTop, and ToxinPred. Allergenicity was assessed for potential cross-reactivity, toxicity for known toxic motifs, and antigenicity for immune response elicitation against MARV. Predicted epitopes underwent further analysis for these properties [9].

2. 5. Docking of proposed epitopes

Docking simulations of proposed epitopes with MARV proteins were conducted using Galaxy Pepp-Dock, a tool for protein-peptide docking. This process predicted binding modes and affinities, revealing potential interaction sites and energies. The study utilized 6v4q () from the Protein Data Bank for docking, exploring interactions with various HLAs to select high-affinity epitopes for vaccine development [10].

2. 6. Molecular dynamic simulation

Molecular dynamic simulations of epitope-protein interactions were conducted using iMods, a web-based platform for protein dynamics analysis. The simulations assessed the stability of epitope-protein complexes over time, providing insights into conformational changes and dynamic interactions. This helped evaluate the suitability of epitopes for vaccine development against MARV [11].

3. Results

3. 1. Retrieval of protein sequences of MARV and determination of most antigenic and non-allergenic protein

The protein sequences of MARV (NP, VP24, VP35, VP30, VP40, GP, and L) were retrieved from UniProtKB. VaxiJen and AllerTop servers assessed these sequences to identify the most antigenic and non-allergenic protein. Antigenicity ranged from 0.4107 to 0.5636, confirming their antigenic nature. VP30 (UniProtKB id: P35258 VP30_MABVM) emerged as the most potent antigenic protein with a score of 0.5636 (**Table 1**). Further analysis focused on this protein for its effectiveness, utilizing it for subsequent investigations.

Table 1

List of Antigenic Protein & their antigenic nature

UniProt Id	MARV Protein Names	Antigenicity	Allergenicity
P27588	Nucleoprotein (NP)	0.4761 (Probable non-antigen)	Probable non-allergen
P35260	Matrix protein (VP40)	0.4107 (Probable antigen)	Probable non-allergen
P35256	Membrane-associated protein (VP24)	0.5423 (Probable antigen)	Probable non-allergen
P35259	Polymerase cofactor (VP35)	0.4360 (Probable antigen)	Probable non-allergen
P35258	Transcriptional activator (VP30)	0.5636 (Probable antigen)	Probable non-allergen
P31352	Polymerase L protein (L)	0.4518 (Probable antigen).	Probable non-allergen
P35253	Glycoprotein (GP)	0.5481 (Probable antigen).	Probable non-allergen

3. 2. Epitope prediction and selection

Epitope prediction began by submitting the filtered protein candidate VP30 sequence to the Immune Epitope Database (IEDB). NetMHCpan EL 4.0 predicted MHC-I binding, utilizing the recommended method against a reference set of HLA alleles. For MHC-II binding, IEDB suggested 2.22 prediction methods, running the binding prediction against the full HLA reference set. AllerTop, Vaxijen, and ToxinPred webservers were used to predict the allergenicity, antigenicity, and toxicity of the selected epitopes, respectively (**Table 2**).

Table 2

Top-ranked T-cell epitopes (MHC-I peptides) of VP30

Epitopes	Antigenicity score	Toxicity	Allergenicity
LSKPPPPPK	0.5344	Non-Toxic	Non-allergen
ESSPTNHIPR	–0.5090	Non- Toxic	Non-allergen
TQLPSKPHY	0.6737	Non- Toxic	Non-allergen
SPQDCGSPSL	0.3780	Non- Toxic	Non-allergen
FEAALWQGW	–0.3350	Non- Toxic	Non-allergen

The epitope “TQLPSKPHY” from the VP30 protein of the Marburg virus stood out as the top T-cell epitope, boasting the highest antigenicity score. This epitope was chosen for detailed computational modeling and docking studies due to its robust antigenicity score of 0.6737, indicating strong immunogenic potential. Assessments confirmed its non-toxic and non-allergenic nature, ensuring safety for further exploration. By focusing on “TQLPSKPHY” for in-depth structural and binding analyses, let’s aim to uncover its interactions with immune receptors, enhancing our understanding of its immunogenic properties and potential as a Marburg virus vaccine candidate.

T-cell epitopes from the VP30 protein of the Marburg virus, focusing on MHC-II binding peptides. Each epitope was evaluated for antigenicity, toxicity, and allergenicity. “YTNYHPRAR” emerged as the top epitope with an antigenicity score of 1.3410, indicating strong immunogenic potential against MARV. It exhibited non-toxic and non-allergenic traits, making it a safe candidate for vaccine development. Due to its high antigenicity and safety profile, “YTNYHPRAR” was chosen for detailed structural modeling and docking studies. These investigations will unveil its three-dimensional structure and interactions with target proteins, aiding in the design of effective MARV vaccines (**Table 3**).

Table 3

Top-ranked T-cell epitopes (MHC-II peptides) of VP 30

Epitopes	Antigenicity score	Toxicity	Allergenicity
IHLDKGGQF	0.5053	Non-Toxic	Non-allergen
INTMTELHM	0.4519	Non- Toxic	Non-allergen
VTPTIYHET	0.7431	Non- Toxic	Non-allergen
YTNYHPRAR	1.3410	Non- Toxic	Non-allergen
YTGIHLDKG	1.1426	Non- Toxic	Non-allergen

B cell epitopes from VP30 of the Marburg virus were predicted using the IEDB web server (**Fig. 2**).

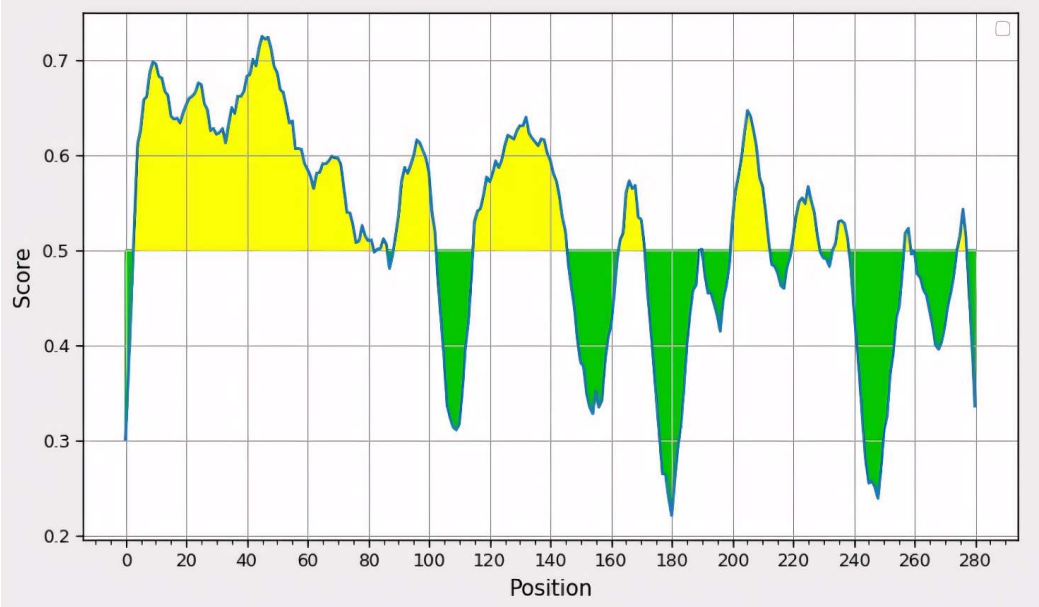


Fig. 2. Bepipred linear epitope prediction of the most antigenic protein VP30

Epitopes larger than 9 peptides were selected, reducing the list to five epitopes in **(Table 4)** with a threshold value of 0.35. Notably, all filtered B cell epitopes were forecasted to be non-allergenic and non-toxic. This selection process ensured a refined set of epitopes for in-depth analysis, emphasizing their safety and immunogenic potential for potential applications in vaccine development against MARV.

Among the selected epitopes, "SEIGKLDET" exhibited the highest antigenicity score of 0.9343, indicating its strong potential to elicit an immune response against MARV. This epitope, along with "IHLDKGGQF" and "PTCNRDHDLDNLTN" demonstrated non-toxic and non-allergenic attributes, ensuring their safety profile for vaccine development. While "MNHENLPQDQNGV" displayed a lower antigenicity score, its inclusion in the analysis provides additional insights into potential B-cell epitopes derived from VP30. Despite its relatively lower score, this epitope may still contribute to the overall immunogenicity of a vaccine candidate.

Table 4

Predicted B cell epitopes from VP30

Epitopes	Antigenicity score	Toxicity	Allergenicity
SEIGKLDET	0.9343	Non-toxic	Non-allergen
IHLDKGGQF	0.5053	Non-toxic	Non-allergen
MNHENLPQDQNGV	0.1026	Non-toxic	Non-allergen
PTCNRDHDLDNLTN	0.4668	Non-toxic	Non-allergen

3. 3. Docking of proposed epitopes

The docking phase has commenced, encompassing MHC I, MHC II, and B cell epitopes. Higher antigenicity scores indicate greater immunogenic potential, reflecting the likelihood of eliciting an immune response. For MHC I, the epitope TQLPSKPHY was chosen due to its optimal docking properties, boasting a non-toxic, non-allergenic nature, and an antigenicity score of 0.6737. Similarly, the epitope YTNYHPRAR was selected for MHC II docking with a high antigenicity score of 1.3410. SEIGKLDET from B cell epitopes, with an antigenicity score of 0.9343, non-toxicity, and non-allergenic attributes, was also chosen for docking studies:

a) the MHC-1 epitope (TQLPSKPHY) was docked to protein 6v4q via Galaxy WEB Pep Dock for binding interaction assessment. This web-based server conducts protein-peptide docking simulations using a computational method to predict complex structure and affinity. The model (2A6I_B)

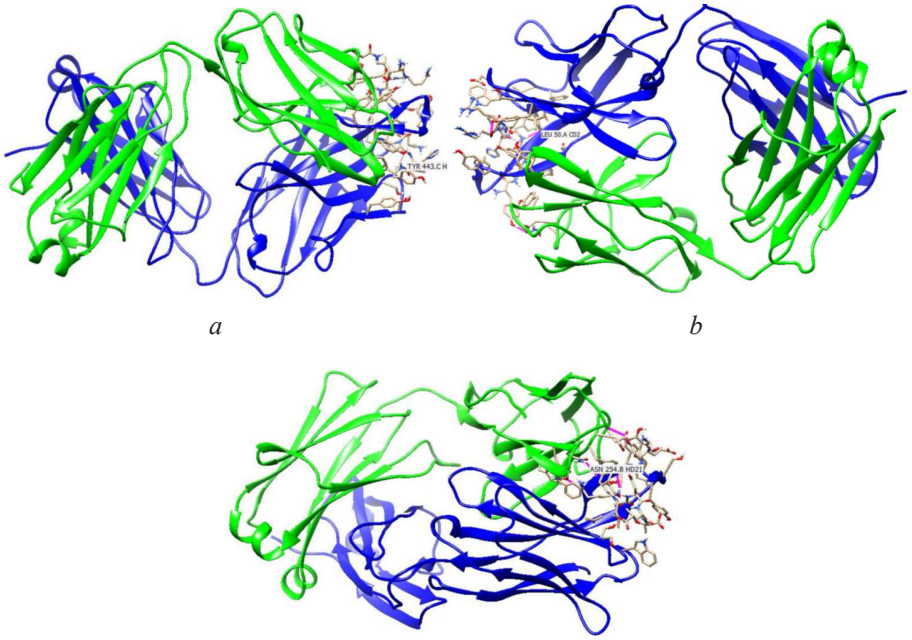
with the highest interaction similarity score, TM-score of 0.916, and 62.0 interaction similarity score was chosen for further analysis, indicating strong structural similarity and potential binding affinity;

b) the MHC-2 epitope (YTNYHPRAR) was docked to protein 6v4q via Galaxy WEBPep Dock to evaluate binding interactions. The model (12A6I_B) with the highest interaction similarity score, TM-score of 0.916, and 77.0 interaction similarity score was chosen for further analysis. This indicates a strong structural similarity and potential binding affinity according to Galaxy WEB PepDock's scoring criteria;

c) the B-cell epitope (SEIGKLDET) was docked to protein 6v4q with Galaxy WEB Pep Dock for interaction assessment. The model (1E4X_L) with the highest interaction similarity score was chosen for further analysis, showing a TM-score of 0.960 and an interaction similarity score of 35.0, indicating good structural similarity but a lower predicted binding affinity based on Galaxy WEB Pep Dock's criteria.

3. 4. Visualization and analysis of docked models

The top-ranked models (2A6I_B for MHC-1 epitope TQLPSKPHY, 2A6I_B for MHC-2 epitope YTNYHPRAR, and PDB ID: 1E4X_L for B-cell epitope SEIGKLDET) from Galaxy WEBPep Dock were visualized in Chimera for detailed structural analysis, exploring protein-peptide interactions. Specific interactions like hydrogen bonds and hydrophobic contacts were examined, providing insights into the binding affinity and orientation within the respective protein complexes (**Fig. 3**).



a *b*

c

Fig. 3. Ribbon structures complex of: *a* – antibody (PDB ID: 2A6I_B) and MHC-I epitope TQLPSKPHY interaction; *b* – antibody (PDB ID: 2A6I_B) and MHC-II epitope YTNYHPRAR interaction; *c* – antibody(PDB ID: 1E4X_L) and B-cell epitope SEIGKLDET interaction

3. 5. Molecular dynamic simulation

To evaluate the stability and dynamics of the multi-epitope vaccine construct, molecular dynamics (MD) simulations were conducted using the iMODS server, a web-based tool for biomolecular simulations. This platform enabled the examination of the construct's behavior in a simulated physiological setting, offering insights into its structural stability and flexibility. These analyses are crucial for assessing the vaccine's efficacy against Marburg virus infections. MD simulations track atom and molecule movements over time, with iMODS analyzing these dynamics to reveal

molecular flexibility and motion, providing valuable information for optimizing the vaccine design and understanding its potential therapeutic impact.

The ribbon representation of a molecular complex enables visual inspection of its overall structure and organization, revealing the spatial arrangement of domains, subunits, or functional elements. The notation "(CG-CA ENM-edNMA Deform-ON)" refers to the CG-CA elastic network model connecting alpha-carbon and beta-carbon atoms; ENM-edNMA normal mode analysis for studying molecular vibrations; and Deform-ON, an iMODS setting allowing protein complex deformation during simulation.

3. 6. Graphical outcomes of molecular dynamic simulation

The graphical outcomes of molecular dynamic simulations from Galaxy WEB PepDock provide visual representations of the structural dynamics and interactions within the simulated molecular systems. These outcomes include detailed visualizations of conformational changes, such as those revealed by principal component analysis (PCA), which identify statistically significant conformations and essential motions for conformational changes. Additionally, graphical outcomes may illustrate trajectories of principal components, highlighting key motions and structural variations over time. These visualizations offer insights into the behavior and dynamics of biomolecules, aiding in the understanding of complex molecular interactions and conformational changes during the simulation process (**Fig. 4**).

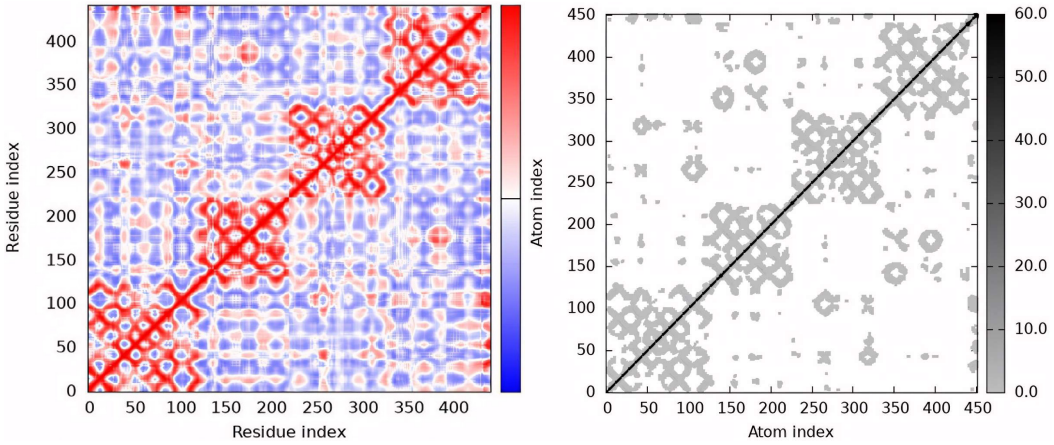


Fig. 4. Molecular dynamic simulation:
a – covariance map; *b* – elastic network of epitope (antigen) and antibody interaction

4. Discussion

To combat the growing number of diseases, rapid vaccine development is crucial. Advancements in sequence-based technologies have provided vast genomic and proteomic data for various viruses, enabling peptide-based vaccine design using bioinformatics tools. While epitope-based vaccines are gaining acceptance, no research has focused on developing Marburg virus vaccines using this approach. This work attempts to create an in-silico epitope-based vaccine against the Marburg virus. Most vaccines target B cell immunity, but T cell epitope-based immunization has accelerated due to the host's potent T cell response against infected cells. T cell-mediated immunity often provides long-lasting protection, as antibody memory can wane over time due to antigenic drift. The proposed epitope against the Marburg virus meets the requirements for an effective epitope-based vaccine candidate.

Allergenicity is a significant obstacle in vaccine development. Screening for adverse allergic reactions involves assessing IgE and Th2 induction. AllerTop scores represent the likelihood of a sequence being a cross-reactive allergen, with a prediction limit of approximately 0.08. The proposed epitope's allergenicity score of 0.06 indicates non-allergenicity. Evaluating immune databases and sequences is necessary for these in silico predictions. Given the experimental validation of such studies, the proposed epitope is expected to elicit a robust immune response as a peptide immunization in vivo.

5. Conclusions

This research employed in silico methods to design a T-cell and B-cell epitope-based peptide vaccine against the Marburg virus. Computational approaches proved valuable for predicting vaccine candidates against pathogens like the Marburg virus, saving time and costs compared to traditional methods. In silico studies guide experimental work, increasing the chances of finding desired solutions with fewer clinical trials. Future research can utilize this work and in-vitro procedures to assess the vaccine's adequacy and efficacy.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Financing

The study was performed without financial support.

Data availability

Data will be made available on reasonable request.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

References

[1] Marburg virus disease. World Health Organization. Available at: https://www.who.int/news-room/fact-sheets/detail/marburg-virus-disease

[2] Brauburger, K., Hume, A. J., Mühlberger, E., Olejnik, J. (2012). Forty-Five Years of Marburg Virus Research. Viruses, 4 (10), 1878–1927. https://doi.org/10.3390/v4101878

[3] Marburg virus disease - Equatorial Guinea. World Health Organization. Available at: https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON444

[4] Pervin, T., Oany, A. R. (2021). Vaccinomics approach for scheming potential epitope-based peptide vaccine by targeting l-protein of Marburg virus. In Silico Pharmacology, 9 (1). https://doi.org/10.1007/s40203-021-00080-3

[5] Flower, D. R. (2008). Bioinformatics for Vaccinology. John Wiley & Sons. https://doi.org/10.1002/9780470699836

[6] Apweiler, R. (2004). UniProt: the Universal Protein knowledgebase. Nucleic Acids Research, 32 (90001), 115D–119. https://doi.org/10.1093/nar/gkh131

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

[8] Gasteiger, E. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research, 31 (13), 3784–3788. https://doi.org/10.1093/nar/gkg563

[9] Debroy, B., Chowdhury, S., Pal, K. (2023). Designing a novel and combinatorial multi-antigenic epitope-based vaccine "MarVax" against Marburg virus – a reverse vaccinology and immunoinformatics approach. Journal of Genetic Engineering and Biotechnology, 21 (1), 143. https://doi.org/10.1186/s43141-023-00575-w

[10] Zhang, Y., Zhang, M., Wu, H., Wang, X., Zheng, H., Feng, J. et al. (2024). A novel MARV glycoprotein-specific antibody with potentials of broad-spectrum neutralization to filovirus. ELife, 12. https://doi.org/10.7554/elife.91181

[11] Lee, H., Heo, L., Lee, M. S., Seok, C. (2015). GalaxyPepDock: a protein–peptide docking tool based on interaction similarity and energy optimization. Nucleic Acids Research, 43 (W1), W431–W435. https://doi.org/10.1093/nar/gkv495

Received date 17.04.2024
Accepted date 19.06.2024
Published date 28.06.2024

How to cite: *Thorle, S. S., Jadhav, M. A., Chavan, D. P., Dhadge, S. S. (2024). Development of epitope-based vaccine to prevent Marburg virus infection: an in silico approach. EUREKA: Life Sciences, 2, 3–10. https://doi.org/10.21303/2504-5695.2024.003411*