Heliyon 10 (2024) e28223

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Computational approach for identifying immunogenic epitopes and optimizing peptide vaccine through *in-silico* cloning against *Mycoplasma genitalium*

Asma Akter [1], Natasha Farhin Ananna [1], Hedayet Ullah, Sirajul Islam, Md. Al Amin, K. M. Kaderi Kibria, Shahin Mahmud [*]

Department of Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh

ARTICLE INFO

Keywords:
Mycoplasma genitalium
Peptide vaccination
In-silico vaccine design
Outer membrane protein P110
Epitope prediction of *Mycoplasma genitalium*
In-silico cloning

ABSTRACT

Mycoplasma genitalium is a pathogenic microorganism linked to a variety of severe health conditions including ovarian cancer, prostate cancer, HIV transmission, and sexually transmitted diseases. A more effective approach to address the challenges posed by this pathogen, given its high antibiotic resistance rates, could be the development of a peptide vaccine. In this study, we used experimentally validated 13 membrane proteins and their immunogenicity to identify suitable vaccine candidates. Thus, based on immunogenic properties and high conservation among other *Mycoplasma genitalium* sub-strains, the P110 surface protein is considered for further investigation. Later on, we identified T-cell epitopes and B-cell epitopes from the P110 protein to construct a multiepitope-based vaccine. As a result, the 'NIAPISFSFTPFTAA' T-cell epitope and 'KVKYESSGSNNISFDS' B-cell epitope have shown 99.53% and 87.50% population coverage along with 100% conservancy among the subspecies, and both epitopes were found to be non-allergenic. Furthermore, focusing on molecular docking analysis showed the lowest binding energy for MHC-I (−137.5 kcal/mol) and MHC-II (−183.3 kcal/mol), leading to a satisfactory binding strength between the T-cell epitopes and the MHC molecules. However, the constructed multiepitope vaccine (MEV) consisting of 54 amino acids demonstrates favorable characteristics for a vaccine candidate, including a theoretical pI of 4.25 with a scaled solubility of 0.812 and high antigenicity probabilities. Additionally, structural analyses reveal that the MEV displays substantial alpha helices and extended strands, vital for its immunogenicity. Molecular docking with the human Toll-like receptors TLR1/2 heterodimer shows strong binding affinity, reinforcing its potential to elicit an immune response. Our immune simulation analysis demonstrates immune memory development and robust immunity, while codon adaptation suggests optimal expression in *E. coli* using the pET-28a(+) vector. These findings collectively highlight the MEV's potential as a valuable vaccine candidate against *M. genitalium*.

1. Introduction

Mycoplasma genitalium (M. genitalium) is a bacterial pathogen of significant concern, particularly in the sphere of sexually

* Corresponding author.
E-mail address: shahin018mbstu@gmail.com (S. Mahmud).
[1] Equal contribution.

https://doi.org/10.1016/j.heliyon.2024.e28223
Received 11 July 2023; Received in revised form 13 March 2024; Accepted 13 March 2024
Available online 28 March 2024

transmitted infections (STIs) [1]. It is notable for being the smallest self-replicating prokaryote, with a genome size of 580 kb [2]. As a member of the Mollicutes class, this bacterium is characterized by its minuscule size and lack of a cell wall, rendering it resistant to many common antibiotics [3,4]. With its ability to cause chronic infections, *M. genitalium* has attracted attention as a challenging, understudied pathogen with substantial public health implications. This microorganism has a flask-like shape and possesses a terminal organelle that is slightly curved [5]. It causes inflammation in the urogenital tract by attaching to host epithelial cells and activating innate immune sensors, resulting in the activation of pro-inflammatory signals and the recruitment of leukocytes to the site of infection, causing tissue damage [6]. It may also be transmitted through sexual contact, leading to the spread of infection between individuals. Adhesion, gliding motility, and cell invasion are just a few of the mechanisms that influence pathogenesis [7]. The flask-shaped form of *M. genitalium* is attributed to the extension of the cellular membrane, which is commonly referred to as the attachment organelle or terminal organelle. The attachment organelle has a complicated structure that is linked *to M. genitalium's* capacity to adhere [8]. Several virulence factors contribute to the pathogenesis of *M. genitalium*. In pathogenesis, adhesion is the first and most crucial stage. P110 and P140 (surface adhesins) are the main adhesion proteins that are responsible for the adhesion process [9]. A complex cytoskeleton with three essential substructures—the terminal button, rod, and wheel complex (or bowl)—maintains *M. genitalium's* adhesion organelle. Gliding motility is an additional consideration. Gliding motility is required for the entry of *M. genitalium* into the mucus layer enveloping the epithelial cells, adhesion to the epithelial cells, and invasion. After attaching to the epithelial cells, the tip structure organelle mediates cell entrance [10].

M. genitalium infection is one of the primary causes of non-gonococcal urethritis in males and a common cause of urethritis and cervicitis in women [11]. *M. genitalium* has been reported to be the cause of 15%–20% of non-gonococcal urethritis (NGU), 20%–25% of non-chlamydial NGU, and 30% of recurrent or persistent urethritis [12]. According to the 2017–2018 National Health and Nutrition Examination Survey, the overall prevalence of urogenital *M. genitalium* among adults aged 14–59 years in the United States is 1.7%. According to the survey, males had a 1.8% frequency, while females had a 1.7% prevalence [13]. The prevalence of *M. genitalium* strains resistant to multiple classes of antibiotics has increased dramatically over the past decade [14]. It is intrinsically resistant to antibiotics that target cell wall formation, such as beta-lactams and penicillin, because it lacks a cell wall [3]. Recently, azithromycin and moxifloxacin have proven less effective against this infection. As a result of this growing resistance to *M. genitalium* medications, therapeutic failures are common, and the need for routine antimicrobial resistance detection has been emphasized [15]. Furthermore, prioritizing the characterization of surface-exposed antigenic epitopes, which are crucial for antibody-mediated elimination, is essential for the potential development of successful drugs and vaccines targeting bacterial pathogens [16].

Reverse vaccinology, a powerful strategy that has transformed the landscape of vaccine development, offers a promising approach to address this urgency [17]. This innovative approach enables the systemic prediction and selection of antigens, primarily focusing on both T-cell and B-cell epitopes, which have the potential to trigger a robust immune response against the pathogen [18].

By employing a computational approach, researchers can unravel the proteome of *M. genitalium*, identify immunogenic epitopes, and design peptide vaccines that target specific antigens, with the ultimate goal of stimulating the host's immune system to mount a potent and enduring defense against the pathogen. In this work, we present a new approach to develop and optimize T-cell and B-cell epitopes capable of eliciting specific immune responses against *M. genitalium*. Additionally, we have performed *in-silico* cloning and immune simulation analysis of the multiepitope vaccine to assess the antibody response.

2. Methods and materials

This Flowchart displays the overall methodology of the research techniques.

Table 1
Membrane-associated proteins used to predict peptide-based vaccine candidates against *M. genitalium*.

No	Protein	Peptide Length	Subcellular Localization	GI Number	Vaxijen Score	Antigen ProScore	Average Score	PsortB
1	P110	1053	Outside	84626170	0.6900	0.9634	0.8276	Outer Membrane(4.69)
2	P140	88	Outside	468399087	0.5372	0.6472	0.5922	Unknown
3	HMW1	1139	Outside	1046012	0.5884	0.9240	0.7562	Outer Membrane(9.49)
4	HMW2	1805	Outside	3844818	0.4466	0.6944	0.5705	Unknown
5	HMW3	599	Outside	3844894	0.6206	0.8393	0.7299	Outer Membrane(9.49)
6	P32	280	Inside	3844895	0.6996	0.8422	0.7709	Cytoplasmic Membrane (9.86)
7	MG491	346	Outside	3844820	0.3604	0.8471	0.6037	Cytoplasmic Membrane (8.96)
8	MG219	148	Outside	3844821	0.8761	0.6675	0.7718	Unknown
9	P65 homolog	372	Outside	3844817	0.4967	0.5780	0.5373	Cytoplasmic Membrane (9.82)
10	P200	1616	Outside	3844975	0.5361	0.9139	0.7250	Unknown
11	MgpA	131	Outside	4220419	0.8496	0.8894	0.8695	Unknown
12	MgpB	75	Outside	1918701277	0.6302	0.5367	0.5834	Unknown
13	MgpC	450	Outside	1918701961	0.6794	0.9500	0.8147	Unknown

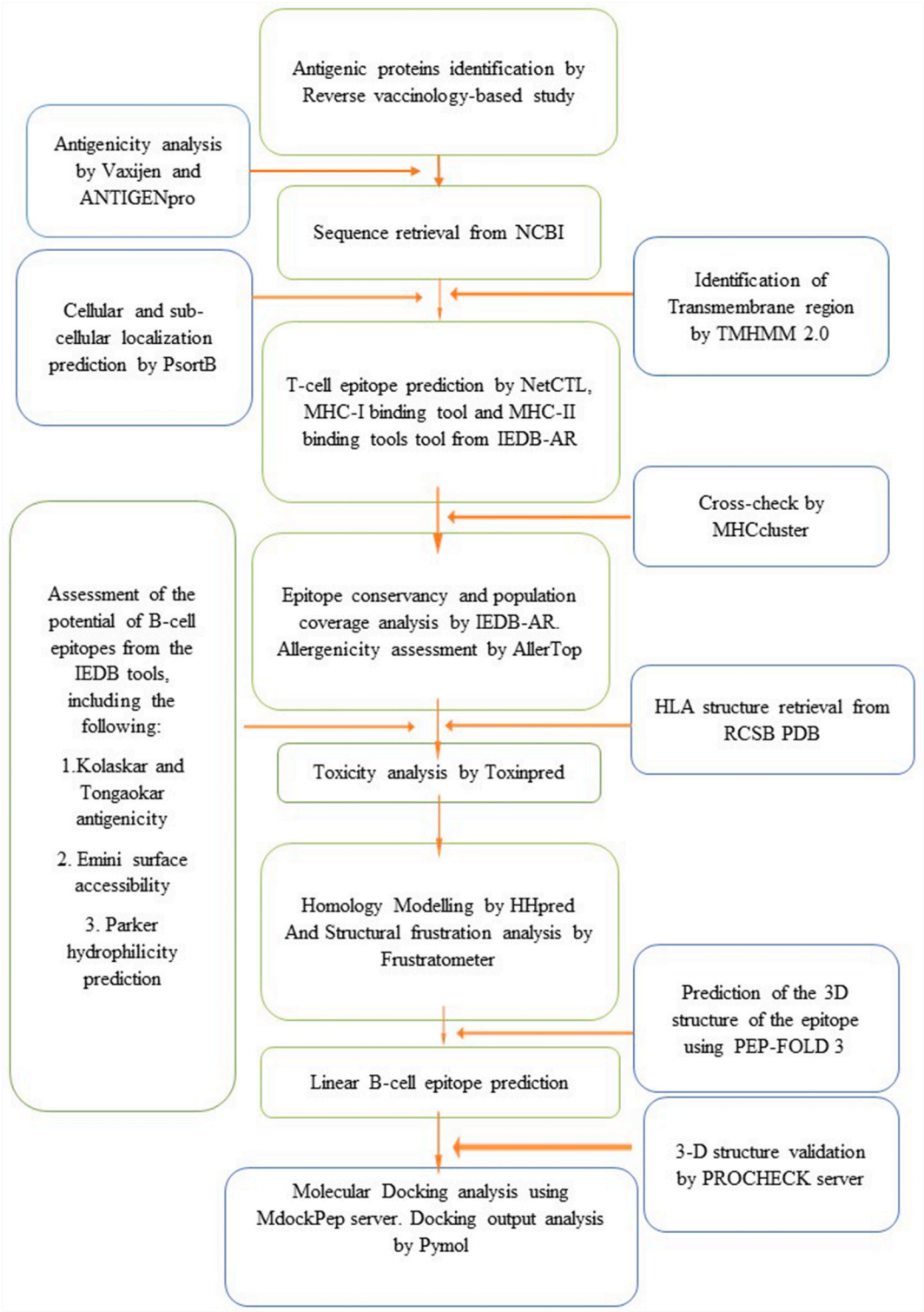

Fig. 1. Visual representation of the workflow, including essential tools and software.

2.1. Selection and retrieval of antigenic protein sequence

At first, we extracted *M. genitalium* membrane-associated proteins from the National Center for Biotechnology Information (NCBI) GenBank (https://www.ncbi.nlm.nih.gov/), along with their GI accession listed in Table 1. Afterward, we identified the most antigenic extracellular proteins, given the significance of outer membrane proteins as viable targets for vaccine design [19]. Through literature review, we listed 13 proteins with extracellular localization and potential antigenicity [4,20,21]. Furthermore, we determined the subcellular localization of these selected proteins as a cross-check. Fig. 1 provides a comprehensive visual representation of the study's workflow, illustrating the sequential steps, essential tools, and software employed throughout the study.

2.2. Identification of antigenic protein

Initially, 13 proteins were identified for our analysis due to their exposure on the cell's surface, these proteins, were identified from a comprehensive background study [4,20,21]. To make sure that the designed vaccine would cause a long-lasting immune reaction, interaction with B and T cell receptors and the antigenicity of the chosen proteins were calculated using the ANTIGENpro server (http://scratch.proteomics.ics.uci.edu/) [22] and the VaxiJen 2.0 server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) [23]. TMHMM-2.0 server (https://services.healthtech.dtu.dk/services/TMHMM-2.0/) [24] and PsortB server (https://www.psort.org/psortb/) [25] were used to predict each protein's subcellular localization for being confirmed of their inheritance.

2.3. T cell epitope prediction

Further, the NetCTL 1.2 server (https://services.healthtech.dtu.dk/services/NetCTL-1.2/) was used to identify T-cell epitopes from highly ranked antigenic proteins, incorporating major histocompatibility complex class I (MHC-I) binding, proteasomal C-terminal cleavage, and TAP transport efficiency prediction [26]. Later, epitopes were then organized by the highest score within each supertype with a score of $\geq$1. However, to assess the binding affinity of the predicted NetCTL epitopes with various MHC class I molecules, we employed the Immune Epitope Database (IEDB) analysis resource database (http://tools.iedb.org/main/), specifically, IC50 value measurement [27]. For MHC-II binding affinity analysis, we extended 9-mer epitopes to 15-mer and determined the IC50 value using the recommended MHC-II binding tool from IEDB. Also, the threshold for MHC-I epitopes was considered at an IC50 value of $\leq$250 nM and for MHC-II epitopes at a value of $\leq$100 nM, respectively [28].

2.4. Allergenicity and toxicity assessment of the predicted epitopes

To evaluate the allergenicity of the epitopes, AllerTOP v.2 servers (http://www.ddg-pharmfac.net/AllerTOP/) utilized auto cross-covariance (ACC) techniques, which included factors such as hydrophobicity, molecular size, and helix forming tendency to determine whether an epitope/peptide exhibits allergenic properties [29]. Additionally, the ToxinPred server (http://crdd.osdd.net/raghava/toxinpred/) is used to assess the toxicity of the epitopes, this server uses multiple SVM algorithms and employs 5-fold cross-validation to enhance the accuracy of the prediction [30]. For our analysis, we employed an SVM-based method with an e-value threshold of 10.

2.5. Cluster evaluation of MHC-restricted alleles

To reevaluate the prediction, MHC-restricted alleles were subjected to cluster analysis. Here, the MHCcluster v2.0 server (https://services.healthtech.dtu.dk/services/MHCcluster-2.0/) was used to cluster MHC class I and MHC class II molecules according to their predicted binding specificity [31]. MHCcluster v2.0 server displays the binding affinity in both graphical tree-based and highly intuitive heat map formats.

2.6. Population coverage and conservancy analysis

Predicted human leukocyte antigens (HLA) are conducted for population coverage to account for the highly polymorphic MHC molecules [32]. The IEDB-AR population coverage tool was used to determine how much of the population is covered by each epitope's binding alleles [33]. The IEDB-AR conservancy tool was also used to assess the conservation of each epitope in different *Mycoplasma* species strains [27].

2.7. B cell epitope prediction

To assess the possibility of 15-mer epitopes as B cell epitopes, we used the IEDB-AR tools Kolaskar & Tongaonkar Antigenicity, Emini Surface Accessibility Prediction, and Parker Hydrophilicity Prediction [27]. Linear B cell epitopes were also identified from the P110 sequence using the ABCpred server (https://webs.iiitd.edu.in/raghava/abcpred/). In this case, the scoring threshold was set at 0.51 and the amino acid length was set to 16 [34]. In the instance of several peptide predictions within this threshold, we picked the best 15 epitopes. The predicted linear epitopes were further checked for potentiality using BepiPred-2.0 [35], BcePred [36], and the BEPITOPE server [37], and were rated as 3 (if more than 75% of amino acids were detected), 2 (51–75%), 1 (26–50%), and 0 (less than 26%) [35]. Then, utilizing the IEDB-AR, AllerTop, VexiJen, and ToxinPred servers, we evaluated the conservancy, allergenicity,

antigenicity, and toxicity of each proposed epitope. We picked the best-fitting epitopes based on this analysis.

2.8. Structure-based virtual screening and molecular docking analysis

2.8.1. Homology modeling and structural validation

To model the 3D structure of the targeted protein P110, we used the HHpred server [38] for protein homology modeling. This server employs a Hidden Markov model that compares pairwise query-template alignments from protein structural databases [39]. Among the available templates, we have chosen the most equivalent and the most accurate alignment as a reference for constructing the homology structure of our targeted protein P110. Further, the PROCHECK server [40] estimated relative quality and residue-residue interactions to evaluate the quality and geometric arrangement of the resulting protein.

2.8.2. HLA allele protein preparation and molecular docking

For the molecular docking study, we carefully selected the most promising MHC-1 and MHC-II peptides from the P110 protein that fulfilled all the previously established criteria as potential vaccine candidates. Likewise, the 3D structure of HLA-C07:02 (PDB ID: 5VGE) MHC-I and HLA-DPB102:01 (PDB ID: 4D8P) MHC-II obtained from RCSB PDB [41], respectively. Similarly, the PEP-FOLD server [42] was used to construct the 3D structure of the peptide candidates. Subsequently, to validate the interaction with HLA, a molecular docking study was conducted using the MDOCKPEP server [43], and the results were then visualized using the PyMol [44].

2.9. Construction of the multi-epitopes vaccine (MEV)

Multiepitope Vaccine (MEV) involves a multi-step process for the development and analysis of the efficacy of the candidate. To construct the MEV, the selected T-cell epitope and B-cell epitope were utilized in conjunction with a suitable adjuvant, the GPI anchor protein of Trypanosoma cruzi (Uniprot ID: P84883) [45,46]. These components were interconnected using a rigid linker (EAAAK) and a flexible linker (GPGPG) to create the MEV [47,48]. The MEV was assembled in the following format: Adjuvant (GPI anchor protein of Trypanosoma cruzi; Uniprot ID: P84883) + EAAAK (Rigid linker) + T-Cell Epitope + GPGPG (Flexible linker) + B-Cell Epitope. Furthermore, the ProtParam tool (https://web.expasy.org/protparam/) is utilized to assess the physiochemical properties of the vaccine candidate. Solubility analysis was conducted using SOLpro [49] and Protein-Sol [50] tools to assess the stability. The secondary structure of the MEV was predicted using the PSIPRED [51] and SOMPA servers [52], which provided valuable insights into its structural characteristics. Additionally, the tertiary structure of the MEV was predicted, refined, and validated using the I-TASER server [53] to ensure its quality and structural integrity. However, molecular docking studies with human Toll-Like receptors (TLR1/2 heterodimer) were performed using the ClusPro 2.0 platform [54] to understand how the vaccine interacts with the host immune system. After the initial tertiary structure prediction, the models of the MEV were further refined using the Galaxy refine server [55]. This step generated refined models. To validate the integrity of MEV, the PDBsum server [56] predicted the conformational quality of the vaccine candidate. This design strategy, incorporating adjuvants and specific linkers for optimal epitope presentation, is a common approach in vaccine development.

2.10. Immune simulation of the MEV

To comprehensively evaluate the immunogenicity and immune response induced by the multi-epitope vaccine, we employed *in-silico* immune simulations conducted using the C-ImmSim server (https://kraken.iac.rm.cnr.it/C-IMMSIM/). C-ImmSim utilizes an agent-based model incorporating a position-specific scoring matrix (PSSM) for immune epitope prediction and employs machine learning techniques to predict immune interactions. The simulations, executed through the C-ImmSim server, encompass the emulation of three anatomical regions crucial for mammalian immune responses: the bone marrow, thymus, and a tertiary lymphatic organ such as a lymph node, tonsil, or spleen. To align with the recommended minimum four-week interval between vaccine injections, we defined time intervals for three injections at time steps 1, 84, and 168. Each time-step corresponds to an 8-h real-time interval, with time-step 1 representing the initial injection at time = 0. The total number of simulation steps equated to 1050. Key simulation parameters, including the random seed, simulation volume (10 μl), and simulation step (1050), were maintained at predefined values, while all other simulation parameters remained at their default settings.

2.11. Codon optimization and SnapGene-associated in-silico cloning

To achieve the highest level of expression in *E. coli*, codon optimization was carried out using the K-12 strain. The Java codon adaptation tool available at http://www.jcat.de/, was utilized to determine the GC content and CAI score of the vector that exhibited the most significant expression. The CAI-value is 1.0 and the GC content of the improved sequence is 54.95%, whereas 35–70% GC content is required for better expression [57]. This optimization process involved meticulous calculations to avoid rho-independent transcription terminators and prokaryotic ribosome binding sites, ensuring optimal and complete protein expression [58]. Finally, the constructed vaccine underwent *in-silico* cloning using SnapGene software (https://www.snapgene.com/).

3. Results

3.1. Antigenic protein identification and sequence retrieval

From the identified 13 proteins, we selected the P110 protein based on the average antigenicity score of the VaxiJen & ANTIGENpro server (Table 1). Among the selected proteins, P110 stands out as a promising candidate, exhibiting a Vaxijen antigenic score of 0.6900 and the highest antigenic ProScore of 0.9634. In comparison, other proteins like HMW1 (0.5884), HMW3 (0.6206), and MG219 (0.8761), although having respectable antigenic scores lag with lower antigen ProScore and average scores of antigenicity. Although MgpA showed competitive results compared to P110 protein, the data availability for MgpA is less than that for P110 protein. In the NCBI, there are more variants of P110 protein available, which enabled us to identify the most conserved P110 protein among a majority of strains. Additionally, we eliminate the P32 protein as its Subcellular localization inside the membrane. The average scores of the rest of the protein are not highly antigenic. Thus, the P110 protein (NCBI GI: 84626170) was chosen based on the average score we got from VaxiJen and the Antigen Pro tool, as well as sequence availability in other *Mycoplasma* species, allergenicity, and population coverage of the corresponding protein's epitopes. Lastly, TMHMM-2.0 server and PsortB server tools were employed to identify the subcellular localization of the selected antigenic protein within *M. genitalium*.

3.2. Prediction of T-cell epitopes and their affinity with MHC class I molecules

Twelve epitopes for each of the twelve HLA supertypes were found in each protein. The NetCTL service provides overall scores for each protein peptide's intrinsic potential utilizing MHC class I binding, proteasomal C-terminal cleavage, and TAP transport efficiency. Table 2 presents selected epitopes along with their positions in the protein sequences.

3.3. MHC class I binding affinity estimation

IEDB-AR tools were used to calculate the potential interacting MHC class I restricted alleles based on the IC50 value (for MHC I $\leq$ 250). MHC allele binding prediction analysis was performed on all top peptides from each NetCTL supertype. Table S1 shows how MHC class I alleles interacted with each P110 epitope, and Table 3 summarizes the results. The overall assessment of MHC I analysis revealed that from P110, **KAYWLLLPF** has interacted with 11 alleles and **NIAPISFSF** interacted with 9 alleles.

3.4. MHCII binding affinity analysis

The MHC class II molecules that interacted with the extended 15-mer epitopes of the P110 protein are shown in Table S2 and summarized in Table 4. Based on the IC50 value and the quantity of interacting MHC class II alleles, we determined the protein epitopes for each one with a strong affinity for MHC class II alleles.

The interacted alleles for both the MHC were reevaluated by cluster analysis to support the functional interaction analysis. The results are shown in Fig. S1 as a dynamic tree and in Fig. S2 as a heat map.

3.5. Conservancy and population coverage analysis

The IEDB Analysis Resource was utilized to evaluate the conservation of all P110 epitopes (Table S4). Using NCBI-accessible sequences, we determined that the peptide sequence 'NIAPISFSFTPFTAA' is 100% conserved.

The IEDB-AR population coverage tool was used to predict MHC class I and MHC class II-based coverage for the entire global population and the various regions of the globe. The P110 protein peptide 'NIAPISFSFTPFTAA' has the highest population coverage of 99.53% in this instance. Population coverage of the binding alleles in different areas of the world is listed in Table S3.

Table 2
The NetCTL server derived T-cell epitopes of P110 based on the combined score. Here, epitopes for all the 12 different HLA supertypes have been presented.

SuperTypes	Epitope (start position)	Combined Score
A1	LTTAVGSVY (988)	2.65
A2	ALPAWVIPV (936)	1.46
A3	QLYYASTSK (239)	1.61
A4	IYKKAYWLL (8)	1.95
A26	DTFSFVVPY (462)	2.21
B7	RPENRGASV (597)	1.72
B8	FPSRVFAGF (926)	1.85
B27	NRGASVTTF (600)	1.34
B39	NQDQATTTL (749)	2.06
B44	VEFADNSVL (842)	1.81
B58	FSIQDTFSF (458)	1.80
B62	IQFLDGNFY (903)	1.50

Table 3
Epitopes for $CD8^+$ T-cell along with the number of interacting MHC class I alleles. The IC-50 range was used at <250 nM.

No.	Epitope (Start position)	Number of Alleles
1	AGFAALPAW (932)	5
2	ALANTFLVK (23)	5
3	DIYKNNNGL (801)	5
4	SLAQLDSSY (66)	7
5	NIAPISFSF (677)	9
6	SPNELRSAK (549)	3
7	KAYWLLLPF (11)	11
8	VFDALGLNY (518)	7
9	EILNSFFRF (868)	3

Table 4
The potential $CD4^+$ T-cell epitopes along with the number of interacting MHC class II alleles with affinity (IC50) < 100 nM.

No.	Epitope (Start position)	Number of Alleles
1	SRVFAGFAALPAWVI (928)	285
2	PFLPLALANTFLVKE (18)	89
3	NKIDIYKNNNGLFEI (798)	17
4	DLVSLAQLDSSYQIA (52)	7
5	NIAPISFSFTPFTAA (677)	116
6	FVYGIVSPNELRSAK (543)	30

3.6. Allergenicity, toxicity and immunogenicity analysis

Based on the physicochemical characteristics of the proteins, the AllerTop v2.0 web server estimated the allergenicity of the putative epitopes. The chosen P110 peptide (NIAPISFSFTPFTAA) was determined as allergy-free (Table 5). We used ToxinPred to assess the toxicity of each peptide, and Vaxijen server to assess the immunogenicity, which resulted in positive scores for immunogenic and negative scores for non-immunogenic peptides (Table S4).

The values on the left and right bars (Fig. 2) represent the proportion of individual and cumulative population coverage of the epitope. The following graph illustrates how the epitopes cover different strains of *M. genitalium*, offering valuable information about their potential applicability.

3.7. Homology modeling: 3D structure prediction and validation

Comparative 3D structures of the selected proteins were predicted using the HHpred server. The server identifies the homologous sequences and aligns them sequentially with the query based on the highest matching. Thus, the P110 sequence of *M. genitalium* has a 100% matching probability with Sugar sugar-binding protein of *M. genitalium G37* (E-value: 0E+00). Based on the highest matching probability, MODELLER modeled the three-dimensional structures of proteins gradually. On the other hand, the Ramachandran plot derived from the PROCHECK server revealed the promising stereo-chemical quality of the predicted model. Here, The P110 model represents that 90.2% of residues are in the most favored regions, 8.9% residues in additional allowed regions, 0.7% residues in generously allowed regions, and 0.1% residues in disallowed regions (Figure S3).

3.8. Localization of the epitopes and structural frustrations analysis

We predicted the transmembrane regions of the protein by using the TMHMM server. The transmembrane region of the P110 was from 13th to 30th, and 939th to 961th amino acid residues. The outside region from 31st to 938th, amino acid residues. Our predicted

Table 5
Population coverage, Allergenicity, Conservancy, toxicity, and Immunogenicity of P110.

no.	Peptide epitope (15 mer) (Start position)	Population coverage (world)	Allergenicity	Conservancy	Toxicity	Immunogenicity
01	SRVFAGFAALPAWVI (928)	99.64%	Probable Non-Allergen	31.25%	Non- Toxin	0.4847
02	PFLPLALANTFLVKE (18)	76.49%	Probable Non-Allergen	31.25%	Non- Toxin	0.7394
03	NKIDIYKNNNGLFEI (798)	26.57%	Probable Non-Allergen	100.00%	Non- Toxin	−0.0371
04	DLVSLAQLDSSYQIA (52)	51.05%	Probable Allergen	93.75%	Non- Toxin	0.0597
05	NIAPISFSFTPFTAA (677)	99.53%	Probable Non-Allergen	100.00%	Non- Toxin	0.9081
06	FVYGIVSPNELRSAK (543)	28.33%	Probable Non-Allergen	100.00%	Non- Toxin	0.5317

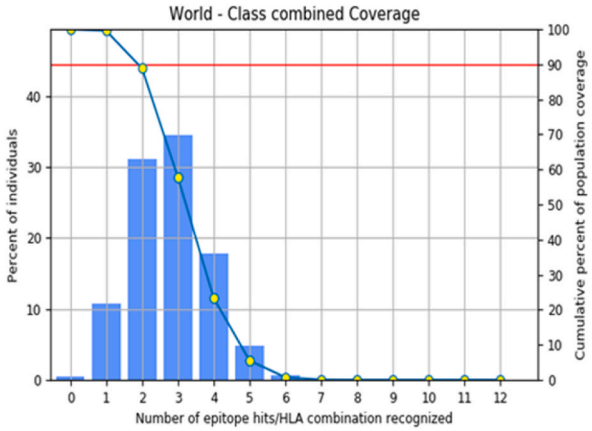

Fig. 2. Analysis of population coverage for the most probable epitopes based on HLA interaction. Figures depicting the global population coverage of a selected 15-mer epitope of P110.

peptide sequences start at the 676th amino acid position of the P110 protein (Fig. S4). So, the epitope is at the outside of the cell and thus it is located at the extracellular parts of the proteins. We also analyzed the position of the final candidate epitope of P110 in their constructed 3D model using PyMOL visualization software. Here we found candidate epitope is positioned at the surface of the respective protein which is the suitable area for epitope recognition and prompt immune response generation by immune mediators (Fig. 3A). In addition to these, we also performed structural frustration studies to analyze the localization of the predicted epitopes. Here, the epitope was in structurally less frustrated regions (Fig. 3B).

3.9. Molecular docking simulation: conformation of binding affinity

Docking simulation provides an eminent understanding of binding interaction between peptide and MHC molecules. The MdockPep server provides different binding patterns of the protein-peptide complex. From different models for each docking simulation, we selected the most compatible one according to their parameters. The cutoff of bRMSD (Å) for the restriction of the peptide conformation value was 5.5(Å) and the exhaustiveness value for the sampling was 100. The server provides ITscorepep which represents the binding score of the peptide-protein complex. A negative binding energy indicates a stronger binding interaction. Here, is the binding energy for the selected model (MHC class I: −137.5; MHC class II: −183.3). Then selected models are analyzed with PyMOL to visualize the specific interaction between different amino acids of MHC alleles that are represented in Fig. 4A, B, and Fig. 5A, B, and

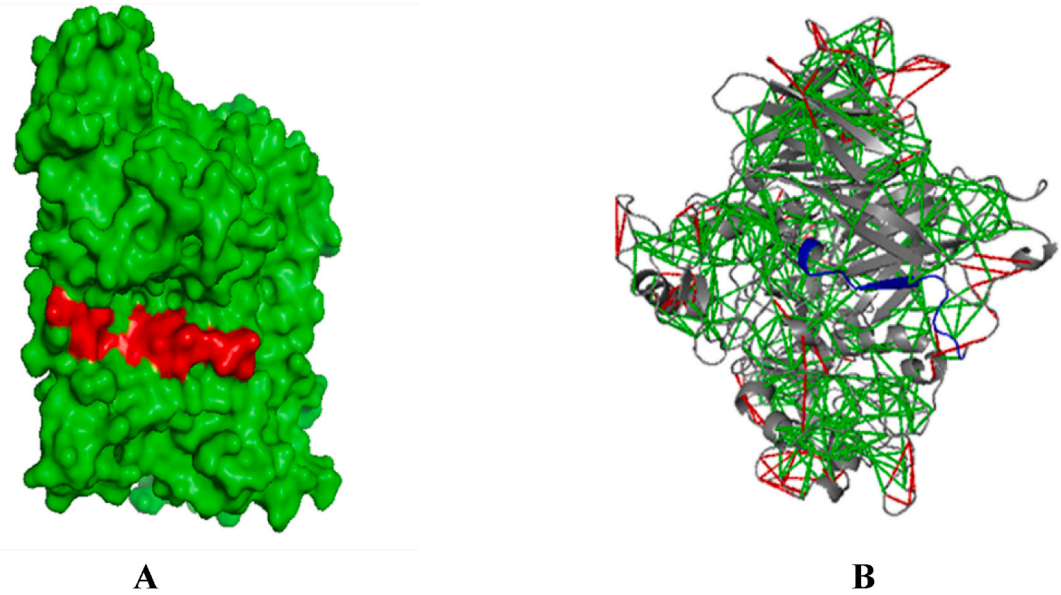

Fig. 3. (A) The three-dimensional model of protein P110. The red region shows the position of the selected most potential epitope. (B) Structural frustration analysis of the proposed epitope of P110. The candidate epitopes are in the green zone signifying that they are not frustrated.

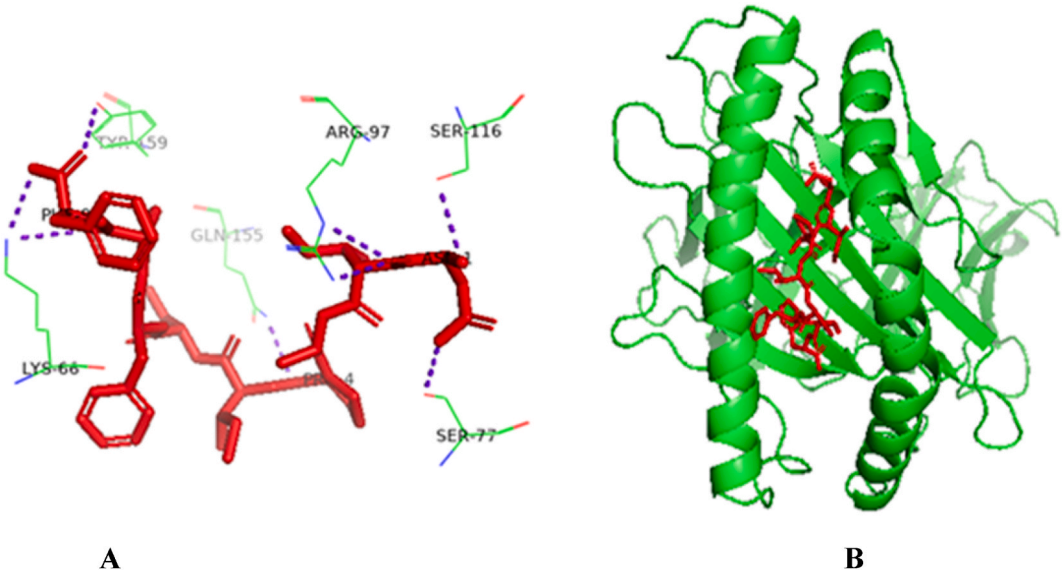

Fig. 4. Docking visualization of P110 predicted peptides (9-mer peptide **NIAPISFSF)** with HLA-C*07:02 allele (PDB ID: 5VGE). Here, (A) the interaction is depicted in cartoon form, and (B) the interacting residues with the peptide ensure precise binding.

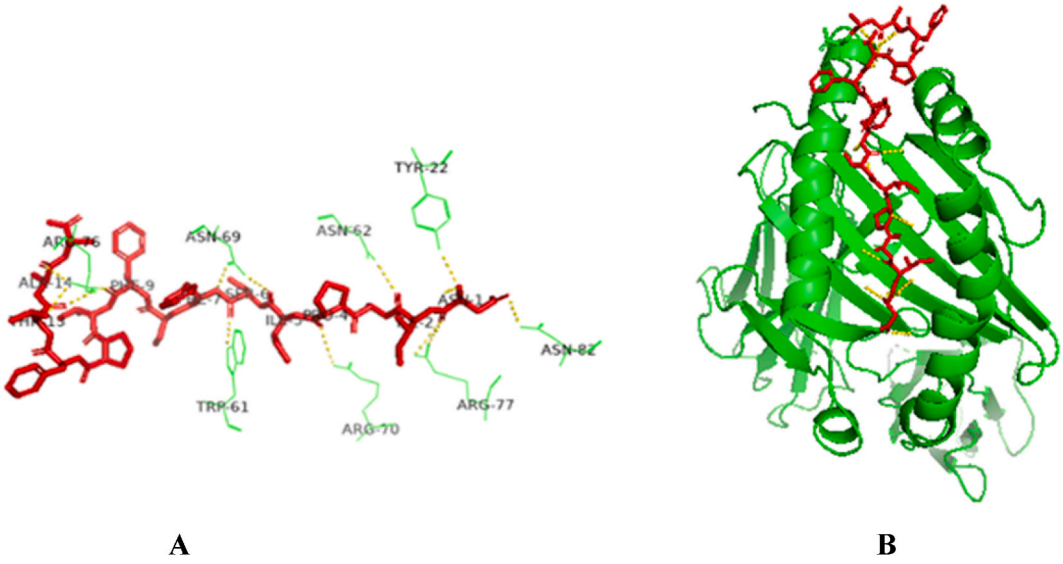

Fig. 5. Docking visualization of P110 predicted peptides (15-mer peptide NIAPISFSFTPFTAA) with HLA-DPB1*02:01 (PDB ID: 4D8P). Here, (A) the interaction is depicted in cartoon form, and (B) the interacting residues with the peptide ensure precise binding.

candidate epitopes listed in Table S7.

3.10. B-cell epitope prediction

Linear B-cell epitope prediction for 15-mer peptides was carried out by evaluating the antigen sequence properties. To predict antigenicity, we selected vaccine candidate peptide **NIAPISFSFTPFTAA** and used the Kolaskar & Tongaonkar antigenicity tool from IEDB-AR. Here the maximum score was 1.078 (Fig. 6A). To strengthen the prediction, the hydrophilicity of the peptides was revealed using the Parker Hydrophilicity scale resulting in a maximum score of 0.386 (Fig. 6C). We also assessed the surface accessibility of the peptides using the Emini Surface Accessibility scale which weighed a maximum score of 1.576 (Fig. 6B).

We also searched out linear B-cell epitopes using several servers. From P110 protein, non-allergen, having higher conservancy, immunogenic, and top scoring epitope was chosen as a potential B-cell epitope. Thus we got the 16-mer B-cell epitope of P110 (**KVKYESSGSNNISFDS)** as a potential B-cell epitope. The predicted B-cell epitopes were also validated by Bepipred, BCEpred, and

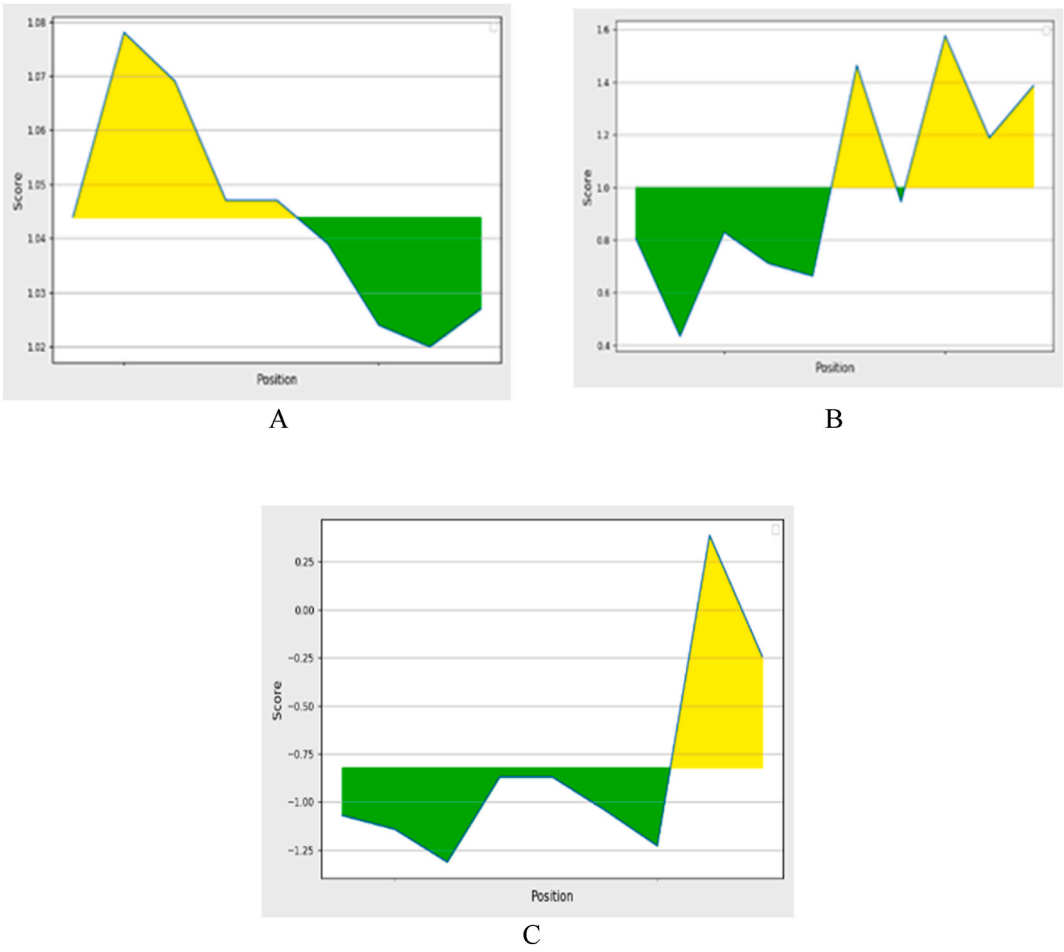

Fig. 6. B-cell epitope prediction in IEDB-AR tool. (A) Kolaskar & Tongaonkar Antigenicity prediction, (B) Emini Surface Accessibility prediction, and (C) Parker Hydrophilicity Prediction.

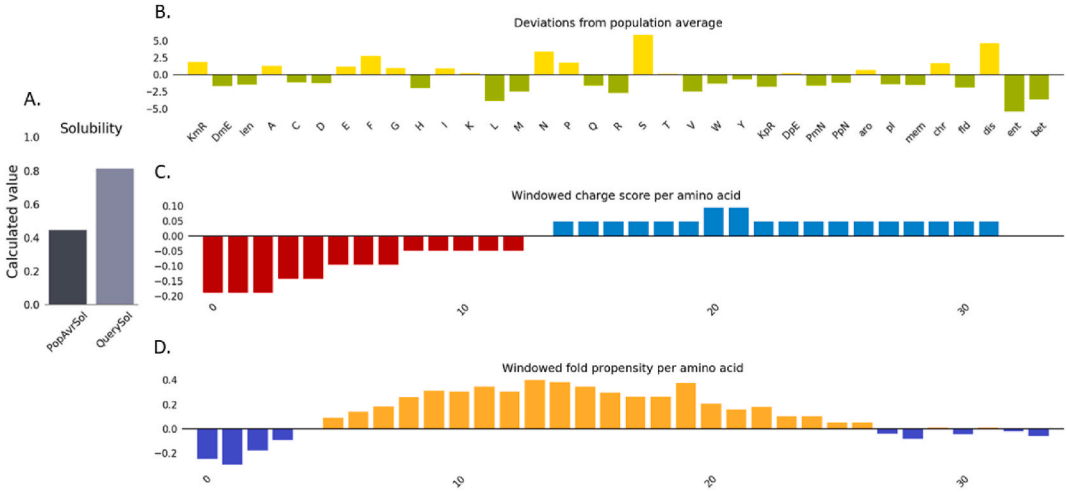

Fig. 7. Solubility analysis of the multiepitope construct was conducted using the protein-sol server. (A) Solubility plot of the MEV alongside the population average for the investigational dataset. (B) deviations from population averages for 35 features. (C) Windowed net charge. (D) Windowed fold propensity.

bepitope servers (Table S5&Table S6).

3.11. Physicochemical properties, antigenicity, allergenicity, and solubility analysis of the MEV

The designed MEV against *M. genitalium* exhibits several promising characteristics, as indicated by various *in-silico* analyses. According to ProtParam tool outputs, the MEV comprises 54 amino acids, with a molecular weight of 5538.89 Da and a theoretical pI of 4.25. The MEV includes 7 negatively charged residues (Asp + Glu) and 3 positively charged residues (Arg + Lys), reflecting its charge distribution. It consists of 754 atoms, which could impact its structural properties. The estimated half-life of the MEV varies across different systems: 4.4 h in mammalian reticulocytes (*in vitro*), >20 h in yeast (*in vivo*), and >10 h in *Escherichia coli* (*in vivo*). The aliphatic index, measuring the relative volume occupied by aliphatic side chains, is 47.22, potentially contributing to the protein's thermostability. The Grand Average of Hydropathicity (GRAVY) score is −0.559, indicating hydrophilicity and solubility, represented in Fig. 7(A-D). The antigenicity of the MEV is highly probable, with a VexiJen score of 1.1526 and an ANTIGENpro probability of 0.623777.

Furthermore, solubility analysis predicted the MEV to be highly soluble with probabilities of 0.951509 and a scaled solubility of 0.812, represented in Fig. 7. These attributes make the MEV a promising candidate for a vaccine against *Mycoplasma genitalium*.

The resulting MEV sequence is as follows:

AQENETNESGSIDEAAAKNIAPISFSFTPFTAAGPGPGKVKYESSGSNNISFDS

3.12. The secondary structure analysis and tertiary structure refinement of the MEV

The secondary structure analysis of the MEV revealed its structural compositions. The results indicate that the MEV contains approximately 14.81% of its sequence in alpha helices and 27.78% in extended strands. Notably, no residues are predicted to form 3_{10} helices, pi helices, beta bridges, beta turns, or bend regions. A significant portion, around 57.41%, is expected to adopt a random coil conformation. Graphically, secondary structure elements are provided in Fig. 8 (A-G). The tertiary structure prediction, refinement, and validation of the MEV offer a comprehensive assessment of the model's quality.

Initially, the I-TASER server was employed to predict the 3D structure of the MEV, selecting Model 1 as the most promising candidate (Fig. 9A). Model 1 displayed a C-score of −2.19, an estimated TM-score of 0.46 ± 0.15, and an estimated RMSD of 7.2 ± 4.2 Å while the other four models (models 2,3,4 and 5) show the C-score of −2.77, −3.20, −2.42 and −4.78. The subsequent refinement of Model 1 using the Galaxy refine server resulted in five refined models (Table 6).

Among them, Model 1 demonstrates a GDT-HA score of 0.9306, low RMSD (0.500), favorable MolProbity (2.186), and a relatively low Clash score (9.2), indicative of structural improvement (Fig. 9B). The validation process, utilizing PdbSum server to generate a Ramachandran plot which provides 72.1%, 20.9%, 2.3%, and 4.7% most favored regions [A, B, L], additional allowed regions [a, b, l, p], Generously allowed regions [~a,~b,~l,~p] and Disallowed regions [XX] respectively (Fig. 9C), further supporting the quality of the selected model. Notably, Model 1 also exhibited a high percentage of favored Ramachandran regions (82.7%) in the GalaxyRefine scoring procedure, with a Z-Score of −1.99, which reinforced the model's reliability (Fig. 9D).

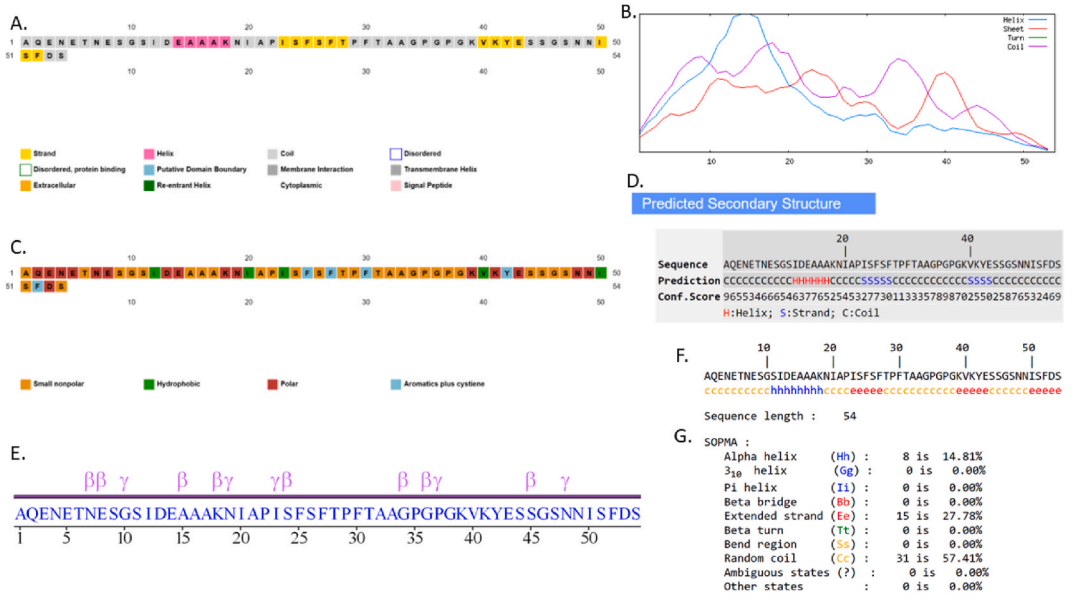

Fig. 8. Secondary Structure Analysis of the MEV: Color-coded annotation grid representing the secondary elements of MEV as produced by PSIPRED 4.0. (A and C), PdbSum (E), I-TASSER (D), and SOMPA (B, F and G).

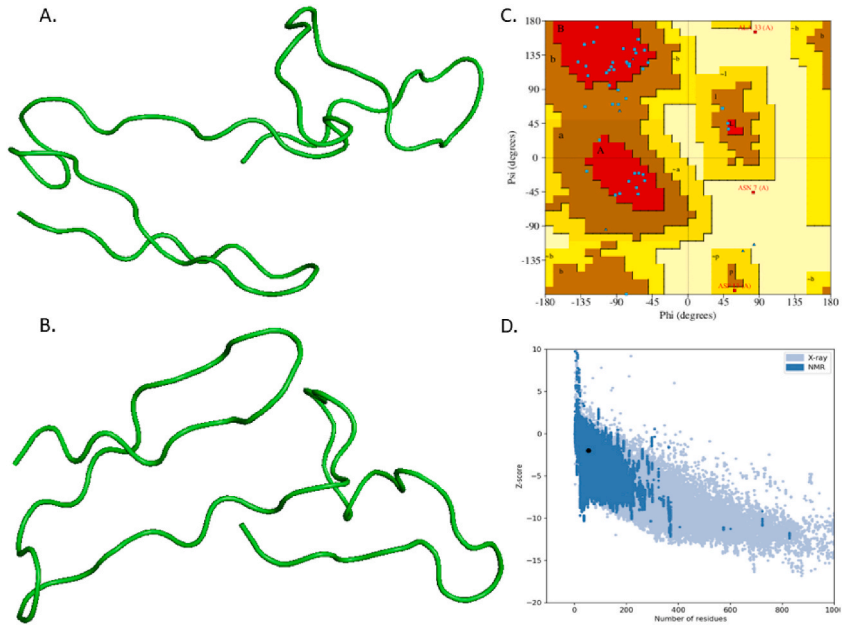

Fig. 9. Modeling, fine-tuning, and quality inspection of 3D models: (A) MEV tertiary model created via I-TASSER; (B) refined model using Galaxyweb; (C) Ramachandran plot of MEV; (D) Z-score (−1.99) as projected by ProSA-Web.

Table 6
Summary of the five refined models generated through the GalaxyWeb server for the MEV.

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	2.655	1.1	14.3	53.8
MODEL 1	0.9306	0.500	2.186	9.2	0.0	82.7
MODEL 2	0.9074	0.566	2.018	7.9	0.0	88.5
MODEL 3	0.9028	0.566	2.215	13.1	0.0	88.5
MODEL 4	0.9259	0.526	2.293	13.1	0.0	84.6

Based on these results and various metrics, Model 1 emerges as the most accurate and reliable 3D structure prediction, making it the optimal choice among the models considered. Therefore, this structure was selected for subsequent studies.

3.13. Molecular docking of the MEV with human toll-like receptors TLR1/2 heterodimer

The molecular docking of the MEV with the human Toll-like receptors TLR1/2 heterodimer was conducted based on insights from the literature suggesting the significant involvement of TLR1/2 in enhancing the immune response against *M. genitalium*. To accomplish this, molecular docking procedures were employed using the ClusPro 2.0 platform. A total of 26 docking models were generated, highlighting potential binding interactions between the human TLR1/2 heterodimer and the MEV.

Subsequently, one model was chosen, primarily based on its larger size and the lowest energy score, which implies a more favorable interaction between the ligand and the receptor. Notably, the selected model, referred to as Model 1 (Fig. 10 A & B), exhibited a notably low energy score of −726.5 and involved 55 docked conformations. These findings collectively suggest a high binding affinity and robust electrostatic interactions between the ligand and the receptor, reinforcing the potential effectiveness of the MEV in eliciting the desired immune response.

3.14. The immune simulation analysis of the MEV

The immune simulation analysis, following the administration of the MEV, reveals intriguing insights as depicted in Fig. 11 With repeated exposures of the host immune system to MEV, there is a pronounced elevation in the levels of secondary and tertiary antibodies, surpassing the detection levels of primary antibodies. Consequently, a swift reduction in antigen concentration, signifying rapid clearance, is evident in Fig. 11A.

The distribution of immune cell populations in various states is portrayed in Fig. 11, showcasing a substantial overall increase in these immune cell populations (Fig. 11B), notably characterized by the development of memory cells. It's worth noting that while T-cell levels initially rise, a subsequent slight decrease is observed in Fig. 11 (C & D). Moreover, the levels of cytokines, particularly IFN-

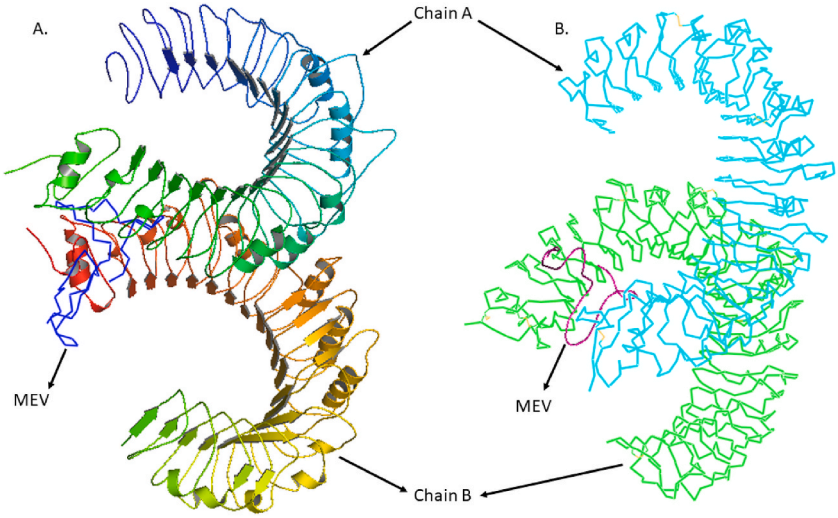

Fig. 10. Molecular docking of the MEV with human toll-like receptors TLR1/2 heterodimer: A. Cartoon Structure B. Ribbon structure.

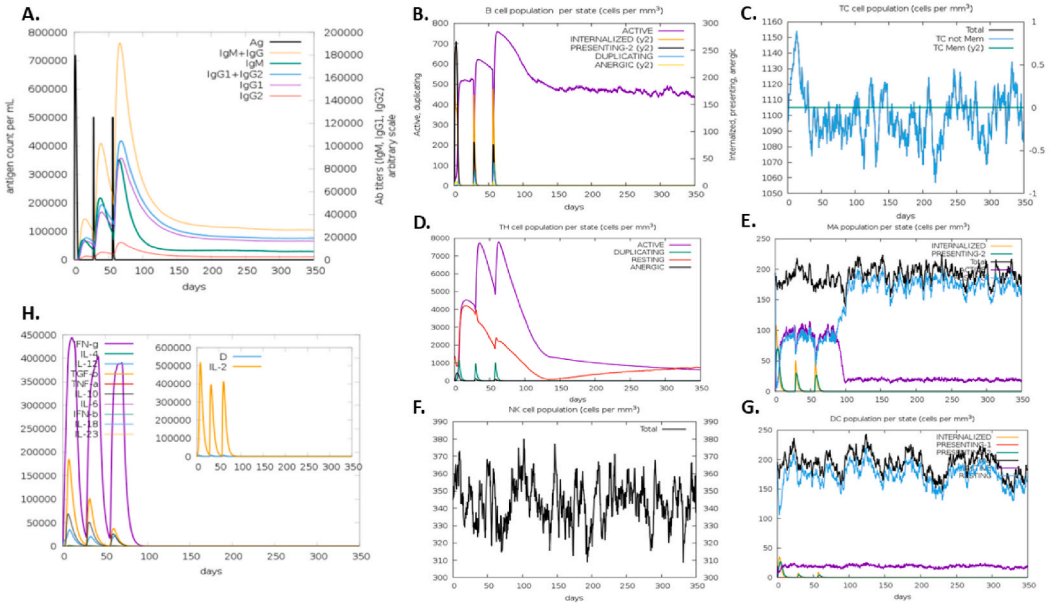

Fig. 11. Immunogenic potential of the MEV. In response to the exposure of MEV after three injections, (A) production of immunoglobulins (B) active B-cell populations/state; (C) cytotoxic T-cell population/state; (D) helper T-cell population/state; (E) activity of macrophage population/state; (F) NK cell population; (G) dendritic cell population per state; (H) cytokine level and interleukins (smaller plot) in different states with the Simpson index (dotted line). All units are in cells/mm3 in three subsequent immune responses.

ϒ, exhibit a significant increase, exceeding 400,000 ng/ml (Fig. 11H). The cellular immune system response to pathogen identification at re-encounter was also robust, including the development of memory cells. The population of the T cell was reported to be > 1150 cells/mm^3 (Fig. 11C). The maximum concentration of 380 cells/mm^3 for the phagocytic natural killer cell population was reported (Fig. 11F) while the dendritic cells and phagocytic macrophage population was reported to be 240 cell/mm^3, 230 cell/mm^3, respectively (Fig. 11 G & E). These compelling findings strongly imply the development of immune memory and underscore the MEV's potential to confer robust immunity against *M. genitalium*.

3.15. Codon adaptation and in-silico cloning

The amino acid sequence of the vaccine was uploaded to JCAT to adapt vaccine sequence codon usage according to the *E. coli* K12 strain. The calculated CAI was 1.00, with the GC contents 47.53%. Such results indicated a better expression level of vaccine construct

in *E. coli* (K12 strain). GC contents between 35 and 70% have been reported for better expression of the gene, as shown in Fig. 12. Finally, the optimized vaccine was cloned into pET28a (+) plasmid using the restriction cloning tool of SnapGene 7.0.0 software.

For this purpose, we selected *Not*I (ATAAGAATGCGGCCGCTAAACTAT) and *Hin*dIII (CCCAAGCTTGGG) restriction sites which show 90% and 75% cleavage at 20 h respectively for restriction digestion of both the target DNA and the vector. Restriction sites for *Not*I and *Hin*dIII were added at the N-terminal and C-terminal of the optimized vaccine construct and cloned in the pET-28a (+) vector which is extracted from pET-28a(+) Sequence and Map (snapgene.com) website. We inserted the target DNA sequence into the UTR region of the pET-28a(+) vector. The size of the cloned pET-28a(+) vector construct with the inserted gene fragment was about 5583 bp.

4. Discussion

Sexually transmitted diseases (STDs) pose a significant global health concern, affecting an estimated one-eighth of the world's population [59]. Among these infections, *M. genitalium* has emerged as a pathogen of particular concern, with associations with a spectrum of diseases, including proctitis, prostatitis, urethritis, cervicitis, and epididymitis [60]. Additionally, *M. genitalium* is also strongly linked to HIV, infertility, ovarian cancer, and prostate cancer [61,62]. The treatment of *M. genitalium* infections historically relied on antibiotics like azithromycin, macrolides, and fluoroquinolones, however, the alarming rise in antibiotic resistance has made it imperative to explore alternative strategies [4].

Our study began with the selection of 13 highly antigenic proteins, including P110 as initial targets (Table 1). Subsequently, the antigenicity of these proteins was thoroughly analyzed, ultimately satisfied us to focus on the P110 protein for further investigation. The p110 protein of *M. genitalium* is a surface-exposed lipoprotein that primarily facilitates adhesion to host cells [63]. Due to its high antigenicity and conservancy (Table 5), making it a suitable target for vaccine development [6]. After analyzing the MHCI and MHCII binding analyses of these core epitopes and their extensions, a total of 50 epitopes with the highest binding to MHCI and MHCII alleles were selected for further analysis based on various criteria [64]. The first criterion to be evaluated for a potential vaccine candidate is

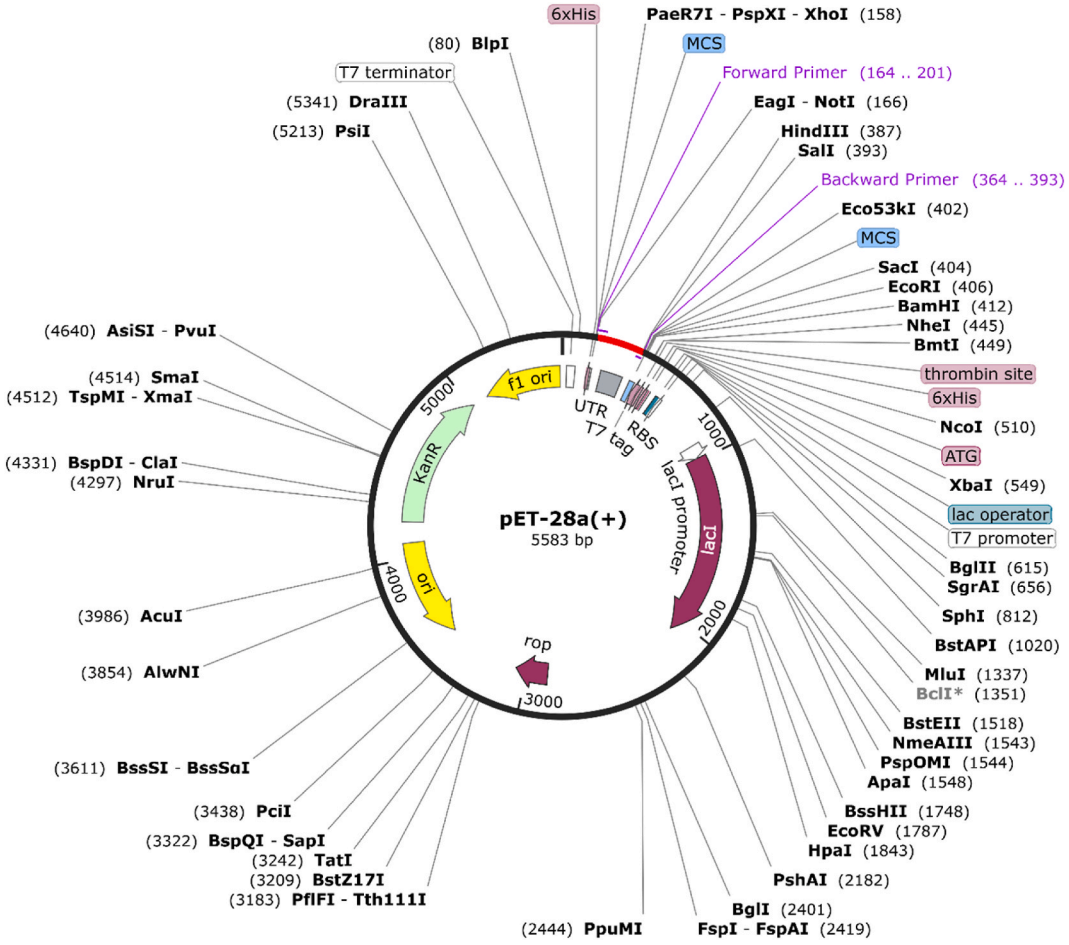

Fig. 12. *In-silico* restriction cloning of the codon-optimized final designed vaccines in the pET-28a (+) vector between the *Not*I (166) and *Hin*dIII (387) restriction enzyme sites (SnapGene software). The final MEV can further be expressed in *E. coli* (strain K12) for efficient vaccine production.

the epitopes' conservation. There are 12 more epitopes that are 100% conserved across all strains of *M. genitalium*. The coverage of each epitope in the world population was then computed. Only NIAPISFSFTPFTAA has a population coverage of over 99.53 percent among the most conserved epitopes (Table 5) [28]. In addition, the population coverage of this epitope is measured in various regions of the globe. In South Asia, Northeast Asia, Europe, East Africa, West Africa, Central Africa, North America, and Oceania, 100 percent of the population is covered. This epitope is nonallergenic and nontoxic as well. In addition, its immunogenicity is 0.9081 (threshold = 0.4), which is exceptionally high.

The predicted epitope is also located in a region of the protein that is neither frustrated nor disordered (Fig. 2). The analysis of molecular docking also revealed a high docking score, low binding energy, and high polar interactions (Figs. 3 and 4). Finally, it was determined that the proposed 15-mer T cell epitope is also a potential B cell epitope, as determined by Emini Surface Accessibility prediction, Kolaskar & Tongaonkar Antigenicity prediction, and Parker Hydrophilicity Prediction method using the IEDB-AR server [27,65,66]. Using numerous B-cell epitope prediction servers, we also predicted linear B-cell epitopes. In addition, we assessed the allergenicity, antigenicity, toxicity, and conservation of all predicted B cell epitopes (Table 5). The epitope KVKYESSGSNNISFDS is a potential B-cell epitope and meets all of the criteria and our requirements [67]. Finally, we can assert that our predicted T cell epitope NIAPISFSFTPFTAA, and B cell epitope, KVKYESSGSNNISFDS, are amazing vaccine candidates, as their sequences have been meticulously predicted. Therefore, the MEV designed in this study exhibits a multitude of promising characteristics, making it a strong candidate for a vaccine against *M. genitalium*. We assessed the antigenicity of our designed vaccine using the Vaxijen server, which provided a score of 1.1526, indicating high antigenic potential. Furthermore, the solubility score, calculated as 0.951509, underscores the favorable solubility characteristics of the vaccine protein [68]. Molecular docking serves as a crucial tool for the exploration of specific interactions and binding affinities between ligands and chosen receptors, following the lock-and-key principle [69]. Numerous immunoinformatic studies have underscored the pivotal role played by these receptors in eliciting an efficient immune response [70–73]. Toll-like receptors (TLRs) are known to have a substantial impact on innate immunity, including immune activation and the initiation of adaptive immune response. Additionally, prior research has delved into the participation of TLR1 and TLR2 in the recognition of viral structural proteins, culminating in the production of inflammatory cytokines [74]. The docking calculation involves the designed MEV and the receptors including MHCs and TLRs, have revealed an excellent fit of the vaccines into the binding pockets of these receptors. Lastly, *in-silico* cloning was performed using the pET28a (+) vector, following codon optimization by the JCAT web server [58]. The codon adaptation indices (CAIs) and GC content of the optimized nucleotides fell within the accepted ranges of 0.8–1.0 and 47.53%, respectively [75,76]. These results confirm that the designed MEVs are reliable for expression in the *E. coli* strain K12. The suggested vaccine exhibits potential in countering immune response, as indicated by analysis results, including allergenicity, antigenicity, population coverage, and immuno-simulation outcomes. It is important to note that a limitation of the study is the need for experimental validation of the final vaccine construct to confirm its ability to enhance immunity and clinical application.

5. Conclusion

Before commencing the tangible synthesis in a wet lab, the adoption of *in-silico* methodologies for vaccine design is frequently embraced. This saves time and money as well as helps guide lab work, reducing errors and increasing the chance of success. In our study, the application of computational methods to predict excellent parts of the P110 protein that could be used as a peptide-based vaccine. Ultimately, we unveiled a coveted T cell and B cell epitope that embodies exceptional immunogenicity, complete conservation, non-allergenicity, non-toxicity, substantial population coverage, and all other necessary prerequisites for a vaccine candidate, where the 'NIAPISFSFTPFTAA' is T-cell epitope and the 'KVKYESSGSNNISFDS' is B-cell epitope. This monumental revelation holds immense potential in combating the recently uncovered *M. genitalium* bacteria, firmly establishing itself as a suitable candidate for vaccine development. Nevertheless, to authenticate the veracity of this *in-silico* work, additional implementation of the wet lab experiment is imperative.

Data availability statement

All data included within the study will be available for everyone as per journal policy.

Funding information

None/No funding source.

CRediT authorship contribution statement

Asma Akter: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Natasha Farhin Ananna:** Writing – original draft, Methodology, Formal analysis, Conceptualization. **Hedayet Ullah:** Validation, Software, Methodology, Investigation, Formal analysis. **Sirajul Islam:** Validation, Methodology, Formal analysis. **Md Al Amin:** Validation, Software, Methodology, Investigation, Formal analysis. **K.M. Kaderi Kibria:** Writing – review & editing, Software, Methodology, Formal analysis. **Shahin Mahmud:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e28223.

References

[1] G. Martín-Saco, A. Tristancho, A. Arias, I. Ferrer, A. Milagro, J.M. García-Lechuz, Mycoplasma genitalium and sexually transmitted infections: evidence and figures in a tertiary hospital, Rev. Española Quimioter. 35 (1) (2022), https://doi.org/10.37201/req/091.2021.
[2] R. Gnanadurai, H. Fifer, Mycoplasma genitalium: A review," Microbiology (United Kingdom) 166 (1) (2020), https://doi.org/10.1099/mic.0.000830.
[3] J.L. Munoz, O.J. Goje, Mycoplasma genitalium: an emerging sexually transmitted infection, Sci. Tech. Rep. 2016 (2016), https://doi.org/10.1155/2016/7537318.
[4] T. Kirby, Mycoplasma genitalium: a potential new superbug, Lancet Infect. Dis. 18 (9) (2018), https://doi.org/10.1016/S1473-3099(18)30506-1.
[5] D. Taylor-Robinson, J.S. Jensen, Mycoplasma genitalium: from chrysalis to multicolored butterfly, Clin. Microbiol. Rev. 24 (3) (2011), https://doi.org/10.1128/CMR.00006-11.
[6] W. Yueyue, X. Feichen, X. Yixuan, L. Lu, C. Yiwen, Y. Xiaoxing, Pathogenicity and virulence of Mycoplasma genitalium: unraveling ariadne's thread, Virulence 13 (1) (2022) 1161–1183, https://doi.org/10.1080/21505594.2022.2095741.
[7] L. García-Morales, L. González-González, E. Querol, J. Piñol, A minimized motile machinery for Mycoplasma genitalium, Mol. Microbiol. 100 (1) (2016), https://doi.org/10.1111/mmi.13305.
[8] C.L. McGowin, P.A. Totten, The unique microbiology and molecular pathogenesis of Mycoplasma genitalium, J. Infect. Dis. 216 (2017), https://doi.org/10.1093/infdis/jix172.
[9] R. Burgos, O.Q. Pich, M. Ferrer-Navarro, J.B. Baseman, E. Querol, J. Piñol, Mycoplasma genitalium P140 and P110 cytadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development, J. Bacteriol. 188 (24) (2006), https://doi.org/10.1128/JB.00978-06.
[10] J.B. Baseman, J.G. Tully, Mycoplasmas: sophisticated, reemerging, and burdened by their notoriety, Emerg. Infect. Dis. 3 (1) (1997), https://doi.org/10.3201/eid0301.970103.
[11] C. Anagrius, B. Loré, J.S. Jensen, Mycoplasma genitalium: prevalence, clinical significance, and transmission, Sex. Transm. Infect. 81 (6) (2005), https://doi.org/10.1136/sti.2004.012062.
[12] S. Harish, Sexually transmitted infection by Mycoplasma genitalium : a short review, J. Ski. Sex. Transm. Dis. 3 (1) (2021) 46–50, https://doi.org/10.25259/jsstd_14_2021.
[13] E.A. Torrone, et al., Prevalence of urogenital Mycoplasma genitalium infection, United States, 2017 to 2018, Sex. Transm. Dis. 48 (11) (2021) e160–e162, https://doi.org/10.1097/OLQ.0000000000001394.
[14] T.E. van der Schalk, J.F. Braam, J.G. Kusters, Molecular basis of antimicrobial resistance in Mycoplasma genitalium, Int. J. Antimicrob. Agents 55 (4) (2020) 2–7, https://doi.org/10.1016/j.ijantimicag.2020.105911.
[15] E. Shipitsyna, et al., Mycoplasma genitalium prevalence, antimicrobial resistance-associated mutations, and coinfections with non-viral sexually transmitted infections in high-risk populations in Guatemala, Malta, Morocco, Peru and South Africa, 2019–2021, Front. Microbiol. 14 (2023), https://doi.org/10.3389/fmicb.2023.1130762.
[16] W.G. Nogueira, et al., Computational identification of putative common genomic drug and vaccine targets in Mycoplasma genitalium, Genomics 113 (4) (2021) 2730–2743, https://doi.org/10.1016/j.ygeno.2021.06.011.
[17] R. Moxon, P.A. Reche, R. Rappuoli, Editorial: reverse vaccinology, Front. Immunol. 10 (2019), https://doi.org/10.3389/fimmu.2019.02776.
[18] S. Khan, et al., Proteome-wide mapping and reverse vaccinology-based B and T cell multi-epitope subunit vaccine designing for immune response reinforcement against Porphyromonas gingivalis, J. Biomol. Struct. Dyn. 40 (2) (2022), https://doi.org/10.1080/07391102.2020.1819423.
[19] M.J.H. Gerritzen, D.E. Martens, R.H. Wijffels, L. van der Pol, M. Stork, Bioengineering bacterial outer membrane vesicles as vaccine platform, Biotechnol. Adv. 35 (5) (2017), https://doi.org/10.1016/j.biotechadv.2017.05.003.
[20] L. Li, D. Luo, Y. Liao, K. Peng, Y. Zeng, Mycoplasma genitalium protein of adhesion induces inflammatory cytokines via cyclophilin A-cd147 activating the ERK-NF-κb pathway in human urothelial cells, Front. Immunol. 11 (2020), https://doi.org/10.3389/fimmu.2020.02052.
[21] S.L. Iverson-Cabral, S.G. Astete, C.R. Cohen, P.A. Totten, mgpB and mgpC sequence diversity in Mycoplasma genitalium is generated by segmental reciprocal recombination with repetitive chromosomal sequences, Mol. Microbiol. 66 (1) (2007), https://doi.org/10.1111/j.1365-2958.2007.05898.x.
[22] C.N. Magnan, et al., High-throughput prediction of protein antigenicity using protein microarray data, Bioinformatics 26 (23) (2010), https://doi.org/10.1093/bioinformatics/btq551.
[23] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinf. 8 (2007), https://doi.org/10.1186/1471-2105-8-4.
[24] A. Krogh, B. Larsson, G. Von Heijne, E.L.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (3) (2001), https://doi.org/10.1006/jmbi.2000.4315.
[25] N.Y. Yu, et al., PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes, Bioinformatics 26 (13) (2010), https://doi.org/10.1093/bioinformatics/btq249.
[26] M.V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction, BMC Bioinf. 8 (2007), https://doi.org/10.1186/1471-2105-8-424.
[27] S.K. Dhanda, et al., IEDB-AR: immune epitope database - analysis resource in 2019, Nucleic Acids Res. 47 (W1) (2019), https://doi.org/10.1093/nar/gkz452.
[28] H. Ullah, S. Mahmud, M.J. Hossain, M.S. Bin Islam, K.M.K. Kibria, Immunoinformatic identification of the epitope-based vaccine candidates from Maltoporin, FepA and OmpW of Shigella Spp, with molecular docking confirmation, Infect. Genet. Evol. 96 (2021), https://doi.org/10.1016/j.meegid.2021.105129.
[29] Y. Kim, et al., Immune epitope database analysis resource, Nucleic Acids Res. 40 (W1) (2012) 525–530, https://doi.org/10.1093/nar/gks438.
[30] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, In silico approach for predicting toxicity of peptides and proteins, PLoS One 8 (9) (2013), https://doi.org/10.1371/journal.pone.0073957.

[31] M. Thomsen, C. Lundegaard, S. Buus, O. Lund, M. Nielsen, MHCcluster, a method for functional clustering of MHC molecules, Immunogenetics 65 (9) (2013), https://doi.org/10.1007/s00251-013-0714-9.

[32] E.T. Abualrous, J. Sticht, C. Freund, Major histocompatibility complex (MHC) class I and class II proteins: impact of polymorphism on antigen presentation, Curr. Opin. Immunol. 70 (2021), https://doi.org/10.1016/j.coi.2021.04.009.

[33] W. Fleri, et al., The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design, Front. Immunol. 8 (MAR) (2017), https://doi.org/10.3389/fimmu.2017.00278.

[34] S. Saha, G.P.S. Raghava, Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, Proteins Struct. Funct. Genet. 65 (1) (2006), https://doi.org/10.1002/prot.21078.

[35] M.C. Jespersen, B. Peters, M. Nielsen, P. Marcatili, BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes, Nucleic Acids Res. 45 (W1) (2017), https://doi.org/10.1093/nar/gkx346.

[36] S. Saha, G.P.S. Raghava, BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties, Lect. Notes Comput. Sci. 3239 (May 2014) (2004) 197–204, https://doi.org/10.1007/978-3-540-30220-9_16.

[37] M. Odorico, J.L. Pellequer, BEPITOPE: predicting the location of continuous epitopes and patterns in proteins, J. Mol. Recogn. 16 (1) (2003), https://doi.org/10.1002/jmr.602.

[38] J. Söding, A. Biegert, A.N. Lupas, The HHpred interactive server for protein homology detection and structure prediction, Nucleic Acids Res. 33 (SUPPL. 2) (2005), https://doi.org/10.1093/nar/gki408.

[39] A. Krogh, M. Brown, I.S. Mian, K. Sjölander, D. Haussler, Hidden Markov Models in computational biology applications to protein modeling, J. Mol. Biol. 235 (5) (1994), https://doi.org/10.1006/jmbi.1994.1104.

[40] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (2) (1993), https://doi.org/10.1107/s0021889892009944.

[41] J.L. Sussman, et al., Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules, Acta Crystallogr., Sect. D: Biol. Crystallogr. 54 (6 I) (1998), https://doi.org/10.1107/S0907444998009378.

[42] P. Thévenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux, P. Tufféry, PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides, Nucleic Acids Res. 40 (W1) (2012), https://doi.org/10.1093/nar/gks419.

[43] X. Xu, X. Zou, MDockPeP: a web server for blind prediction of protein–peptide complex structures, in: Methods in Molecular Biology, vol. 2165, 2020.

[44] M.A. Lill, M.L. Danielson, Computer-aided drug design platform using PyMOL, J. Comput. Aided Mol. Des. 25 (1) (2011), https://doi.org/10.1007/s10822-010-9395-8.

[45] A. Bateman, et al., UniProt: the universal protein knowledgebase in 2021, Nucleic Acids Res. 49 (2021), https://doi.org/10.1093/nar/gkaa1100. D1.

[46] J.I. MacRae, A. Acosta-Serrano, N.A. Morrice, A. Mehlert, M.A.J. Ferguson, Structural characterization of NETNES, a novel glycoconjugate in Trypanosoma cruzi epimastigotes, J. Biol. Chem. 280 (13) (2005), https://doi.org/10.1074/jbc.M412939200.

[47] V.P. Reddy Chichili, V. Kumar, J. Sivaraman, Linkers in the structural biology of protein-protein interactions, Protein Sci. 22 (2) (2013), https://doi.org/10.1002/pro.2206.

[48] X. Chen, J.L. Zaro, W.C. Shen, Fusion protein linkers: property, design and functionality, Adv. Drug Deliv. Rev. 65 (10) (2013), https://doi.org/10.1016/j.addr.2012.09.039.

[49] C.N. Magnan, A. Randall, P. Baldi, SOLpro: accurate sequence-based prediction of protein solubility, Bioinformatics 25 (17) (2009), https://doi.org/10.1093/bioinformatics/btp386.

[50] M. Hebditch, M.A. Carballo-Amador, S. Charonis, R. Curtis, J. Warwicker, Protein-Sol: a web tool for predicting protein solubility from sequence, Bioinformatics 33 (19) (2017), https://doi.org/10.1093/bioinformatics/btx345.

[51] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics 16 (4) (2000), https://doi.org/10.1093/bioinformatics/16.4.404.

[52] C. Geourjon, G. Deléage, Sopma: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, Bioinformatics 11 (6) (1995), https://doi.org/10.1093/bioinformatics/11.6.681.

[53] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinf. 9 (2008), https://doi.org/10.1186/1471-2105-9-40.

[54] D. Kozakov, et al., The ClusPro web server for protein-protein docking, Nat. Protoc. 12 (2) (2017), https://doi.org/10.1038/nprot.2016.169.

[55] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, Nucleic Acids Res. 41 (2013), https://doi.org/10.1093/nar/gkt458. Web Server issue.

[56] R.A. Laskowski, J. Jabłońska, L. Pravda, R.S. Vařeková, J.M. Thornton, PDBsum: structural summaries of PDB entries, Protein Sci. 27 (1) (2018), https://doi.org/10.1002/pro.3289.

[57] M. Ali, R.K. Pandey, N. Khatoon, A. Narula, A. Mishra, V.K. Prajapati, Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection, Sci. Rep. 7 (1) (2017), https://doi.org/10.1038/s41598-017-09199-w.

[58] A. Grote, et al., JCat: a novel tool to adapt codon usage of a target gene to its potential expression host, Nucleic Acids Res. 33 (SUPPL. 2) (2005), https://doi.org/10.1093/nar/gki376.

[59] M.O. Thorpe, R.R. Clipstein, Global epidemiology, Br. Med. J. 1 (5274) (1962) 332, https://doi.org/10.1136/bmj.1.5274.332.

[60] P.J. Horner, D.H. Martin, Mycoplasma genitalium infection in men, J. Infect. Dis. 216 (2017), https://doi.org/10.1093/infdis/jix145.

[61] M.B. Rogers, Mycoplasma and cancer: in search of the link, Oncotarget 2 (4) (2011) 271–273, https://doi.org/10.18632/oncotarget.264.

[62] S. Huang, J.Y. Li, J. Wu, L. Meng, C.C. Shou, Mycoplasma infections and different human carcinomas, World J. Gastroenterol. 7 (2) (2001) 266–269, https://doi.org/10.3748/wjg.v7.i2.266.

[63] D. Aparicio, et al., Mycoplasma genitalium adhesin P110 binds sialic-acid human receptors, Nat. Commun. 9 (1) (2018), https://doi.org/10.1038/s41467-018-06963-y.

[64] M.S. Bin Islam, M. Miah, M.E. Hossain, K.M.K. Kibria, A conserved multi-epitope-based vaccine designed by targeting hemagglutinin protein of highly pathogenic avian H5 influenza viruses, 3 Biotech 10 (12) (2020), https://doi.org/10.1007/s13205-020-02544-3.

[65] S.S. Bappy, et al., Extensive immunoinformatics study for the prediction of novel peptide-based epitope vaccine with docking confirmation against envelope protein of Chikungunya virus: a computational biology approach, J. Biomol. Struct. Dyn. 39 (4) (2021), https://doi.org/10.1080/07391102.2020.1726815.

[66] T. Kar, et al., A candidate multi-epitope vaccine against SARS-CoV-2, Sci. Rep. 10 (1) (2020), https://doi.org/10.1038/s41598-020-67749-1.

[67] W. Li, M.D. Joshi, S. Singhania, K.H. Ramsey, A.K. Murthy, Peptide vaccine: progress and challenges, Vaccines 2 (3) (2014), https://doi.org/10.3390/vaccines2030515.

[68] A. Singh, M. Thakur, L.K. Sharma, K. Chandra, Designing a multi-epitope peptide based vaccine against SARS-CoV-2, Sci. Rep. 10 (1) (2020), https://doi.org/10.1038/s41598-020-73371-y.

[69] J. Fan, A. Fu, L. Zhang, Progress in molecular docking, Quantitative Biology 7 (2) (2019), https://doi.org/10.1007/s40484-019-0172-y.

[70] L.A. Kerepesi, O. Leon, S. Lustigman, D. Abraham, Protective immunity to the larval stages of Onchocerca volvulus is dependent on toll-like receptor 4, Infect. Immun. 73 (12) (2005), https://doi.org/10.1128/IAI.73.12.8291-8297.2005.

[71] K.M. Pfarr, K. Fischer, A. Hoerauf, Involvement of Toll-like receptor 4 in the embryogenesis of the rodent filaria Litomosoides sigmodontis, Med. Microbiol. Immunol. 192 (1) (2003), https://doi.org/10.1007/s00430-002-0159-5.

[72] S. Ahmad, et al., Design of a novel multi epitope-based vaccine for pandemic coronavirus disease (COVID-19) by vaccinomics and probable prevention strategy against avenging zoonotics, Eur. J. Pharmaceut. Sci. 151 (2020), https://doi.org/10.1016/j.ejps.2020.105387.

[73] K. Vijay, Toll-like receptors in immunity and inflammatory diseases: past, present, and future, Int. Immunopharm. 59 (2018), https://doi.org/10.1016/j.intimp.2018.03.002.

[74] T. Compton, et al., Human cytomegalovirus activates inflammatory cytokine responses via CD14 and toll-like receptor 2, J. Virol. 77 (8) (2003), https://doi.org/10.1128/jvi.77.8.4588-4596.2003.
[75] M. Wang, et al., The growth performance, intestinal digestive and absorptive capabilities in piglets with different lengths of small intestines, Animal 14 (6) (2020), https://doi.org/10.1017/S175173111900288X.
[76] F.J. Liu, et al., Evaluation of a recombinant tetanus toxin subunit vaccine, Toxicon 187 (2020), https://doi.org/10.1016/j.toxicon.2020.08.001.