Scholars Journal of Engineering and Technology

OPEN ACCESS

Abbreviated Key Title: Sch J Eng Tech
ISSN 2347-9523 (Print) | ISSN 2321-435X (Online)
Journal homepage: https://saspublishers.com

Application of Immunoinformatics for Efficient Epitope-Based Peptide Vaccine Development

Simiao Yu[1], Xinhui Yu[1], Yantong Zhou[1], Nan Chen[1], Peng Wang[1], Shaojie Bi[1*]

[1]College of Life Science and Technology, Heilongjiang Bayi Agricultural University, Xinfeng Road 5, Daqing 163319, China

DOI: https://doi.org/10.36347/sjet.2024.v12i10.001 | **Received:** 26.08.2024 | **Accepted:** 02.10.2024 | **Published:** 04.10.2024

***Corresponding author:** Shaojie Bi
College of Life Science and Technology, Heilongjiang Bayi Agricultural University, Xinfeng Road 5, Daqing 163319, China

Abstract **Review Article**

As global public health continues to face evolving challenges, developing effective and safe vaccines has become increasingly crucial, with the design of epitope-based vaccines representing a significant breakthrough in the field. The design of epitope vaccines generally involves utilizing immunoinformatics tools and techniques to analyze known or newly identified amino acid sequences, allowing for pre-selecting and identifying potential dominant epitopes. Subsequently, peptide vaccines containing these epitopes are synthesized through chemical or genetic engineering. The rapid progress in Immunoinformatics has greatly expanded its application in vaccinology, making it the most efficient approach for developing epitope-based peptide vaccines. The review outlines the general workflow of Immunoinformatics in the design and validation of epitope vaccines, highlights the key Immunoinformatics tools used in epitope vaccine development, and discusses the specific applications of Immunoinformatics in epitope vaccine design. These insights aim to provide valuable references for the rational design and development of potent epitope vaccines.

Keywords: Immunoinformatics, Epitope, Vaccine.

Copyright © 2024 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

INTERDACTION

As the threat of emerging and re-emerging infectious diseases continues to rise, there is an urgent need to develop innovative vaccine strategies that are both effective and safe to safeguard global public health [1]. In recent years, the continual mutation of bacterial and viral pathogens has brought vaccines to the forefront of public health discussions. Vaccines elicit humoral and cellular immune responses within the body, facilitating rapid immune defenses that lower infection rates and reduce morbidity and mortality. Vaccine development is successful largely depending on its capacity to elicit robust and appropriate immune responses [2]. Though capable of generating effective neutralizing antibodies and protective immune responses, traditional vaccine development encounters challenges such as high production costs, potential allergic reactions, and the need to cultivate pathogens in vitro, raising concerns over vaccine safety. In contrast, peptide vaccines can selectively and precisely activate immune responses without requiring in vitro culture, offering a higher biosafety level. Additionally, the long development timelines and substantial financial burdens associated with traditional vaccines can be alleviated by integrating modern information technologies, molecular biology advancements, and genomic database tools. Computer-assisted immunoinformatics allows the prediction of optimal immunogenic elements to rationally design antigens while enabling the simulation and evaluation of vaccine candidates before biological testing [3]. The approach significantly enhances the efficiency of vaccine development by minimizing side effects and reducing both the time and cost involved. The review explores the widely-used immunoinformatics tools for epitope vaccine design, describes the general workflow and validation process for epitope vaccine development, and reviews the progress made in applying immunoinformatics to epitope vaccine design, offering a reference for future research in the evolving field.

1. Immune Informatics

Immune informatics, also known as immunoinformatics, is an interdisciplinary field that integrates informatics principles with modern immunology to address complex immunological challenges [4]. The field employs bioinformatics tools, computational models, and experimental immunological data to study the mechanisms of immune signal recognition and transmission, particularly focusing on antigen signals. Immune informatics involves identifying the patterns and regularities in how immune signals are processed and predicting immunological

phenomena, which are then validated through experimental studies. The approach addresses practical issues in immunodiagnostics, immunotherapy, and vaccine development. As a tool in vaccine design, immune informatics streamlines research processes, reduces associated costs, and enables the effective screening of candidate epitopes for vaccine development, facilitating previously difficult or impossible advancements. The analytical approach is increasingly important in modern immunology and crucial for developing immunotherapeutic drugs and vaccines.

2. General Process and Vaccine Validation of Immune Informatics in Epitope Vaccine Development

2.1 The General Workflow of Immunoinformatics in Epitope Vaccine Development

The rapid development of immunoinformatics has brought about revolutionary changes in vaccine development by providing advanced computational tools that facilitate the rational design of vaccines, with a particular emphasis on improving antigenicity and safety [5]. The progress addresses the limitations of traditional vaccine development by enhancing the precision of antigen selection and minimizing adverse immune reactions. The key steps in epitope-based vaccine design involve several essential processes: identifying potential antigens, analyzing protein structures to understand functional regions, predicting epitopes recognized by T cells and B cells, characterizing the immunogenicity and stability of these epitopes, and evaluating the interactions between proteins and epitopes at the molecular level to ensure efficacy. These steps allow researchers to design vaccines more likely to elicit protective immune responses while minimizing risks. The standard workflow for designing epitope vaccines using immunoinformatics approaches is illustrated in Figure 1, which highlights each stage of analysis and validation essential for creating targeted, safe, and effective vaccines.

2.1.1 Proper Selection of Pathogen Antigens

Choosing appropriate target antigens is critical as it influences the effectiveness of epitope prediction and ultimately determines the vaccine's ability to generate a robust immune response [6]. Selection should extend beyond conserved regions of structural proteins to include highly conserved areas of structural and non-structural proteins, increasing the likelihood of identifying potent antigenic and immunogenic epitopes. The process typically involves aligning homologous pathogen genome sequences to produce amino acid sequences in FASTA format, followed by antigenicity testing to select the most immunogenic proteins for further epitope prediction.

2.1.2 Epitope Prediction and Screening

Epitope prediction identifies T and B cell epitopes, which are pivotal in developing epitope vaccines [7]. Effective utilization of Immunoinformatics tools is essential for comprehensive epitope prediction. Moreover, predicting both T and B cell epitopes is advantageous as incorporating both into vaccine design enhances overall immunogenicity. After identifying cellular epitopes, further analyses using Immunoinformatics tools are necessary to select the most promising epitopes.

2.1.3 Vaccine Design and Assessment of the Structure and Relevance of Vaccine Constructs

The design of the vaccine construct significantly influences its immunogenic potential within the host [8]. Fusion proteins are typically synthesized by linking selected epitopes using specific linkers. Key considerations in linker design include the length, sequence, hydrophobicity, secondary structure, and protease sensitivity, as well as how these features impact the interaction between epitopes and immunogenic constructs. If the vaccine construct shows inadequate immunogenicity, incorporating suitable adjuvants can enhance immune responses. Careful linker and adjuvant design optimization is crucial in developing vaccines with strong and effective immunogenic profiles.

2.2 Common Validation Methods in Epitope Vaccine Development

The practical application of vaccine constructs derived from rational design and epitope prediction serves as the true standard for determining the success of vaccine design and development. Before clinical application, vaccines must undergo in vivo testing to evaluate both their immunogenicity and toxicity. An important consideration in preclinical studies is how to design immunological experiments that thoroughly assess the functionality and safety of the vaccine.

3 Immunoinformatics Tools Involved in Epitope Vaccine Development

3.1 Protein Sequence Retrieval

The primary structure of a protein dictates its higher-order structures, and understanding these structures depends on accurate analysis of protein sequences. The UniProt database (Universal Protein Resource) (https://www.uniprot.org/) is the most widely utilized protein database worldwide [9]. It includes three primary databases for protein sequences: The PIR International Protein Sequence Database (http://pir.georgetown.edu/) is maintained collaboratively by the Protein Information Resources (PIR), the Munich Information Center for Protein Sequences (MIPS), and the Japan International Protein Information Database (JIPID). It remains the largest publicly accessible protein sequence repository. The SWISS-PROT database [10] (http://www.ebi.ac.uk/swissprot/), a curated collection maintained by the European Bioinformatics Institute (EBI) (https://www.ebi.ac.uk/) and the Swiss Institute of Bioinformatics (SIB) (https://www.sib.swiss/). The TrEMBL database

(http://www.bioinfo.pte.hu/more/TrEMBL.htm), jointly managed by EBI and SIB, includes computationally annotated sequences encompassing all translated protein-coding nucleotide sequences from the EMBL nucleotide database [11].

3.2 Epitope Prediction

3.2.1 T-Cell Epitope Prediction

The prediction of T-cell epitopes is centered on identifying the shortest peptide fragments within antigens that possess immunogenic properties and can specifically stimulate CD4+ or CD8+ T cells. The process primarily relies on predicting the binding affinity between these peptide fragments and major histocompatibility complex (MHC) molecules, a critical factor in determining the strength of the immune response. Various computational tools have been developed to facilitate T-cell epitope prediction, with various online resources available for researchers. An overview of the key online tools commonly used for T-cell epitope prediction is presented in Table 1, highlighting their functionalities and strengths across various prediction models. These tools play a crucial role in refining epitope selection, thereby improving the accuracy and effectiveness of vaccine design.

3.2.2 B-Cell Epitope Prediction

B-cell epitopes are pivotal in the adaptive immune response, making them essential elements in vaccine design [12]. These epitopes are categorized as either linear or conformational. Linear B-cell epitopes consist of continuous sequences of amino acids and are often used as antigen substitutes in immunization and antibody production [13]. Their significance in immunotherapy has drawn considerable attention. In the body, B-cell antibodies are naturally selected to target external antigen proteins, with their corresponding binding sites referred to as conformational B-cell epitopes. These epitopes are critical for designing peptide- and protein-based vaccines and developing immunotherapeutic agents. Conformational epitope prediction typically requires the 3D structure of the protein, making it more complex than linear epitope prediction. Various online tools for B-cell epitope prediction are provided in Table 2.

3.3 Structural Assessment and Prediction of Vaccine Constructs

Once an optimal combination of epitopes and linkers has been designed, the vaccine construct's structure must undergo evaluation. Tools such as VaxiJen v2.0, AllerTOP v2.0, Protein Sol, ToxinPred, and ProtParam are commonly used to assess the antigenicity, allergenicity, solubility, toxicity, and physicochemical properties of the construct to ensure its effectiveness [14-16]. The structural prediction of vaccine constructs typically involves predicting both their secondary and tertiary structures, followed by structural optimization. The online tools commonly used for vaccine construct structural prediction are listed in Table 3.

3.4 Molecular Docking

Antigen molecules can elicit an immune response only when they bind to specific immune receptors in the host. Molecular docking is a computational method to predict non-covalent interactions between molecules, such as binding proteins and ligands [17, 18]. By simulating molecular interactions, the method helps researchers understand the binding patterns and sites between antigens and receptors, providing theoretical support and guidance for drug design and vaccine development. Molecular docking is widely used in the biomedical field, not only aiding in the understanding of immune response mechanisms but also contributing to the screening of potential drugs, optimization of molecular structures, and prediction of protein functions. Commonly used tools for molecular docking research are listed in Table 4. These tools have gained wide adoption in scientific studies due to their diverse computational approaches, algorithmic efficiency, and user-friendliness.

3.5 Molecular Dynamics Simulations and Immune Simulations

Molecular dynamics (MD) simulations are used to study the physical behavior of molecular docking complexes. The ABMER 20 software supports MD simulations to evaluate the stability of protein structures and the free binding energy between interacting molecules. C-ImmSim (http://150.146.2.1/C-IMMSIM/index.php) facilitates immune simulation studies, using position-specific scoring matrices for epitope prediction and machine learning models to assess immune interactions[19]. The system can simulate immune responses in various contexts, including HIV infection, hypersensitivity reactions, cancer immunotherapy, inflammation, and type 2 diabetes.

3.6 Codon Optimization and Computational Cloning

The Java Codon Adaptation Tool (JCat) server (http://www.jcat.de/) supports reverse translation and codon optimization for vaccine constructs, ensuring that CAI and GC content values fall within acceptable ranges [20]. NEBcutter (http://nc2.neb.com/NEBcutter2/) is then used to analyze the nucleotide sequences from JCat, and SnapGene (https://www.snapgene.com/) aids in virtual verification and restriction site-based cloning.

4. Application of Immune Informatics in Epitope Vaccine Design

4.1 Application in Bacterial Vaccine Development

In recent years, significant progress has been made in the design of multi-epitope vaccines (MEVs) with the advancement of immunoinformatics and reverse vaccinology. By using computational tools to predict T-cell and B-cell epitopes and incorporating various methods for vaccine construction, researchers have been able to enhance the immunogenicity and stability of

vaccines effectively. Numerous studies have already achieved important breakthroughs in developing vaccines against different bacteria pathogens.

Tuberculosis is a highly contagious disease caused by Mycobacterium tuberculosis (Mtb), which is one of the prominent reasons for the death of millions worldwide. Bibi *et al*., [21], utilized the IEDB to predict B-cell and T-cell epitopes from different Mycobacterium tuberculosis (Mtb) antigens. Subsequently, various servers docked these epitopes with toll-like receptor 3 (TLR-3). The constructed vaccine was refined based on allergenicity, antigenicity, solubility, various physicochemical properties, and molecular docking scores. Computer-based immune simulations demonstrated a significant response from immune cells. Guneswar *et al*., [22], applied immunoinformatics strategies to generate a multi-epitope immune response against Mtb, utilizing five antigenic proteins. Potential B-cell, cytotoxic T lymphocyte, and helper T lymphocyte epitopes were predicted from the target proteins and conjugated with the 50S ribosomal L7/L12 adjuvant to construct the vaccine. The physicochemical properties of the vaccine indicated antigenicity, solubility, and non-allergenicity. Additionally, docking, molecular dynamics simulations, and essential dynamics analysis revealed strong interactions between the multi-epitope vaccine structure and Toll-like receptors (TLR2 and TLR3). MM-PBSA analysis was conducted to determine the system's intermolecular binding free energy precisely. The immunoinformatics analysis indicated that the multi-epitope vaccine could induce a specific immune response, making it a potential candidate for combating Mtb.

In the development and application of vaccines for human and animal brucellosis, Li *et al*., [23], used various bioinformatics tools, including IEDB, NetCTLpan1.1, NetMHCIIpan4.0, ABCpred, and BCPREDS, to analyze the epitopes of Omp10, Omp25, Omp31, and BtpB. Epitope linking required linkers, where B-cell and CTL epitopes were connected with an AAY linker and HTL and CTL epitopes were connected with a GPGPG linker. The physicochemical properties of the multi-epitope vaccine were analyzed, and its tertiary structure was constructed, followed by molecular dynamics analysis of the MEV-TLR4 complex. Finally, animal experiments were conducted to validate the predicted results. The study provides theoretical and data support for further research and offers new methods and insights for developing MEV targeting *Brucella* in sheep.

Helicobacter pylori (*H. pylori*) is a Gram-negative spiral bacterium that infects half of the world's population and has been classified as a Type I carcinogen by the World Health Organization. Naz *et al*., [24], applied reverse vaccinology to predict and analyze surface antigen epitopes from all candidate proteins in the *H. pylori* genome, identifying potential epitopes such as vacA, babA, sabA, fecA, and omp16. The work lays the foundation for developing a multi-component *H. pylori* vaccine. Ma *et al*., [25], performed computational immunogenicity calculations on 381 protein sequences of *H. pylori* and tested the immunogenicity of selected top-scoring proteins. B-cell and T-cell epitopes from three well-performing protein (UreB, PLA1, and Omp6) were assembled randomly into six multi-epitope vaccine constructs. Structural stability and the solvent-accessible surface area (SASA) were analyzed to assess the exposure of B-cell epitopes on the vaccine surface and select the optimal construct. Construct S1 was ultimately chosen, and molecular docking demonstrated its potential to bind to TLR2, TLR4, and TLR9, suggesting it could elicit a strong immune response. Additionally, Anand *et al*., [26], modeled a multi-chimeric subunit vaccine candidate targeting *H. pylori* by screening for secreted/outer membrane proteins of the bacterium. They identified B-cell epitopes, MHC-II epitopes, and IFN-γ-inducing epitopes within these proteins. Population coverage, antigenicity, physicochemical properties, and secondary structure were evaluated using various computational tools, suggesting that the vaccine could be a potent and effective candidate. The 3D structure was predicted, refined, validated, and docked with TLRs. Molecular docking and immuno-simulation studies were conducted to verify the stability of interactions, and the epitope sequences were computationally cloned into the pET28b (+) plasmid vector. The study demonstrated that the designed multi-epitope antigens possess significant immunological potential, making them promising biomarker candidates for developing immunodiagnostic assays and vaccines targeting *H. pylori* infection.

4.2 Applications in Virus Vaccine Development

Significant progress has been made in the design of multi-epitope vaccines targeting various viruses and pathogens. Researchers have employed bioinformatics tools to predict T-cell and B-cell epitopes and have integrated multiple approaches to design novel vaccines aimed at enhancing immunogenicity and broadening the coverage of the vaccines.

The SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) is responsible for the COVID-19 outbreak. Srivastava *et al*., [27], identified antigenic regions from the ORFs of the novel coronavirus proteome using the immune epitope database and defined overlapping clusters of epitopes as antigen patches (Ag-patches). Using these Ag-patches, they developed multi-patch vaccines (MPVs), which cover more epitopes and HLA alleles than MEVs. The study identified 73 CTL epitopes and 49 HTL epitopes from the novel coronavirus proteome, with the Ag-patches containing 768 antigenic epitopes targeting 55 HLA alleles, covering 99.98% of the global population with high conservation. These epitopes also formed stable complexes with TLR3. The research highlights highly immunogenic antigen regions across the coronavirus proteome, making it a strong candidate for a new vaccine. Muhammad *et al*., [28], downloaded protein

sequences from GenBank and applied a combination of immunoinformatics and computational methods to predict B-cell and T-cell epitopes from highly antigenic structural proteins of SARS-CoV-2, aiming to design an effective multi-epitope subunit vaccine (MESV). The predicted epitopes exhibited high antigenicity, conservation, strong interactions with HLA alleles, and a global population coverage of 88.40%. By connecting 3 CTL, 6 HTL, and 4 B-cell epitopes with appropriate adjuvants and linkers, a 276-amino acid long MESV was designed. The MESV construct was non-allergenic, stable, and highly antigenic. Molecular docking showed that the MESV had a stable and high binding affinity with human pathogenic TLR3. Additionally, computational immune simulations revealed a significant immunogenic response induced by the MESV. Ashkan *et al*., [29], used immunoinformatics tools to select immunodominant regions of six non-structural SARS-CoV-2 proteins (nsp7, nsp8, nsp9, nsp10, nsp12, nsp14) to stimulate T-cell responses. Immunodominant fragments of the spike protein (residues 400-510) were also chosen to induce neutralizing antibodies. Selected regions were linked using RVRR linkers. β-defensin, an agonist of the TLR-4/MD complex, was added to the N-terminus via an (EAAAK) linker. The PADRE epitope was fused to the C-terminus using GPGPG and A(EAAAK)2A linkers. The construct showed immunogenicity, non-allergenicity, and non-functionality, with overlapping CD8+ and CD4+ epitopes. Molecular docking indicated stable interactions with TLR-4/MD. HLA class I and II coverage was 96.2% and 97.1%. Further experimental validation is needed despite high in silico efficacy.

The Orf virus (ORFV) is an acute, contact-transmissible, epitheliotropic, zoonotic, double-stranded DNA virus that causes significant economic losses in livestock. Pang *et al*., [30], selected multiple immunodominant CTL, HTL, and B-cell epitopes from the B2L, F1L, and 080 proteins of ORFV. They constructed a multi-epitope subunit vaccine by linking these epitopes with short linkers. The adjuvant β-defensin was added to the N-terminus of the vaccine via an EAAAK linker to enhance immunogenicity. The vaccine demonstrated good antigenicity and solubility and lacked allergenicity and toxicity. The 3D structure of the vaccine was subsequently predicted, optimized, and validated. The results indicated that the designed multi-epitope subunit vaccine exhibited high stability, could induce robust humoral and cellular immune responses, and is a promising ORFV vaccine candidate.

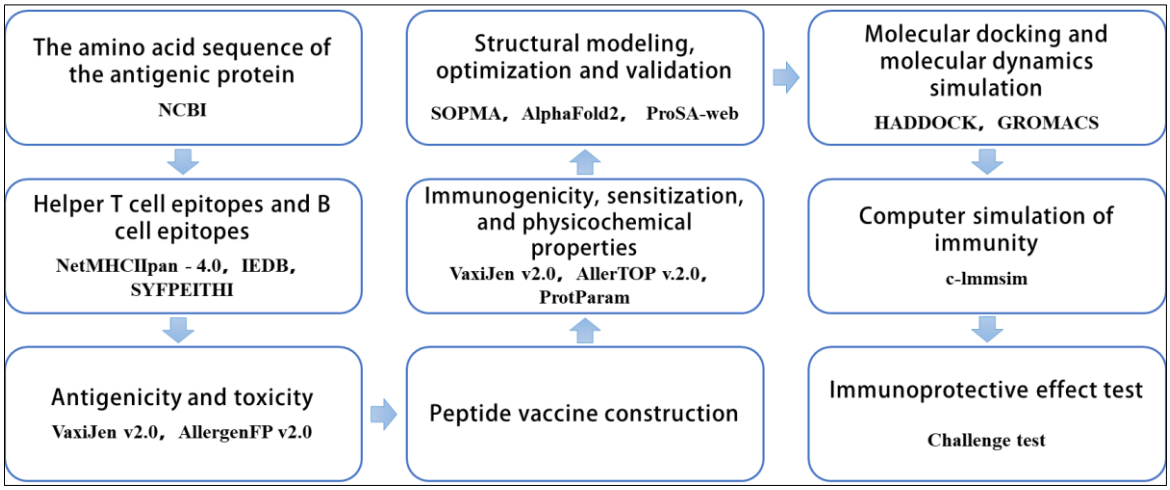

Fig. 1: Standard workflow for epitope vaccine design using immunoinformatics

Table 1: Online tools commonly used for T-cell epitope prediction

Tool	website	Description
NetMHCpan	http://www.cbs.dtu.dk/services/NetMHCpan/	The peptide binding capacity of multiple MHC alleles was predicted using artificial neural networks, focusing on class I MHC molecules.
IEDB MHC I Binding	http://www.cbs.dtu.dk/services/NetMHCpan/	Various methods, including artificial neural networks, steady-state epitope binding data, predicted the binding affinity of class I MHC molecules to peptides.
IEDB MHC II Binding	http://tools.iedb.org/mhcii/	Similar to the MHC I tool, which targets class II MHC molecules, the binding ability of peptides to MHC molecules is predicted.
SYFPEITHI	http://www.syfpeithi.de/	The binding affinity between peptides and MHC molecules is predicted based on a known library of MHC-binding peptides, using sequence information and binding site scoring.

NetMHCIIpan	http://www.cbs.dtu.dk/services/NetMHCIIpan/	Like NetMHCpan, but specifically focused on MHC class II molecules, machine learning models predict peptide binding affinity to MHC class II molecules.
TepiTool	http://tools.iedb.org/tepitool/	The IEDB prediction tool combines multiple algorithms to assess the binding affinity of peptides to both MHC class I and II molecules.
MHCflurry	https://openvax.github.io/mhcflurry/	Deep learning algorithms predict peptide binding affinity to MHC class I molecules, supporting personalized MHC allele analysis.
ProPred	http://www.imtech.res.in/raghava/propred/	The binding affinity between peptides and MHC molecules is evaluated using a matrix model, and antigenicity prediction is provided for MHC class II molecules.
PickPocket	http://www.cbs.dtu.dk/services/PickPocket/	Based on structural and sequence information, peptide binding to various MHC class I alleles is predicted through MHC binding pocket simulations.
EPIMHC	http://tools.immuneepitope.org/advanced_tools	Sequence information and statistical methods predict T-cell epitopes and peptide-MHC binding capacity.

Table 2: Online tools commonly used for B-cell epitope prediction

Tool	website	Description
BepiPred	https://services.healthtech.dtu.dk/services/BepiPred-3.0/	A widely used tool based on linear sequence features and machine learning algorithms, such as hidden Markov models, to predict epitopes. It can identify potential B-cell epitopes within protein sequences.
ABCpred	https://webs.iiitd.edu.in/raghava/abcpred/	Utilizes artificial neural network (ANN) algorithms to predict B-cell epitopes. Capable of predicting linear B-cell epitopes of a fixed length (16 amino acids) by assessing the antigenicity of the protein sequence to identify possible epitope regions.
BcePred	http://ailab-projects2.ist.psu.edu/bcpred/index.html	It uses physicochemical properties (such as hydrophobicity, flexibility, surface accessibility, etc.) to predict potential B-cell epitopes.
iBCE-EL	http://211.239.150.230/iBCE-EL/	Combines machine learning with physicochemical properties to predict linear B-cell epitopes. Employs multiple algorithms (such as support vector machines and random forests) to enhance prediction accuracy.
Ellipro	http://tools.iedb.org/ellipro/	It uses three-dimensional structural information to predict B-cell epitopes. Bases predictions on protein 3D models, considering surface accessibility and antigenicity to identify potential epitopes within protein structures.
DiscoTope	https://services.healthtech.dtu.dk/services/DiscoTope-3.0/	A structure-based epitope prediction tool. Combines surface data and amino acid characteristics to predict conformational B-cell epitopes.
EPMLR	https://maayanlab.cloud/datasets2tools/landing/tool/EPMLR	Uses multiple linear regression models for B-cell epitope prediction. Integrates amino acid composition and sequence information, offering accurate predictions for linear epitopes.
SVMTrip	http://sysbio.unl.edu/SVMTriP/	It uses a support vector machine (SVM) model to predict epitopes in tripeptide fragments. Combines sequence features and physicochemical properties to predict linear epitopes.

Table 3: Online tools commonly used for structural prediction in vaccine design

Tool	website	Description
VaxiJen v2.0	http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html	Sequence-based antigenicity prediction tools distinguish antigens from non-antigens by leveraging automatically generated physicochemical properties.
AllerTOP v2.0	https://www.ddg-pharmfac.net/AllerTOP/	Used for predicting whether a protein is an allergen, The tool employs self-organizing maps and multivariate analysis to classify proteins, utilizing amino acid composition and sequence features to assess allergenicity.
Protein-Sol	https://protein-sol.manchester.ac.uk/	It predicts protein solubility based on physicochemical properties derived from the protein sequence, including amino acid composition and polarity.
ToxinPred	http://crdd.osdd.net/raghava/toxinpred/	It uses machine learning models to evaluate toxicity based on amino acid sequence features, physicochemical properties, and statistical methods to predict protein toxicity and design low-toxicity mutants.
ProtParam	https://web.expasy.org/protparam/	It calculates various physicochemical properties of protein sequences, including molecular weight, isoelectric point, amino acid composition, and peptide bond count, primarily in basic protein characterization.

Table 4: Online tools commonly used for molecular docking

Tool	website	Description
AutoDock	autodock.scripps.edu	It uses the Lamarckian Genetic Algorithm to predict binding affinities.
Vina	vina.scripps.edu	The improved version of AutoDock employs an iterated local search method.
DOCK	dock.compbio.ucsf.edu	It uses shape-based docking to match ligand shape with protein binding site.
MOE (Molecular Operating Environment)	chemcomp.com/MOE-Molecular_Operating_Environment.htm	It uses a force-field-based approach to predict binding energies.
GOLD (Genetic Optimization for Ligand Docking)	ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/	Utilizes genetic algorithms to explore binding conformations.
Glide	schrodinger.com/glide	It uses a hierarchical search method to predict ligand binding to proteins.
SwissDock	swissdock.ch	Based on the EADock algorithm, it uses energy minimization and binding free energy prediction.
FlexX	biosolveit.de/FlexX	Utilizes incremental construction of ligands to optimize docking results.

5. Outlook

Immunoinformatics methods and tools are extensively employed in vaccine development to analyze conserved sequences and predict potential epitopes. The development of novel vaccines utilizing epitope-based antigens is anticipated to be a major advancement focus. The approach markedly reduces costs, accelerates development timelines, and lessens the experimental workload involved in epitope screening, presenting a promising strategy for epitope discovery. As immunoinformatics tools and recombinant DNA technology continue to evolve, the design and production of epitope vaccines will become increasingly sophisticated. Nevertheless, many vaccines remain unvalidated in clinical settings despite the extensive application of immunoinformatics-driven strategies for predicting and constructing epitope vaccines against viruses, bacteria, and other pathogens. A significant challenge in vaccine development is the variability in antigens and genetic polymorphism, which are crucial for precise vaccine design and assessment. Vaccines based on a single or limited number of epitopes might also demonstrate reduced immune efficacy. Therefore, it is critical to utilize immunoinformatics tools to identify highly immunogenic epitope combinations and select appropriate adjuvants and linkers to enhance vaccine effectiveness. Current research in epitope vaccines is focused on addressing these challenges. Thus, further research is needed to optimize the integration of immunoinformatics with other immunological experimental methods in the design of epitope vaccines.

Acknowledgment

The work was supported by the school-directed Training (Introduction of talents) Scientific Research start-up Fund project (XYB202318), the Guiding Science and Technology Program of Daqing City (zd-2024-38), and Heilongjiang University Students Innovation and Entrepreneurship Training Program project (S202410223008).

REFERENCES

1. Lancet, T. (2020). Zoonoses: beyond the human-animal-environment interface. *Lancet (London, England)*, *396*(10243), 1.

2. Tripathi, T. (2023). Advances in vaccines: revolutionizing disease prevention. *Scientific Reports*, *13*(1), 11748.
3. Bravi, B. (2024). Development and use of machine learning algorithms in vaccine target selection. *npj Vaccines*, *9*(1), 15.
4. Tomar, N., & De, R. K. (2014). Immunoinformatics: a brief review. *Immunoinformatics*, 23-55.
5. Oli, A. N., Obialor, W. O., Ifeanyichukwu, M. O., Odimegwu, D. C., Okoyeh, J. N., Emechebe, G. O., ... & Ibeanu, G. C. (2020). Immunoinformatics and vaccine development: an overview. *ImmunoTargets and therapy*, 13-30.
6. Soria-Guerra, R. E., Nieto-Gomez, R., Govea-Alonso, D. O., & Rosales-Mendoza, S. (2015). An overview of bioinformatics tools for epitope prediction: implications on vaccine development. *Journal of biomedical informatics*, *53*, 405-414.
7. Sanchez-Trincado, J. L., Gomez-Perosanz, M., & Reche, P. A. (2017). Fundamentals and methods for T-and B-cell epitope prediction. *Journal of immunology research*, *2017*(1), 2680160.
8. Meyer, C. U., & Zepp, F. (2022). Principles in Immunology for the Design and Development of Vaccines. *Vaccine Design: Methods and Protocols, Volume 1. Vaccines for Human Diseases*, 27-56.
9. Bowler-Barnett, E. H., Fan, J., Luo, J., Magrane, M., Martin, M. J., & Orchard, S. (2023). UniProt and mass spectrometry-based proteomics—a 2-way working relationship. *Molecular & Cellular Proteomics*, *22*(8).
10. Gasteiger, E., Jung, E., & Bairoch, A. (2001). SWISS-PROT: connecting biomolecular knowledge via a protein database. *Current issues in molecular biology*, *3*(3), 47-55.
11. Calderón-González, K. G., Hernández-Monge, J., Herrera-Aguirre, M. E., & Luna-Arias, J. P. (2016). Bioinformatics tools for proteomics data interpretation. *Modern Proteomics–Sample Preparation, Analysis and Practical Applications*, 281-341.
12. Zheng, D., Liang, S., & Zhang, C. (2023). B-Cell Epitope Predictions Using Computational Methods, *Methods Mol Biol*, *2552*, 239-254.
13. Galanis, K. A., Nastou, K. C., Papandreou, N. C., Petichakis, G. N., Pigis, D. G., & Iconomidou, V. A. (2021). Linear B-cell epitope prediction for in silico vaccine design: A performance review of methods available via command-line interface. *International journal of molecular sciences*, *22*(6), 3210.
14. Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC bioinformatics*, *8*, 1-7.
15. Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014). AllerTOP v. 2—a server for in silico prediction of allergens. *Journal of molecular modeling*, *20*, 1-6.
16. Khan, M. T., Islam, M. J., Parihar, A., Islam, R., Jerin, T. J., Dhote, R., ... & Halim, M. A. (2021). Immunoinformatics and molecular modeling approach to design universal multi-epitope vaccine for SARS-CoV-2. *Informatics in medicine unlocked*, *24*, 100578.
17. Pishesha, N., Harmand, T. J., & Ploegh, H. L. (2022). A guide to antigen processing and presentation. *Nature Reviews Immunology*, *22*(12), 751-764.
18. Rawal, K., Sinha, R., Abbasi, B. A., Chaudhary, A., Nath, S. K., Kumari, P., ... & Bottazzi, M. E. (2021). Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. *Scientific reports*, *11*(1), 17626.
19. Rapin, N., Lund, O., & Castiglione, F. (2011). Immune system simulation online. *Bioinformatics*, *27*(14), 2013-2014.
20. Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C., & Jahn, D. (2005). JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic acids research*, *33*(suppl_2), W526-W531.
21. Bibi, S., Ullah, I., Zhu, B., Adnan, M., Liaqat, R., Kong, W. B., & Niu, S. (2021). In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. *Scientific reports*, *11*(1), 1249.
22. Sethi, G., Varghese, R. P., Lakra, A. K., Nayak, S. S., Krishna, R., & Hwang, J. H. (2024). Immunoinformatics and structural aided approach to develop multi-epitope based subunit vaccine against Mycobacterium tuberculosis. *Scientific Reports*, *14*(1), 15923.
23. Li, M., Zhu, Y., Niu, C., Xie, X., Haimiti, G., Guo, W., ... & Zhang, F. (2022). Design of a multi-epitope vaccine candidate against Brucella melitensis. *Scientific Reports*, *12*(1), 10146.
24. Naz, A., Awan, F. M., Obaid, A., Muhammad, S. A., Paracha, R. Z., Ahmad, J., & Ali, A. (2015). Identification of putative vaccine candidates against Helicobacter pylori exploiting exoproteome and secretome: a reverse vaccinology based approach. *Infection, Genetics and Evolution*, *32*, 280-291.
25. Ma, J., Qiu, J., Wang, S., Ji, Q., Xu, D., Wang, H., ... & Liu, Q. (2021). A novel design of multi-epitope vaccine against Helicobacter pylori by immunoinformatics approach. *International journal of peptide research and therapeutics*, *27*, 1027-1042.
26. Keshri, A. K., Kaur, R., Rawat, S. S., Arora, N., Pandey, R. K., Kumbhar, B. V., ... & Prasad, A. (2023). Designing and development of multi-epitope chimeric vaccine against Helicobacter pylori by exploring its entire immunogenic epitopes: an immunoinformatic approach. *BMC bioinformatics*, *24*(1), 358.
27. Srivastava, S., Verma, S., Kamthania, M., Agarwal, D., Saxena, A. K., Kolbe, M., ... & Pandey, K. C.

(2022). Computationally validated SARS-CoV-2 CTL and HTL Multi-Patch vaccines, designed by reverse epitomics approach, show potential to cover large ethnically distributed human population worldwide. *Journal of Biomolecular Structure and Dynamics*, *40*(5), 2369-2388.

28. Tahir ul Qamar, M., Shahid, F., Aslam, S., Ashfaq, U. A., Aslam, S., Fatima, I., ... & Chen, L. L. (2020). Reverse vaccinology assisted designing of multiepitope-based subunit vaccine against SARS-CoV-2. *Infectious diseases of poverty*, *9*, 1-14.
29. Safavi, A., Kefayat, A., Mahdevar, E., Abiri, A., & Ghahremani, F. (2020). Exploring the out of sight antigens of SARS-CoV-2 to design a candidate multi-epitope vaccine by utilizing immunoinformatics approaches. *Vaccine*, *38*(48), 7612-7628.
30. Pang, F., Long, Q., & Liang, S. (2024). Designing a multi-epitope subunit vaccine against Orf virus using molecular docking and molecular dynamics. *Virulence*, *15*(1), 2398171.