Manuscript title

VaccineDesigner: A Web-based Tool for Streamlined Epitope-based Vaccine Design

Authors

Dimitrios Trygoniaris[1,*], Anna Korda[2], Anastasia Paraskeva[2], Esmeralda Dushku[3], Georgios Tzimagiorgis[1,4], Minas Yiangou[2,4], Charalampos Kotzamanidis[3] and Andigoni Malousi[1,4,*]

[1] School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece
[2]Department of Genetics, Development & Molecular Biology, School of Biology, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece
[3]Veterinary Research Institute of Thessaloniki, Campus of Thermi, Thermi, 57001, Greece
[4] Genomics and Epigenomics Translational Research Group, Center for Interdisciplinary Research and Innovation, Thessaloniki, 57001, Greece

* To whom correspondence should be addressed. Tel: [+302310999163; Fax: [+302313092347; Email: [dtrygoni@auth.gr]

Present Address: [Dimitrios Trygoniaris], School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, 54142, Greece

Abstract

Epitope-based vaccine design is a promising alternative to conventional methods, focusing on selected antigenic epitopes and molecular fragments that can interact with the immune system and elicit appropriate immune responses. Computational approaches are integral parts of the epitope-based vaccine design, implementing a wide spectrum of immune-related analyses, through distributed and heterogeneous interfaces that are technically hard to orchestrate into workflows. To address this challenge, we developed a comprehensive platform, called VaccineDesigner, that streamlines the design of epitope-based vaccines by seamlessly integrating methods for B-cell, CTL, and HTL epitope prediction. VaccineDesigner incorporates additional functionalities, such as multi-epitope vaccine generation, estimation of population coverage, molecular mimicry, and proteasome cleavage that are transparently integrated into the modular architecture, providing a single-access point for rationalized and time/cost-effective multi-epitope vaccines. The source code is freely available under the academic license at: https://github.com/BioIApps/VaccineDesigner and the Web-based interface is accessible at: http://bioinformatics.med.auth.gr/VaccineDesigner

Introduction

Developing effective vaccines against pathogens is a primary research priority in immunology and public health. Vaccination, a cornerstone of preventive medicine, has historically transformed global healthcare and reduced disease burdens by saving lives[1]. Conventional vaccine development relies on live attenuated pathogens, inactivated agents, subunit formulations, or recombinant protein antigens, facing challenges, such as pathogen cultivation and safety concerns[2]. Epitope-based vaccine design offers a promising alternative, focusing on the identification of antigenic epitopes and molecular fragments that engage with the immune system, including B-cell epitopes for humoral responses, cytotoxic T lymphocyte epitopes for cellular immunity, and helper T lymphocyte epitopes for immune regulation[3]. B-cell epitopes consist of surface-accessible clusters of amino acids and activate immune responses through antibody production[4]. T-cell epitopes consist of small peptide fragments and activate immune responses by presentation to antigen-presenting cells (APCs) through MHC class I or class II molecules. Cytotoxic T cells respond to MHC class I-restricted peptides, called CTL epitopes,

whereas Helper T cells target MHC class II-restricted peptides, referred to as HTL epitopes[5]. Vaccines combining these components have the unique ability to simultaneously activate both humoral and cellular immunity while avoiding the presence of elements that could provoke harmful reactions[6].

Computational approaches have revolutionized our ability to predict and strategically harness epitopes. During the past decade, plenty of methods have been developed across various facets of the immune response, encompassing B-cell, CTL, and HTL epitope prediction[7].A typical analysis for the prediction of epitopes includes a series of computational steps that are implemented by complementary tools and are successively executed, while in several cases, the analyses can be executed in parallel. To further reinforce the quality of a prediction it is generally recommended to combine the outputs of multiple tools performing the same type of analysis. Consequently, researchers face the challenge to infiltrate information from multiple sources that is overall hard to manageable and difficult to assess. Despite the obvious advances of an automated workflow, the orchestration of relevant tools is technically challenging as most of them are accessible through heterogeneous interfaces, lacking interoperability features.

Herein, we present VaccineDesigner, a Web-based tool that aims to provide a flexible framework for rational epitope-based vaccine design. VaccineDesigner integrated computational methods for B-cell, CTL, and HTL epitope prediction and incorporates additional functionalities, such as multi-epitope vaccine generation, estimation of population coverage, molecular mimicry, and proteasome cleavage that are transparently integrated into the modular architecture.

Methods

The main components of the workflow involve epitope prediction for B-cell and T-cell lymphocytes, as crucial indicators of protein antigenicity[8] (Fig. 1). The application leverages Bepipred for B-cell epitope prediction[9], NetMHCpan for MHC class I and II[10], or Consensus IEDB tool for binding prediction[11]. The predicted epitopes are further assessed for their antigenicity, toxicity, and allergenicity, according to user-defined criteria, as shown below.

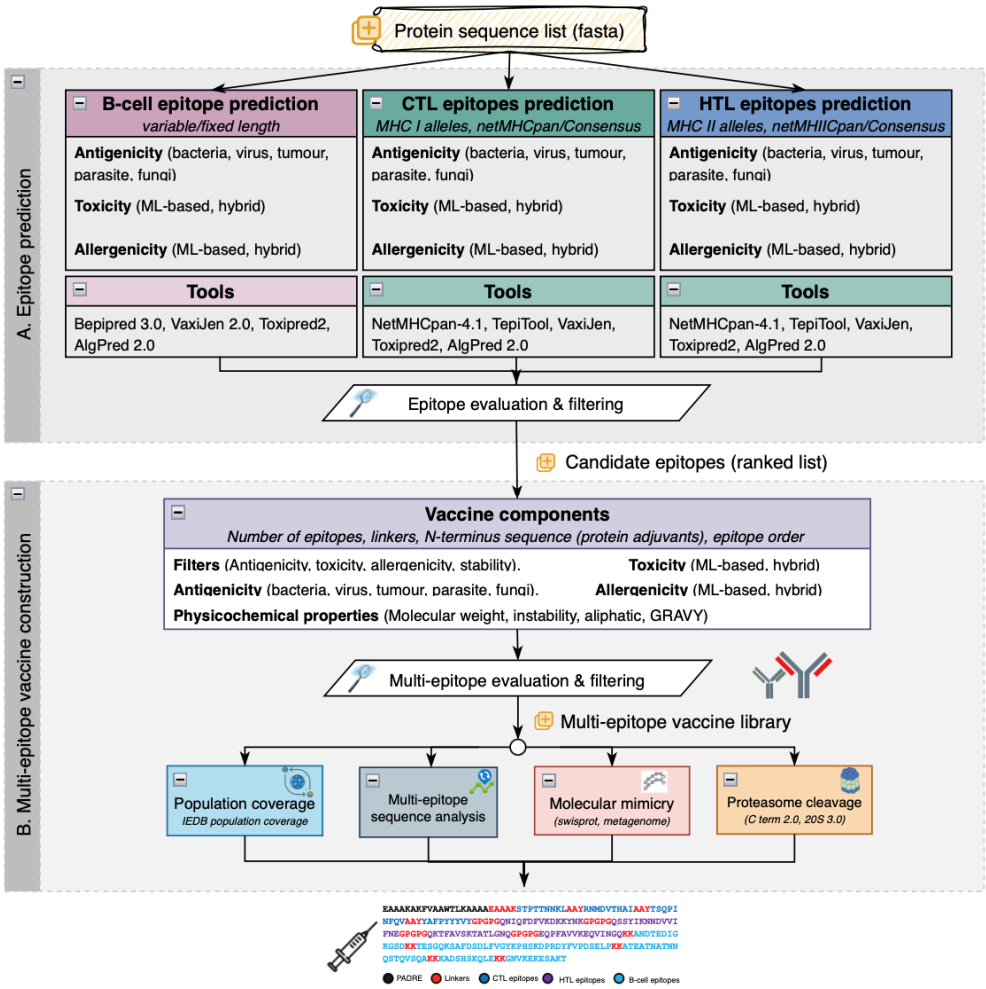

Figure 1: The main functional components of the streamlined process implemented by VaccineDesigner. **A.** The *Epitope prediction module* starts with the upload of a fasta protein file, then performs epitope prediction (B cell, CTL, HTL), and ends up with the evaluation and filtering of epitope sequences, based on their physicochemical properties. **B.** The *Multi-epitope sequence generation module* takes as input the tabular-formatted results containing the candidate epitopes and generates vaccine sequences based on user-defined parameters. To assess the quality of the candidate vaccine sequences the users can evaluate a vaccine sequence of their preference for population coverage, molecular mimicry, proteasome cleavages, and re-analyse the construct using the supported epitope prediction modules.

Epitope Prediction and Evaluation

VaccineDesigner initiates the process with a set of user-selected protein sequences (Fig. 1). In order to proceed with the epitope prediction, the users are asked to select a protein of interest. The analysis is initiated as an integral component of epitope-based vaccine design that harnesses well-established algorithms, such as BepiPred 3.0, NetMHCpan, and NetMHCIIpan, to serve various epitope prediction purposes.

B-cell epitope prediction. B-cell epitope prediction uses BepiPred 3.0. BepiPred systematically scans protein sequences and assesses key physicochemical properties and sequence patterns to identify regions that can induce robust antibody responses[9]. VaccineDesigner uses the amino acid scores and generates B-cell epitopes, adhering to user-defined epitope length parameters. Two types of analysis can be deployed. The first type involves the formation of B-cell epitopes in high-scoring regions. The lower length limit and the maximum number of amino acids between the predicted epitopes are

customizable, enabling the merge of separate epitopes into larger fragments. To avoid the selection of low-scoring amino acids, VaccineDesigner includes a user-defined threshold for the minimum amino acid score. The second type of analysis pertains to the generation of B-cell epitopes of user-defined lengths. In this analysis, a fixed-length scanning window is defined to assess the scores of all amino acids. Epitopes are defined as continuous regions of nice or more amino acids, wherein each position must fulfill a certain filtering threshold.

CTL epitope prediction. VaccineDesigner performs CTL epitope prediction using two methods, NetMHCpan 4.1[10] and the Consensus Method by IEDB Class I prediction[11]. NetMHCpan is an algorithm renowned for its accuracy in identifying high-affinity regions for MHC class I alleles. The output includes both strong and weak binders counts, along with information about the associated MHC alleles. On the other hand, the IEDB Consensus method stands as a reputable algorithm in forecasting CTL epitopes. This method seamlessly integrates various bioinformatics tools, harnessing their combined capabilities to elevate the accuracy of epitope prediction. The CTL epitopes are ranked based on the number of alleles interacting with each respective epitope.

HTL epitope prediction. The HTL epitope prediction module is implemented based on the NetMHCIIpan 4.0 framework. NetMHCIIpan identifies regions characterized by strong binding affinity for MHC class II molecules[10]. The results, akin to CTL epitopes, include strong and weak binders counts, along with details of the MHC class II alleles. Additionally, HTL epitope prediction can be executed using the IEDB Consensus method for class II binding prediction, employing the same framework as for the Class I module[11]. HTL epitopes are ranked based on the number of interacting alleles, following the approach used for CTL epitopes.

Given a set of candidate B-cell, CTL and HTL epitopes, VaccineDesigner implements a quality control step that ensures the meticulous selection of epitopes with exceptional antigenicity and safety profiles, adhering to user-defined thresholds. VaxiJen[12] is used to assess the antigenicity of each epitope, prioritizing highly immunogenic candidates. ToxinPred[13] identifies potential toxins within epitope sequences, serving as a crucial safety measure and AlgPred[14] predicts allergenicity, minimizing the risk of allergic reactions in vaccine recipients. Epitopes that meet high-quality criteria for antigenicity, toxicity, and allergenicity are considered more likely to improve the overall efficacy and safety of multi-epitope constructs.

Multi-epitope Vaccine Sequences

The synthesis of multi-epitope constructs is considered a more favorable approach to mitigate junctional immunogenicity and efficient epitope separation and presentation to T-cell and B-cell receptors[3]. VaccineDesigner elaborates further on the candidate epitopes to craft multi-epitope vaccine sequences. Using the most prominent epitopes, VaccineDesigner builds larger constructs by merging individual epitopes with appropriate linker sequences (default: GPGPG/HTL, AAY/CTL, KK/B-cell)[3,15-20]. The analysis is implemented in a two-step procedure:

Vaccine Sequences Construction. Users are prompted to define the number B-cell, CTL, and HTL epitopes to be included in the construct, while also specifying user-defined linker sequences[17] between the epitopes and the N-terminus sequence. A variety of different options of protein adjuvants are included for the synthesis of the multi-epitope vaccine sequences[21-24]. Additionally, the order of epitope combination can be defined, enabling further customization, such as arranging B-cell, CTL, and HTL epitopes in a specific sequence, exemplified as B-C-H (B cell - CTL - HTL). VaccineDesigner generates candidate vaccine sequences, that form a versatile library of epitope sequence combinations. Researchers are free to explore diverse epitope combinations, with a single constraint that is the maximum number (up to five) of epitopes from each category (B-cell, CTL, and HTL).

Vaccine Sequences Evaluation and Selection. In the final step, the constructed candidate vaccine sequence library undergoes rigorous evaluation. The algorithm assesses each sequence against predefined filters, including VaxiJen for antigenicity, ToxinPred for toxin prediction, Algpred for allergenicity, and ProtParam for sequence analysis and stability[25].

User-defined parameters, such as the desired number of vaccine sequences, serve as thresholds that are used to keep track of the number of candidates that meet all quality criteria. The algorithm exports the final multi-epitope sequence constructs once the preferred number of candidates is reached. This dynamic approach ensures the selection of vaccine sequences that meet strict standards of immunogenicity, safety, and biochemical attributes, offering researchers the most promising candidates for further development and validation. The final set of candidates is ranked based on the cumulative sum of individual antigenicity, toxicity, allergenicity, and stability rankings.

The newly generated sequences can be further examined for the population coverage, molecular mimicry, proteasome cleavages, while re-analysis for epitope prediction is also supported. Population coverage analysis is enabled for human, using the IEDB Population Coverage[26] algorithm. To address molecular mimicry users can conduct protein similarity searches against SwissProt[27] or metagenomic proteins (env_nr) using the BLASTp[28] algorithm. The analysis estimates whether the vaccine sequence shares similar domains with a known host protein, thereby ensuring its antigenicity and minimizing the risk of autoimmune reactions. Furthermore, proteasome cleavages can be identified on the protein sequences using NetChop 3.1[29]. The results are provided in both tabular format and graphical representations as shown in Figure 2A.

A multi-epitope vaccine sequence can be further examined for the presence of epitopes, including both those already predicted by also new ones. Figure 2B includes an example visual representation, depicting new epitopes on the final sequence construct.

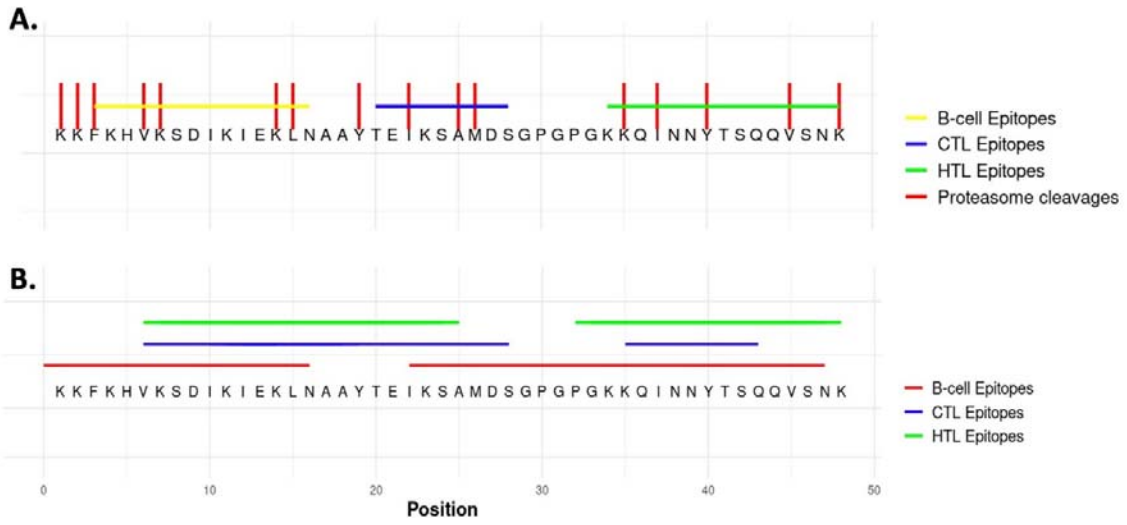

Figure 2: Visualization of multi-epitope vaccine based on the epitope prediction module (A), and the proteasome cleavage functionality (B).

A usage scenario for Staphylococcus aureus pathogens

To demonstrate the functionalities of the overall workflow, we implemented an example scenario based on the *Atl* (Bifunctional Autolysin, UniProt: P0C5Z8) and the *IsdA* (Iron-regulated surface

determinant protein A, UniProt: Q7A152) proteins of *Staphylococcus aureus*. 123 B-cell epitopes were predicted on the *Atl* protein and 136 on the *IsdA* using a fixed sequence length of 10 amino acids. 13 of these epitopes were considered high-quality candidates based on their quantified toxicity, allergenicity, and antigenicity measures. 96 and 18 CTL epitopes were identified in *Atl* and *IsdA,* respectively. These were predicted as strong or weak binders to MHC class I alleles, including *HLA-A0101, HLA-A0102, HLA-A0103, HLA-A0104, HLA-A0106, HLA-A0107, HLA-A0108* and *HLA-A0109*. The predicted epitopes underwent toxicity, allergenicity, and antigenicity filtering, resulting in six favourable candidates. The number of HTL epitopes that were considered strong or weak binders to MHC class II alleles were 193 from *Atl* and 50 from *IsdA* in total, including *DBR1_0101*, *DBR1_0102*, *DBR1_0103*, *DBR1_0104*, *DBR1_0105*, *DBR1_0106*, *DBR1_0107* and *DBR1*_0108. A significant number of epitopes was disqualified due to the high allergenicity and toxicity, resulting in 19 high-quality HTL epitopes. To construct multi-epitope vaccines, candidates were combined following the C-H-B order (CTL, HTL, and B-cell). Four B-cell epitopes and three CTL and HTL epitopes were used for the synthesis of the constructs resulting in 864 vaccines sequences. Table 1a and 1b list the best scoring sequences and the quantified measures for antigenicity, toxicity, and relative properties, respectively.

Table 1: (a) The top five multi-epitope sequences include a shared domain of 98 amino acids and a variable region colored by the amino acid polarity. **(b)**: Antigenicity and toxicity alongside their antigenicity, toxicity, allergenicity, instability, and GRAVY scores of the top five multi-epitope sequence candidates.

(a)

Shared domain	AAYQVNSSINDYAAYVSDNKSQQTAAYYLRSHNYSYGPGPGINGEISY MKNNYQNAGPGPGMDDYMQHPGKVIKQNGPGPGNGEISYMKNNYQNAF
Alignment of the variable multi-epitope domain	KKSTSTTAPKTNKKSDNKSQQTNKKKDTRANQSATTKKVSDNKSQQTN---------- KKDTRANQSATTKKSDNKSQQTNKKKSTSTTAPKTNKKVSDNKSQQTN---------- KKDTRANQSATTKKSTSTTAPKTNKKSDNKSQQTNKKKVSDNKSQQTN---------- KKSTSTTAPKTNKKDTRANQSATTKKSDNKSQQTNKKKVSDNKSQQTN---------- KK----------DTRANQSATTKKSDNKSQQTNKKKVSDNKSQQTNKKSTSTTAPKTN

(b)

Seq. ID	Antigenicity	Toxicity	Allergenicity	Instability	GRAVY
1	1.0539	0.245	0.262	38.62	-1.393
2	1.0527	0.245	0.262	38.62	-1.393
3	1.0408	0.245	0.262	38.62	-1.393
4	1.0443	0.245	0.262	38.62	-1.393
5	1.06	0.245	0.262	38.62	-1.393

Discussion

In the evolving landscape of infectious diseases, epitope-based vaccine design has emerged as a pivotal strategy, offering rationalized detection of potent immune responses. Epitope-based methods focus on the identification of antigenic determinants in pathogenic proteins, enabling the development of vaccines that reduce the risk of adverse reactions and ensure broad coverage across various pathogenic strains. VaccineDesigner integrates a broad spectrum of analyses for B-cell, CTL, and HTL epitope and generates reliable multi-epitope vaccine candidates. Compared to existing approaches, VaccineDesigner differs in both the depth and coverage of the supported functionalities. Reverse vaccinology platforms such as Vaxign2[30] focus on predicting vaccine candidates from pathogen genomes, while VaccineDesigner supports the analysis of user-defined target proteins, facilitating the prediction of immunogenic regions for tailored vaccine development. Compared to other approaches such as iVAX Suite[31], Vacceed[32] and the Immune Epitope Database and Analysis

Resource (IEDB)[33]. VaccineDesigner combines favorable features, such as comprehensive epitope assessment, based on a broad array of alleles, and technical advantages such as transparent execution of cascading tasks and user-friendly access that eliminates the need for local software installations and special requirements for computer resources. As our knowledge of the complex host-pathogen interactions evolves, we anticipate that VaccineDesigner will contribute to building real-world applications and will be used as a generic framework for the development of safer and more effective vaccines, through rationalized knowledge synthesis.

Data availability

VaccineDesigner is an open-source tool freely available under the academic license at https://github.com/BiolApps/VaccineDesigner[34]

Competing interests

The authors declare that they have no competing interests.

References

1. Thomas S, Luxon BA. Vaccines based on structure-based design provide protection against infectious diseases. Expert Rev Vaccines. 2013;12(11):1301-1311. doi:10.1586/14760584.2013.840092

2. Webster RG, Govorkova EA. Continuing challenges in influenza. Ann N Y Acad Sci. 2014;1323(1):115-139. doi:10.1111/nyas.12462

3. Ayyagari, V. S., T C, V., K, A. P., and Srirama, K. (2022). Design of a multi-epitope-based vaccine targeting M-protein of SARS-CoV2: an immunoinformatics approach. Journal of biomolecular structure and dynamics, 40(7), 2963–2977.

4. Potocnakova, L., Bhide, M., and Pulzova, L. B. (2016). An Introduction to B-Cell Epitope Mapping and In Silico Epitope Prediction. Journal of immunology research, 2016, 6760830.

5. Shepherd FR, McLaren JE. T Cell Immunity to Bacterial Pathogens: Mechanisms of Immune Control and Bacterial Evasion. Int J Mol Sci. 2020;21(17):6144. Published 2020 Aug 26. doi:10.3390/ijms21176144

6. Zhang L. Multi-epitope vaccines: a promising strategy against tumors and viral infections. Cell Mol Immunol. 2018;15(2):182-184. doi:10.1038/cmi.2017.92

7. Feng Y, Jiang H, Qiu M, et al. Multi-Epitope Vaccine Design Using an Immunoinformatic Approach for SARS-CoV-2. Pathogens. 2021;10(6):737. Published 2021 Jun 11. doi:10.3390/pathogens10060737

8. Raoufi E, Hemmati M, Eftekhari S, et al. Epitope Prediction by Novel Immunoinformatics Approach: A State-of-the-art Review. Int J Pept Res Ther. 2020;26(2):1155-1163. doi:10.1007/s10989-019-09918-z

9. Clifford JN, Høie MH, Deleuran S, Peters B, Nielsen M, Marcatili P. BepiPred-3.0: Improved B-cell epitope prediction using protein language models. Protein Sci. 2022;31(12):e4497. doi:10.1002/pro.4497

10. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. 2020;48(W1):W449-W454. doi:10.1093/nar/gkaa379

11. Paul S, Sidney J, Sette A, Peters B. TepiTool: A Pipeline for Computational Prediction of T Cell Epitope Candidates. Curr Protoc Immunol. 2016;114:18.19.1-18.19.24. Published 2016 Aug 1. doi:10.1002/cpim.12

12. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8:4. Published 2007 Jan 5. doi:10.1186/1471-2105-8-4

13. Sharma N, Naorem LD, Jain S, Raghava GPS. ToxinPred2: an improved method for predicting toxicity of proteins. Brief Bioinform. 2022;23(5):bbac174. doi:10.1093/bib/bbac174

14. Sharma N, Patiyal S, Dhall A, Pande A, Arora C, Raghava GPS. AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes. Brief Bioinform. 2021;22(4):bbaa294. doi:10.1093/bib/bbaa294

15. Yang Y, Sun W, Guo J, et al. In silico design of a DNA-based HIV-1 multi-epitope vaccine for Chinese populations. Hum Vaccin Immunother. 2015;11(3):795-805. doi:10.1080/21645515.2015.1012017

16. Yano A, Onozuka A, Asahi-Ozaki Y, et al. An ingenious design for peptide vaccines. Vaccine. 2005;23(17-18):2322-2326. doi:10.1016/j.vaccine.2005.01.031

17. Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev. 2013;65(10):1357-1369. doi:10.1016/j.addr.2012.09.039

18. Livingston B, Crimi C, Newman M, et al. A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes. J Immunol. 2002;168(11):5499-5506. doi:10.4049/jimmunol.168.11.5499

19. Li X, Guo L, Kong M, et al. Design and Evaluation of a Multi-Epitope Peptide of Human Metapneumovirus. Intervirology. 2015;58(6):403-412. doi:10.1159/000445059

20. Willimsky G, Beier C, Immisch L, et al. In vitro proteasome processing of neo-splicetopes does not predict their presentation in vivo. Elife. 2021;10:e62019. Published 2021 Apr 20. doi:10.7554/eLife.62019

21. Díaz-Dinamarca DA, Salazar ML, Castillo BN, et al. Protein-Based Adjuvants for Vaccines as Immunomodulators of the Innate and Adaptive Immune Response: Current Knowledge, Challenges, and Future Opportunities. Pharmaceutics. 2022;14(8):1671. Published 2022 Aug 11. doi:10.3390/pharmaceutics14081671

22. Toussi DN, Carraway M, Wetzler LM, Lewis LA, Liu X, Massari P. The amino acid sequence of Neisseria lactamica PorB surface-exposed loops influences Toll-like receptor 2-dependent cell activation. Infect Immun. 2012;80(10):3417-3428. doi:10.1128/IAI.00683-12

23. Ghaffari-Nazari H, Tavakkol-Afshari J, Jaafari MR, Tahaghoghi-Hajghorbani S, Masoumi E, Jalali SA. Improving Multi-Epitope Long Peptide Vaccine Potency by Using a Strategy that Enhances CD4+ T Help in BALB/c Mice. PLoS One. 2015;10(11):e0142563. Published 2015 Nov 10. doi:10.1371/journal.pone.0142563

24. Del Giudice G. Hsp70: a carrier molecule with built-in adjuvanticity. Experientia. 1994;50(11-12):1061-1066. doi:10.1007/BF01923462

25. Wilkins MR, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol. 1999;112:531-552. doi:10.1385/1-59259-584-7:531

26. Dhanda SK, Mahajan S, Paul S, et al. IEDB-AR: immune epitope database-analysis resource in 2019. Nucleic Acids Res. 2019;47(W1):W502-W506. doi:10.1093/nar/gkz452

27. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res. 2000;28(1):45-48. doi:10.1093/nar/28.1.45

28. McGinnis S, Madden TL. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res. 2004;32(Web Server issue):W20-W25. doi:10.1093/nar/gkh435

29. Saxová P, Buus S, Brunak S, Keşmir C. Predicting proteasomal cleavage sites: a comparison of available methods. Int Immunol. 2003;15(7):781-787. doi:10.1093/intimm/dxg084

30. Ong E, Cooke MF, Huffman A, et al. Vaxign2: the second generation of the first Web-based vaccine design program using reverse vaccinology and machine learning. Nucleic Acids Res. 2021;49(W1):W671-W678. doi:10.1093/nar/gkab279

31. Moise L, Gutierrez A, Kibria F, et al. iVAX: An integrated toolkit for the selection and optimization of antigens and the design of epitope-driven vaccines. Hum Vaccin Immunother. 2015;11(9):2312-2321. doi:10.1080/21645515.2015.1061159

32. Goodswen SJ, Kennedy PJ, Ellis JT. Vacceed: a high-throughput in silico vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. Bioinformatics. 2014;30(16):2381-2383. doi:10.1093/bioinformatics/btu300

33. Vita R, Mahajan S, Overton JA, et al. The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. 2019;47(D1):D339-D343. doi:10.1093/nar/gky1006

34. Trygoniaris D, Malousi A. VaccineDesigner (Version Version3). Zenodo. 2023 https://doi.org/10.5281/zenodo.10090612