VaccineDesigner: A Web-based Tool for Streamlined Multi-epitope Vaccine Design

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] Lab. of Biological Chemistry, 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; email: andigoni@auth.gr)

Keywords: Multi-epitope vaccines, reverse vaccinology, immunoinformatics, pathogen genomes, B-cell, CTL, and HTL epitope prediction, Web-based tool

Abstract

Background

Multi-epitope vaccines have become the preferred strategy for protection against infectious diseases by integrating multiple MHC-restricted T-cell and B-cell epitopes that elicit both humoral and cellular immune responses against pathogens. Computational methods address various aspects independently, yet their orchestration is technically challenging as most bioinformatics tools are accessible through heterogeneous interfaces and lack interoperability features. The present work proposes a novel framework for rationalized multi-epitope vaccine design that streamlines end-to-end analyses through an integrated Web-based environment.

Results

VaccineDesigner is a comprehensive Web-based framework that streamlines the design of protective epitope-based vaccines by seamlessly integrating computational methods for B-cell, CTL, and HTL epitope prediction. VaccineDesigner incorporates single epitope prediction and evaluation as well as additional analyses, such as multi-epitope vaccine generation, estimation of population coverage, molecular mimicry, and proteasome cleavage. The functionalities are transparently integrated into a modular architecture, providing a single access point for rationalized, multi-epitope vaccine generation, time and cost-effectively.

Conclusions

VaccineDesigner is a Web-based interface for developing multi-epitope vaccines. Given a protein sequence, VaccineDesigner identifies and evaluates candidate B-cell, CTL, and HTL epitopes and constructs a library of multi-epitope vaccines that combine strong immunogenic responses, safety, and broad population coverage. The source code is available under the academic license at: https://github.com/BiolApps/VaccineDesigner. VaccineDesigner is freely accessible at: http://bioinformatics.med.auth.gr/VaccineDesigner.

1. Background

Developing effective and safe vaccines against pathogens is a leading research priority in immunology and public health. Vaccine development is typically based on live attenuated pathogens, inactivated agents, subunit formulations, or recombinant protein antigens. However, these methods raise several safety concerns, struggle with pathogen cultivation, and are time-consuming, therefore impractical particularly during pandemics[1]. During the last two decades, computational methods have become indispensable in trying to improve efficacy, safety, and population coverage, rationalizing vaccine design. Reverse Vaccinology (RV)[2] is fundamentally based on computational methods to sift through the genetic information of a pathogen and identify efficient vaccine candidates that engage with the immune system. However, the identification of suitable proteins in pathogen proteomes is only the first step in computational vaccine design. Novel computational approaches have revolutionized our ability

to predict and strategically harness epitopes[3] i.e., tiny peptide fragments of antigens, capable of inducing protective immunity against pathogens. Since epitope-based approaches have been systematically used for some years now[4] a reasonable advancement would be the identification of immunodominant epitopes in vaccine candidates and their combination in efficient multi-epitope peptide constructs.

Multi-epitope vaccines constitute a preferable strategy for several reasons, including the increased specificity over traditional methods, customizability, stability and convenient production[5]. In addition, multi-epitope vaccines can include different types of epitopes, epitopes of different antigens, epitopes of different pathogens for cross protection, or epitopes of different protein isoforms, for broader population coverage. On the antipode, multi-epitope vaccines are often inadequately immunogenic, and optimization is difficult as it is not feasible to exhaustively examine all possible peptide combinations[5]. Current methods on multi-epitope peptide construction are known from fusion protein technology[6] and the strategic design of multi-epitope vaccines offers a promising approach for addressing diseases against a wide spectrum of pathogens, including parasites (visceral leishmaniasis[7] and onchocerciasis[8]), viruses (COVID-19[9-10] and Ebola[11]) and bacterial infections (Helicobacter pylori[12] and Klebsiella pneumoniae[13]).

The components of a multi-epitope vaccine include B-cell epitopes for humoral responses (B-cell), cytotoxic T lymphocyte epitopes for cellular immunity (CTL), and helper T lymphocyte epitopes (HTL) for immune regulation[14]. B-cell epitopes consist of surface-accessible clusters of amino acids and activate immune responses through antibody production[15]. 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[16]. A multi-epitope vaccine has the unique ability to simultaneously activate both humoral and cellular immunity while avoiding the presence of elements that could provoke harmful reactions[17]. The final multi-epitope construct is important to include components with adjuvant capacity, to maximize

immunogenicity as well as suitable linkers/spacers between epitopes, for optimal proteasome processing and minimization of junctional epitopes[18].

Over the last decades, epitope-based computational approaches have revolutionized vaccine design. BepiPred 3.0[19], BCEPRED[20], and ABCpred[21] rely on machine learning methods to predict linear B-cell epitopes, while CBTOPE[22] integrates structural data to predict conformational epitopes. DiscoTope 3.0[23] and ElliPro[24] focus on the prediction of discontinuous B-cell epitopes based on the spatial arrangements of amino acids. To address T-cell epitope prediction, NetMHCpan-4.0[25] and NetMHCIIpan-4.0[25] quantify peptide-MHC binding affinities using a neural network architecture while PROPRED[26] is primarily focused on identifying promiscuous MHC class-II binding regions. A wide range of tools have been developed to evaluate the properties of the predicted epitopes. ProtParam[27], VaxiJen[28], and Vaxign-ML[29] assess the physicochemical characteristics, and the antigenic, protegenicity potential of the predicted epitopes, respectively. IFNepitope[30] predicts IFN-γ inducing epitopes, while AlgPred 2.0[31] and ToxinPred2[32] assess the allergenic and toxic potential. Novel approaches attempt to combine outputs of multiple complementary tools in order to provide more effective and generalizable solutions. Vaxign2[33], iVax[34], and Vacceed[35] are examples of workflows facilitating the discovery of vaccine targets in bacterial and eukaryotic pathogens. In support of these analyses The Immune Epitope Database and Analysis Resource (IEDB-AR)[36] provides a robust and particularly valuable repository and analysis tool for experimentally validated immune epitopes and predictions.

Despite the advances in building integrated software tools[37], the orchestration of RV resources remains technically challenging[38] due to their structural heterogeneities and lack of interoperability[39]. Furthermore, existing frameworks have functional and technical limitations in terms of their ability to address epitope prediction for all cell types, ability to construct multi-epitope vaccines, and ease of use. To cope with these issues, we introduce a novel Web-based framework for rationalized multi-epitope vaccine design called VaccineDesigner. Rather than combining the outputs of individual tools, VaccineDesigner integrates and seamlessly execute cascading tasks for B-cell, CTL, and HTL epitope

prediction through a user-friendly and fully customizable graphical interface. VaccineDesigner incorporates important functionalities such as multi-epitope vaccine generation, prioritization, population coverage estimation, molecular mimicry, and proteasome cleavage. The overall goal is to provide a comprehensive tool that supports end-to-end analyses, facilitating the prioritization of immunogenic epitopes into multi-epitope constructs likely to elicit protective immune responses.

2. Implementation

VaccineDesigner supports epitope prediction for B-cell and T-cell lymphocytes, as crucial indicators of protein antigenicity[40]. The architectural components are split into two discrete modules (Fig. 1). Module A *(Epitope prediction)* includes a set of tools for B-cell, CTL, and HTL epitope prediction, coupled with antigenicity, toxicity and allergenicity assessment for the candidate epitopes of all cell types. Module B *(Multi-epitope vaccine construction)* gets the individual tool predictions of Module A and builds a high-quality multi-epitope vaccine library that includes relevant vaccine components and quality filtering tools. VaccineDesigner applies the default parameter configuration of each tool in modules A and B, yet the parameters and thresholds of each tool are configurable according to user preferences.

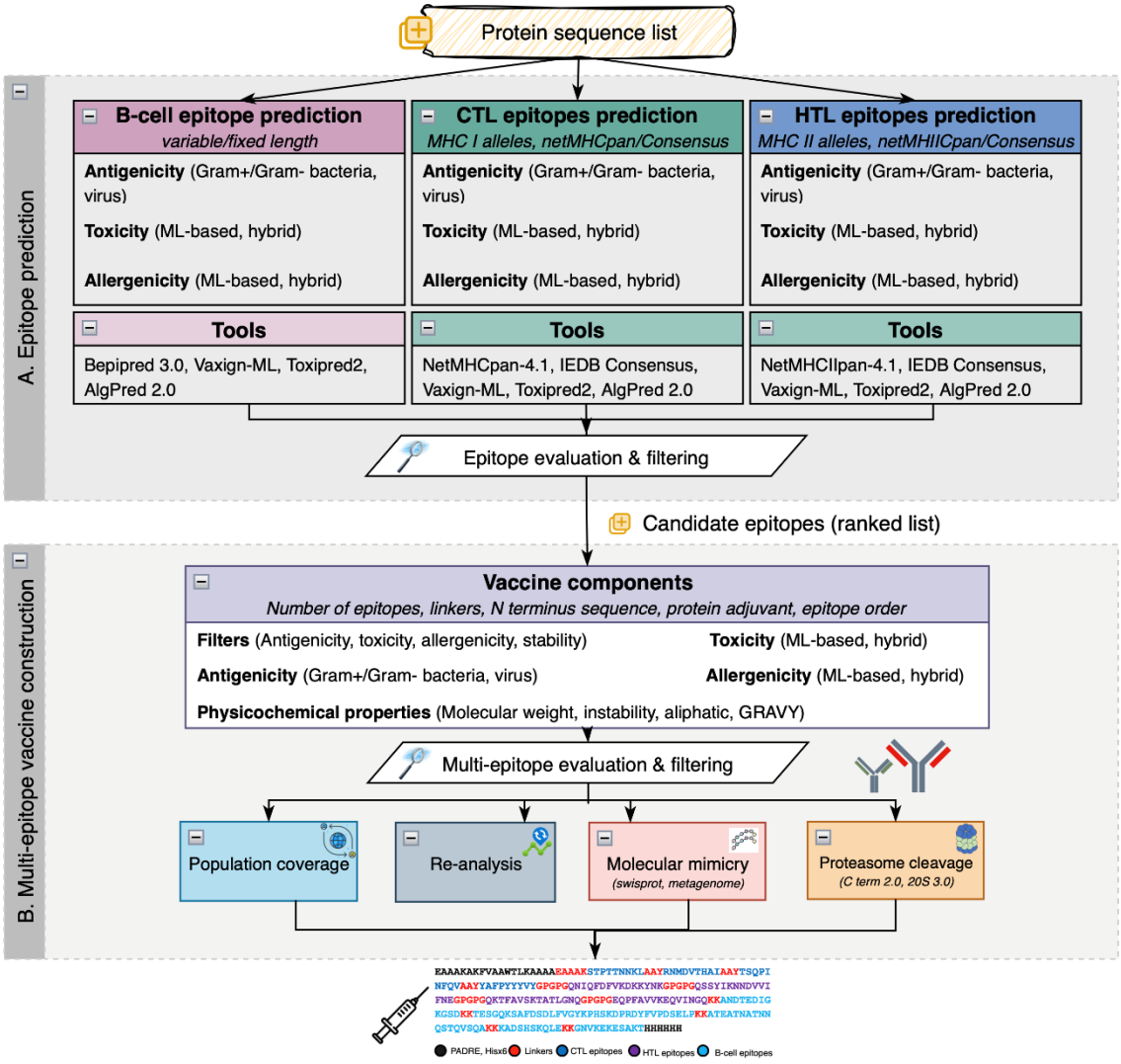


Figure 1: The main functional components of the streamlined process implemented by VaccineDesigner. **A.** The *Epitope prediction module* gets a fasta protein file, performs epitope prediction (B-cell, CTL, HTL), and ends up with the evaluation and filtering of epitope sequences, based on their antigenicity,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

2.1 Epitope Prediction and Evaluation

2.1.1 B-cell epitope prediction

B-cell epitopes are predicted using BepiPred 3.0. BepiPred 3.0 systematically scans protein sequences and assesses key physicochemical properties and sequence patterns to identify regions that can induce robust antibody responses[19]. Based on the amino acid scores VaccineDesigner generates B-cell epitopes,

adhering to user-defined epitope length parameters. Two alternative analyses can be applied. The first 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 can be modified, enabling the combination of individual epitopes into larger immunogenic fragments. The second option pertains to the B-cell epitope generation. Users can define a fixed-length scanning window to assess the scores of all amino acids. Eligible epitopes are set as continuous amino acid regions with a user-defined length, wherein each position must fulfill certain filtering criteria.

2.1.2 CTL epitope prediction

VaccineDesigner performs CTL epitope prediction using NetMHCpan 4.1[25] and IEDB Consensus Method[41]. NetMHCpan implements an algorithm that identifies high-affinity regions for MHC class I alleles. Both methods predict strong and weak binder counts and export various information for the associated MHC alleles. The IEDB Consensus method combines the results of multiple prediction methods into a single consensus outcome providing a more reliable assessment, as prediction algorithms may produce varying results due to differences in underlying models and assumptions. The final ranking of CTL epitopes is determined by the number of alleles interacting with each epitope.

2.1.3 HTL epitope prediction

HTL epitope prediction relies on the NetMHCIIpan 4.0 framework. NetMHCIIpan 4.0 identifies regions characterized by strong binding affinity for MHC class II molecules[25]. The results, akin to CTL epitopes, include strong and weak binders counts, along with details of the MHC class II alleles. HTL epitope prediction can be performed by the IEDB Consensus method for class II binding prediction, employing the same framework as for the Class I module[41]. The ranking of the HTL epitopes is determined by the number of interacting alleles, similar to the approach used for CTL epitopes.

2.1.4 Epitope evaluation

Given a set of candidate B-cell, CTL, and HTL epitopes, VaccineDesigner applies various quality metrics to filter out epitopes with suboptimal antigenicity and safety profiles, adhering to user-defined thresholds. Vaxign-ML[29] is used to comparatively assess the level of antigenicity for each epitope,

prioritizing those that are highly immunogenic. To cope with potential safety issues, ToxinPred2[32] is integrated into Module A to identify toxins within epitope sequences, and AlgPred 2.0[31] to minimize the risk of allergic reactions in vaccine recipients. The epitopes that meet high-quality criteria for antigenicity, toxicity, and allergenicity are considered more likely to improve the overall efficacy and safety of the candidate multi-epitope constructs.

2.2 Multi-epitope Vaccine Sequences

The synthesis of multi-epitope constructs is considered a favorable approach to mitigate junctional immunogenicity and efficient epitope separation and presentation to T-cell and B-cell receptors[14]. VaccineDesigner expands the analysis of candidate epitopes to craft multi-epitope vaccine sequences. Using the most prominent epitopes, VaccineDesigner fuses larger constructs by merging individual epitopes with appropriate linker sequences (default: GPGPG/HTL, AAY/CTL, KK/B-cell)[6,14,42-46] in a two-step procedure as follows:

2.2.1 Vaccine Sequences Construction

The multi-epitope vaccine construct is built by a variable, user-defined number of B-cell, CTL, and HTL epitopes, as well as linker sequences between the epitopes and the N-terminus sequence. A variety of protein adjuvants are available for the synthesis of the multi-epitope vaccine sequences[47-50]. In addition, the order of epitope combination can be defined in arrangements of B-cell, CTL, and HTL epitopes, exemplified as B-C-H (B-cell - CTL - HTL). VaccineDesigner generates candidate vaccine sequences, that form a versatile library of epitope sequence combinations. Researchers can then explore diverse epitope combinations containing up to five epitopes from each category (B-cell, CTL, and HTL).

2.2.2 Vaccine Candidate 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 Vaxign-ML for antigenicity, ToxinPred2 for toxin prediction, Algpred 2.0 for allergenicity, and ProtParam for physicochemical sequence analysis and stability[27]. User-defined parameters, such as the desired number of vaccine sequences are used to keep track of the number of candidates that meet all quality criteria. The algorithm

exports the final multi-epitope sequence constructs when the preferred number of candidates is reached. This dynamic approach ensures the selection of vaccine sequences that meet strict measures for immunogenicity, safety, and biochemical attributes, offering the most promising candidates for further development and validation.

The candidate multi-epitope library is short-listed based on the ranking of the antigenicity, allergenicity, stability, and physicochemical properties of each construct. In addition, each candidate multi-epitope vaccine is scored by the deep neural network algorithm employed by DeepVacPred[51]. The analysis exports the list of candidate multi-epitope sequences once the user-defined maximum number of candidates is reached. The final candidate sequences are ranked based on the weighted sum of individual antigenicity, toxicity, allergenicity, stability, and the DeepVacPred rankings as follows: Let S be the vector of scores for each parameter, with $S_i \in \mathbb{R}$ and $R(S)$ the corresponding rankings with $R(S_i) \in \mathbb{Z}$ and $R(S)$ being a permutation of rankings, such that higher scores receive lower positions. In addition, let W be a vector of user-defined weights with $W_i \in \mathbb{R}$. The best-scoring multi-epitope construct is the one that yields the smallest possible sum of weighted rankings, across the n ranking criteria i.e.,

$$min\left(\sum_{i=1}^{5} W_i \times R(S_i)\right).$$

The newly generated sequences can be re-analyzed to verify the favorable properties of individual epitopes, in the whole construct, and can be further validated for population coverage, molecular mimicry, and the presence of putative proteasome cleavages. Population coverage is a highly recommended analysis for human vaccines and is implemented by the IEDB Population Coverage[36] algorithm. To address molecular mimicry VaccineDesigner conducts protein similarity searches against SWISS-PROT[52] or metagenomic proteins (env_nr) using the BLASTp[53] algorithm. The analysis estimates whether a 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[54]. The results are provided in both graphical and tabular formats as shown in Figure 2A.

3. Results

3.1 Usage scenario

To demonstrate the functionalities of the overall workflow, we implemented an example usage scenario on the *Atl* (UniProt: P0C5Z8) and the *IsdA* (UniProt: Q7A152) proteins of *Staphylococcus aureus*. *Atl* is a bifunctional autolysin playing a crucial role in the physiological processes, virulence, and pathogenesis of *S. aureus*. Autolysins are promising vaccine targets as they combine favourable features such as surface exposure, immunogenicity, and increased conservation among different strains. Iron-regulated surface determinant protein A (*IsdA*) is a virulence factor that is involved in the acquisition of iron, bacterial growth, and adherence to host tissues. *IsdA* combines desired vaccine properties including surface exposure on the cell wall of *S. aureus* and induction of immune responses in infected individuals. *IsdA* together with other cell wall-anchored surface proteins has been reported to induce significant protective immunity in murine models[55].

In the usage scenario, 123 B-cell epitopes were predicted on the *Atl* and 136 on the *IsdA* protein using a fixed sequence length of 10 amino acids. 16 of these epitopes were considered high quality 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-A*01:01, HLA-A*01:02, HLA-A*01:03, HLA-A*01:04, HLA-A*01:06, HLA-A*01:07, HLA-A*01:08* and *HLA-A*01:09*. The predicted epitopes underwent toxicity, allergenicity, and antigenicity filtering, resulting in 13 candidates. 193 and 50 HTL epitopes were considered strong or weak binders to MHC class II alleles for the *Atl* and the *IsdA* protein respectively, including *HLA-DRB1*01:01, HLA-DBR1*01:02, HLA-DBR1*01:03, HLA-DBR1*01:04, HLA-DBR1*01:05, HLA-*

*DBR1*01:06*, *HLA-DBR1*01:07* and *HLA-DBR1**01:08. Allergenic and toxic epitopes accounted for a significant number of epitopes resulting in 23 high-quality HTL epitopes for both proteins.

The list of the top-scoring epitopes of each category (B-CTL-HTL) was further narrowed down to four to reduce the computational cost. The candidates combined in the C-H-B order (CTL, HTL, and B-cell). Two epitopes of each cell type were fused in the final constructs, resulting in 1,728 candidate vaccine sequences. *Table 1a* presents the alignments of the newly generated multi-epitope vaccine sequences. *Table 1b* contains a ranking list of the first five multi-epitope sequences that passed all thresholds based on the level of antigenicity, toxicity, and relative properties.

The final ranking of the high-quality, multi-epitope sequences (combined ranking) is defined by the collective sum of individual rankings. The best scoring sequence (seq2) including linker sequences is shown in Figure 2. Figure 2A shows the outcome of the proteasome cleavage analysis in which the predicted cleavages are plotted for all the detected epitopes. Figure 2B depicts the multi-epitope sequence and its newly predicted B-cell, CTL and HTL epitope regions after re-evaluating the synthesized construct in the multi-epitope prediction module.

Table 1: (a) A conserved domain of 98 amino acids in the five high-scoring multi-epitope sequences (top) and the aligned variable region of the 46 amino acid residues. (b): Combined ranking of the five multi-epitope sequence candidates along with their predicted antigenicity, toxicity, allergenicity, and instability index.

(a)

Conserved domain of the multi-epitope sequences
AAYQVNSSINDYAAYVSDNKSQQTAAYYLRSHNYSYGPGPGINGEISYMKNNYQNAGPGPGMDDYMQHPGKVIKQNGPGP GNGEISYMKNNYQNAFKK

Alignment of the variable multi-epitope region
seq1 STSTTAPKTNKKDTRANQS----------ATTKKSDNKSQQTNKKKVSDNKSQQTN----------
seq2 STSTTAP----------KTNKKSDNKSQQTNKKKDTRANQSATTKKVSDNKSQQTN
seq3 ----------DTRANQSATTKKSTSTTAPKTNKKSDNKSQQTNKKKVSDNKSQQTN----------
seq4 ----------DTRANQS----------ATTKKSDNKSQQTNKKKSTSTTAPKTNKKVSDNKSQQTN
seq5 ----------DTRANQS----------ATTKKSDNKSQQTNKKKVSDNKSQQTNKKSTSTTAPKTN

(b)

Combined ranking	**Sequence identifier**	**Antigenicity** *(thr=50)*	**Toxicity** *(thr=0.6)*	**Allergenicity** *(thr=0.3)*	**Instability index** *(thr=30)*
1	seq2	79.29	0.298	0.281	39.09
2	seq3	80.33	0.304	0.291	35.83
3	seq4	78.88	0.347	0.282	38.35

4	seq5	80.01	0.345	0.283	37.91
5	seq1	90.86	0.304	0.291	34.86

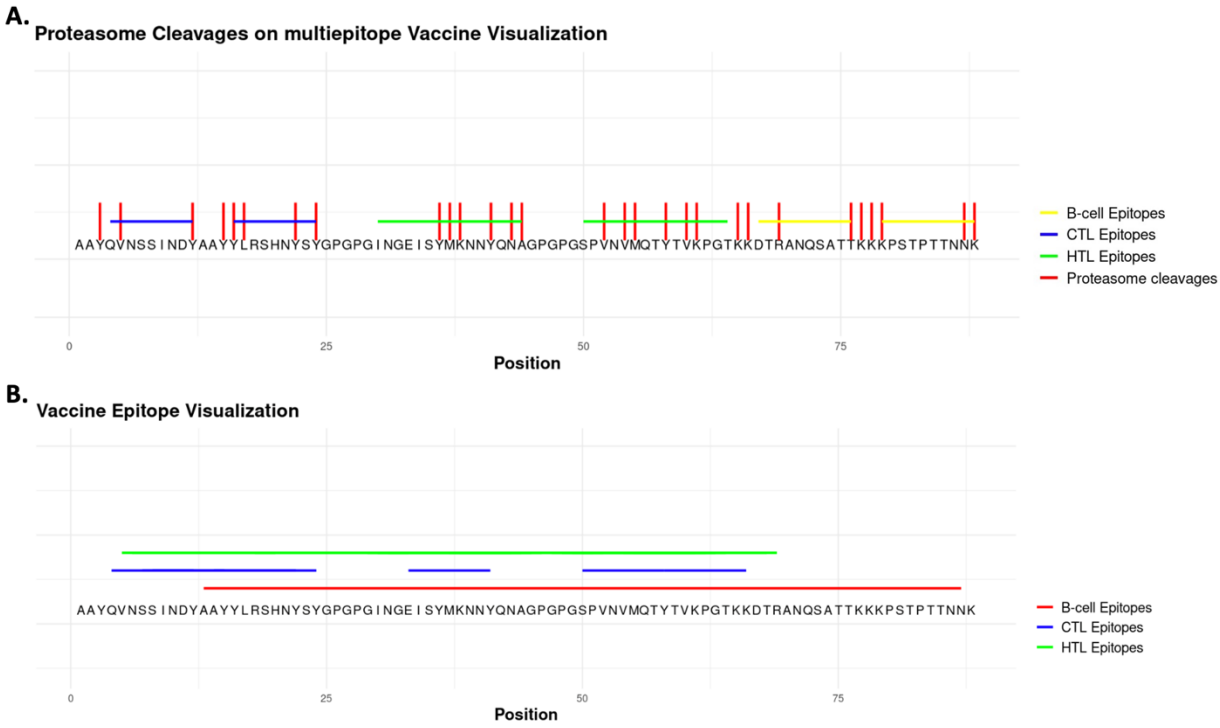


Figure 2: Visualization of multi-epitope vaccine based on the proteasome cleavage functionality (A), and multi-epitope sequence re-evaluation with the epitope prediction module (B).

3.2 Evaluation against experimentally validated epitopes

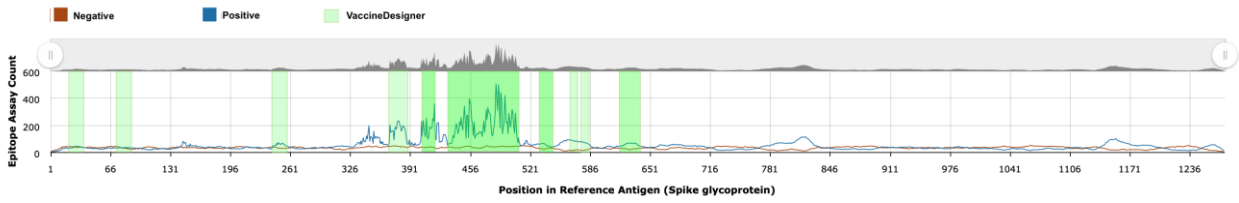
To assess the competence of VaccineDesigner we sought to compare the predictions of linear B-cell epitopes and, cytotoxic and helper T-cell epitopes of SARS-CoV-2 Spike glycoprotein (UniProt ID: P0DTC2) against experimentally validated sequences provided by the IEDB database.

Testing procedure: The configuration parameters of each tool were set using the integrated graphical environment of VaccineDesigner. To quantify the level of antigenicity, Vaxign-ML was configured for viruses[21] and filtered based on the protegenicity of each epitope (>40%). Since viruses do not produce toxins and their pathogenicity arises from their infection and replication cycle[56], ToxinPred2 can be omitted or applied with high thresholds for testing purposes. In this example, ToxinPred2 and AlgPred 2.0 thresholds were set to 0.8 and 0.45, respectively. BepiBred 3.0 was used to predict B-cell epitopes of variable length. The threshold was set to 0.1, lower than the default 0.1512 to identify a broader range

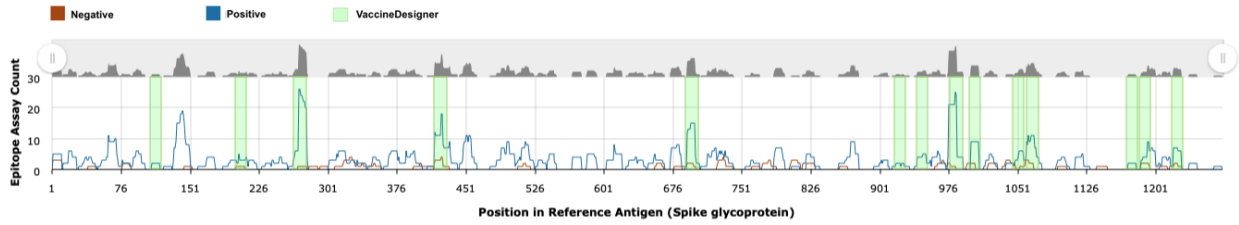
of epitopes and subsequently narrow down the predictions either by applying more stringent thresholds or by ranking the peptides based on other properties. The parameter *Subthreshold Amino Acid Inclusion Count* was set to 1 to enable the synthesis of larger continuous epitope regions that are separated by a single low-score amino acid (*Secondary threshold*: 0.08). CTL and HTL epitopes were predicted on the most common HLA alleles with positive assays that were collected from the IEDB database (MHC I molecules: *HLA-A*02:01*, *HLA-A*02:03*, *HLA-A*02:02*, *HLA-A*68:02*, *H2-Kb*, *HLA-A*24:02*, *HLA-A*11:01*, *HLA-E*01:01*, and MHC II molecules: *HLA-DRA*01:01/DRB1*04:01*, *HLA-DRB1*15:01*, *HLA-DRB3*02:02*, *HLA-DQA1*01:02/DQB1*06:02*, *HLA-DRB1*07:01*, *HLA-DRB5*01:01*). VaccineDesigner was set to run netMHCpan-4.1 and netMHCIIpan-4.0 with peptide length of 9 and 15, respectively.

Results: VaccineDesigner exported seventeen variable-length B-cell epitopes of which seven were filtered out due to the unqualified toxicity and/or antigenicity levels (Suppl. file). All predictions overlap with on average 40 experimentally validated epitopes. The highest scoring B-cell epitope is predicted in a non-toxic and non-allergen 74 amino acids sequence that matches the region including the most positive assay counts. Figure 3A shows a significant overlap between the predicted B-cell epitopes with the experimentally validated ones. Similarly, of the 242 CTL and 346 HTL predicted epitopes we selected the top fifteen sequences with the highest number of strong and weak binders for comparison against validated IEDB epitopes (Suppl. file). Figures 3B and 3C show the residual overlaps between the predicted MHC-I and MHC-II epitopes and the validated 263 MHC-I and 1,783 MHC-II epitopes from the IEDB positive assays (IEDB v.3.10.0). Evidently, most CTL and HTL predictions are collocated with known epitopes. CTL predictions are on average more accurate compared to HTL which is possibly related to the significantly higher number of positive assays implying stronger experimental evidence (on average 11.4 CTL vs 1.6 positive assay counts). However, further comparative analyses are needed to assess how different parameter configurations can affect the overall performance which is not realistic to implement manually. By enabling transparent execution of the pipelined processes VaccineDesigner reduces the execution time which is particularly important when a candidate protein must be examined exhaustively under different configurations.

A.



B.



C.

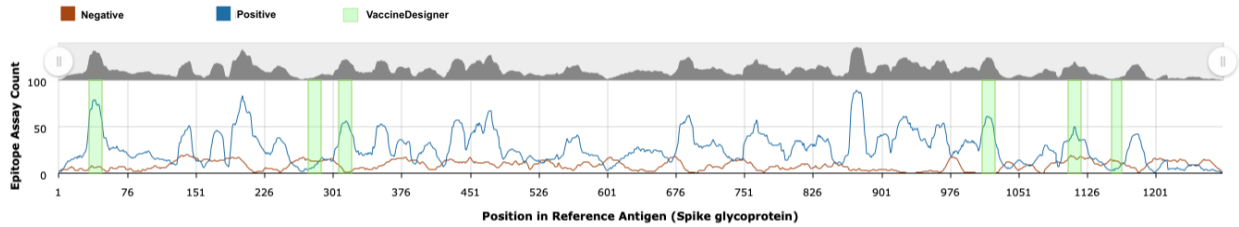


Figure 3: Overlapping residual loci between the top-scoring B-cell (A), CTL (B), and HTL (C) predictions and the experimentally validated positive and negative assays (IEDB v.3.10.0).

3.3 Comparison with other RV pipelines

During the last two decades, several RV pipelines have been developed. Table 2 lists the key attributes of the most popular RV implementations including the availability of graphical interfaces, streamlined processes, support for different epitope types, prediction capabilities for non-human species, multi-epitope sequence construction, pathogen protein selection, primary methods for vaccine candidate evaluation, and source code availability. Most of the RV tools predict epitopes following various approaches however, VaccineDesigner is the only tool that predicts and combines epitopes and then filters the multi-epitope constructs to select only the best epitope combinations. In addition, VaccineDesigner has the unique flexibility to enable

modifications of the number of epitopes from each category and select epitopes from different proteins based on their quality and user preference.

Table 2: Main features of RV pipelines.

Name/ Tool	Graphical User Interface (GUI) (Yes/No)	Streamlined process (Yes/No)	Prediction for different epitope types (BP/TP/BP+TP)	Epitope prediction for non-human species (Yes/No)	Multi-epitope sequence construction (Yes/No)	Pathogen protein selection (Yes/No)	Vaccine Candidate Evaluation (*)	Source code availability (Yes/No)
NERVE[57]	No	Yes	No	No	No	Yes	SL, AP, HS	Yes
iVAX[34]	Yes	Yes	TP	Yes	Yes	Yes	I, C	No
OptiVax[58]	No	Yes	TP	Yes	Yes	Yes	I, PC, SP	Yes
ReVac[59]	No	Yes	BP + TP	Yes	No	Yes	A, C, SL, CR	Yes
IEDB-AR[36]	Yes	No	BP + TP	Yes	No	No	No	Yes
Vacceed[35]	No	Yes	TP	Yes	No	Yes	SL, SP, TH, T	Yes
VacSOL[60]	Yes	Yes	BP+TP	No	No	Yes	SL, V, TH, A, B+T	Yes
Vaxign2[33]	Yes	Yes	TP	No	No	Yes	SL, TH, AP, B&T	Yes
VaccineDesigner	No	Yes	BP+TP	Yes	Yes	No	B+T, A, CR, S, PC	Yes

***SL**: Subcellular Localization, **AP**: Adhesin Probability, **HS**: Host Similarity, **I**: Immunogenicity, **C**: Conservation, **PC**: Population Coverage, **A**: Antigenicity, **CR**: Cross Reactivity, **SP**: Signal Peptides, **TH**: Transmembrane Helices, **V**: Virulence, **BP**: B-cell prediction, **TP**: T-cell prediction, **BP+TP:** B-cell and T-cell prediction, **B+T**: Presence of B-cell and T-cell epitopes, **T**: Presence of T-cell epitopes, **S**: Stability*

4. Discussion

VaccineDesigner was developed to build protective multi-epitope constructs using in silico B-cell, CTL, and HTL epitope prediction methods coupled with toxicity, allergenicity, and antigenicity assessment. Putative multi-epitope vaccines are further examined to estimate population coverage, molecular mimicry, and proteasome cleavage to maximize protective immune responses. Technically, VaccineDesigner aims to address the reported underutilization of automated workflows in reverse vaccinology by enabling pipelined analyses through a user-friendly graphical environment that includes preconfigured software tools and effortless I/O data flow management[61].

VaccineDesigner extends RV frameworks by introducing several technical and functional advancements: 1) Implements a stepwise procedure that starts with a list of candidate antigenic proteins, and progressively leads to the most likely effective multi-epitope vaccines. 2) Leverages the collective strengths of each tool and combines their output in unified tabular and graphical reports. 3) Applies a weighted scoring scheme to prioritize multi-

epitope constructs based on individual properties. 4) Ensures time-efficient and effortless application due to the seamless data flow between cascading tasks implemented by individual tools. 5) Is independent of local installations and operating system restrictions as it is available through a fully customizable Web-based interface.

Considering the architectural components, VaccineDesigner is built on a foundation of well-validated methodologies recognized and trusted within the scientific community. The selection of the tools was primarily based on benchmarking studies[29,61], citation-based recognition by the scientific community, and also on the technical constraints that are related to the programmatic access and availability of the source code. A feasible enhancement enabled by the modular architecture of VaccineDesigner is to combine evidence from multiple tools performing the same analysis. However, the lack of interoperable tools puts additional technical challenges as it is hard to standardize inputs, parameters, and outputs and most importantly to provide sophisticated and more effective consensus algorithms. Considering the significant number of informative properties that are produced for each candidate protein, a reasonable enhancement would be to mine discriminative properties through machine learning models to predict peptides that induce protective immune responses. Still, the impact of these models is limited by the inaccuracies of the positive and negative training data and the lack of a definitive list of analyses that all RV pipelines should include[61].

It should be noted that VaccineDesigner does not address the initial protein selection stage. Practically, the starting point of the pipeline is a set of proteins that have already been selected as potentially immunogenic. To incorporate this analysis several challenges must be addressed including the selection of proteins based on the subcellular localization, the exclusion of homologs to avoid autoimmune reactions in the host, and the use of pangenomes to select core genome components that reduce the selective pressure[62]. Moving to more holistic approaches VaccineDesigner will pursue the integration of 3D modelling algorithms to predict the folded 3D structure of the multi-epitope vaccine constructs and to identify exposed peptides that are more likely to interact with immune cells. Alphafold 3 revolutionizes research in this field and is expected to accelerate research in precision vaccine discovery by predicting the structures of DNA, RNA, and molecules like ligands, which are essential to drug discovery[63].

5. Conclusions

In the evolving landscape of infectious diseases, multi-epitope vaccine design has emerged as a pivotal strategy to identify antigenic determinants in pathogenic proteins, while also reducing the risk of adverse reactions and insufficient coverage against diverse pathogenic strains. VaccineDesigner is the only reverse vaccinology framework that combines an automated procedure for the prediction and assessment of protective multi-epitope constructs with a graphical and fully customizable Web-based environment. As our insights into the host-pathogen interactions advance, we anticipate that VaccineDesigner will provide a valuable tool for rationalized knowledge synthesis and will contribute to real-world applications as a generic framework for developing safer and more effective vaccines.

Declarations

Abbreviations

APCs: antigen-presenting cells
CTL: Cytotoxic T cells
HTL: Helper T cells
IEDB: Immune Epitope Database
IEDB-AR: Immune Epitope Database and Analysis Resource
MHC: Major Histocompatibility Complex
RV: Reverse Vaccinology

Ethics approval and consent to participate

Not applicable

Consent for publication

All authors gave their consent for publication

Availability of data and materials

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

Project name: VaccineDesigner
Project home page: http://bioinformatics.med.auth.gr/VaccineDesigner
Operating system(s): Platform independent

Programming language: R, Python
Other requirements: Web browser
License: Academic Free License
Any restrictions to use by non-academics: license needed

Competing Interests

Not applicable

Funding

Not applicable

Authors' contributions

D.T. contributed to: design of the work, creation of new software, preparation of the manuscript A.K. contributed to: design of the work, acquisition, analysis, preparation of the manuscript A.P. contributed to: acquisition, analysis, interpretation of data E.D. contributed to: acquisition, analysis, interpretation of data G.T. contributed to: interpretation of data, manuscript revision M.Y. contributed to: interpretation of data, manuscript revision C.K. contributed to: design of the work, acquisition, analysis, manuscript revision A.M. contributed to: conception, design of the work, creation of new software, preparation of the manuscript

All authors have approved the submitted version and have agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work

Acknowledgments

Not applicable

Supplementary file: List of predicted B-cell, CTL, and HTL epitopes in Spike S glycoprotein (P0DTC2) of SARS-CoV-2 and the number of validated matching epitopes from IEDB database.

References

1. Sette, A., & Rappuoli, R. (2010). Reverse vaccinology: developing vaccines in the era of genomics. Immunity, 33(4), 530–541. https://doi.org/10.1016/j.immuni.2010.09.017
2. Rappuoli R. (2000). Reverse vaccinology. Current opinion in microbiology, 3(5), 445–450. https://doi.org/10.1016/s1369-5274(00)00119-3

3. 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. https://doi.org/10.1016/j.jbi.2014.11.003

4. Suhrbier, A. (1997b). Multi‐epitope DNA vaccines. Immunology and Cell Biology, 75(4), 402–408. https://doi.org/10.1038/icb.1997.63

5. Negahdaripour, M., Nezafat, N., Eslami, M., Ghoshoon, M. B., Shoolian, E., Najafipour, S., Morowvat, M. H., Dehshahri, A., Erfani, N., & Ghasemi, Y. (2018c). Structural vaccinology considerations for in silico designing of a multi-epitope vaccine. Infection, Genetics and Evolution, 58, 96–109. https://doi.org/10.1016/j.meegid.2017.12.008 /

6. Chen, X., Zaro, J. L., & Shen, W. (2013). Fusion protein linkers: Property, design and functionality. Advanced Drug Delivery Reviews, 65(10), 1357–1369. https://doi.org/10.1016/j.addr.2012.09.039

7. Khatoon, N., Pandey, R. K., & Prajapati, V. K. (2017). Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. Scientific reports, 7(1), 8285. https://doi.org/10.1038/s41598-017-08842-w

8. Shey, R. A., Ghogomu, S. M., Esoh, K. K., Nebangwa, N. D., Shintouo, C. M., Nongley, N. F., Asa, B. F., Ngale, F. N., Vanhamme, L., & Souopgui, J. (2019). In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. Scientific reports, 9(1), 4409. https://doi.org/10.1038/s41598-019-40833-x

9. Enayatkhani, M., Hasaniazad, M., Faezi, S., Gouklani, H., Davoodian, P., Ahmadi, N., Einakian, M. A., Karmostaji, A., & Ahmadi, K. (2021). Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an in silico study. Journal of biomolecular structure & dynamics, 39(8), 2857–2872. https://doi.org/10.1080/07391102.2020.1756411

10. Aziz, S., Almajhdi, F. N., Waqas, M., Ullah, I., Salim, M. A., Khan, N. A., & Ali, A. (2022). Contriving multi-epitope vaccine ensemble for monkeypox disease using an immunoinformatics approach. Frontiers in immunology, 13, 1004804. https://doi.org/10.3389/fimmu.2022.1004804

11. Alizadeh, M., Amini-Khoei, H., Tahmasebian, S., Ghatrehsamani, M., Ghatreh Samani, K., Edalatpanah, Y., Rostampur, S., Salehi, M., Ghasemi-Dehnoo, M., Azadegan-Dehkordi, F., Sanami,

S., & Bagheri, N. (2022). Designing a novel multi-epitope vaccine against Ebola virus using reverse vaccinology approach. Scientific reports, 12(1), 7757. https://doi.org/10.1038/s41598-022-11851-z

12. Khan, M., Khan, S., Ali, A. et al. Immunoinformatics approaches to explore Helicobacter Pylori proteome (Virulence Factors) to design B and T cell multi-epitope subunit vaccine. Sci Rep 9, 13321 (2019). https://doi.org/10.1038/s41598-019-49354-z

13. Dar, H. A., Zaheer, T., Shehroz, M., Ullah, N., Naz, K., Muhammad, S. A., Zhang, T., & Ali, A. (2019). Immunoinformatics-Aided Design and Evaluation of a Potential Multi-Epitope Vaccine against Klebsiella Pneumoniae. Vaccines, 7(3), 88. https://doi.org/10.3390/vaccines7030088

14. 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.

15. 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.

16. 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

17. Zhang, L. (2017b). Multi-epitope vaccines: a promising strategy against tumors and viral infections. Cellular & Molecular Immunology/Cellular & Molecular Immunology, 15(2), 182–184. https://doi.org/10.1038/cmi.2017.92

18. Sette, A., Livingston, B., McKinney, D., Appella, E., Fikes, J., Sidney, J., Newman, M., & Chesnut, R. (2001). The development of multi-epitope vaccines: epitope identification, vaccine design and clinical evaluation. Biologicals, 29(3–4), 271–276. https://doi.org/10.1006/biol.2001.0297

19. Clifford, J. N., Høie, M. H., Deleuran, S., Peters, B., Nielsen, M., & Marcatili, P. (2022). BepiPred‐3.0: Improved B‐cell epitope prediction using protein language models. Protein Science, 31(12). https://doi.org/10.1002/pro.4497

20. Saha, S., & Raghava, G. P. S. (2004). BCEPRED: Prediction of continuous B-Cell epitopes in antigenic sequences using physico-chemical properties. In Lecture notes in computer science (pp. 197–204). https://doi.org/10.1007/978-3-540-30220-9_16

21. Saha, S., & Raghava, G. P. S. (2006). Prediction of continuous B‐cell epitopes in an antigen using recurrent neural network. Proteins, 65(1), 40–48. https://doi.org/10.1002/prot.21078

22. Ansari, H., & Raghava, G. P. (2010). Identification of conformational B-cell Epitopes in an antigen from its primary sequence. Immunome Research, 6(1), 6. https://doi.org/10.1186/1745-7580-6-6

23. Høie, M. H., Gade, F. S., Johansen, J. M., Würtzen, C., Winther, O., Nielsen, M., & Marcatili, P. (2024). DiscoTope-3.0: improved B-cell epitope prediction using inverse folding latent representations. Frontiers in Immunology, 15. https://doi.org/10.3389/fimmu.2024.1322712

24. Ponomarenko, J., Bui, H., Li, W., Fusseder, N., Bourne, P. E., Sette, A., & Peters, B. (2008). ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics, 9(1). https://doi.org/10.1186/1471-2105-9-514

25. Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020b). 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 Research, 48(W1), W449–W454. https://doi.org/10.1093/nar/gkaa379

26. Singh, H., & Raghava, G. P. S. (2001). ProPred: prediction of HLA-DR binding sites. Bioinformatics, 17(12), 1236–1237. https://doi.org/10.1093/bioinformatics/17.12.1236

27. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (2003). Protein identification and analysis tools in the ExPASY server. In Humana Press eBooks (pp. 531–552). https://doi.org/10.1385/1-59259-584-7:531

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

29. Ong, E., Wang, H., Wong, M. U., Seetharaman, M., Valdez, N., & He, Y. (2020). Vaxign-ML: supervised machine learning reverse vaccinology model for improved prediction of bacterial

protective antigens. Bioinformatics, 36(10), 3185–3191. https://doi.org/10.1093/bioinformatics/btaa119

30. Dhanda, S. K., Vir, P., & Raghava, G. P. (2013). Designing of interferon-gamma inducing MHC class-II binders. Biology Direct, 8(1). https://doi.org/10.1186/1745-6150-8-30

31. Sharma, N., Patiyal, S., Dhall, A., Pande, A., Arora, C., & Raghava, G. P. S. (2020). AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes. Briefings in Bioinformatics, 22(4). https://doi.org/10.1093/bib/bbaa294

32. Sharma, N., Naorem, L. D., Jain, S., & Raghava, G. P. S. (2022). ToxinPred2: an improved method for predicting toxicity of proteins. Briefings in Bioinformatics, 23(5). https://doi.org/10.1093/bib/bbac174

33. Ong, E., Cooke, M. F., Huffman, A., Xiang, Z., Wong, M. U., Wang, H., Seetharaman, M., Valdez, N., & He, Y. (2021). Vaxign2: the second generation of the first Web-based vaccine design program using reverse vaccinology and machine learning. Nucleic Acids Research, 49(W1), W671–W678. https://doi.org/10.1093/nar/gkab279

34. Moise, L., Gutierrez, A., Kibria, F., Martin, R., Tassone, R., Liu, R., Terry, F., Martin, B., & De Groot, A. S. (2015). iVAX: An integrated toolkit for the selection and optimization of antigens and the design of epitope-driven vaccines. Human Vaccines & Immunotherapeutics, 11(9), 2312–2321. https://doi.org/10.1080/21645515.2015.1061159

35. Goodswen, S. J., Kennedy, P. J., & Ellis, J. T. (2014). Vacceed: a high-throughput in silico vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. Bioinformatics, 30(16), 2381–2383. https://doi.org/10.1093/bioinformatics/btu300

36. Dhanda, S. K., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M. C., Jurtz, V., Andreatta, M., Greenbaum, J. A., Marcatili, P., Sette, A., Nielsen, M., & Peters, B. (2019). IEDB-AR: immune epitope database—analysis resource in 2019. Nucleic Acids Research, 47(W1), W502–W506. https://doi.org/10.1093/nar/gkz452

37. Dalsass, M., Brozzi, A., Medini, D., & Rappuoli, R. (2019). Comparison of Open-Source Reverse Vaccinology Programs for Bacterial vaccine antigen discovery. Frontiers in Immunology, 10. https://doi.org/10.3389/fimmu.2019.00113

38. Srivastava, P., & Jain, C. K. (2023). Computer aided Reverse Vaccinology: A game-changer approach for vaccine development. Combinatorial Chemistry & High Throughput Screening, 26(10), 1813–1821. https://doi.org/10.2174/1386207325666220930124013

39. Schubert, B., De La Garza, L., Mohr, C., Walzer, M., & Kohlbacher, O. (2017). ImmunoNodes – graphical development of complex immunoinformatics workflows. BMC Bioinformatics, 18(1). https://doi.org/10.1186/s12859-017-1667-z

40. 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.

41. 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

42. 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

43. 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

44. 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

45. 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

46. 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

47. 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

48. 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

49. 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

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

51. Yang, Z., Bogdan, P., & Nazarian, S. (2021). An in silico deep learning approach to multi-epitope vaccine design: a SARS-CoV-2 case study. Scientific reports, 11(1), 3238. https://doi.org/10.1038/s41598-021-81749-9

52. 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

53. 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

54. 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

55. Stranger-Jones YK, Bae T, Schneewind O. Vaccine assembly from surface proteins of Staphylococcus aureus. Proc Natl Acad Sci U S A. 2006;103(45):16942-16947. doi:10.1073/pnas.0606863103.

56. Baron S, Fons M, Albrecht T. Viral Pathogenesis. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 45. https://www.ncbi.nlm.nih.gov/books/NBK8149/

57. Vivona, S., Bernante, F., & Filippini, F. (2006). NERVE: new enhanced reverse vaccinology environment. BMC biotechnology, 6, 35. https://doi.org/10.1186/1472-6750-6-3

58. Liu, G., Carter, B., Bricken, T., Jain, S., Viard, M., Carrington, M., & Gifford, D. K. (2020). Computationally Optimized SARS-CoV-2 MHC Class I and II Vaccine Formulations Predicted to Target Human Haplotype Distributions. Cell systems, 11(2), 131–144.e6. https://doi.org/10.1016/j.cels.2020.06.009

59. D'Mello, A., Ahearn, C. P., Murphy, T. F., & Tettelin, H. (2019). ReVac: a reverse vaccinology computational pipeline for prioritization of prokaryotic protein vaccine candidates. BMC genomics, 20(1), 981. https://doi.org/10.1186/s12864-019-6195-y

60. Rizwan, M., Naz, A., Ahmad, J., Naz, K., Obaid, A., Parveen, T., Ahsan, M., & Ali, A. (2017). VacSol: a high throughput in silico pipeline to predict potential therapeutic targets in prokaryotic pathogens using subtractive reverse vaccinology. BMC bioinformatics, 18(1), 106. https://doi.org/10.1186/s12859-017-1540-0

61. Goodswen, S. J., Kennedy, P. J., & Ellis, J. T. (2023). A guide to current methodology and usage of reverse vaccinology towards in silico vaccine discovery. FEMS microbiology reviews, 47(2), fuad004. https://doi.org/10.1093/femsre/fuad004

62. Naz, K., Naz, A., Ashraf, S. T., Rizwan, M., Ahmad, J., Baumbach, J., & Ali, A. (2019). PanRV: Pangenome-reverse vaccinology approach for identifications of potential vaccine candidates in microbial pangenome. BMC bioinformatics, 20(1), 123. https://doi.org/10.1186/s12859-019-2713-9

63. Abramson J, Adler J, Dunger J, et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 630(8016):493-500. doi:10.1038/s41586-024-07487-w