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 epitope-based vaccine design provides an efficient strategy for immunogenic and safe vaccines; however, the implementation of this approach requires the integration of heterogeneous tools with limited interoperability. VaccineDesigner addresses this issue by offering a user-friendly platform that integrates methods for predicting and evaluating B-cell, CTL, and HTL epitopes. The goal is to provide a Web-based interface for integrated analyses, empowering fast, cost-effective, and rationalized multi-epitope vaccine design in a streamlined process. VaccineDesigner is based on a modular architecture that seamlessly integrates B-cell, CTL, and HTL epitope prediction tools and scoring algorithms, along with multi-epitope vaccine sequence generation. VaccineDesigner is an open-source tool freely available under the academic license at: https://github.com/BiolApps/VaccineDesigner. The Web-based application is accessible at: http://bioinformatics.med.auth.gr/VaccineDesigner

INTRODUCTION

Developing efficacious vaccines against a range of pathogens is a primary research priority in the fields of 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 use of selected antigenic epitopes and molecular fragments that engage with the immune system. These epitopes include 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).

Bioinformatics is an indispensable ally in advancing our understanding of immunology and vaccine development by providing the means to predict and strategically harness epitopes. Over the past decade, a comprehensive set of tools and methods has been developed across various facets of the immune response, encompassing B-cell, CTL, and HTL epitope prediction (7).

Although considerable efforts have been made to harness Artificial Intelligence in the multi-epitope sequence design process, there is still a notable gap in the field of bioinformatics tools (8). Several tools and algorithms have been developed, each targeting specific aspects of epitope prediction, such as B-cell, CTL, or HTL epitopes (9,10). These tools are available through heterogeneous interfaces and most of them lack interoperability features. Consequently, researchers face the challenge of combining the functionalities of several standalone tools, each with different interface and input requirements, to design multi-epitope vaccine sequences. In this context, the need to orchestrate these tools through a unified and user-friendly interface would offer a highly desirable potential for innovation and advancement in epitope-based vaccine development.

68 This work presents VaccineDesigner, a Web-based tool that aims to bring flexible and rational
69 epitope-based vaccine design applications to a broader audience. VaccineDesigner offers a
70 streamlined and user-friendly experience, requiring researchers to provide nothing more than
71 FASTA files containing their meticulously selected proteins. This simplified input process
72 eliminates the need for intricate programming and data manipulation. Under the hood,
73 VaccineDesigner processes protein sequence data and seamlessly integrates the predictive
74 prowess of B-cell, CTL, and HTL epitope methodologies.

75

76 **MATERIAL AND METHODS**

77 The main components of the workflow involve epitope prediction for B-cell and T-cell
78 Lymphocytes (Fig. 1), as crucial indicators of protein antigenicity (9). The application leverages
79 Bepipred for B-cell epitope prediction (10), NetMHCpan for MHC class I and II (11), or
80 Consensus IEDB tool for binding prediction (12). The predicted epitopes undergo subsequent
81 filtering steps, assessing their antigenicity, toxicity, and allergenicity according to user-defined
82 criteria, as shown below.

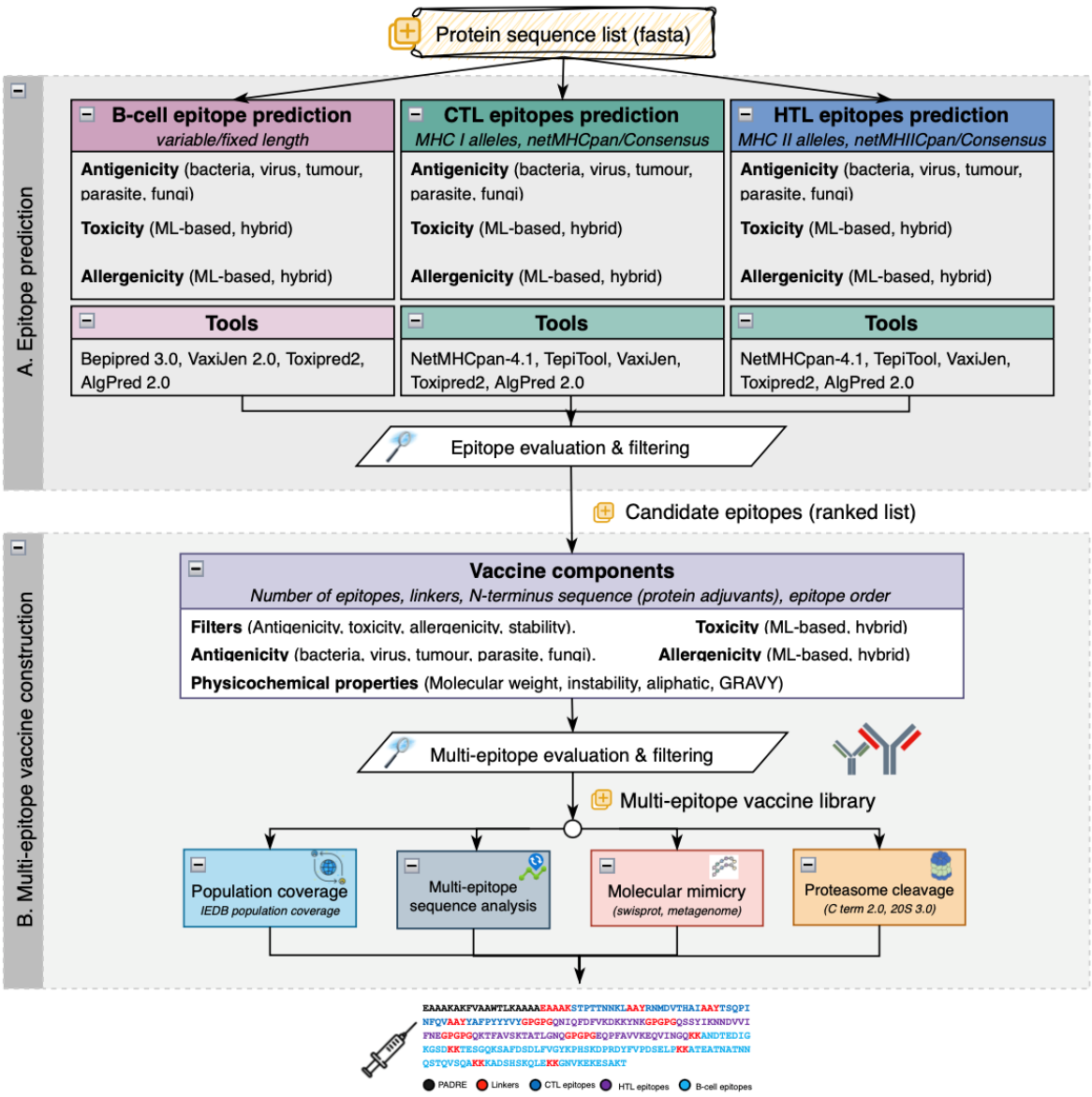

Figure 1

Epitope Prediction

VaccineDesigner starts by uploading a FASTA protein file with user-selected proteins (Fig. 1). Users are then prompted to select a protein of interest to proceed with epitope prediction. In the VaccineDesigner workflow, the epitope prediction module 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.

92 *B-cell epitope prediction.* B-cell epitope prediction uses BepiPred 3.0. BepiPred systematically
93 scans protein sequences and assesses key physicochemical properties and sequence patterns
94 to identify regions that can induce robust antibody responses (10). VaccineDesigner uses the
95 amino acid scores and generates B-cell epitopes, adhering to user-defined epitope length
96 parameters. Two types of analysis can be deployed. The first type involves the formation of B-
97 cell epitopes in high-scoring regions with no upper length limit but with low length, tuned by
98 the *Minimum Epitope Length* parameter. The *Subthreshold Amino Acid Inclusion Count*
99 parameter determines the maximum number of amino acids between predicted epitopes to
100 be included in the output, allowing the merging of separate epitopes into a larger epitope
101 region. To avoid low-scoring amino acids from being included in the results, VaccineDesigner
102 uses the *Secondary Threshold* parameter that sets a minimum amino acid score to be included
103 in the epitope prediction. The second type of analysis pertains to the generation of B-cell
104 epitopes of user-defined lengths. Upon user input, a scanning window of the specified epitope
105 length assesses the scores of all amino acids. Epitopes are defined as continuous regions of 9
106 or more amino acids, wherein each position possesses a score equal to or exceeding the
107 threshold value. The B-cell epitopes are presented in ranked order based on the mean
108 Bepipred score of the epitope region, arranged in decreasing order.

109 *CTL epitope prediction*. CTL epitope prediction in VaccineDesigner is implemented by
110 NetMHCpan 4.1 (11) or the Consensus Method by IEDB Class I prediction (12). NetMHCpan is
111 an algorithm renowned for its accuracy in identifying high-affinity regions for MHC class I
112 alleles selected by the user. The output includes both strong and weak binders counts along
113 with detailed information about the associated MHC alleles. On the other hand, the IEDB
114 Consensus method stands as a reputable algorithm in forecasting CTL epitopes. This method
115 seamlessly integrates various bioinformatics tools, harnessing their combined capabilities to
116 elevate the accuracy of epitope prediction. The CTL epitopes are presented in descending
117 order based on the number of alleles interacting with each respective epitope.

118 *HTL epitope prediction.* The HTL epitope prediction module relies on NetMHCIIpan 4.0, that
119 employs the same machine learning framework as NetMHCpan 4.1. NetMHCIIpan identifies
120 regions characterized by strong binding affinities for MHC class II molecules, customizable to
121 specific user-defined MHC class II alleles (11). The results, akin to CTL epitopes, include strong
122 and weak binders counts, along with details of the MHC alleles. Additionally, HTL epitope

prediction can be executed using the IEDB Consensus method for class II binding prediction, employing the same framework as the Class I module. HTL epitopes are similarly ranked based on the number of interacting alleles, mirroring the approach used for CTL epitopes.

Epitope Evaluation and Filtering

After epitope prediction, VaccineDesigner includes an optional step that performs epitope evaluation and filtering with VaxiJen (13), ToxinPred (14), and AlgPred (15). VaccineDesigner ensures the meticulous selection of epitopes with exceptional antigenicity and safety profiles, adhering to user-defined thresholds. With VaxiJen, a software tool known for its high prediction accuracy rates, VaccineDesigner assesses the antigenicity of each epitope, prioritizing highly immunogenic candidates. ToxinPred identifies potential toxins within epitope sequences, serving as a crucial safety measure. AlgPred predicts allergenicity, minimizing the risk of allergic reactions in vaccine recipients. This comprehensive evaluation guarantees the selection of epitopes based on user-defined criteria for antigenicity and safety, thereby improving the overall efficacy and safety of candidate multi-epitope constructs.

Multi-epitope Vaccine Sequences

VaccineDesigner includes a module that employs epitope predictions to craft multi-epitope vaccine sequences. Using the most prominent epitopes, VaccineDesigner builds larger constructs by merging these epitopes with appropriate linker sequences (default: GPGPG/HTL, AAY/CTL, KK/B-cell) (3,16–21). These multi-epitope constructs are known to mitigate junctional immunogenicity and efficient epitope separation and presentation to T-cell and B-cell receptors (3). The analysis is implemented in a two-step procedure:

Vaccine Sequences Construction. VaccineDesigner grants control over the development of multi-epitope vaccine sequences. Users are prompted to define the number of epitopes for B-cell, CTL, and HTL, while also specifying user-defined linker sequences (18) between the epitopes and the N-terminus sequence. A variety of different options of protein adjuvants are provided for the user to select and include in the multi-epitope vaccine sequences (22–25).

151 Additionally, the order of epitope combination can be defined, enabling precise customization,
152 such as arranging B-cell, CTL, and HTL epitopes in a specific sequence, exemplified as B-C-H
153 (B cell - CTL - HTL). VaccineDesigner enables the generation of candidate vaccine sequences,
154 which form a versatile library constructed through epitope sequence combinations.
155 Researchers have the flexibility to explore diverse epitope combinations, with the only
156 constraint being a maximum limit of five epitopes from each category (B-cell, CTL, and HTL).

157 *Vaccine Sequences Evaluation and Selection*. In the final step, the constructed candidate vaccine
158 sequence library undergoes rigorous evaluation. The algorithm assesses each sequence
159 against predefined filters, including VaxiJen for antigenicity, ToxinPred for toxin prediction,
160 Algpred for allergenicity, and ProtParam for sequence analysis and stability (26).

161 User-defined parameters, such as the desired number of vaccine sequences, serve as
162 thresholds that are used to keep track of the number of candidates that meet all quality criteria.
163 The algorithm exports the final multi-epitope sequence constructs once the preferred number
164 of candidates is reached. This dynamic approach ensures the selection of vaccine sequences
165 that meet strict standards of immunogenicity, safety, and biochemical attributes, offering
166 researchers the most promising candidates for further development and validation. Following
167 this, the final sequences are ranked based on the collective sum of individual rankings of
168 antigenicity, toxicity, allergenicity, and stability in decreasing order, facilitating the
169 identification of the optimal sequences for further analysis.

170 Users can evaluate their chosen sequence for several parameters, including human population
171 coverage, molecular mimicry, proteasome cleavages, and re-analysis for epitope prediction on
172 the newly generated sequence.

173 The population coverage analysis is enabled for human using the IEDB Population Coverage
174 (27) algorithm, which gets user-selected alleles and evaluates the coverage of CTL and HTL
175 epitopes on the vaccine sequence across regions defined by the user.

176 To address molecular mimicry users can conduct protein similarity searches against SwissProt
177 (28) or metagenomic proteins (env_nr) using the BLASTp (29) algorithm. Users can determine
178 whether the vaccine sequence shares similar domains with a protein, thereby ensuring its
179 antigenicity and minimizing the risk of autoimmune reactions.

180 Furthermore, users can evaluate their preferred vaccine sequence using the NetChop 3.1 (30)
181 module, which identifies proteasome cleavages on protein sequences. Users have the flexibility
182 to adjust the default threshold and algorithm settings for the module. Upon execution,
183 NetChop generates results in tabular format and a comprehensive visualization is provided,
184 depicting the positions of epitopes and proteasome cleavages together.

185 Finally, users have the option to evaluate the multi-epitope vaccine sequence for the presence
186 of the initial epitopes or new ones. Users can select the method for analysis within this module.
187 Upon execution, the results include a data table listing new epitopes, a visual plot depicting
188 new epitopes on the sequence, and a comparison table between the initial epitopes and the
189 predicted ones, along with any mismatches identified between them.

190 **RESULTS**

191 **Multi-epitope Vaccine Sequences Prediction Example**

192 To demonstrate VaccineDesigner capabilities, we implemented an example scenario using the
193 Atl (Bifunctional Autolysin, UniProt: P0C5Z8) and the IsdA, (Iron-regulated surface determinant
194 protein A, UniProt: Q7A152) protein sequences from *Staphylococcus aureus* as targets. These
195 protein sequences can be selected from the "Upload" tab panel after uploading the
196 corresponding FASTA file for of each epitope prediction module.

197 *B-cell epitope prediction.* 123 B-cell epitopes were predicted from Atl and 136 from IsdA using
198 Bepipred scores with a standard amino acid length of 10. The parameters used for epitope
199 prediction were set to "Analysis 2" to obtain standard-length epitopes, with the threshold set
200 to 0.1512, the default value. 13 of these epitopes were considered candidates for inclusion in
201 the vaccine sequences based on their quantified toxicity, allergenicity and antigenicity
202 measures.

203 *CTL epitope prediction*. 96 CTL epitopes from Atl and 18 from IsdA were predicted as strong or
204 weak binders to MHC class I alleles, including *HLA-A0101, HLA-A0102, HLA-A0103, HLA-A0104,*
205 *HLA-A0106, HLA-A0107, HLA-A0108* and *HLA-A0109*. These predicted epitopes underwent
206 toxicity, allergenicity and antigenicity filters, leading to the identification of six promising
207 candidates.

208 *HTL epitope prediction*. The number of epitopes considered strong or weak binders to MHC
209 class II alleles, including *DBR1_0101*, *DBR1_0102*, *DBR1_0103*, *DBR1_0104*, *DBR1_0105*,
210 *DBR1_0106*, *DBR1_0107* and *DBR1_0108* totalled 193 from Atl and 50 from IsdA. After applying
211 the filtering step for allergenicity and toxicity, the resulting epitopes were reduced to 19.

212 *Multi-epitope vaccine sequences*. After uploading the tabular-formatted results of the
213 candidate epitopes, the sequences were combined following the C-H-B order (CTL, HTL and
214 B-cell). The number of B cell epitopes included in the vaccine sequence was set to 4 and the
215 number of both CTL and HTL epitopes included was set to 3, resulting in 864 vaccine sequences.
216 Table 1a lists the first 10 sequences that passed the antigenicity and toxicity filtering along
217 with their properties (Table 1b).

218 End-to-end execution is implemented without special server-side memory and processing
219 requirements and time-efficiently, as each building module requires about one minute of
220 runtime with minor deviations depending on the number of input sequences.

221

222 **Table 1a**

Seq. ID	Sequences
1	AAYQVNSSINDYAAYVSDNKSQQTAAYYLRSHNYSYGPGPGINGEISYMKNNYQNAGPGPGMDDYMQHPGKVIKQNGPGPGNGEISYMKNNYQNAFKKSTSTTAPKTNKKDTRANQSATTKKSDNKSQQTNKKKVSDNKSQQTN
2	AAYQVNSSINDYAAYVSDNKSQQTAAYYLRSHNYSYGPGPGINGEISYMKNNYQNAGPGPGMDDYMQHPGKVIKQNGPGPGNGEISYMKNNYQNAFKKSTSTTAPKTNKKSDNKSQQTNKKKDTRANQSATTKKVSDNKSQQTN
3	AAYQVNSSINDYAAYVSDNKSQQTAAYYLRSHNYSYGPGPGINGEISYMKNNYQNAGPGPGMDDYMQHPGKVIKQNGPGPGNGEISYMKNNYQNAFKKDTRANQSATTKKSTSTTAPKTNKKSDNKSQQTNKKKVSDNKSQQTN
4	AAYQVNSSINDYAAYVSDNKSQQTAAYYLRSHNYSYGPGPGINGEISYMKNNYQNAGPGPGMDDYMQHPGKVIKQNGPGPGNGEISYMKNNYQNAFKKDTRANQSATTKKSDNKSQQTNKKKSTSTTAPKTNKKVSDNKSQQTN
5	AAYQVNSSINDYAAYVSDNKSQQTAAYYLRSHNYSYGPGPGINGEISYMKNNYQNAGPGPGMDDYMQHPGKVIKQNGPGPGNGEISYMKNNYQNAFKKDTRANQSATTKKSDNKSQQTNKKKVSDNKSQQTNKKSTSTTAPKTN

223

224 **Table 1b**

Seq. ID	Vaxijen Score	Toxinpred Score	Algpred Score	Instability index	GRAVY score
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

225

226 VaccineDesigner stands out from similar platforms in several ways. Firstly, Firstly, reverse-
227 vaccinology platforms like Vaxign2 (31) focus on predicting vaccine candidates from pathogen
228 genomes. Our tool differs since it offers users the flexibility to analyze proteins of their
229 preference, facilitating the prediction of immunogenic regions for tailored vaccine
230 development. Secondly, compared to the iVAX Suite (32), we include B-cell epitope prediction
231 and utilize a broader array of alleles, offering a more comprehensive assessment. Additionally,
232 our platform is freely accessible and supports multi-epitope vaccine design, unlike Vacceed
233 (33), which is limited to command-line use for automating the process of identifying vaccine
234 candidates from thousands of protein sequences of the target pathogen. Furthermore, our
235 integrated workflow unifies in a user-friendly interface the process of epitope vaccine
236 development, distinguishing us from the Immune Epitope Database and Analysis Resource
237 (IEDB) (34).

238 **DISCUSSION**

239 In the evolving landscape of infectious diseases, epitope-based vaccine design has emerged
240 as a pivotal strategy, offering rationalized detection of potent immune responses. Epitope-
241 based methods focus on the identification of antigenic determinants in pathogenic proteins,
242 enabling the development of vaccines that reduce the risk of adverse reactions and ensure
243 broad coverage across various pathogenic strains. VaccineDesigner serves this goal by
244 seamlessly integrating B-cell, CTL, and HTL epitope prediction tools and quality assessment
245 algorithms, and by generating reliable multi-epitope vaccine sequences that are further
246 assessed for population coverage and molecular mimicry.

247 VaccineDesigner has a user-friendly Web interface that streamlines processes for the discovery
248 and optimization of multi-epitope constructs and relies on a modular architecture that ensures
249 continuous improvements toward immunogenic and safer vaccines. Therefore, as our
250 knowledge of the complex molecular interactions evolves VaccineDesigner can both benefit
251 by integrating new functionalities and promote incremental research through knowledge
252 synthesis, broad usage, and adaptability in real-world applications.

253 **DATA AVAILABILITY**

254 VaccineDesigner is an open-source tool freely available under the academic license at
255 https://github.com/BiolApps/VaccineDesigner (35)

256 **REFERENCES**

257 1. Thomas, S., and Luxon, B. A. (2013). Vaccines based on structure-based design provide
258 protection against infectious diseases. Expert review of vaccines, 12(11), 1301–1311.

259 2. Webster, R. G., and Govorkova, E. A. (2014). Continuing challenges in influenza. Annals
260 of the New York Academy of Sciences, 1323(1), 115–139.

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

264 4. Potocnakova, L., Bhide, M., and Pulzova, L. B. (2016). An Introduction to B-Cell Epitope
265 Mapping and In Silico Epitope Prediction. Journal of immunology research, 2016,
266 6760830.

267 5. Shepherd, F. R., and McLaren, J. E. (2020). T Cell Immunity to Bacterial Pathogens:
268 Mechanisms of Immune Control and Bacterial Evasion. International journal of
269 molecular sciences, 21(17), 6144.

270 6. Zhang L. (2018). Multi-epitope vaccines: a promising strategy against tumors and viral
271 infections. Cellular and molecular immunology, 15(2), 182–184.

272 7. Feng, Y., Jiang, H., Qiu, M., Liu, L., Zou, S., Li, Y., Guo, Q., Han, N., Sun, Y., Wang, K., Lu, L.,
273 Zhuang, X., Zhang, S., Chen, S., and Mo, F. (2021). Multi-Epitope Vaccine Design Using
274 an Immunoinformatic Approach for SARS-CoV-2. Pathogens (Basel, Switzerland), 10(6),
275 737.

276 8. Yang, Z., Bogdan, P. and Nazarian, S. (2021) An in silico deep learning approach to multi-
277 epitope vaccine design: a SARS-CoV-2 case study. Sci Rep 11, 3238.

278 9. Raoufi, E., Hemmati, M., Eftekhari, S., Khaksaran, K., Mahmodi, Z., Farajollahi, M. M., and
279 Mohsenzadegan, M. (2020). Epitope Prediction by Novel Immunoinformatics Approach:
280 A State-of-the-art Review. International journal of peptide research and therapeutics,
281 26(2), 1155–1163.

282 10. Clifford, J. N., Høie, M. H., Deleuran, S., Peters, B., Nielsen, M., and Marcatili, P. (2022).
283 BepiPred-3.0: Improved B-cell epitope prediction using protein language models.
284 Protein science: a publication of the Protein Society, 31(12), e4497.

285 11. Reynisson, B., Alvarez, B., Paul, S., Peters, B., and Nielsen, M. (2020). NetMHCpan-4.1 and
286 NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent
287 motif deconvolution and integration of MS MHC eluted ligand data. Nucleic acids
288 research, 48(W1), W449–W454.

289 12. Paul, S., Sidney, J., Sette, A., and Peters, B. (2016). TepiTool: A Pipeline for Computational
290 Prediction of T Cell Epitope Candidates. Current protocols in immunology, 114, 18.19.1–
291 18.19.24.

292 13. Doytchinova, I. A., and Flower, D. R. (2007). VaxiJen: a server for prediction of protective
293 antigens, tumour antigens and subunit vaccines. BMC bioinformatics, 8, 4.

294 14. Sharma, N., Naorem, L. D., Jain, S., and Raghava, G. P. S. (2022). ToxinPred2: an improved
295 method for predicting toxicity of proteins. Briefings in bioinformatics, 23(5), bbac174.

296 15. Sharma, N., Patiyal, S., Dhall, A., Pande, A., Arora, C., and Raghava, G. P. S. (2021). AlgPred
297 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes.
298 Briefings in bioinformatics, 22(4), bbaa294.

299 16. Yang, Y., Sun, W., Guo, J., Zhao, G., Sun, S., Yu, H., Guo, Y., Li, J., Jin, X., Du, L., Jiang, S.,
300 Kou, Z., and Zhou, Y. (2015). In silico design of a DNA-based HIV-1 multi-epitope vaccine
301 for Chinese populations. Human vaccines & immunotherapeutics, 11(3), 795–805.

302 17. Yano, A., Onozuka, A., Asahi-Ozaki, Y., Imai, S., Hanada, N., Miwa, Y., and Nisizawa, T.
303 (2005). An ingenious design for peptide vaccines. Vaccine, 23(17-18), 2322–2326.

304 18. Chen, X., Zaro, J. L., and Shen, W. C. (2013). Fusion protein linkers: property, design and
305 functionality. Advanced drug delivery reviews, *65*(10), 1357–1369.

306 19. Livingston, B., Crimi, C., Newman, M., Higashimoto, Y., Appella, E., Sidney, J., and Sette,
307 A. (2002). A rational strategy to design multiepitope immunogens based on multiple Th
308 lymphocyte epitopes. Journal of immunology (Baltimore, Md. : 1950), 168(11), 5499–
309 5506.

310 20. Li, X., Guo, L., Kong, M., Su, X., Yang, D., Zou, M., Liu, Y., and Lu, L. (2015). Design and
311 Evaluation of a Multi-Epitope Peptide of Human Metapneumovirus. Intervirology, 58(6),
312 403–412.

313 21. Willimsky, G., Beier, C., Immisch, L., Papafotiou, G., Scheuplein, V., Goede, A., Holzhütter,
314 H. G., Blankenstein, T., and Kloetzel, P. M. (2021). In vitro proteasome processing of neo-
315 splicetopes does not predict their presentation in vivo. eLife, 10, e62019.

316 22. Díaz-Dinamarca, D. A., Salazar, M. L., Castillo, B. N., Manubens, A., Vasquez, A. E., Salazar,
317 F., and Becker, M. I. (2022). Protein-Based Adjuvants for Vaccines as Immunomodulators
318 of the Innate and Adaptive Immune Response: Current Knowledge, Challenges, and
319 Future Opportunities. Pharmaceutics, 14(8), 1671.

320 23. Toussi, D. N., Carraway, M., Wetzler, L. M., Lewis, L. A., Liu, X., and Massari, P. (2012). The
321 amino acid sequence of Neisseria lactamica PorB surface-exposed loops influences Toll-
322 like receptor 2-dependent cell activation. Infection and immunity, 80(10), 3417–3428.

323 24. Ghaffari-Nazari, H., Tavakkol-Afshari, J., Jaafari, M. R., Tahaghoghi-Hajghorbani, S.,
324 Masoumi, E., and Jalali, S. A. (2015). Improving Multi-Epitope Long Peptide Vaccine
325 Potency by Using a Strategy that Enhances CD4+ T Help in BALB/c Mice. PloS one,
326 10(11), e0142563.

327 25. Del Giudice G. (1994). Hsp70: a carrier molecule with built-in adjuvanticity. Experientia,
328 50(11-12), 1061–1066.

329 26. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and
330 Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server.
331 Methods in molecular biology (Clifton, N.J.), 112, 531–552.

332 27. Dhanda, S. K., Mahajan, S., Paul, S., Yan, Z., Kim, H., Jespersen, M. C., Jurtz, V., Andreatta,
333 M., Greenbaum, J. A., Marcatili, P., Sette, A., Nielsen, M., and Peters, B. (2019). IEDB-AR:
334 immune epitope database-analysis resource in 2019. Nucleic acids research, 47(W1),
335 W502–W506.

336 28. Bairoch, A., and Apweiler, R. (2000). The SWISS-PROT protein sequence database and
337 its supplement TrEMBL in 2000. Nucleic acids research, 28(1), 45–48.

338 29. McGinnis, S., and Madden, T. L. (2004). BLAST: at the core of a powerful and diverse set
339 of sequence analysis tools. Nucleic acids research, 32(Web Server issue), W20–W25.

340 30. Saxová, P., Buus, S., Brunak, S., and Keşmir, C. (2003). Predicting proteasomal cleavage
341 sites: a comparison of available methods. International immunology, 15(7), 781–787.

342 31. Ong, E., Cooke, M. F., Huffman, A., Xiang, Z., Wong, M. U., Wang, H., Seetharaman, M.,
343 Valdez, N., and He, Y. (2021). Vaxign2: the second generation of the first Web-based
344 vaccine design program using reverse vaccinology and machine learning. Nucleic acids
345 research, 49(W1), W671–W678.

346 32. Moise, L., Gutierrez, A., Kibria, F., Martin, R., Tassone, R., Liu, R., Terry, F., Martin, B., and
347 De Groot, A. S. (2015). iVAX: An integrated toolkit for the selection and optimization of

348 antigens and the design of epitope-driven vaccines. Human vaccines &
349 immunotherapeutics, 11(9), 2312–2321.

350 33. Goodswen, S. J., Kennedy, P. J., and Ellis, J. T. (2014). Vacceed: a high-throughput in silico
351 vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse
352 vaccinology. Bioinformatics (Oxford, England), 30(16), 2381–2383.

353 34. Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., Wheeler, D.
354 K., Sette, A., and Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update.
355 Nucleic acids research, 47(D1), D339–D343.

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

358

359 **TABLE AND FIGURES LEGENDS**

360 **Table 1a**. The top five multi-epitope sequences candidates that passed the filtering step of
361 antigenicity and toxicity alongside with their respective ID.

362

363 **Table 1b.** The top five multi-epitope sequence candidates that passed the filtering step of
364 antigenicity and toxicity alongside with their antigenicity, toxicity, allergenicity, GRAVY scores
365 and instability indexes.

366 **Figure 1**. The main functional components of the streamlined process implemented by
367 VaccineDesigner.
368 A. Epitope prediction module initiates with the upload of a fasta protein file, performs epitope
369 prediction (B cell, CTL, HTL) and ends up with the evaluation and filtering of epitope sequences
370 based on their physicochemical properties. B. Multi-epitope sequence generation module
371 takes as input the CSV results of epitope sequences and generates vaccine sequences based
372 on user-defined parameters. The next step is the evaluation and selection of the vaccine
373 sequences that meet the quality criteria. Finally, the user are able to evaluate the vaccine
374 sequence of their preference for population coverage, molecular mimicry, proteasome
375 cleavages and re-analyse with epitope prediction modules.