JOURNAL OF TROPICAL LIFE SCIENCE
2024, Vol. 14, No. 3, 501 – 508
http://dx.doi.org/10.11594/jtls.14.03.07

Research Article

Design of Epitope-Based Vaccine Against SARS-CoV-2: An Immuno-Informatics Study

Kavana Hafil Kusuma [1], Muhammad Hermawan Widyananda [1], Wirdatun Nafisah [1], Rahmat Grahadi [1], Yuyun Ika Christina [1], Dinia Rizqi Dwijayanti [1], Dewi Mustikaningtyas [2], Nashi Widodo [1], Muhammad Sasmito Djati [1]*

[1] Department of Biology, Faculty of Mathematics and Natural Sciences, Brawijaya University, Malang 65145, Indonesia
[2] Department of Biology, Faculty of Mathematics and Natural Sciences, State University of Semarang, Semarang 50229, Central Java, Indonesia

Article history:
Submission September 2023
Revised February 2024
Accepted February 2024

**Corresponding author:*
E-mail: msdjati@ub.ac.id

ABSTRACT

This study aimed to develop an epitope-based vaccine of SARS-CoV-2 S protein through an immuno-informatics study. The whole genome of SARS-CoV-2 sequences was obtained from the GISAID database and then trimmed to obtain the S protein sequences. The alignment was done by Clustal-W of MEGA software. Epitope prediction and modeling were performed by Discotope BepiPred and the PepFold3 web server. The allergic responses and physicochemical characteristics of predicted epitopes were analyzed using the AlgPred and ProtParam from ExPASy. Molecular docking and dynamic stimulation were performed using AutoDock Vina and YASARA. Biovia Discovery Studio 2019 was used to visualize the molecular docking results. The study predicted 3 potential epitopes, including ‘GDEVRQIAPGQTGKIADYNYKLP’ (epitope 1), ‘YTMSLGAENSVAYSNN’ (epitope 2), and ‘VNNSYECDIPI’ (epitope 3) located in the spike head specifically RBD region. The epitopes did not show an allergen reaction based on IgE epitope mapping. The suitable overexpression for the host of epitopes was mammalian cells. Only epitopes 1 and 2 were stable (instability index above 40). Epitopes 1, 2, and 3 interacted with BCR with binding affinity values -6.6, -7.8, and -7.5 kcal/mol. Epitope 2 wasere stable when interacting with the BCR. Therefore, three epitopes were predicted to have high potency as the SARS-CoV-2 epitope-based vaccine.

*Keywords**:** Epitope, Immuno-informatic, SARS-CoV-2, Vaccine*

Introduction

Severe Acute Respiratory Syndrome Coronavirus-2, also known as SARS-CoV-2, is a new variant of coronavirus that arose from the Wuhan seafood market in late 2019 and has spread rapidly and globally become a pandemic. The outbreak caused more than 772 million cases and 6.9 million deaths estimated worldwide, reported by WHO on 20 December 2023. SARS-CoV-2 is classified as beta coronavirus strain B that can exploit angiotensin-converting enzyme 2 (ACE2), a receptor used to infect human cells [1]. The structural protein of coronavirus includes spike (S), envelope (E), membrane (M), and nucleocapsid (N) protein. The coronavirus spike protein consists of two domains: the S1 domain acts as a receptor-binding domain, and the S2 domain mediates the fusion of the membrane [2].

The interaction of S protein and ACE2, an integral component of the angiotensin renin-2 system, becomes critical in mediating SARS-CoV entry into cells. Both spike proteins of SARS-CoV and SARS-CoV-2 share an almost identical 3D structure of RBD that preserves van der Waals forces, especially at Asn501 against ACE2 [2, 3]. However, the interaction mechanisms of ACE2-SARS-CoV and ACE2-SARS-CoV-2 are non-identical. S glycoprotein of SARS-CoV modulates virus recognition and fusion, where the RBD of

How to cite:
Kusuma KH, Widyananda MH, Nafisah W, et al. (2024) Design of Epitope-Based Vaccine Against SARS-CoV-2: An Immuno-Informatics Study. Journal of Tropical Life Science 14 (3): 501 – 508. doi: 10.11594/jtls.14.03.07.

SARS-CoV-2 proposed concurrent binding of an ACE2 dimer by two S glycoprotein trimers [4]. Previous studies stated that SARS-CoV and SARS-CoV-2 have different and unique strategies for stabilizing their binding to ACE2 [5]. The interaction of its two, ACE2 and S proteins, mediates the initiation of several events cascade that terminate in severe acute respiratory syndrome.

The essential role of the S protein could become a target for a blocking strategy against its interaction with ACE2. Epitope-based vaccines become one of the most promising strategies since no anti-viral drug is approved to treat this virus. Compared with traditional vaccines, it has allergic and autoimmune responses and side effects, their stability is low, and they need to be stored at low temperatures. Moreover, the epitope-based vaccine targets specific epitopes avoiding the response of allergic and autoimmune [6] and activates the immune response to the immunogenic epitopes of the SARS-CoV-2 S protein [7]. Therefore, the study aimed to develop an S protein epitope-based vaccine for SARS-CoV-2 therapy through immuno-informatics.

Material and Methods

Data retrieval and alignment

The study used all variants of the whole genome of SARS-CoV-2 obtained from the GISAID database of the National Center of Biotechnology Information (https://www.gisaid.-org/) from December 2019 to July 2020. The whole genome was collected from the human sample, which was then trimmed on the spike protein of SARS-CoV-2. The selected amino acid sequences have 28.000-29.000 residues with a lack of N content and a gap in the spike protein. Alignment of protein sequences and structure, both in two dimensions (2D) and three dimensions (3D), was done to identify the spike domain. The alignment was done separately for various types of protein based on their structure using Clustal-W of MEGA X software.

Epitope prediction and epitope structure modeling

The potential epitope was predicted through epitope mapping through the Discotope BepiPred tool from Immune Epitope Data Base and Analysis Resource (IEDB). This was a semi-empirical prediction method using physico-chemical properties of amino acids (hydrophilicity, surface accessibility, beta-turns, and flexibility) and their frequencies of occurrence to predict linear B cell epitopes. The modeling process for the predicted epitope structure was carried out using the PepFold3 web server (https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/) [8].

Allergy response analysis and physicochemical properties

Potential allergen and map IgE epitopes prediction was done using AlgPred software [9]. The physicochemical properties of the epitope and its half-life in various cell types were done by predicting the epitope stability using ProtParam tools from ExPASy. ProtParam is an in-silico method to calculate the physical properties of epitopes [10]. The physicochemical properties include molecular weight, theoretical PI, grand average of hydrophobicity, instability, and aliphatic index were determined.

Molecular docking and dynamic simulation

To simulate the interaction of epitopes against B cell receptors, a molecular docking was done. The 3D epitope structure was modeled using the PEP FOLD 3 webserver (https://bioserv.rpbs-.univ-paris-diderot.fr/services/PEP-FOLD3/). The 3D Fab domain BCR structure (1igy) was retrieved from the RCSB PDB database (https-://www.rcsb.org/). Epitopes were prepared by minimizing conformational energy using the Open Babel plugin integrated into the PyRx software [11]. The preparation of BCR protein was done by deleting water molecules and unnecessary ligands using the Biovia Discovery Studio 2019 software [12]. Molecular docking was carried out using AutoDock Vina integrated with PyRx software [13] with a blind docking method. Biovia Discovery Studio 2019 software was used to visualize the docking results. Then, Yet Another Scientific Artificial Reality Application (YASARA) software was used to simulate molecular dynamics [14] with parameters according to the cell's physiology conditions (pH 7.7, temperature 310K, salt content 0.9%, pressure 1 atm for 20 ns and autosave every 25 ps.

Results and Discussion

Sequences alignment

The alignment has reached 56 countries with a total of 5354 sequences, consisting of Italy,

Greece, Bosnia and Herzegovina, Serbia, Sweden, Denmark, Argentina, Brazil, Canada, England, Indonesia, Iran, Japan, Jordan, Kazakhstan, Malaysia, Mexico, Mongolia, Pakistan, Puerto Rico, Russia, United States of America, Saudi Arabia, Thailand, United Arab Emirates Vietnam, India, Sri Lanka, Czech Republic, Slovakia, Slovenia, Hungary, Germany, Austria, Spain, Portugal, France, Turkey, Lithuania, Latvia, Georgia, Belarus, Australia, New Zealand, Guam, Egypt, Democratic Republic of Congo, Zambia, South Africa, Algeria, Tunisia, Morocco, Nigeria, Ghana, South Korea, China, and Australia. Based on the distribution of data that depends on the number of sequences submitted on the GISAID Database from each country, Europe showed the highest number of submitted data (2211 sequences or 41.2%), followed by Australia and Oceania (1323 sequences, 24.7%), Asia (18%), Africa (13.2%), and America and Pacific (3.3%).

Epitope prediction and modeling

B-cell epitopes provide guidance and are used for the identification and activation of B-cells as the immune response against viral infection. The linear epitopes, antigenicity, hydrophilicity, accessibility of surface, beta-turn, and flexibility prediction were all factors in this recognition. The

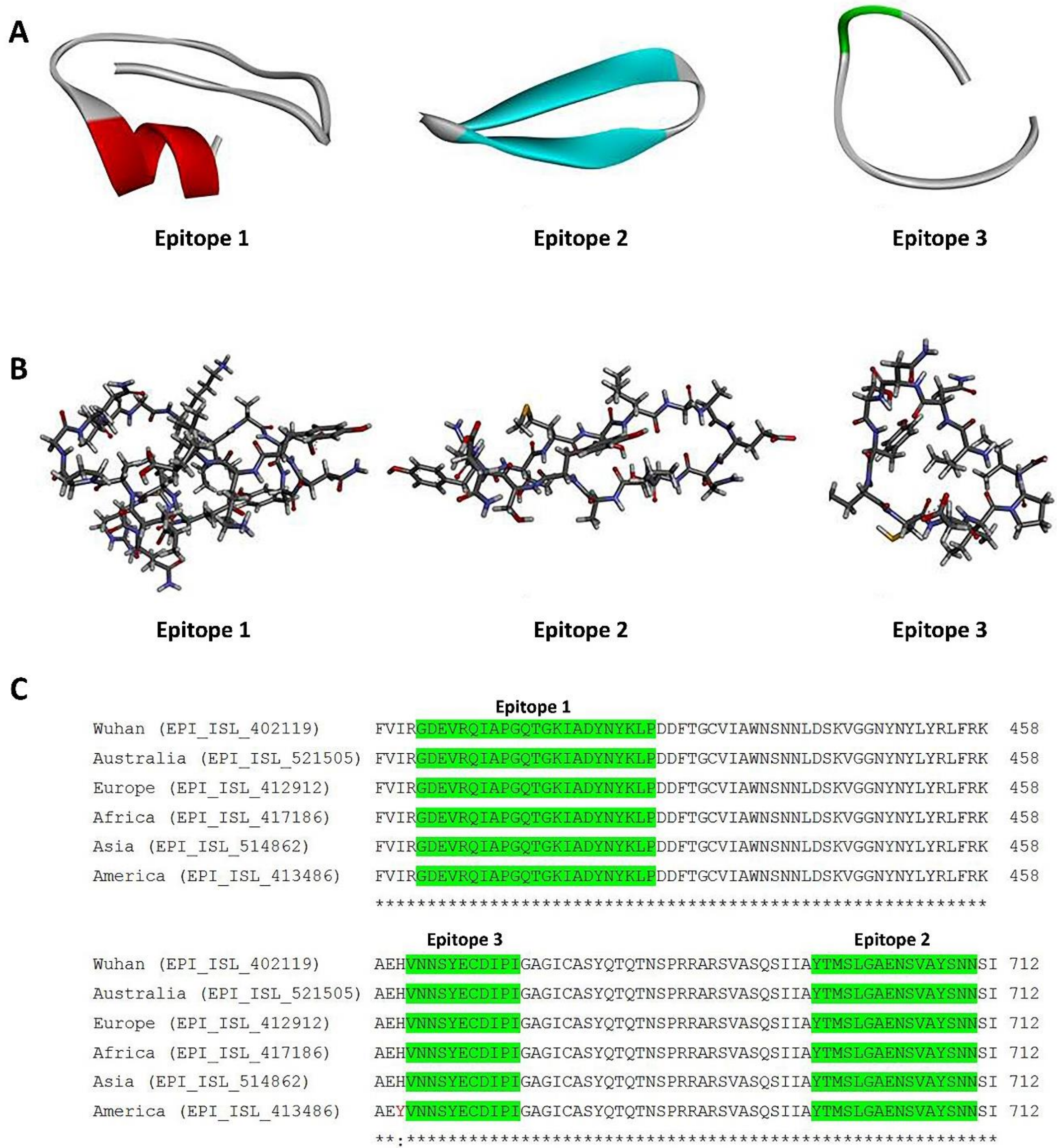

Figure 1. SARS-CoV-2 S protein epitopes in ribbon and stick structure model (A&B). SARS-CoV-2 S protein amino acid sequences in 5 representative continents, the green highlight showed a conserved region of the epitopes (C).

S and M proteins of B-cell epitopes were predicted by default setting methods of Bepipred in IEDB [15]. Based on the results for S protein, 34 linear epitopes were identified. TThe 3D structure of the SARS-CoV-2 S protein was mapped to the identified linear B-cell epitopes, demonstrating that 3 epitopes were conserved regions, including 'GDEVRQIAPGQTGKIADYNYKLP' (epitope 1), 'YTMSLGAENSVAYSNN' (epitope 2), and 'VNNSYECDIPI' (epitope 3) located in the spike head's RBD region, which was displayed as being the most exposed area. (Figure 1). Epitope 1 was reported to considerably overlap with the surface of the binding site of ACE-2 against RBD [16, 17], demonstrating the potential of blocking the mechanism of antibodies against the viral entry to the cells. Based on Noorimotlagh *et al.* [18], epitopes 2 and 3 have similar characteristics to antigenic peptides and the potential to be a vaccine with multiple epitopes.

Allergy response of predicted epitopes

Based on AlgPred analysis, this study found that the epitopes were non-allergenic and lacked an IgE epitope based on IgE epitope mapping analysis. However, the epitopes were identified as an allergen-by-allergen prediction using amino acid and di-epitope composition (Table 1). IgE epitope mapping was used to determine the potential of an allergen of protein due to the responsible of IgE epitope in inducing antibodies that cause allergenic reactions in animal bodies [19]. Allergic response during vaccination often occurs by involving several mechanisms, including the reaction of Fcε receptor-1 type I by mast cell activation, the interaction of IgE antibodies against vaccine particularly, and the occurrence is in minutes or up to four hours [20]. Thus, IgE becomes the most important thing to be considered in vaccine development. These findings suggested that since our novel epitope lacked an IgE epitope, it was not an allergen.

Physicochemical properties

Physical properties analysis of epitopes used to identify an appropriate production cell type. The results of the physicochemical properties of this study suggested that the suitable over-expression of the host of epitope 1 to 3 was in mammals, demonstrating the most stable half-life in this cell (Table 2). The determination of the half-life of an epitope in the different hosts was done according to the "N-end rule" [9]. The molecular weight of epitope 1 was higher than epitopes 2 and 3, caused by a greater number of amino acids. Besides, based on the instability index value, epitopes 1 and 2 are predicted as stable by values of the instability index below 40. However, the instability index value for epitope 3 was greater than 40, which indicated that the epitope

Table 1. Allergic response prediction analysis results

Epitope	Analysis	Result
1	IgE epitope mapping	None
	Amino acid composition	Allergen
	Di-epitope composition	Allergen
2	IgE epitope mapping	None
	Amino acid composition	Allergen
	Di-epitope composition	Allergen
3	IgE epitope mapping	None
	Amino acid composition	Allergen
	Di-epitope composition	Allergen

Table 2. Allergic response analysis results

Characters	Epitope 1	Epitope 2	Epitope 3
Molecular weight	2,533.82 Da	1,720.83 Da	1,266.39 Da
Theoretical PI	6.12	4.00	3.67
The estimated half-life in mammals	30 hours	2.8 hours	100 hours
The estimated half-life in yeast	> 20 hours	10 minutes	> 20 hours
The estimated half-life in *E. coli*	> 10 hours	2 minutes	> 10 hours
Instability index	16.23	26.22	57.07
Aliphatic index	72.17	55.00	97.27
Grand average of hydropathicity	-0.887	-0.412	-0.182

might be unstable and not viable [10].

BCR-peptides interactions

Molecular docking validated the B cell epitopes prediction result by interacting epitopes 1, 2, and 3 with the B Cell Receptor (BCR) fragment. The docking results showed that the binding affinity formed by BCR-Epitope 1 was -6.6 kcal/mol, BCR-epitope 2 -7.8 kcal/mol, and BCR-Epitope 3 -7.5 kcal/mol (Table 3). The amount of energy needed for the ligand to bind to the protein receptor is represented by the binding affinity value. The more negative the binding affinity value, the higher the protein affinity for a ligand [21]. The interaction of epitopes 1, 2, and 3 with BCR formed 20, 25, and 13 hydrogen bonds, respectively (Table 3). Protein-ligand interactions are more stable due in part to hydrogen bonds. The interaction becomes more stable as more hydrogen bonds form [22]. The three epitopes formed hydrogen bonds with the same amino acids, namely Gln38, Glu150, and Thr108 (Figure 2). All three epitopes have low binding affinity and high hydrogen bonds when interacting with BCR. Therefore, the three epitopes were predicted to interact strongly with BCR.

The stability of the BCR-peptide interaction in the human cell was simulated using molecular dynamics simulation using YASARA software. The simulation was performed by adjusting the parameters to resemble the conditions in the human cell with pH 7.8, temperature 310K, pressure of 1 atm, and salt content of 0.9%. Protein-ligand complexes are stable if they have an RMSD value of not more than 4 Å with a slight fluctuation [23, 24]. The BCR-Epitope 1 and BCR-Epitope 3 complexes were constant throughout 10 ns simulations, but the RMSD value increased after 10 ns. Meanwhile, the BCR-Epitope 2 complex was stable, which was shown by the RMSD value in the simulation entirely (Figure 3). The RMSD value of the BCR-Epitope complex showed that epitope 2 was at a stable state when interacting with BCR

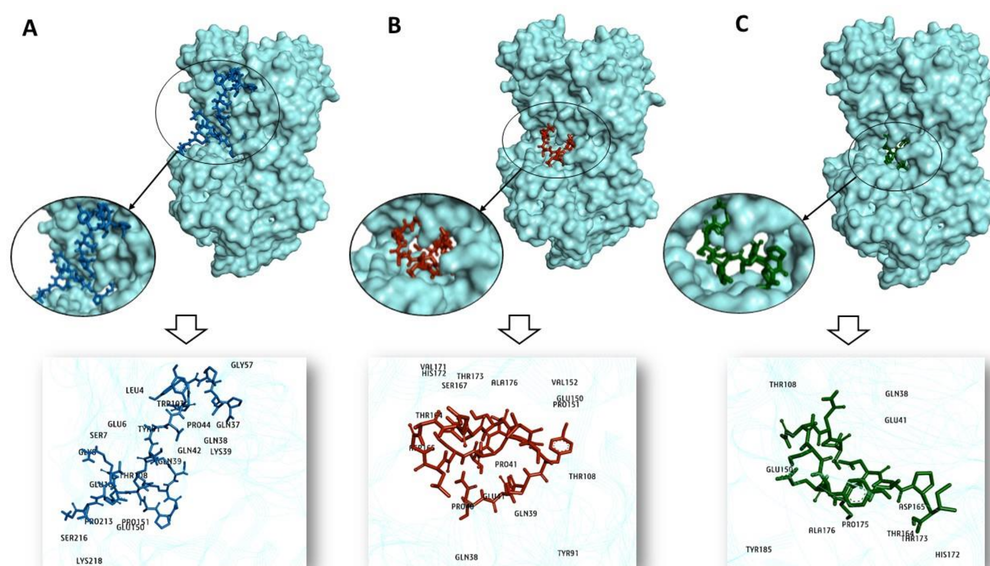

Figure 2. Interaction between BCR (Light blue) and peptides. A. The interaction between BCR and epitope 1 (blue) formed hydrogen bonds with 20 amino acids. B. BCR-epitope 2 (red) formed hydrogen bonds with 17 amino acids. C. BCR-epitope 3 (green) formed hydrogen bonds with 11 amino acids.

Table 3. Binding affinity and the number of hydrogen bonds in the BCR-Peptide complex.

Protein	Epitope	Binding Affinity (Kcal/mol)	Hydrogen bonds
BCR	Epitope 1	-6.6	20
	Epitope 2	-7.8	25
	Epitope 3	-7.5	13

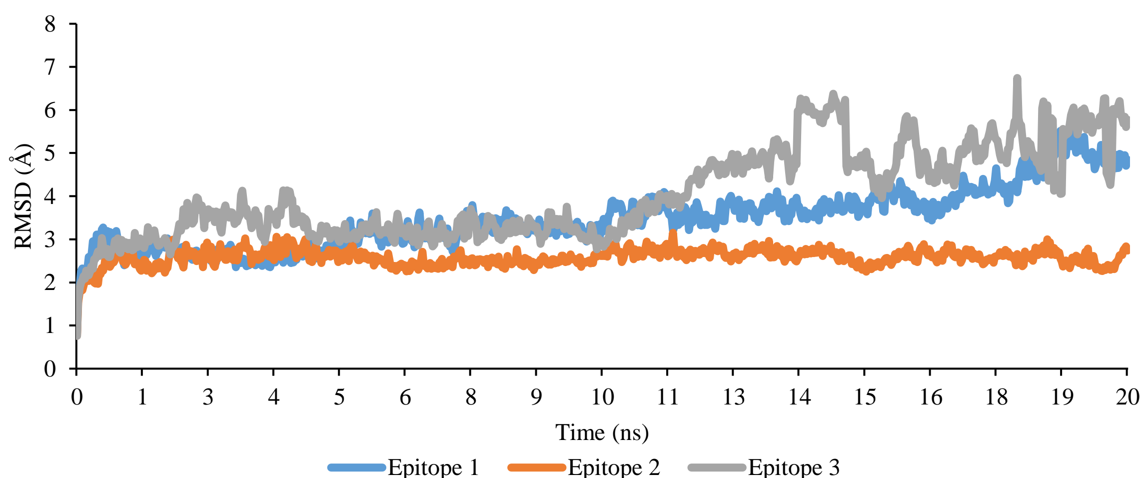

Figure 3. The stability of the BCR-peptide interaction represented by the RMSD value

and had potential as an epitope-based vaccine candidate.

Conclusion

This study predicted three epitopes from SARS-CoV-2 Spike Protein sequences. These epitopes were conserved, not allergen, and had a long half-life in a mammalian cell. However, based on their stability, only epitope 2 was predicted as stable and could interact strongly with BCR. Additional research is required to confirm the viability and the potency of the epitope to be a vaccine candidate.

Acknowledgment

The authors thank the National Research and Innovation Agency (BRIN) and Indonesia Endowment Funds for Education (LPDP) for funding this research through the Research and Innovation Funding Program for Advanced Indonesia (grant no. 1758/UN.C10/TU/2022).

References

1. Zhou P, Yang X Lou, Wang XG et al. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579 (7798): 270–273. doi: 10.1038/s41586-020-2012-7.
2. Satarker S, Nampoothiri M (2020) Structural Proteins in Severe Acute Respiratory Syndrome Coronavirus-2. Archives of Medical Research 51 (6): 482–491. doi: 10.1016/j.arcmed.2020.05.012.
3. Xu X, Han M, Li T et al. (2020) Effective treatment of severe COVID-19 patients with tocilizumab. Proceedings of the National Academy of Sciences of the United States of America 117 (20): 10970–10975. doi: 10.1073/pnas.2005615117.
4. Yang J, Petitjean SJL, Koehler M et al. (2020) Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. Nat Commun. doi: 10.1038/s41467-020-18319-6
5. Ali A, Vijayan R (2020) Dynamics of the ACE2–SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. Sci Rep. doi: 10.1038/s41598-020-71188-3
6. Di Natale C, La Manna S, De Benedictis I et al. (2020) Perspectives in Peptide-Based Vaccination Strategies for Syndrome Coronavirus 2 Pandemic. Front Pharmacol. doi: 10.3389/fphar.2020.578382
7. Chukwudozie OS, Gray CM, Fagbayi TA et al. (2021) Immuno-informatics design of a multimeric epitope peptide-based vaccine targeting SARS-CoV-2 spike glycoprotein. PLoS One. doi: 10.1371/journal.pone.0248061
8. Lamiable A, Thevenet P, Rey J et al. (2016) PEP-FOLD3: faster denovo structure prediction for linear peptides in solution and in complex. Nucleic Acids Research 44 (1): W449–W454. doi: 10.1093/nar/gkw329.
9. Widodo, Veronica Margarecaesha Anyndita N, Dluha N et al. (2018) Designing and overproducing a tandem epitope of gp350/220 that shows a potential to become an EBV vaccine. Heliyon 4 (3): e00564. doi: 10.1016/j.heliyon.2018.e00564.
10. Wilkins MR, Gasteiger E, Bairoch A et al. (1999) Protein identification and analysis tools in the ExPASy server. Methods in molecular biology (Clifton, NJ) 112 531–552. doi: 10.1385/1-59259-584-7:531.
11. Dallakyan S, Olson AJ (2015) Small-molecule library screening by docking with PyRx. Methods in Molecular Biology 1263 243–250. doi: 10.1007/978-1-4939-2269-7_19.
12. BIOVIA DS (2019) Discovery Studio 2019.
13. Trott O, Olson AJ (2009) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry 31 (2): NA-NA. doi: 10.1002/jcc.21334.
14. Krieger E, Vriend G (2014) YASARA View - molecular graphics for all devices - from smartphones to workstations. Bioinformatics (Oxford, England) 30 (20): 2981–2982. doi: 10.1093/bioinformatics/btu426.
15. Peters B, Sidney J, Bourne P et al. (2005) The immune epitope database and analysis resource: From vision to blueprint. In: PLoS Biol. pp 0379–0381.

16. Lan J, Ge J, Yu J et al. (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 581 (7807): 215–220. doi: 10.1038/s41586-020-2180-5.
17. Lin L, Ting S, Yufei H et al. (2020) Epitope-based peptide vaccines predicted against novel coronavirus disease caused by SARS-CoV-2. Virus Res. doi: 10.1016/j.virusres.2020.198082
18. Noorimotlagh Z, Karami C, Mirzaee SA et al. (2020) Immune and bioinformatics identification of T cell and B cell epitopes in the protein structure of SARS-CoV-2: A systematic review. Int Immunopharmacol. doi: 10.1016/j.intimp.2020.106738
19. Sharma N, Patiyal S, Dhall A et al. (2021) AlgPred 2.0: An improved method for predicting allergenic proteins and mapping of IgE epitopes. Brief Bioinform. doi: 10.1093/bib/bbaa294
20. Kounis NG, Koniari I, de Gregorio C et al. (2021) Allergic reactions to current available covid-19 vaccinations: Pathophysiology, causality, and therapeutic considerations. Vaccines 9 (3): 1–19. doi: 10.3390/vaccines9030221.
21. Kim R, Skolnick J (2008) Assessment of programs for ligand binding affinity prediction. Journal of Computational Chemistry 29 (8): 1316–1331. doi: 10.1002/jcc.20893.
22. Majewski M, Ruiz-Carmona S, Barril X (2019) An investigation of structural stability in protein-ligand complexes reveals the balance between order and disorder. Commun Chem. doi: 10.1038/s42004-019-0205-5
23. Ghosh R, Chakraborty A, Biswas A, Chowdhuri S (2021) Identification of polyphenols from Broussonetia papyrifera as SARS CoV-2 main protease inhibitors using in silico docking and molecular dynamics simulation approaches. Journal of Biomolecular Structure and Dynamics 39 (17): 6747–6760. doi: 10.1080/07391102.2020.1802347.
24. Wibowo S, Widyarti S, Sabarudin A, Soeatmadji DW, Sumitro SB (2019) The the role of astaxanthin compared with metformin in preventing glycated human serum albumin from possible unfolding: a molecular dynamic study. Asian Journal of Pharmaceutical and Clinical Research 276–282. doi: 10.22159/ajpcr.2019.v12i9.34617.

This page is intentionally left blank.