JURNAL KIMIA VALENSI
p-ISSN: 2460-6065; e-ISSN: 2548-3013
Journal homepage: https://journal.uinjkt.ac.id/index.php/valensi
Research Article

Papain-like Protease Peptides as Construction Material for the SARS-CoV-2 Vaccine Design Candidate: In-silico Study

Richa Mardianingrum[1*], Nur Ihsani Pertiwi[1], Lina Rahmawati Rizkuloh[1], Nur Kusaira Khairul Ikram[2], Ruswanto Ruswanto[3]

[1]*Pharmacy Department, Faculty of Health Sciences, Perjuangan University, 177 Peta Street, Tasikmalaya, 46115, Indonesia*
[2]*Institute of Biological Sciences, Faculty of Science, Universiti Malaya, 50603 Kuala Lumpur Malaysia*
[3]*Faculty of Pharmacy, Universitas Bakti Tunas Husada, Jl. Cilolohan 36 Tasikmalaya, 46196, Indonesia*

Email: richamardianingrum@unper.ac.id

Article Info
Received: Dec 4, 2024
Revised: June 14, 2025
Accepted: June 16, 2025
Online: June 21, 2025

Citation:
Mardianingrum, R., Pertiwi, N. I., Rizkuloh, L. R., Ikram, N. K. K., & Ruswanto. (2025). Papain-like Protease Peptides as Construction Material for the SARS-CoV-2 Vaccine Design Candidate: In-silico Study. *Jurnal Kimia Valensi*, 11(1), 138-149.

Doi: 10.15408/jkv.v11i1.42748

Abstract

COVID-19 remains a major global health threat. In addition to implementing health protocols and consuming supplements, proactive prevention strategies are essential to limit the spread of the virus. One of the most promising approaches is the use of vaccines, particularly peptide-based vaccines, which are under active development. This study aimed to design a peptide vaccine derived from the SARS-CoV-2 papain-like protease (PLpro) and evaluate its interaction with key components of the human immune system, namely Toll-like receptor 3 (TLR3), major histocompatibility complex class I (MHC-I), and class II (MHC-II). The research employed an immunoinformatics approach utilizing NetCTL, IEDB Tepitool, PEP-FOLD3, trRosetta, HDOCK, GalaxyRefine2, and other molecular modeling tools. The designed vaccine construct was visualized in 3D using trRosetta and validated through ERRAT2, achieving a 100% quality score, indicating excellent structural integrity. The docking simulations demonstrated stable interactions between the vaccine and the immune receptors, suggesting strong immunogenic potential. In conclusion, the in silico-designed peptide vaccine based on SARS-CoV-2 PLpro shows promise in triggering immune responses through stable binding with TLR3, MHC-I, and MHC-II, highlighting its potential as a candidate for further experimental validation in COVID-19 vaccine development.

Keywords: Bioinformatics, COVID-19, immunoinformatic, papain-like protease, vaccine design

1. INTRODUCTION

Since its emergence in late 2019, COVID-19 has continued to be a major global health crisis. Caused by the novel coronavirus 2019-nCoV—officially named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)—the pandemic has significantly impacted health systems, economies, and societies worldwide [1]. As of June 2025, the World Health Organization (WHO) reported over 775 million confirmed cases and more than 7 million deaths globally. Despite the global rollout of over 13.6 billion vaccine doses, several regions, including Southeast Asia, have recently experienced a rise in positivity rates, reaching 11%. In contrast, Indonesia has seen a marked decline in new infections, with only three new cases reported in the latest update and no active cases as of August 6, 2023 [2, 3].

Initial treatment efforts have centered on repurposing antiviral agents, such as Remdesivir, Oseltamivir, Lopinavir, Favipiravir, Ritonavir, Nelfinavir, Atazanavir, Baloxavir, Tipranavir, Umifenovir, Darunavir, and Saquinavir [4]. In parallel, the use of vitamins C, B, D, A, E, K, and F has been widely adopted to provide antioxidant, anti-inflammatory, and immunomodulatory support in COVID-19 therapy [5]. However, pharmacological approaches alone are insufficient to suppress transmission on a large scale. Consequently, vaccination remains the most effective strategy for long-term pandemic control.

To date, five primary vaccine platforms have been developed: DNA vaccines (e.g., Takara Bio), RNA vaccines (e.g., Pfizer/BioNTech and Moderna), inactivated virus vaccines (e.g., Sinovac and Sinopharm), protein subunit vaccines (e.g., Novavax), and viral vector-based vaccines (e.g., CanSino, Janssen, and AstraZeneca). Each platform exhibits strengths and limitations. For instance, DNA and RNA vaccines raise concerns regarding safety and long-term stability. Subunit vaccines may generate immune responses to non-essential antigens, while vector vaccines risk reduced efficacy due to preexisting immunity to the vector [6].

Most of the existing vaccines target the spike (S) glycoprotein using mRNA or protein subunit platforms. However, recent innovations have explored peptide-based vaccines composed of 9–15 amino acid residues. These offer advantages in specificity, safety, manufacturing simplicity, and immunogenicity [7]. Previous studies have demonstrated that peptide-based vaccines can induce both humoral and cellular immunity through interaction with immune receptors such as MHC-I, MHC-II, B-cell receptors, and TLR3[8].

Despite a significant focus on spike protein and main protease targets, limited research has explored internal non-structural proteins, such as the papain-like protease (PLpro). PLpro is essential for viral replication through polyprotein cleavage and plays a key role in immune evasion by removing ubiquitin and ISG15 from host proteins, thereby disrupting innate antiviral responses [9]. These dual roles make PLpro a highly promising yet underexplored target for vaccine design.

This study addresses this research gap by designing a multi-epitope peptide vaccine derived from PLpro and evaluating its immunogenic potential. Although TLR3 is classically associated with double-stranded RNA recognition, emerging evidence suggests it may contribute to peptide-based vaccine responses through indirect signaling and immune pathway cross-talk. An immunoinformatic approach was employed using tools such as NetCTL, IEDB Tepitool, trRosetta, HDOCK, and C-IMMSIM to predict epitopes, model 3D vaccine structures, and simulate immune interactions.

The novelty of this study lies in targeting PLpro rather than the commonly used spike protein. Structural validation using ERRAT2 yielded a high-quality model with 100% accuracy, while molecular docking demonstrated strong binding affinity with TLR3 (–393.44 kcal/mol). These findings indicate a stable and potentially immunogenic vaccine candidate. Thus, this research lays the foundation for future *in vitro* and *in vivo* validation of a next-generation peptide vaccine with enhanced specificity, safety, and efficacy.

2. RESEARCH METHODS

Instruments and Materials

The tools used are hardware and software. Hardware in the form of a PC set for in silico studies with specifications LAPTOP-0P90N39G Intel® Celeron® N4120 CPU @ 1.10 GHz, 64-bit operating system, x64-based processor, 4.00 GB RAM and DESKTOP-G33S3N3 Intel® CoreTM i5-7500 CPU @ 3.40 GHz 3.41 GHz, 64-bit operating system, x64-based processor, 16.0 GB RAM. The software used is Notepad, MEGA XI, database, and web server, which are presented in **Table 1.**

The research material used was the papain-like protease receptor PDB ID 6WX4 [10], 4M0W, 7D6H, 7D47, and 7SQE. HLA-A 11*01 allele (MHC-I) [11] PDB ID 1X7Q, HLA-DR1 allele (MHC-II) PDB ID 4I5B [8], TLR3 PDB ID 1ZIW [8] and adjuvant 50S ribosomal protein L7/L12 Uniprot ID P0A7K2 [8].

Table 1. Research database and web server

Database/web server	Website
RCSB-PDB	https://www.rcsb.org/
NCBI	https://www.ncbi.nlm.nih.gov/
UniProt	https://www.uniprot.org/
BLASTp	https://blast.ncbi.nlm.nih.gov/Blast.cgi
VaxiJen 2.0	http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html
NetCTL 1.2	https://services.healthtech.dtu.dk/services/NetCTL-1.2/
IEDBAR	http://tools.iedb.org/mhci
Tepitool IEDB	http://tools.iedb.org/tepitool/
HDOCK	http://hdock.phys.hust.edu.cn/
trRosetta	http://yanglab.nankai.edu.cn/trRosetta/
ProtParam	https://web.expasy.org/protparam/
AllerTOP v. 2.0	https://www.ddg-pharmfac.net/AllerTOP/
SOPMA	https://npsa-prabi.ibcp.fr/NPSA/npsa_sopma.html
GalaxyRefine2	http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE2
ERRAT dan PROCHECK	https://saves.mbi.ucla.edu/
PEP-FOLD3	https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/

Protein data search and phylogenetic analysis

The protein used was obtained from the RCSB-PDB database [12-15] and the National Center of Biotechnology Information (NCBI). The obtained protein is then stored in FASTA format in the Notepad software. Function, structure, and sequence evaluation are based on homologous sequence testing using BLASTp, then carrying out Multiple Sequence Alignment and phylogenetic analysis using MEGA XI software to determine the mutations that occur in the SARS-CoV-2 virus [11]

Protein sequence antigenicity prediction

The results obtained were analyzed for antigenicity using the VaxiJen 2.0 web server with a threshold value ≥0.4 to increase the specificity [9].

T-cell epitope prediction and molecular docking analysis between epitopes and alleles

The NetCTL 1.2 web server was used to identify T-cell epitopes with a threshold setting of 0.95 to maintain sensitivity and specificity. This web server extends predictions for 12 Major Histocompatibility Complex class I (MHC-I) supertypes and combines predicted MHC-I peptide binding, proteasome C-terminal cleavage with TAP transport efficiency, and predicted affinity for MHC-I using the IEDBAR program. Epitopes with IC_{50} values <200 nM and positive immunogenicity were followed by docking analysis [11].

CD4+ T-cell epitope prediction was carried out using the Tepitool IEDB web server. Predictions were carried out using the IEDB Recommended Method and epitope selection criteria using the 7-allele method. The size of the predicted epitope is set to 15-mer. The top five epitopes based on quartile values were selected for further analysis. The predicted epitope was then predicted for its antigenicity value using the Vaxijen 2.0 web server with a threshold value of 0.4. Epitopes with antigenicity values above the threshold are used for further analysis [11].

Epitope and allele molecular docking analysis using the HDOCK web server. The HLA-A 11*01 and HLA-DR1 alleles are considered receptor proteins, while the epitopes are ligands. Molecular docking results are presented as a negative score in Kcal/mol, expressing the binding affinity value between the ligand and the receptor [8,11].

B-cell epitope prediction

Strong cellular and humoral immune responses are necessary for a vaccination to be effective [16] B-cell epitopes were further examined using IEDB for analysis [9]. Epitopes with a length of 12-mer were selected for analysis of immunogenicity, antigenicity, and toxicity [11].

Population coverage prediction

Epitopes from MHC-I and MHC-II were taken, and then population coverage was predicted to calculate the fraction of individuals who were predicted to respond to a given set of epitopes based on HLA genotype frequencies and based on MHC binding data and/or T-cell restrictions. Population coverage was analyzed for populations in Indonesia [11].

Vaccine construction and 3D visualization of vaccine design

Epitopes that have adequate antigenicity, toxicity, allergenicity, and immunogenicity values are connected with linkers. The GPGPG linker links B-cell linear epitopes and MHC-II. The MHC-I epitope is linked to the AAY linker. At the N-terminus, 50S ribosomal protein L7/L12 (UniProt ID: P0A7K2) was added as an adjuvant and connected using an EAAAK linker. The 6xHis tail is added to the C-terminus[8]. The adjuvant and epitope sequences were then visualized in 3D with the trRosetta web server as a vaccine design candidate [11]. In the construction of the multi-epitope vaccine, different linkers were strategically used to enhance expression, immunogenicity, and structural stability. The GPGPG linker was inserted between B-cell and MHC-II epitopes to promote proper folding and presentation while also reducing junctional immunogenicity. This glycine-proline–rich linker is known to enhance the solubility and flexibility of multi-domain constructs. The AAY linker was used between MHC-I epitopes to facilitate proteasomal cleavage and enhance the epitope presentation via the MHC-I pathway, a strategy commonly applied in CTL epitope-based vaccine designs.

At the C-terminus, a 6xHis tag (HHHHHH) was added to assist in downstream purification and identification processes using affinity chromatography. This tag enables efficient detection and purification of the recombinant vaccine protein, especially in experimental and expression systems.

Primary, secondary, tertiary, and validation structure analysis

Primary structure analysis begins with BLASTp analysis to determine the similarity of the vaccine design to human proteins [8, 11]. Next, an analysis of the physicochemical properties of the vaccine design was conducted using the ProtParam web server [17] and tested for allergenicity with AllerTOP 2.0 and antigenicity with VaxiJen 2.0. Secondary structure analysis and vaccine design construction were carried out using SOPMA [11]. Tertiary structure analysis was carried out with the GalaxyRefine2 web server in 3D structure [16]. Ramachandran Plot Analysis using PROCHECK was carried out to analyze the quality of the structure [8,11] while validating the vaccine design structure using the ERRAT web server [11].

Interaction analysis between vaccine design and immune System

Epitope and allele molecular docking analysis using the HDOCK webserver. The immune system in the form of TLR3, MHC-I, and MHC-II are considered as receptor proteins, while the vaccine design is as a ligand. Docking results are presented as a negative score in Kcal/mol because the binding affinity of the ligand to its receptor is calculated as a negative value [8, 11].

3. RESULTS AND DISCUSSION

Protein data search and phylogenetic analysis

The proteins used were obtained from RCSB-PDB and NCBI. The data obtained is presented in **Table 2.**

Table 2. Protein data for papain-like protease and non-structural protein 3

No.	Sequence Name	Accession	Resolution (Å)	Number of amino acids
1.	Chain D, Non-stuctural protein 3	6WX4_D	1.655	326
2.	Chain A, Replicase polyprotein 1a	4M0W_A	1.4	327
3.	Chain A, Papain-like protease	7D6H_A	1.6	320
4.	Chain B, Non-stuctural protein 3	7D47_B	1.97	317
5.	Chain C, Papain-like protease	7SQE_C	2	318

Protein data searches were carried out on the RCSB-PDB and NCBI databases. Look for proteins in the form of papain-like protease in SARS-Cov-2 with different accession numbers, or PDB-ID. The consideration for selecting PDB ID sequences is that the resolution meets the requirements of 1 Å to 3. Structures with a resolution value of 1 Å or more are very regular and easy to see in electron density maps. Meanwhile, structures with a resolution of more than >3 Å have low density or are sparse [18] 6WX4 is the PDB ID of a papain-like protease; therefore, this protein is the standard for selecting other proteins [10]. As in the selection of the number of amino acids, sequences are chosen that do not have much difference in the number of amino acids with papain-like protease 6WX4 in order to facilitate and maximize the alignment results in the MEGA XI application.

Testing homologous sequences by BLASTp, the purpose of BLASTp [11] is to evaluate homologous sequences by comparing the sequences we have with sequences in the database. All selected sequences indicated BLASTp, and the results obtained were that each sequence yielded 100 comparison sequences with several parameters being recorded, including E-value, percentage of identity, and accession. The top 5 homologous sequences are presented in supplementary **Table S6.**

The E-value must be close to 0, the percentage of identity is the percentage of a set of homologous sequences; the higher the percentage, the more similar our sequence is to the comparison sequence, and the accession is a series of unique numbers to identify in database records [19]. In line with this, when the E-value (expect value) is lower, it means that the homologous nature between our sequence and the comparison sequence is more significant [20]. The homologous nature between the two sequences is expressed in the percentage of identity [19]. Sequences that have a low E-value and a high percentage of identity are the best protein sequences [21]. Referring to the references, the results show that all sequences have a low E-value and a high percentage of identity.

The sequences in **Table S6** were performed using multiple sequence alignment and phylogenetic analysis by running 1000 bootstrap replicates. Explained that 1000 replications are often selected, while the default MEGA result is 500 replications [22]. The phylogenetic tree is presented in **Figure 1.**

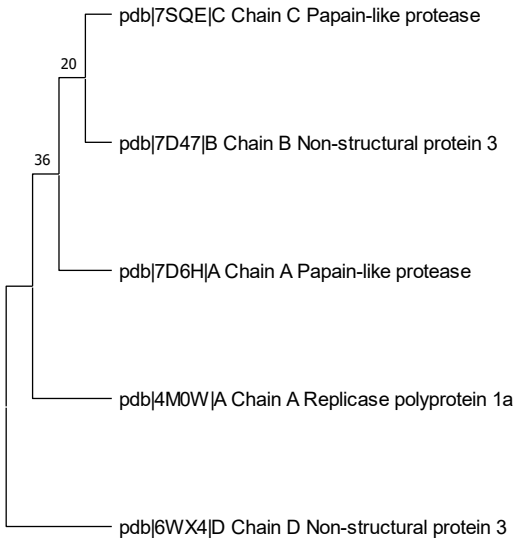


Figure 1. Results of phylogenetic analysis

The 6WX4_D sequence is used as the standard; the phylogenetic tree consists of 3 branches, where the closest branch is occupied by 4M0W_A. Numbers 36 and 20 show bootstrap values. 7SQE_C and 7D47_B are called sister taxa or sibling taxa because they are very closely related to each other [23].

Protein sequence antigenicity prediction

The ability of an antigen to stimulate the formation of antibodies is called antigenicity[7].

Antigenicity prediction results by VaxiJen 2.0 with a threshold of 0.4 are presented in **Table 3.**

The threshold 0,4 is set [24]. All sequences trigger antigenicity with a value greater than the threshold, but only the sequence with the highest antigenicity can be used as material for a vaccine design candidate, which is chosen 6WX4_D. The higher the antigenicity, the higher the antibody titer to fight infection [11].

T-cell epitope prediction and molecular docking analysis between epitopes and alleles

Epitope prediction of CD8+ T-cell with sequence ID 6WX4 was performed on 12 supertypes. However, the B7 supertype did not find selected epitopes because the score did not exceed the NetCTL regulatory threshold. The threshold used is 0.95 to obtain 0.74 sensitivity and 0.98 specificity [25]. A total of 81 CD8+ T-cell epitopes predicted by NetCTL were obtained from 11 supertypes. Epitopes exceeding the threshold on NetCTL were tested for antigenicity with a threshold of 0.4. Four epitopes with the highest antigenicity that exceeded the 0.4 threshold for each supertype were selected as the epitopes that would predict their immunogenicity [11]. The results are presented in Supplementary **Table S2.**

The selected epitopes were predicted for their immunogenicity. The results must be positive to trigger an immune response [7] which are presented in in Supplementary **Table S3.**

Affinity prediction was carried out on IEDBAR with the IC_{50}<200 nM criterion using the NN-align (Neural Network Alignment) algorithm. NN consists of a collection of sample units that form a communication network like neurons in the brain. The unit will respond if there is an incoming signal, which acts as the unit, which is the amino acid residue from MHC-I, while the signal is the predicted protein residue (SARS-CoV-2). This algorithm works by imitating the brain's working system [7]. The results are presented in **Table 4.**

Affinity is classified as IC_{50}<200 nM. A 50% inhibition concentration value means that 50% of the peptide portion is bound to the MHC[7]. Epitopes that met IC_{50}<200 nM and positive immunogenicity were continued for the next prediction.

Next, predict CD4+ or MHC-II T-cell epitopes using the Tepitool IEDB web server with the consensus method. The results of the top 5 percentile ranks were selected for further testing and are presented in **Table 5.**

Table 3. Predicted antigenicity of protein sequences

No.	Accession	Antigenicity Values	Description
1.	6WX4_D	0.5755	Antigen
2.	4M0W_A	0.5645	Antigen
3.	7D6H_A	0.5743	Antigen
4.	7D47_B	0.5720	Antigen
5.	7SQE_C	0.5621	Antigen

Table 4. Affinity prediction results for selected epitopes (IC_{50}<200 nM)

No.	MHC-I Epitopes	Supertypes	Allele	IC_{50} (nM)	Percentile Rank
1.	YMSALNHTK	A3	HLA-A*68:01	29.97	0.26
			HLA-A*11:01	55.21	0.34
			HLA-A*03:01	143.30	0.55
2.	YYHTTDPSF	A24	HLA-A*23:01	25.55	0.08
			HLA-A*24:02	55.32	0.1
	SYLFQHANL		HLA-A*23:01	153.57	0.39
	SYTTTIKPL		HLA-A*23:01	952.76	1.5
			HLA-A*24:02	1854.20	2.1
3.	TTVDNINLH	A26	HLA-A*26:01	1913.61	0.7
4.	TLKGVEAVM	B8	HLA-B*08:01	5674.25	5.9
7.	ARAGEAANF	B27	HLA-B*27:05	1653.30	3.6
	KRVLNVVCK		HLA-B*27:05	91.49	0.28
8.	FTTVDNINL	B39	HLA-B*39:01	7100.59	2.7
	GEAANFCAL		HLA-B*39:01	419.48	0.25
9.	GQQFGPTYL	B44	HLA-B*40:01	552.06	0.76
	GEAANFCAL		HLA-B*40:01	4.89	0.02
			HLA-B*40:02	19.34	0.03
			HLA-B*44:03	193.77	0.28
11.	VQQESPFVM	B62	HLA-B*15:02	1503.31	0.43

Description: Colored epitopes are selected epitopes that will advance to the next prediction

Table 5. Epitope prediction and antigens prediction results in MHC-II

No.	MHC-II Epitopes	Median Consensus Percentile	Antigenicity Values	Description
1.	GVQIPCTCGKQATKY	0.0003	0.8291	Antigen
2.	ARAGEAANFCALILA	0.0005	0.8328	Antigen
3.	CKTCGQQQTTLKGVE	0.0006	0.4617	Antigen
4.	DSCKRVLNVVCKTCG	0.0007	-0.0242	Non-Antigen
5.	AANFCALILAYCNKT	0.0008	0.4842	Antigen

Description: Coloured epitopes are selected epitopes that will advance to the next prediction

The selection of epitopes as vaccine candidates is based on the top 5 scores (percentile rank), referring to the reference that binding will be better if the percentile rank is smaller [7]. The selected epitope was then tested for antigenicity, and the results showed that epitope number 4 in **Table 5** could not advance to the next prediction because it could not trigger antigenicity.

It can be concluded that 5 MHC-I epitopes and 4 MHC-II epitopes can be used as candidates for vaccine design epitopes and predicted for the next stage. Because each epitope is still in the form of an amino acid sequence, 3D structure predictions must be carried out first in PEP-FOLD3. The epitope was then tested for docking using HDOCK. The results obtained are presented in Supplementary **Tables S4** and **Table S5**.

Epitope as the ligand and the allele as the receptor [11]. The HLA-A 11*01 allele was chosen because it is one of the alleles that binds to the MHC-I epitope. Additionally, this allele is available in the RCSB-PDB with PDB ID 1X7Q. The results show that all epitopes produce good binding affinity because the bonds are weak (negative). Furthermore, the molecular docking of the MHC-II epitope with the HLA-DR1 allele is presented in **Table S5**.

Binding of the MHC-II epitopes with the receptor HLA-DR1 allele PDB ID 4I5B according to the reference[8]. HLA-DRB1 is an allele for the Asian race epitope [7]. The yield of binding affinity is very good; the epitopes can be predicted for the next stage as a vaccine design candidate.

B-cell epitope prediction

Prediction of B-cell epitopes in IEDB with several models the results are presented in **Table 6** (detailed in supplementary **Table S6**).

The predicted B-cell epitope that was tested further was an epitope with a length of 12-mer [11]. Based on **Table 6**, of the 31 B-cell epitopes, only 2 epitopes met the requirements, namely those predicted by the Kolaskar and Tangaonkar models, and advanced to the next prediction. Kolaskar and Tongaonkar prediction method provides an accuracy of approximately 75% and is the most commonly used method in B-cell prediction [8]. Selected B-cell epitopes are then predicted for antigenicity, toxicity, and allergenicity. The results are presented in **Table 7**.

The results in **Table 7** show that two epitopes passed the prediction of antigenicity and toxicity; however, only one selected epitope can be considered a candidate B-cell epitope in vaccine design, as the other epitope may trigger an allergic response. Allergenicity is the ability of a material to trigger allergic properties[7]. The results must be non-allergenic and non-toxic.

Table 6. Prediction of B-cell epitopes

No.	Models	Start	End	Epitopes	Length
1.	Kolaskar and Tongaonkar	19	24	HTQVVD	6
2.		68	73	VEAFEY	6
3.		112	126	NCYLATALLTLQQIE	15
4.		148	159	NFCALILAYCNK	12
5.		183	194	CKRVLNVVCKTC	12
6.		221	229	GVQIPCTCG	9
7.		234	241	KYLVQQES	8
8.		283	290	TLYCIDGA	8

Description: Colored epitopes are selected epitopes that will advance to the next prediction

Table 7. Predictions of antigenicity, toxicity, and allergenicity of selected B-cell epitopes

No.	B-cell epitopes	Length	Antigenicity	Toxicity	Allergenicity
1.	NFCALILAYCNK	12	0.4682	Non-Toxic	Allergen
2.	CKRVLNVVCKTC	12	0.5281	Non-Toxic	Non-Allergen

Description: Coloured epitopes are selected epitopes that will advance to the next prediction

Population coverage prediction

This prediction was made for the Indonesian population. A combination of 5 MHC-I epitopes and 4 MHC-II epitopes and their alleles is input on the server; the results are presented in **Figure 2.**

The population coverage results in **Figure 2** are 68.29%, as indicated by the yellow dot at the highest peak of the graph, which does not meet the requirements. Results that meet the requirements are in the range of >90% according to the red mark limit on the graph [26]. Population coverage was carried out to determine the coverage between epitope interactions and MHC-I and MHC-II allele coverage so that fluctuations in allele distribution in each geographic region are different. The selection of epitopes and HLA MHC-I and MHC-II alleles is very important and considered so that references for allele selection are needed to cover various genetic variations in humans so that population coverage is maximized [27].

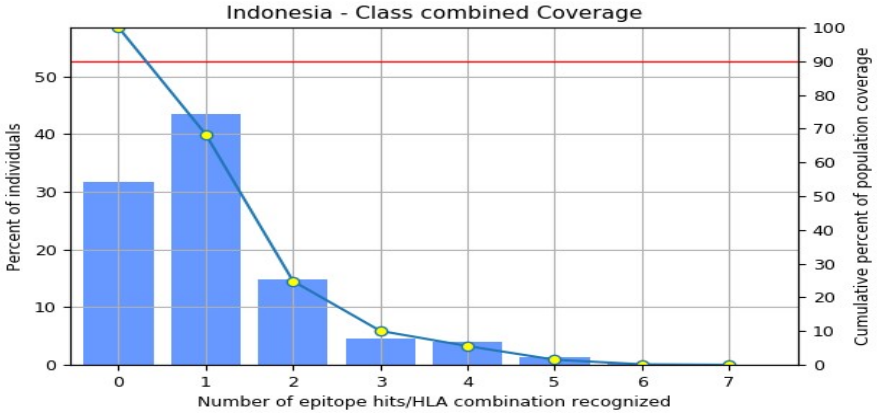


Figure 2. Population coverage prediction results in Indonesia

Vaccine construction and 3D visualization of vaccine design

The vaccine is constructed based on the scheme [8], which is presented in **Figure 3.** Adjuvant vaccines can enhance antigenicity. The 50S ribosomal protein L7/L12 has potential as an adjuvant as suggested by several studies. Apart from that, this protein is an agonist of TLR4, which is believed to have potential as an adjuvant by triggering increased antigen presentation by Antigen-Presenting Cell (APC). Other TLR can also interact with this protein [8].

The use of glycine-rich linkers such as GPGPG can increase solubility [28]. The advantage of using AAY and GPGPG linkers is that they prevent the formation of connections between epitopes, which is the a crucial aspect in the construction of multiepitope vaccines. Another advantage is that they can enhance immunogenicity and epitope presentation [29]. The 6xHis HHHHHH at the C-terminus is a histidine residue that can bind immobilized ions, making it easier for the sequence to work when in buffer conditions [9, 30]. The complete sequence of the vaccine design, along with the 3D visualization results, is presented in **Figure 4.**

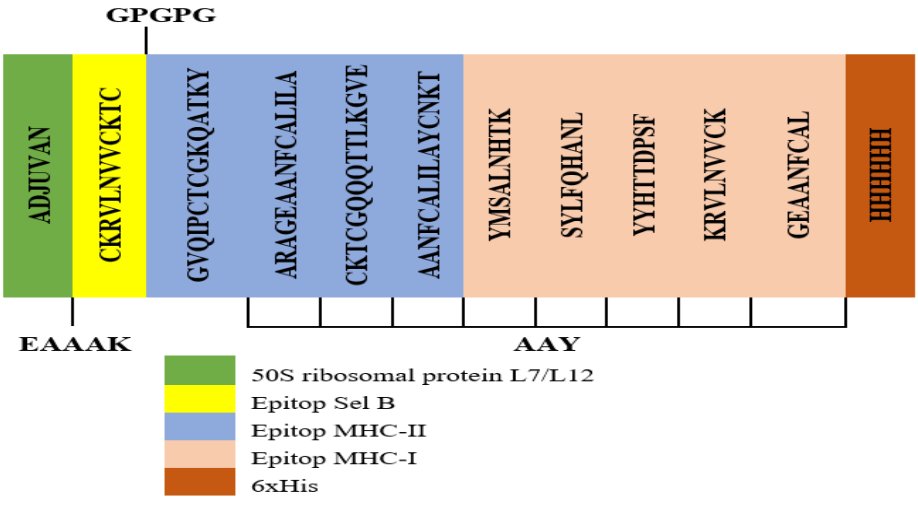
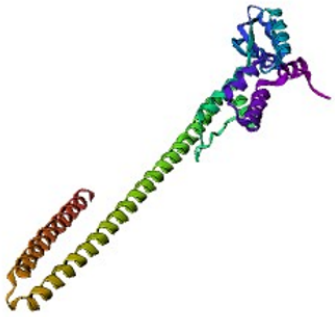


Figure 3. Schematic diagram of the constructed vaccine

MSITKDQIIEAVAAMSVMDVVELISAMEEKFGVSAAAAVAVAAGPVEAAEEKTEFD
VILKAAGANKVAVIKAVRGATGLGLKEAKDLVESAPAALKEGVSKDDAEALKKAL
EEAGAEVEVK**EAAAK**CKRVLNVVCKTC**GPGPG**GVQIPCTCGKQATKY**AAY**ARAG
EAANFCALILA**AAAY**CKTCGQQQTTLKGVE**AAY**AANFCALILAYCNKT**AAY**YMSAL
NHTKAAYSYLFQHANL**AAY**YYHTTDPSF**AAY**KRVLNVVCK**AAY**GEAANFCAL**AA**
YHHHHHH



(a) (b)

Figure 4. (a) Vaccine design sequence, (b) 3D visualization of vaccine design (N-terminus/purple & C-terminus/red)

Primary, secondary, tertiary, and validation structure analysis

Primary structure analysis with BLASTp showed no significant similarities to human or non-homologous proteins. This is an excellent result, BLASTp results should not be the same as proteins in humans to avoid autoimmunity [24].

Furthermore, the vaccine design was analyzed for its physicochemical properties with Protparam. The resulting vaccine consists of 281 amino acids. The amino acid composition consists of Thr:13, Ala:65, Pro:6, Arg:4, Asn:9, Asp:7, Cys:12, Gln:7, Glu:20, Gly:17, His:9, Ile:9, Leu:19, Lys:24, Met:5, Phe:7, Ser:9, Tyr:15, Val:24.

The molecular weight of the vaccine design is 29454.92 Da. This molecular weight is smaller when compared to the vaccine design, which is 32928.31 Da with a total of 309 amino acids [9].

The theoretical pI value is 7.52, which means that the protein charge is positive. The theoretical results of pI >7 indicate a positively charged protein [9].

The estimated half-life in mammalian reticulocytes in vitro is 30 hours; this result is similar to that of a long half-life and is expected to result in longer exposure to the human body. While the half-life in yeast in vivo is >20 hours and in Escherichia coli is >10 hours [31].

The structure of the vaccine design is predicted to be stable, with an instability index of 32.73. In accordance, the index instability is predicted to be stable with a value <40 [32]. Aliphatic index 86.76 and Grand Average of Hydropathicity (GRAVY) 0.165. Negative GRAVY means that the vaccine design is hydrophilic [31]. So the GRAVY resulting from the vaccine design this time is hydrophobic [29].

The next test is the allergenicity, antigenicity, and toxicity of the vaccine design. It was found to be non-allergenic, with an antigenicity value of 0.5491, indicating non-toxic properties. This meets all vaccine design requirements, meaning the vaccine design triggers antigens, does not trigger allergies, and is not toxic to humans.

Analysis of the secondary structure of the vaccine design found that it consisted of 211 Alpha helix (hh) amino acids with 75.09%, 15 Extended strand (Ee) amino acids with 5.34%, 15 β-turn (Tt) amino acids with 5.34 % and Random coil (Cc) 40 amino acids with 14.23%.

Analysis of the tertiary structure of the vaccine design using the GalaxyRefine2 web server, RMSD obtained 1.211 Å with the structure in **Figure 5.**

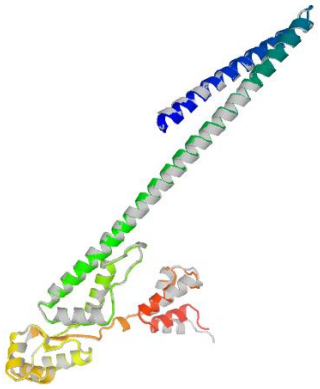


Figure 5. The 3D result of vaccine design in tertiary structure

Ramachandran plot analysis of vaccine design using the PROCHECK webserver with Job ID 1380418. Results are presented in Supplementary **Figure S1**.

The results show that residues in the most favored regions are 96.5%, additional allowed regions are 3.5%, and disallowed regions are 0.0%. The structure is considered good if the amino acid residues in the most favored region are more than 90% [33]. Additionally, the results of the PROCHECK vaccine design model revealed zero errors, indicating that the results met the requirements. Structural validation using ERRAT2 with Job ID 1380418 obtained 100% results, as presented in Supplementary **Figure S2.**

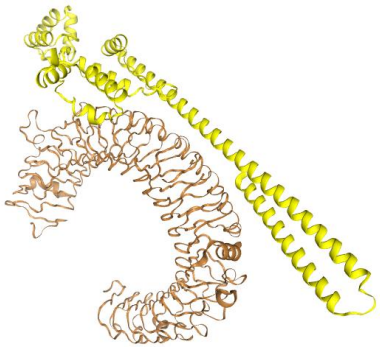
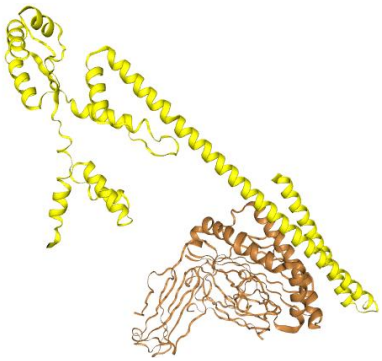
The vaccine design ERRAT validation result has a fairly good quality score, namely 79.268%. Better-quality structural models have higher quality scores [8]. Based on this, the vaccine design has very good quality.

Interaction analysis between vaccine design and immune system

Molecular docking analysis Vaccine design with the immune system as a receptor using HDOCK is presented in **Table 8** (detailed in Supplementary **Table S7**).

The receptors used are those of the immune system, each docking to a vaccine design that acts as a ligand. The results obtained for all binding affinity showed excellent results.

Table 8. Analysis of interactions between vaccine design and immune system

No.	Immune system	Interaction between vaccine design and immune system	Binding affinity (Kcal/mol)
1.	TLR3		-393.44
2.	HLA-A 11*01		-308.45

Description: Vaccine design (yellow) and immune system (brown)

4. CONCLUSIONS

Based on the research, the interaction results between the ligand-protein in the form of a vaccine design targeting the SARS-CoV-2 Papain-like protease receptor and the receptor proteins in the form of the immune system's TLR3, MHC-I, and MHC-II have an in silico stable interaction. As indicated by the results of docking simulations, molecular design reveals that vaccines form molecular complexes with immune receptors and exhibit strong binding energy. The results of the design and simulation of the vaccine design from the SARS-CoV-2 papain-like protease peptide in the form of sequences visualized in 3D structures on tr-Rosetta were then validated with very good structural results, namely 100% on ERRAT2.

As for the advice given, future researchers are expected to modify the epitopes and consider using different HLA alleles as vaccine construction materials to achieve more satisfactory results and increase population coverage in Indonesia, aiming for a coverage rate of over 90%. To maximize the vaccine design results, it can be used as a vaccine candidate and tested in vitro and in vivo at a later stage.

ACKNOWLEDGMENTS

The author thanks the Perjuangan Tasikmalaya University for providing facilities for this research.

REFERENCES

1. Aditia A. COVID-19 : Epidemiologi, Virologi, Penularan, Gejala Klinis, Diagnosa, Tatalaksana, Faktor Risiko Dan Pecegahan. *Jurnal Penelitian Perawat Profesional*. 2021;3(4):653-660. Accessed January 19, 2023. https://jurnal.globalhealthsciencegroup.com/index.php/JPPP/article/view/574/410
2. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. February 16, 2023. Accessed August 30, 2023. https://covid19.who.int/

3. Kemkes RI. COVID-19. Kementrian Kesehatan RI. January 18, 2023. Accessed August 30, 2023. https://infeksiemerging.kemkes.go.id/dashboard/covid-19
4. Lukito JI. Tinjauan Antivirus untuk Terapi COVID-19. CDK-286. 2020;47(5):340-345. Accessed January 19, 2023. http://download.garuda.kemdikbud.go.id/article.php?article=2468337&val=23521&title=Tinjauan%20Antivirus%20untuk%20Terapi%20COVID-19
5. Pisoschi AM, Pop A, Iordache F, et al. Antioxidant, anti-inflammatory and immunomodulatory roles of vitamins in COVID-19 therapy. *Europian Journal of Medicinal Chemistry*. 2022;232:1-24. doi: https://doi.org/10.1016/j.ejmech.2022.114175
6. Dong Y, Dai T, Wei Y, Zhang L, Zheng M, Zhou F. A systematic review of SARS-CoV-2 vaccine candidates. Signal Transduct Target Ther. 2020;5:1-14. doi: . https://doi.org/10.1038/s41392-020-00352-y
7. Rezaldi F, Taupiqurrohman O, Fadillah MF, Rochmat A, Humaedi A, Fadhilah F. Identifikasi Kandidat Vaksin COVID-19 Berbasis Peptida dari Glikoprotein Spike SARS CoV-2 untuk Ras Asia secara In Silico. *Jurnal Biotek Medisiana Indonesia* . 2021;10(1):77-85. Accessed January 19, 2023. https://ejournal2.litbang.kemkes.go.id/index.php/jbmi/article/view/5031/2299
8. Syakuran L 'Abdan. Desain Kandidat Vaksin SARS-CoV-2 Menggunakan Pendekatan Imunoinformatika. ResearchGate. Published online August 26, 2020:1-11. doi:https://doi.org/10.13140/RG.2.2.33453.31202
9. Naveed M, Tehreem S, Arshad S, et al. Design of a novel multiple epitope-based vaccine: An immunoinformatics approach to combat SARS-CoV-2 strains. *J Infect Public Health*. 2021;14(7):938-946. doi: https://doi.org/10.1016/j.jiph.2021.04.010
10. Islam MdJ, Islam NN, Alom MdS, Kabir M, Halim MA. A Review on Structural, Non-structural, and Accessory Proteins of SARS-CoV-2: Highlighting Drug Target Sites. Immunobiology. Published online January 2023:1-24. doi: https://doi.org/10.1016/j.imbio.2022.152302
11. Firmansyah MA, Susilo A, Haryanti SD, Herowati R. Desain Vaksin Berbasis Epitop dengan Pendekatan Bioinformatika untuk Menekan Glikoprotein Spike SARS-CoV-2. *Jurnal Farmasi Indonesia*. 2021;18(2):82-96. Accessed January 19, 2023. http://ejurnal.setiabudi.ac.id/ojs/index.php/farmasi-indonesia/article/view/1351/809
12. Ruswanto. Desain Dan Studi Interaksi senyawa N'-(3,5-Dinitrobenzoyl)-Isonicotinohydrazide pada Mycobacterium Tuberculosis Enoyl-Acyl Carrier Protein Reductase (INHA). *Jurnal Kesehatan Bakti Tunas Husada.* 2014;12(1):192-201. Accessed January 19, 2023. https://ejurnal.universitas-bth.ac.id/index.php/P3M_JKBTH/article/view/79/79.
13. Septian AD, Wardani GA, Mardianingrum R, Ruswanto R. The Virtual Screening of Flavonoid Derivatives on Progesterone, Estrogen, and HER-2 Receptor for Breast Cancer Treatment Candidate. *Jurnal Kimia Valensi*. 2023 Jun 28;9(1):163-82.
14. Mardianingrum R, Susilawati D, Ruswanto R. Computational Study of 1-(3-Nitrobenzoyloxymethyl)-5-Fluorouracil Derivatives as Colorectal Cancer Agents. *Jurnal Kimia Valensi*. 2022 Nov 24;8(2):211-20.
15. Badrunanto B, Asoka SF, Wahyuni WT, Farid M, Wahyudi ST, Batubara I. Component Discrimination and Anti-skin-aging Potency of Emprit and Red Ginger Essential Oil: Chemometric, Molecular Docking and Molecular Dynamics Study. *Jurnal Kimia Valensi*. 2023 Dec 15;9(2):183-94.
16. Khairkhah N, Aghasadeghi MR, Namvar A, Bolhassani A. Design of novel multiepitope constructs-based peptide vaccine against the structural S, N and M proteins of human COVID-19 using immunoinformatics analysis. PLoS One. 2020;15(10):1-28. doi:https://doi.org/10.1371/journal.pone.0240577

17. Bibi S, Ullah I, Zhu B, et al. In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. *Sci Rep*. 2021;11(1):1-16. doi: https://doi.org/10.1038/s41598-020-80899-6
18. RCSB PDB. (2023, July 1). Resolution. RCSB PDB-101. https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/crystallographic-data
19. NCBI. Glossary. National Library of Medicine. 2023. Accessed August 31, 2023. https://www.ncbi.nlm.nih.gov/books/NBK470040/
20. Syngai G, Barman P, Bharali R, Dey S. BLAST: An introductory tool for students to Bioinformatics Applications. *Keanean Journal of Science*. 2013;2:67-76. https://www.researchgate.net/publication/267332265
21. Anwar M, Nurjanah S, Rahayu WP. Aplikasi Basic Local Alignment Search Tool (Blast) NCBI Pada Penelitian Molekuler Salmonella SPP. Syntax Literate: *Jurnal Ilmiah Indonesia*. 2022;7(11):15446-15464. Accessed July 7, 2023. https://doi.org/10.36418/syntax-literate.v7i10
22. Russo CA de M, Selvatti AP. Bootstrap and rogue identification tests for phylogenetic analyses. Molecular Biology and Evolution, 35(9), 2327–2333. https://doi.org/10.1093/molbev/msy118.
23. Eernisse DJ. Introduction to Phylogeny: What is a Sister Taxon? Biology 404 - Evolution. 2023. Accessed August 31, 2023. http://biology.fullerton.edu/biol404/phylo/sisters.html
24. Rawal, K., Sinha, R., Abbasi, B. A., Chaudhary, A., Nath, S. K., Kumari, P., Preeti, P., Saraf, D., Singh, S., Mishra, K., Gupta, P., Mishra, A., Sharma, T., Gupta, S., Singh, P., Sood, S., Subramani, P., Dubey, A. K., Strych, U., … Bottazzi, M. E. (2021). Identification of vaccine targets in pathogens and design of a vaccine using computational approaches. *Scientific Reports*, 11(1). https://doi.org/10.1038/s41598-021-96863-x.22
25. NetCTL. (2023, July 1). NetCTL - 1.2 Predection of CTL epitopes in protein sequences. DTU Health Tech. https://services.healthtech.dtu.dk/service.php?NetCTL-1.2.
26. IEDBAR. (2023, July 1). IEDB Analysis Resource. IEDBAR. http://tools.iedb.org/main/
27. Ong E, He Y, Yang Z. Epitope promiscuity and population coverage of Mycobacterium tuberculosis protein antigens in current subunit vaccines under development. Infection, Genetics and Evolution. 2020;80:1-6. doi: https://doi.org/10.1016/j.meegid.2020.104186
28. Kar, T., Narsaria, U., Basak, S., Deb, D., Castiglione, F., Mueller, D. M., & Srivastava, A. P. (2020). A candidate multi-epitope vaccine against SARS-CoV-2. Scientific Reports, 10(1), 10895. https://doi.org/10.1038/s41598-020-67749-1.
29. Qamar MT ul, Rehman A, Tusleem K, et al. Designing of a next generation multiepitope based vaccine (MEV) against SARS-COV-2: Immunoinformatics and in silico approaches. PLOS ONE. 2020;15(12):1-25. doi: https://doi.org/10.1371/journal.pone.0244176.
30. Mulyani S, Egel E, Kittel C, Turkanovic S, Wohlleben W, Süssmuth RD, van Pée KH. The Thioesterase Bhp is Involved in the Formation of β-Hydroxytyrosine during Balhimycin Biosynthesis in Amycolatopsis balhimycina. *ChemBioChem*. 2010 Jan 25;11(2):266-71.
31. Sanami, S., Alizadeh, M., Nosrati, M., Dehkordi, K. A., Azadegan-Dehkordi, F., Tahmasebian, S., Nosrati, H., Arjmand, M. H., Ghasemi-Dehnoo, M., Rafiei, A., & Bagheri, N. (2021). Exploring SARS-COV-2 structural proteins to design a multi-epitope vaccine using immunoinformatics approach: An in silico study. *Computers in Biology and Medicine*, 133. https://doi.org/10.1016/j.compbiomed.2021.104390
32. Walker JM, Gasteiger E, Hoogland C, et al. The Proteomics Protocols Handbook Edited Protein Identification and Analysis Tools on the ExPASy Server. (Walker J, ed.).

Springer Science & Business Media; 2005. http://www.expasy.org/tools/

33. Renadi S, Pratita ATK, Mardianingrum R, Ruswanto R. The Potency of Alkaloid Derivates as Anti-Breast Cancer Candidates: In Silico Study. *Jurnal Kimia Valensi*. 2023;9(1):89-108. doi: https://doi.org/10.15408/jkv.v9i1. 31481