

Dissection of Capsid Protein HPV 52 to Rationalize Vaccine Designs Using Computational Approaches Immunoinformatics and Molecular Docking

Moh Egy Rahman Firdaus^{1,2}, Apon Zaenal Mustopa^{1*}, Lita Triratna³, Gita Syahputra⁴, Maritsa Nurfatwa¹

Abstract

Background: Human Papillomavirus type 52 (HPV 52) is considered one of the threatening HPV types inducing cervical cancer worldwide. This study was conducted to address strategies of an effective vaccine against cervical cancer using computational approaches immuno-informatics and molecular docking. **Methods:** Major capsid protein L1 and L2 HPV 52 (L1 and L2 HPV 52) sequences were investigated by multiple analyses including B and T cell epitope, toxicity, allergenicity, Immunogenicity, epitope conservancy, population coverage, and molecular docking. **Results:** L1 and L2 HPV 52 showed a conserved sequence among amino acid levels. Q307K, S383D/N, and D473E are found as major mutations in L1, while mutations in L2 are S122T, Q247H, L247S, and E365D. Multiple epitopes were identified and elicited strong immune responses against cross types of HPV in various HLA populations. To enhance vaccine effectiveness that allows having cross-protection over HPV types, N terminus HPV L2 was analyzed suggesting multi-candidates chimeric L1/L2 vaccine design. **Conclusion:** This study shed a light on a useful pipeline with robust analysis for effective vaccine production.

Keywords: L1- L2- HPV 52- multi-epitope vaccine- Immuno-informatics

Asian Pac J Cancer Prev, **23** (7), 2243-2253

Introduction

Human papillomaviruses (HPV) play a key role in causing 85% of cervical death in developing countries. They are responsible for 604,127 new cases with 341,831 deaths worldwide in 2020 (Sherris et al., 2001; Bañuelos-Villegas et al., 2021). Among them, HPV 52 is the top three oncogenic HPV, which is considered one of the threatening types that spread throughout the Indonesian female population (<http://www.hpvcntr.net/>). In addition, another study showed that HPV 52 was most commonly found in HIV patients receiving antiretroviral treatment in Indonesia (Risnawati et al., 2020). Thus the introduction of HPV vaccination and massive screening has become a national agenda to reduce the burden of cervical cancer in the future. It was reported that more than 100 countries have established HPV vaccination as a national immunization program (Markowitz and Schiller, 2021).

Currently, HPV vaccine designs are dominated by HPV

types 16 and 18 as the main contributors to oncogenic HPV (Markowitz and Schiller, 2021). However, the distribution of the virus and vaccine effectiveness may differ in different areas. Thus this study was conducted to propose effective vaccine design strategies against cervical cancer that focus on the Indonesian population which can also cover populations around the world.

The prophylactic HPV vaccine plays an important role in reducing cancer by suppressing viral infection (Garbuglia et al., 2020). Currently, much of the research on the HPV vaccine focused on L1 capsid protein, which can self-assemble to form virus-like particles (VLP). It was reported to be highly immunogenic, induce high neutralizing antibody titers, and is also safe (Stanley, 2007; Nooraei et al., 2021). The conserved region of the L2 cross-type and its function in binding to the secondary receptor of the host cell (Wang et al., 2015; Mikuličić et al., 2021), make the design of chimeric vaccines more consistently studied (Huber et al., 2021). The chimeric vaccine has the potential to enhance the effectiveness of

¹Research Center for Genetic Engineering, National Research and Innovation Agency (BRIN), Bogor, 16911, West Java, Indonesia.

²Present address: The International Institute of Molecular Mechanisms and Machines, Polish Academy of Sciences, Warsaw, Poland.

³Research Center for Applied Zoology, National Research and Innovation Agency (BRIN), Bogor, 16911, West Java, Indonesia. ⁴Research Center for Vaccine and Drug, National Research and Innovation Agency (BRIN), Bogor, 16911, West Java, Indonesia. *For Correspondence: azae.mustopa@gmail.com

the vaccine against a broad range of HPV types.

The prophylactic multivalent vaccine-based VLP Gardasil, as well as Cervarix®, have been licensed and shown to be highly effective in preventing viral infection and associated diseases (Schiller et al., 2008; Pineo et al., 2013). Although many approaches have been used to produce HPV vaccines, the high cost of vaccine production is currently unaffordable for most developing countries (Moscicki, 2008).

To overcome that issue, several approaches have been used including peptide-based vaccines which are considered easy to design and manufacture, stable, and also safe (Yang et al., 2016; Jabbar et al., 2018). This study was conducted to address effective vaccine strategies against HPV infection using computational approaches. The current prophylactic and therapeutic progress of epitope-based vaccine, in particular, L2 HPV were described elsewhere (Olczak and Roden, 2020).

Despite the traditional protein characterization in a laboratory setting, Immuno-informatics offers a shorter time and lower production cost (Chukwudozie et al., 2021). We applied immuno-informatics to dissect the sequences of a major capsid protein HPV 52 as a major antigenic component that potentially induces high titer of the humoral and cellular immune system. The polymorphism analysis, B and T cell epitopes predictions, epitope conservancy, antigenicity, toxicity, and allergenicity were carried out in this study, suggesting several strategies for vaccine designs. The designed vaccine was focused on Indonesian and worldwide coverage points of view, which leads to the design of a potent vaccine for certain areas that have wide coverage. This study highlights a useful pipeline with robust analysis for the production of an effective vaccine.

Materials and Methods

Retrieval of Capsid Protein L1 and L2 Sequences

Major capsid proteins L1 and L2 from the worldwide sequence of HPV 52 isolates were retrieved from NCBI (Supplementary file 1). The consensus sequence was determined by alignment using BioEdit 7.2.5 (www.mbio.ncsu.edu/BioEdit/BioEdit.html) RRID: SCR_007361. The L2 HPV 52 was referred to Japanese isolate sequence (BBA19628.1) with the L1 HPV 52 sequences (K1202) showing 99% similarity to the Indonesian isolate (KF225497). Polymorphism analysis of L2 was performed in the full sequence, but only 120 amino acids are exposed on the virus surface that was used for further design.

Linear and Conformational B-Cell Epitope Prediction

ElliPro (<http://tools.immuneepitope.org/ellipro/>) was implemented to predict linear and conformation B cell epitopes (Ponomarenko et al., 2008). To cover entire sequences that were not elucidated in protein structure, A BepiPred-2.0: Sequential B-Cell Epitope Predictor (<http://tools.iedb.org/bcell/>) was used in this analysis. The prediction can be determining epitopes and non-epitopes from the crystal structures (Jespersen et al., 2017). In addition, Kolaskar and Tongaonkar B cell predictions were also used to strengthen the prediction of other B

cell epitopes. This method is a semi-empirical method that has been used by a large number of proteins and is considered to provide 75% better accuracy than other methods (Kolaskar and Tongaonkar, 1990). The Predicted B cell epitopes were mapped to the available L1 HPV 52 tertiary structure (PDB: 6IGF) highlighting the most common mutation.

T Cell Epitopes - MHC Binding Prediction

MHCI and MHCII binding prediction from IEDB (<http://tools.immuneepitope.org/main/tcell/>) was used to predict the L1 and L2 HPV 52 T cell epitopes. The prediction used an IEDB NetMHCpan EL 4.1 recommendation method, with a cutoff $\leq 1\%$ of a percentile rank for each MHC allele (Moutaftsi et al., 2006). In terms of MHC-II, the binding prediction used the IEDB recommendation method. The method uses a consensus approach, combining NN-align, SMM-align, CombLib, and Sturniolo. A consensus percentile rank of 10% was used as a cutoff for selected epitopes that were considered to have a strong binding affinity (Wang et al., 2010) (Southwood et al., 1998).

Toxicity, Allergenicity, Population coverage, and Epitope Conservancy Prediction

Antigenicity, toxicity, and allergenicity of the epitope sequences were respectively evaluated by Vaxijen v.2.0 (Doytchinova and Flower, 2007), Toxinpred (Gupta et al., 2013), and AllerTOP v. 2.0 (Dimitrov et al., 2014). Vaccine efficacy, population coverage, and vaccine effectiveness in specific populations are also important issues to be addressed. Thus population coverage (<http://tools.iedb.org/population/>) (Bui et al., 2006) and epitope conservation (<http://tools.iedb.org/conservancy/>) (Bui et al., 2007) from IEDB were used for those examinations.

Epitope Shortlisting

The best epitopes were selected based on the strongest binding affinity, overlapping profiles of B and T cells, and profiles of non-toxic and non-allergenic agents. In addition, the best epitope was selected for further docking analysis, based on its cross-protection profile and population coverage.

Docking and Interaction Analysis

A 3D structure of epitope was modeled using the PEP FOLD 3 method which is available online (Lamiabile et al., 2016). The interaction of epitope and HLA proteins was evaluated by docking analysis using Swarmdock (<https://bmm.crick.ac.uk/~svc-bmm-swarmdock/index.html>) (Torchala et al., 2013; Torchala and Bates, 2014). This was performed by focusing on the binding pocket of the most HLA prevalence in Indonesia. The strongest binding affinity was selected following cluster standard at Swarmdock. The antigen-receptor interaction was analyzed in more depth using BIOVIA Discovery Studio 2016 (Dassault Sys-temes, BIOVIA Corp., San Diego, CA, USA). All proteins were visualized using PyMOL software.

Identification of candidate L2 Chimeric epitope

The known neutralizing and non-neutralizing L2 HPV16 epitopes described by (Karanam et al., 2009) were used as a reference. The L2 HPV 52 consensus sequence was generated using BioEdit 7.2.5 (www.mbio.ncsu.edu/BioEdit/BioEdit.html) RRID: SCR_007361 and then used to analyze epitope conservancy by IEDB (http://tools.iedb.org/conservancy/) (Bui et al., 2007).

Results

Sequences Determination and Polymorphism Analysis

Both L1 and L2 HPV 52 showed a highly conserved sequence among amino acid levels. Of all the polymorphism found, three and four L1 and L2 HPV 52 mutations respectively showed major mutations with a prevalence of more than 5% (Figure 1a and Figure 1c). It was then analyzed by mapping each position to predicted B cell epitopes, resulting in Q307K, S383D/N, and D473E considered the major mutations in L1. In terms of L2, S122T, Q247H, L247S, and E365D were found to be the main mutations, with E365D mutations accounting for 40% of the sequences. The mutation was not found in the Indonesia HPV 52 isolate sequences, except Q307K which was most commonly found in Canada and was phylogenetically different from Indonesia sequences (data not shown).

Linear and Conformational B-Cell Epitope Prediction

The predicted B cell epitopes of BepiPred-2.0, Kolaskar Tongaonkar, and Elipro are described in Figures 1b and 1d. Most of the L1 and L2 epitopes overlapped with T cell epitopes, with the B cell epitope scattered between the protein structures, mainly in the loops (Figure 2a). These B cell epitopes are exposed to the exterior of the protein and potentially interact with immune cells (Figure 2b). The polymorphisms of L1 and L2 HPV 52 were found across the predicted B cell epitope with the three major mutations shown in Figure 2a.

Cytotoxic T-Cell Epitope Prediction

MHC-I and MHC-II binding epitopes were analyzed using the most frequently available alleles in IEDB. Analysis using the frequent alleles supports the rationalization of vaccine designs across populations and ethnicities. All the top predicted MHC epitopes showed overlapping B and T cell epitopes. The predicted epitopes could recognize not only the adaptive immune but also the innate immune system.

Epitope Shortlisting

The top affinity lists of predicted epitopes that potentially protect the broad type of HPV are described in Table 1 and Table 2. The top 5 predicted epitopes of each capsid protein sequence showed high population coverage, with 82.04% and 97.37% respectively for Indonesia and the world (referred to HLA I databases). In terms of HLAI, the predicted epitope also showed a high population covering up to 96.65% and 99.65% of Indonesia and the world population respectively (Figures 4a and 4b).

Molecular Docking and Interaction Analysis

Out of the top 5 analyzed epitopes, the best epitopes that showed highly bound to HLA-11 (PDB ID: 6JP3) were FVTVVDTTR, and GVFFGGLGI; while GRKFLQAGLQARPK and TTIADQLLKYGSLGV were bound to HLA DRB1-15 (PDB ID: 2wbj) (Figure 3). Interactions of HLA HLA-11 with FVTVVDTTR and GVFFGGLGI were shown that amino acids T73, N66, and Q156 of the receptor interacted with both ligands by hydrogen bonds. It is generating electrostatic interactions between ligand-receptor. Amino acids V76 and A150 are two amino acids that mainly contributed to hydrophobic interactions.

In terms of TTIADQLLKYGSLGV, GRKFLQAGLQARPK, and HLA DRB1-15 interaction,

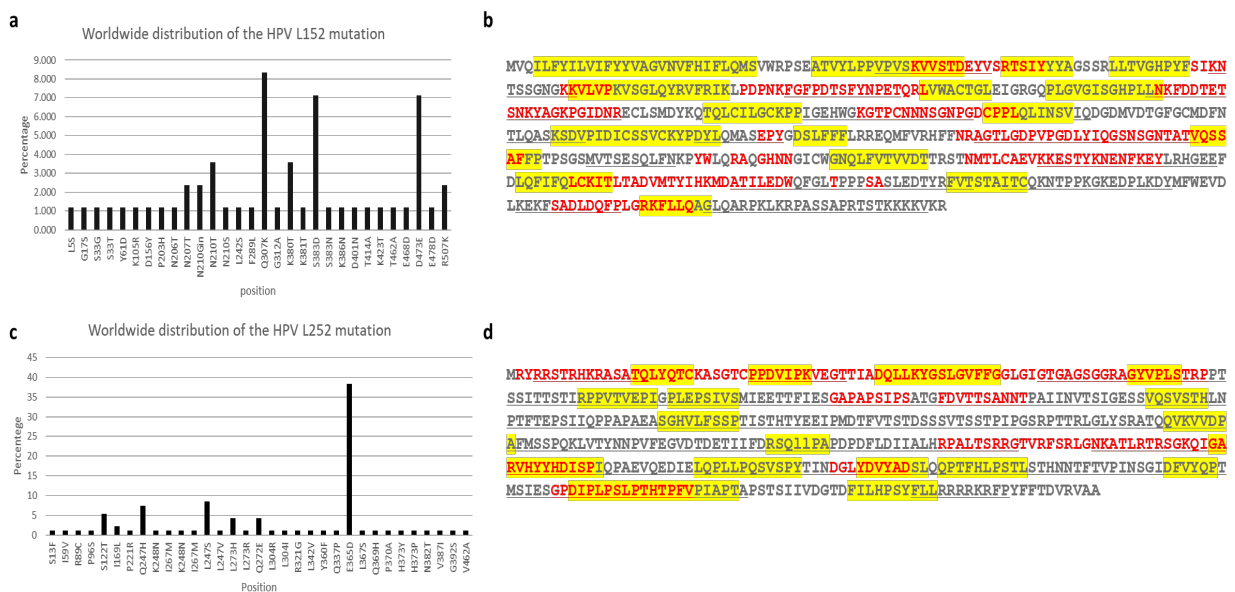


Figure 1. HPV 52 Capsid Protein Mutation Distribution Worldwide and Consensus Sequence of L1 HPV 52 (a-b) and L2 HPV 52 (b-c). Highlighted yellow indicates predicted B cell epitopes using Kolaskar and Tongaonkar. Red color exhibits conformational epitope prediction by Elipro. And the underlined indicates linear B cell prediction by Bbrep.

Table 1. The MHC I Epitope Prediction, HLA Coverage, and Percent Conservancy of Top L1 and L2 HPV 52 Epitopes

No	Sequences	Allele	HPV type conservansy	persen conservansy
1	LPDPNKEGF	HLA-B*35:01, HLA-B*53:01, HLA-B*07:02, HLA-B*51:01, HLA-B*08:01	HPV39,68 (77.78%); HPV18,45,51,56,59,66,70 (88.89%); HPV52,16,31,33,34,35,58 (100%)	100.00% (16/16)
2	VLVPKVSGL	HLA-A*02:03, HLA-A*02:01, HLA-A*02:06, HLA-B*08:01	HPV45, 59 (55.56%); HPV31, 34, 35 (77.78%); HPV16, 33 (88.89%); HPV31, 58 (100.00%)	56.25% (9/16)
3	YVSRISIVY	HLA-A*01:01, HLA-B*35:01, HLA-A*30:02, HLA-A*26:01, HLA-B*15:01, HLA-A*68:01, HLA-A*11:01, HLA-B*58:01, HLA-A*03:01	HPV51, 68 (66.67%); 16,18,31,34,35,39,56,59,66,70 (77.78%); HPV45 (88.89%); 52, 33, 58 (100%)	100.00% (16/16)
4	FVTVDITTR	HLA-A*68:01, HLA-A*33:01	HPV51 (77.78%); HPV34,39,59,68,70 (88.89%); HPV52,16,18,31,33,35,45,56,58,66 (100.00%)	100.00% (16/16)
5	LEIGRGQPL	HLA-B*40:01, HLA-B*44:03, HLA-B*44:02	HPV16,34,35,39,51 (77.78%); HPV18,31,45,56,59,66,68,70 (88.89%); 52,33,58 (100.00%)	100.00% (16/16)
6	ATILEDWQF	HLA-B*57:01, HLA-B*58:01, HLA-A*32:01, HLA-A*23:01, HLA-A*24:02, HLA-B*53:01, HLA-A*26:01	HPV34,45,56,66,68 (55.56%); HPV18,31,35,51 (66.67%); HPV16,33,58,59 (77.78%); HPV52(100.00%)	87.50% (14/16)
7	GLQARPKLK	HLA-A*03:01, HLA-A*30:01	HPV16, 34, 35, 59 (55.56%); HPV31, 58 (66.67%); HPV33 (77.78%); HPV52 (100%)	50.00% (8/16)
8	TSTATCQK	HLA-A*68:01, HLA-A*11:01, HLA-A*03:01	HPV45,59,70 (66.67%); HPV16,18,34,35,39,56,66,68 (77.78%); HPV31,33,58 (88.89%); HPV52(100%)	93.75% (15/16)
9	KVSGLQYRV	HLA-A*02:06, HLA-A*02:01, HLA-A*02:03, HLA-A*32:01, HLA-A*68:02, HLA-A*30:01	HPV18,39,45,51,56,59,66,68,70 (77.78%); HPV52,16,18,31,33,34,35,58 (100.00%)	100.00% (16/16)
10	SASLEDITYR	HLA-A*68:01	HPV16,18,34,45,56,59 (66.67%); HPV35,59 (77.78%); HPV31,33,39,51,58,68,70 (88.89%); HPV52(100%)	100.00% (16/16)
L2 HPV52				
1	LLKYGSLGV	HLA-A*02:03, HLA-A*02:01	HPV16, 31, 35 (66.67%), HPV34, 58 (77.78%), HPV33 (88.89%), HPV52 (100%)	43.75% (7/16)
2	VTVRPIGPL	HLA-A*68:02, HLA-A*26:01, HLA-A*02:06	HPV16,18,31,34,35,39,45,59,68,70 (55.56%), HPV33,58 (66.67%), HPV52 (100%)	81.25% (13/16)
3	ETTHESGA	HLA-A*68:02	HPV34, 58, 66 (55.56%), HPV16, 56 (66.67%), HPV33 (77.78%), HPV35 (88.89%), HPV52 (100%)	50.00% (8/16)
4	STRHKRASA	HLA-A*30:01, HLA-B*08:01	HPV18,51 (55.56%), HPV16,31,34,35,39,45,59,68,70 (66.67%), HPV56,66 (77.78%), HPV33,58 (88.89%), HPV52(100%)	100.00% (16/16)
5	AGSGGRAGY	HLA-A*30:02	HPV16,18,31,34,35,39,58,59,68,70 (66.67%), HPV33,45,51 (77.78%), HPV56,66 (88.89%), HPV52 (100%)	100.00% (16/16)
6	VPLSTRPPT	HLA-B*07:02	HPV34, 39,45,58,70 (55.56%), HPV33, 56, 66 (66.67%), HPV16, 35 (77.78%), HPV31 (88.89%), HPV52 (100%)	75.00% (12/16)
7	RPPVTVEPI	HLA-B*07:02, HLA-B*51:01	HPV31,56,59,66 (55.56%), HPV16,18,33,39,45,58,68,70 (66.67%), HPV35(88.89%), HPV52 (100%)	87.50% (14/16)
8	GVEFGGLGI	HLA-A*32:01	HPV 56,66 (66.67%), HPV18,39,45,51,59,68,70 (77.78%), HPV35 (88.89%), HPV52,16,31,33,34,58 (100%)	100.00% (16/16)

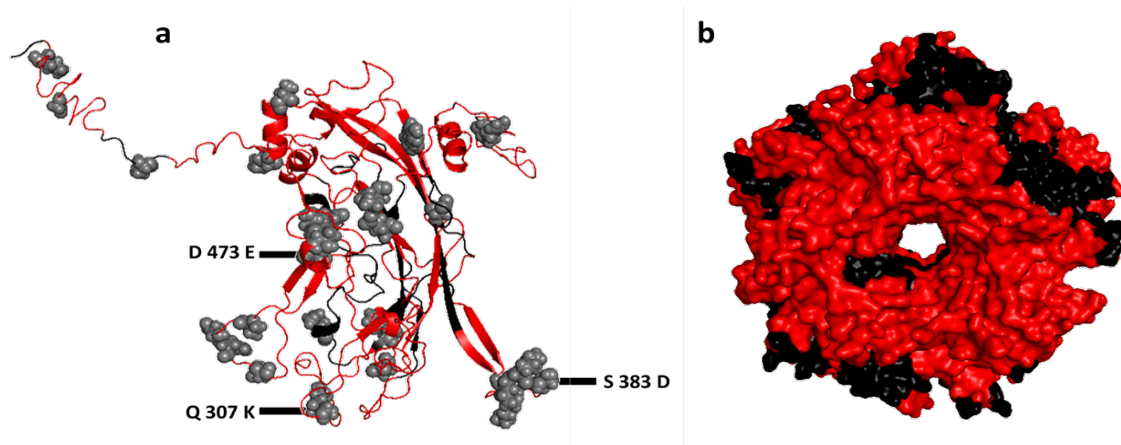


Figure 2. B Cell Epitope Mapping on L1 and L2 Protein Structure and Its Mutation Location (a). Surface pentamer of L1 HPV 52 (b). The red color shows the predicted B cell epitope, the grey ball shows the polymorphism site.

the Q7, W232, and D242 (amino acids in the receptors) interacted with both ligands through hydrogen bonds as well. Other interactions were identified as hydrophobic interactions via I243, Y236, V63, A66, and W237. The analysis found no similarity among all ligands in the electrostatic interactions (Table 3).

Candidate L2 Chimeric epitope

Three main candidates were predicted as neutralizing and non-neutralizing antibody epitopes (Figure 4). Interestingly all of the predicted neutralizing epitopes showed more than 50% identity in the template with

16-QLYQTCKASGTCPPDVIPKV-35, showing the highest similarity up to 80% (Table 4). All predicted epitopes potentially induce a high immune response against a wider range of HPV types.

Discussion

In this study, we emphasize vaccine design through three different approaches including VLP, chimeric, and peptide-based vaccines. The polymorphism of capsid protein provides valuable information to determine

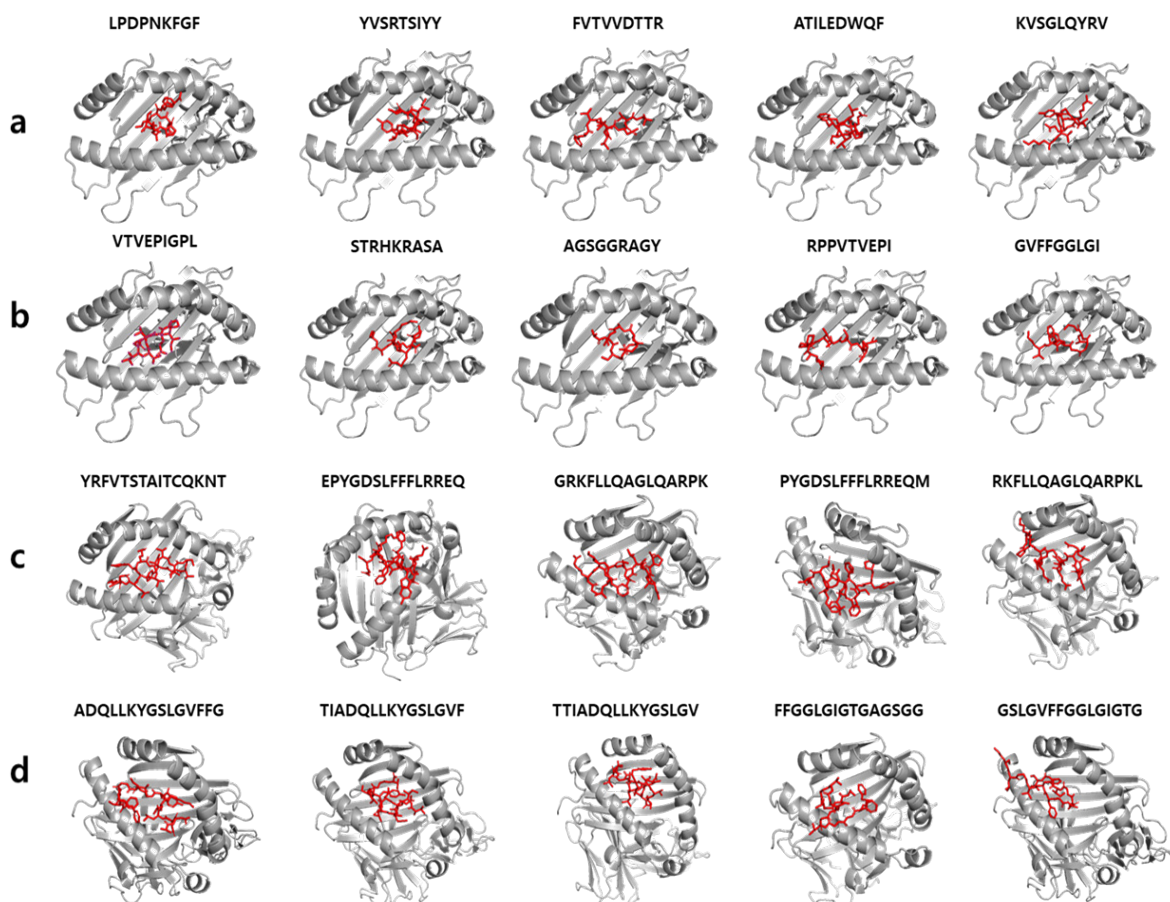


Figure 3. Molecular Docking of Top Five Predicted Epitopes with HLA Receptor. HLA 11-L1 (a) HLA 11-L2 (b) HLA DRB1-15-L1 (c) and HLA DRB1-15-L2 (d)

Table 2. The MHC II Epitope Prediction, HLA Coverage, and Percent Conservancy of Top L1 and L2 HPV 52 Epitopes

No	Sequences	Allele	HPV type conservansy	persen conservansy
L1 HPV52				
1	YRFVTSIATTCQKNT	HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPB1*14:01, HLA-DRB3*02:02, HLA-DRB1*04:01, HLA-DQAI*01:02, HLA-DQB1*06:02, HLA-DRB1*04:05, HLA-DRB5*01:01, HLA-DRB1*01:01	HPV70 (53.33%), HPV39,51,66,68 (60%), HPV34,35,56 (66.67%), HPV18,45,59 (73.33%), HPV16,31,33,58 (80%), HPV52 (100%)	100.00% (16/16)
2	VOILFYLVIFYVVA	HLA-DQAI*05:01, HLA-DQB1*02:01, HLA-DQAI*01:01, HLA-DQB1*05:01, HLA-DRB1*15:01, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DPA1*03:01, HLA-DPB1*04:02, HLA-DRB1*04:05, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPA1*02:01, HLA-DPB1*05:01, HLA-DRB1*11:01, HLA-DRB3*01:01, HLA-DRB4*01:01	HPV52 (100%)	6.25% (1/16)
3	MVQILFYLVIFYVYV	HLA-DQAI*05:01, HLA-DQB1*02:01, HLA-DQAI*01:01, HLA-DQB1*05:01, HLA-DRB1*15:01, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPA1*03:01, HLA-DRB1*04:02, HLA-DRB3*01:01, HLA-DRB4*01:01	HPV52 (100%)	6.25% (1/16)
4	EPYGDLSLFFFLRREQ	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPA1*02:01, HLA-DPB1*05:01, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DQAI*01:01, HLA-DQB1*05:01, HLA-DRB1*11:01	HPV34,51,56,66,70 (66.67%), HPV39,59,68 (73.33%), HPV18,45 (80%), HPV31,35 (86.67%), HPV16 (93.33%), HPV52,33,58 (100%)	100.00% (16/16)
5	GRRKFLQAGLQARRPK	HLA-DPA1*02:01, HLA-DPB1*14:01, HLA-DRB1*01:01, HLA-DRB3*02:02, HLA-DRB4*01:01, HLA-DRB1*11:01, HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DRB1*04:01	HPV56 (53.33%), HPV59 (60%), HPV18, 39, 51, 70 (66.67%), HPV34, 45, 58, 68 (73.33%), HPV16, 31, 33, 35 (86.67%), HPV52 (100%)	93.75% (15/16)
6	ILFYLVIFYVYVAGV	HLA-DRB1*15:01, HLA-DQAI*01:01, HLA-DQB1*05:01, HLA-DQAI*05:01, HLA-DQB1*02:01, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DPA1*03:01, HLA-DPB1*04:02, HLA-DRB1*04:05, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPA1*02:01, HLA-DPB1*05:01, HLA-DRB1*11:01, HLA-DRB3*01:01, HLA-DRB1*13:02	HPV52 (100%)	6.25% (1/16)
7	PYGDLSLFFFLRREQM	HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPA1*02:01, HLA-DPB1*05:01, HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DQAI*01:01, HLA-DQB1*05:01, HLA-DRB1*11:01	HPV51,56,66,70 (66.67%), HPV34,39,59,68 (73.33%), HPV18,45 (80%), HPV31,35 (86.67%), HPV16 (93.33%), HPV52,33,58 (100%)	100.00% (16/16)
8	RKFLQAGLQARRPKL	HLA-DPA1*02:01, HLA-DPB1*14:01, HLA-DRB1*01:01, HLA-DRB3*02:02, HLA-DRB4*01:01, HLA-DRB1*11:01, HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DRB1*04:01	HPV59 (53.33%), HPV18, 39, 51, 70 (60%), HPV45, 68 (66.67%), HPV34, 58 (73.33%), HPV16, 31, 35 (80%), HPV86 (86.67%), HPV52 (100%)	87.50% (14/16)
9	FYLVIFYVYVAGVNV	HLA-DRB1*01:01, HLA-DRB1*15:01, HLA-DQAI*01:01, HLA-DQB1*05:01, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DRB1*13:02, HLA-DPA1*03:01, HLA-DPB1*04:02, HLA-DRB1*07:01, HLA-DRB1*08:02, HLA-DRB1*04:05, HLA-DRB1*11:01	HPV56 (53.33%), HPV52 (100%)	12.50% (2/16)
10	GKKVLYPKVSGLQYR	HLA-DRB1*08:02, HLA-DRB1*12:01	HPV56, 59 (53.33%), HPV18, 39,66,68, 70 (60%), HPV45 (66.67%), HPV34, 35 (73.33%), HPV16, 31 (80%), HPV33 (86.67%), HPV58 (93.33%), HPV32 (100%)	93.75% (15/16)
9	SLGVFFGGLGIGTGA	HLA-DQAI*05:01/DQB1*03:01	HPV35, 51,56,66 (73.33%), HPV18,31,34,39,45,59,68,70 (80%), HPV16 (86.67%), PV33,58 (93.33%), HPV52 (100%)	100.00% (16/16)
10	GSLGVFFGGLGIGTGTG	HLA-DQAI*05:01/DQB1*03:01	HPV51 (73.33%), HPV18,35,39,45,56,59,66,68,70 (80%), HPV31,34 (86.67%), HPV16 (93%), HPV52,33,58 (100%)	100.00% (16/16)

Table 2. Continued

No	Sequences	Allele	HPV type conservansy	persen conservansy
L2 HPV52				
1	ADQLLKYSLSGVFVG	HLA-DRB1*15:01, HLA-DRB1*12:01, HLA-DRB1*13:02, HLA-DRB1*04:05, HLA-DRB1*07:01	HPV18, 39, 45, 59, 66, 68, 70 (53.33%), HPV16 (73.33%), HPV31, 35 (80%), HPV34, 58 (86.67%), HPV33 (93.33%), HPV52 (100%)	93.75% (15/16)
2	TIADQLLKYSLSGVF	HLA-DRB1*15:01, HLA-DRB1*12:01	HPV18, 39, 45, 59, 68, 70 (53.33%), HPV16, 35 (73.33%), HPV31, 34 (80%), HPV58 (86.67%), HPV33 (93.33%), HPV52 (100%)	81.25% (13/16)
3	TTIADQLLKYSLSGV	HLA-DRB1*15:01, HLA-DRB1*12:01	HPV18, 39, 45, 59, 68, 70 (53.33%), HPV16, 35 (66.67%), HPV34 (73.33%), HPV31 (80%), HPV33, 58 (86.67%), HPV52 (100%)	81.25% (13/16)
4	FFGGLGIGTGAGSGG	HLA-DQA1*05:01/DQB1*03:01	HPV18, 31, 34, 35, 39, 59, 68, 70 (80%), HPV16, 45, 51, 56, 58, 66 (88.67%), HPV33 (93.33%), HPV52 (100%)	100.00% (16/16)
5	GVFPGGLGIGTGAGS	HLA-DQA1*05:01/DQB1*03:01	HPV18, 35, 39, 56, 59, 66, 68, 70 (73.33%), HPV31, 34, 45, 51 (80%), HPV16, 58 (86.67%), HPV33 (93.33%), HPV52 (100%)	100.00% (16/16)
6	GAGSGGRAGYVPLST	HLA-DQA1*05:01/DQB1*03:01	HPV18, 39, 59, 68, 70 (60%), HPV16, 51, 58 (66.67%), HPV33, 34, 35, 45 (73.33%), HPV31, 56, 66 (80%), HPV52 (100%)	100.00% (16/16)
7	VFFGGLGIGTGAGSG	HLA-DQA1*05:01/DQB1*03:01	HPV18, 39, 59, 68, 70 (73.33%), HPV31, 34, 35, 45, 51, 56, 66 (80%), HPV16, 58 (86.67%), HPV33 (93.33%), HPV52 (100%)	100.00% (16/16)
8	LGVFFGGLGIGTGAG	HLA-DQA1*05:01/DQB1*03:01	HPV35, 56, 66 (73.33%), HPV18, 31, 34, 39, 45, 51, 59, 68, 70 (80%), HPV16 (86.67%), HPV33, 58 (93.33%), HPV52 (100%)	100.00% (16/16)
9	SLGVFFGGLGIGTGGA	HLA-DQA1*05:01/DQB1*03:01	HPV35, 51, 56, 66 (73.33%), HPV18, 31, 34, 39, 45, 59, 68, 70 (80%), HPV16 (86.67%), PV33, 58 (93.33%), HPV52 (100%)	100.00% (16/16)
10	GSLGVFFGGLGIGTG	HLA-DQA1*05:01/DQB1*03:01	HPV51 (73.33%), HPV18, 35, 39, 45, 56, 59, 66, 68, 70 (80%), HPV31, 34 (86.67%), HPV16 (93%), HPV52, 33, 58 (100%)	100.00% (16/16)

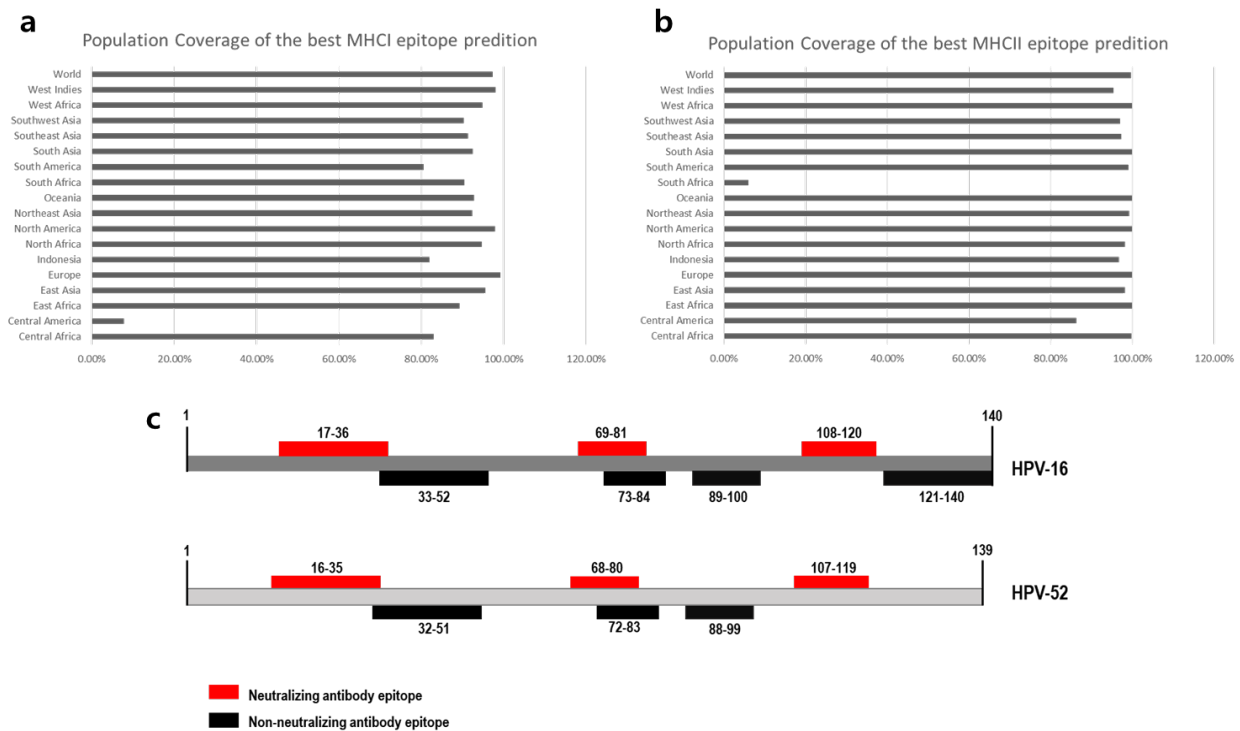


Figure 4. Population Coverage of the Top List MHC I Epitopes (a) and MHC II epitopes (b) over different populations. The putative binding site of neutralizing and non-neutralizing epitope of L2 HPV 52 (c).

Table 3. Ligands Receptor Interaction of the best L1 and L2 HPV 52 Epitopes

Receptors	Ligands	Binding Affinity, ΔG (kcal/mol)	Amino acids involved and distance (Å)		
			Hydrogen-Binding Interaction	Hydrophobic Interaction	Electrostatic Interaction
6JP3	FVTVDITTR	-17.63	T73 (2.74), N66 (2.44; 2.90, 2.86), Y99 (3.03), Q156 (3.05; 2.79), R114 (3.38)	V76 (4.97), A150 (3.67), A152 (4.76), W147 (4.74)	Y159 (4.72), R163 (93.46)
6JP3	GVFFGGLGI	-21.23	T73(2.82), N66 (2.83; 3.15), R163 (2.67), Q156 (2.78), W147 (3.90),	K146 (5.01), V76 (4.81), A150 (4.66), Y99 (5.29), Y159 (4.99)	W147 (3.90), R114 (3.64), K146 (5.01)
2WBJ	TTIADQLLYGSLGV	-27.53	Q7 (2.77; 2.77), D234 (3.63), W232 (2.82), Q246 (2.75; 3.09), D242 (2.74; 3.34)	A252 (5.39), I243 (4.73), Y236 (5.64), W237 (5.48), A66 (3.83), V63 (3.73; 4.83; 5.48), A59 (5.43)	D242 (5.59), W232 (4.82), R74 (4.45), D204 (5.11), E9 (5.39)
2WBJ	GRKFLQAGLQARPK	-25.51	D242 (2.87), Q246 (2.92), W232 (2.78), E53(3.19), F52(4.05), Q7 (3.53)	A250 (4.68), I243 (5.45), W237 (3.81; 5.30; 5.24; 5.31), Y236 (4.90), I70 (4.48), A66 (3.72), V63 (5.16; 4.26)	F52 (4.05), Y254 (4.14), D233 (4.53)

Table 4. The Homologous Monoclonal Antibody Region of L2 HPV16 in HPV 52

No	Type	Position	Sequences	Identity
1	HPV16	17-36	QLYKTCKQAGTCCPPDIIPKV	80.00%
	HPV52	16-35	QLYQTCKASGTCPPDVIPKV	
2	HPV16	33-52	IPKVEGKTIAEQILQYGSMG	75.00%
	HPV52	32-51	IPKVEGTTIADQLLYGSLG	
3	HPV16	69-81	RTGYIPLGTRPPT	76.92%
	HPV52	68-80	RAGYVPLSTRPPT	
4	HPV16	73-84	IPLGTRPPTATD	58.33%
	HPV52	72-83	VPLSTRPPTSSI	
5	HPV16	89-100	VRPLTVDPVGP	66.67%
	HPV52	88-99	IRPPTVEPIGP	
6	HPV16	108-120	LVEETSFIDAGAP	61.54%
	HPV52	107-119	MIEETTFIESGAP	
7	HPV16	121-140	TSVPSIPDPVSGFSITSTD	35.00%
	HPV52	419-438	PFVPIAPTAPSTSIHVDGTD	

immune recognition for vaccine design in particular areas (Wilson et al., 2016). VLP resembles a virus without harming the genetic material that is indistinguishable from its native form by the immune system (Chackerian, 2007). HPV VLP-based vaccines have also been shown to provide strong immune memory and produce high antibody titers (Stanley, 2007). In addition, the VLP system is reported to be the best platform not only for producing HPV vaccines but can also be used as a delivery vehicle platform (Qian et al., 2020).

L1 protein has a critical role in initial attachment to the host cell receptor and protects all genetic material from degradation during infection. it facilitates endocytosis of the viral body (Buck et al., 2013). L1 is present on the exterior surface as VLP, exposed and targeted by the immune system (Norman et al., 2013). Another capsid protein that is also important in viral invasion is L2. The conserved L2 region and its function in binding to

the secondary receptor of the host cell (after the initial contact) have been exploited to enhance the immune response (Wang and Roden, 2013). L2 is more conserved in a wide variety of HPV types, where the N terminus can enhance the immune responses via cross-neutralizing and non-neutralizing antibodies across HPV types (Rubio et al., 2011).

Combining recombinant proteins as a cocktail in a certain formulation, or generating a chimeric form are useful strategies to achieve optimal vaccine production with high immunogenicity (Greenstone et al., 1998; Chabeda et al., 2019; Arribillaga et al., 2020). This strategy has been used by several studies, showing that it has a good impact on overcoming certain problems (Tumban et al., 2012; Wang et al., 2015).

Despite focusing on prophylactic vaccines via neutralizing antibodies, T cell-mediated immunity does play a central role in controlling persistent viral infections such as human immunodeficiency virus, cytomegalovirus, and hepatitis C virus (Panagiotti et al., 2018). Through peptide-based vaccine design, selecting the most immunogenic sequences and combining them with specific adjuvant is also being continuously studied (Malonis et al., 2020; Topuzo-Gullari et al., 2020; Stephens et al., 2021).

Antibody production is stimulated by exposed epitopes that are recognized by the immune system. The Exposed epitopes have specific profiles including accessibility, flexibility, and hydrophilicity (Haste Andersen et al., 2006). The common prediction method was designed based on a profile that matches the native state using amino acid propensity scales (Potocnakova et al., 2016). The epitopes and immune interaction required strong affinity. This is facilitated by hydrogen bond, hydrophobic, and electrostatic interactions as the main key to ligand-receptor interaction (Patil et al., 2010; Chen et al., 2016; Erbaş et al., 2018). This study found that although all the ligands did not share the same electrostatic interaction, the strong affinity between epitopes and HLA molecules was still possible.

Since Indonesia has a diverse society with about 633 ethnicities, humoral and cellular immune responses likely vary among populations. In addition, several studies reported that HLA was considered an important aspect of vaccine design (Ovsyannikova et al., 2008; Stern and Calvo-Calle, 2009). HLA has been known to have a thousand polymorphisms that correspond to certain populations across the world. The current Study-related COVID-19 pandemics reported that HLA alleles can affect susceptibility as well as the severity of viral infections (Tavasolian et al., 2021). Epitope coverage and epitope conservancy analysis in this study showed that the selected epitopes were a great candidate vaccine component. It showed high population coverage for Indonesia and worldwide.

The ability to induce cross-neutralizing antibodies makes the vaccine more potent and more considered. Sharing epitopes among HPV types is an essential point to produce board effective antibodies (Nakagawa et al., 2015). It has the potential to reduce production costs, particularly for developing countries. The selected

consensus sequence and epitopes in this study could be further subjected to in vitro and in vivo studies to continue developing an effective vaccine against HPV type 52. Thus this study highlights a useful pipeline with robust analysis for effective vaccine production.

Author Contribution Statement

FMER and MAZ designed the study. FMER, MAZ, TL, SG, and NM conducted laboratory work and data analysis. FMER, SG, and TL writing, reviewing, and editing. All authors read and approved the manuscript.

Acknowledgments

The authors would like to express their gratitude to Muhammad Yusuf, Ph.D., Research Center for Molecular Biotechnology and Bioinformatics, Universitas Padjajaran for technical guidance and data quality control, and Mrs. Pramukti Nawar Ra'idah Firdaus, M.Si, Organic Chemistry Division, Institut Teknologi Bandung in supporting data analysis.

Funding

This study was supported by a grant from Lembaga Pengelola Dana Pendidikan (LPDP) No 105/E1/PRN/2020. The funder had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Ethics approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Data availability

The datasets supporting the conclusions of this article are included within the article and its supplementary information file.

Conflict of interest

The authors declare that they have no conflict of interest

References

- Arribillaga L, Echeverria I, Belsue V, et al (2020). Bivalent therapeutic vaccine against HPV16/18 genotypes consisting of a fusion protein between the extra domain A from human fibronectin and HPV16/18 E7 viral antigens. *J Immunother Cancer*, **8**, e000704.
- Bañuelos-Villegas EG, Pérez-yPérez MF, Alvarez-Salas LM (2021). Cervical Cancer, Papillomavirus, and miRNA Dysfunction. *Front Mol Biosci*, **8**, 758337.
- Buck CB, Day PM, Trus BL (2013). The papillomavirus major capsid protein L1. *Virology*, **445**, 169-74.
- Bui HH, Sidney J, Dinh K, et al (2006). Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinform*, **7**, 153.
- Bui HH, Sidney J, Li W, et al (2007). Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinform*, **8**, 361.
- Chabeda A, van Zyl AR, Rybicki EP, et al (2019). Substitution of Human Papillomavirus Type 16 L2 Neutralizing Epitopes

- Into L1 Surface Loops: The Effect on Virus-Like Particle Assembly and Immunogenicity. *Front Plant Sci*, **10**, 779.
- Chackerian B (2007). Virus-like particles: flexible platforms for vaccine development. *Expert Rev Vaccines*, **6**, 381-90.
- Chen D, Oezguen N, Urvil P, et al (2016). Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci Adv*, **2**, e1501240-e.
- Chukwudozie OS, Duru VC, Ndiribe CC, et al (2021). The Relevance of Bioinformatics Applications in the Discovery of Vaccine Candidates and Potential Drugs for COVID-19 Treatment. *Bioinform Biol Insights*, **15**, 11779322211002168.
- Dimitrov I, Bangov I, Flower DR, et al (2014). AllerTOP v.2--a server for in silico prediction of allergens. *J Mol Model*, **20**, 2278.
- Doytchinova IA, Flower DR (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform*, **8**, 4.
- Erbas A, de la Cruz MO, Marko JF (2018). Effects of electrostatic interactions on ligand dissociation kinetics. *Phys Rev E*, **97**, 022405.
- Garbuglia AR, Lapa D, Sias C, et al (2020). The Use of Both Therapeutic and Prophylactic Vaccines in the Therapy of Papillomavirus Disease. *Front Immunol*, **11**, 188.
- Greenstone HL, Nieland JD, de Visser KE, et al (1998). Chimeric papillomavirus virus-like particles elicit antitumor immunity against the E7 oncoprotein in an HPV16 tumor model. *Proc Natl Acad Sci U S A*, **95**, 1800-5.
- Gupta S, Kapoor P, Chaudhary K, et al (2013). In Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS One*, **8**, e73957.
- Haste Andersen P, Nielsen M, Lund O (2006). Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci*, **15**, 2558-67.
- Huber B, Wang JW, Roden RBS, et al (2021). RG1-VLP and Other L2-Based, Broad-Spectrum HPV Vaccine Candidates. *J Clin Med*, **10**, 1044.
- Jabbar B, Rafique S, Salo-Ahen OMH, et al (2018). Antigenic Peptide Prediction From E6 and E7 Oncoproteins of HPV Types 16 and 18 for Therapeutic Vaccine Design Using Immunoinformatics and MD Simulation Analysis. *Front Immunol*, **9**, 3000.
- Jespersen MC, Peters B, Nielsen M, et al (2017). BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*, **45**, W24-w9.
- Karanam B, Jagu S, Huh WK, et al (2009). Developing vaccines against minor capsid antigen L2 to prevent papillomavirus infection. *Immunol Cell Biol*, **87**, 287-99.
- Kolaskar AS, Tongaonkar PC (1990). A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*, **276**, 172-4.
- Lamiable A, Thévenet P, Rey J, et al (2016). PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res*, **44**, W449-54.
- Malonis RJ, Lai JR, Vergnolle O (2020). Peptide-Based Vaccines: Current Progress and Future Challenges. *Chem Rev*, **120**, 3210-29.
- Markowitz LE, Schiller JT (2021). Human Papillomavirus Vaccines. *J Infect Dis*, **224**, S367-s78.
- Mikuličić S, Strunk J, Florin L (2021). HPV16 Entry into Epithelial Cells: Running a Gauntlet. *Viruses*, **13**, 2460.
- Moscicki AB (2008). HPV Vaccines: today and in the future. *J Adolesc Health*, **43**, S26-S40.
- Moutafsi M, Peters B, Pasquetto V, et al (2006). A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. *Nat Biotechnol*, **24**, 817-9.
- Nakagawa M, Greenfield W, Moerman-Herzog A, et al (2015). Cross-Reactivity, Epitope Spreading, and De Novo Immune Stimulation Are Possible Mechanisms of Cross-Protection of Nonvaccine Human Papillomavirus (HPV) Types in Recipients of HPV Therapeutic Vaccines. *Clin Vaccine Immunol*, **22**, 679-87.
- Nooraee S, Bahrulolum H, Hoseini ZS, et al (2021). Virus-like particles: preparation, immunogenicity and their roles as nanovaccines and drug nanocarriers. *J Nanobiotechnol*, **19**, 59.
- Norman I, Hjerpe A, Andersson S (2013). High-risk HPV L1 capsid protein as a marker of cervical intraepithelial neoplasia in high-risk HPV-positive women with minor cytological abnormalities. *Oncol Rep*, **30**, 695-700.
- Olczak P, Roden RBS (2020). Progress in L2-Based Prophylactic Vaccine Development for Protection against Diverse Human Papillomavirus Genotypes and Associated Diseases. *Vaccines (Basel)*, **8**, 568.
- Ovsyannikova IG, Jacobson RM, Dhiman N, et al (2008). Human leukocyte antigen and cytokine receptor gene polymorphisms associated with heterogeneous immune responses to mumps viral vaccine. *Pediatrics*, **121**, e1091-e9.
- Panagiotti E, Klenerman P, Lee LN, et al (2018). Features of Effective T Cell-Inducing Vaccines against Chronic Viral Infections. *Front Immunol*, **9**, 276.
- Patil R, Das S, Stanley A, et al (2010). Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of drug-designing. *PLoS One*, **5**, e12029-e.
- Pineo CB, Hitzeroth, II, Rybicki EP (2013). Immunogenic assessment of plant-produced human papillomavirus type 16 L1/L2 chimaeras. *Plant Biotechnol J*, **11**, 964-75.
- Ponomarenko J, Bui H-H, Li W, et al (2008). ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform*, **9**, 514.
- Potočnakova L, Bhide M, Pulzova LB (2016). An Introduction to B-Cell Epitope Mapping and In Silico Epitope Prediction. *J Immunol Res*, **2016**, 6760830-.
- Qian C, Liu X, Xu Q, et al (2020). Recent Progress on the Versatility of Virus-Like Particles. *Vaccines*, **8**, 139.
- Risnawati, Soejoenoes A, Hadisaputro S, et al (2020). Hpv 52 most dominant human papillomavirus (Hpv) genotype in women infected by human immunodeficiency virus (hiv) that get antiretroviral treatment (art). *Int J Pharm Res*, **12**, 2568-73.
- Rubio I, Seitz H, Canali E, et al (2011). The N-terminal region of the human papillomavirus L2 protein contains overlapping binding sites for neutralizing, cross-neutralizing and non-neutralizing antibodies. *Virology*, **409**, 348-59.
- Schiller JT, Castellsagué X, Villa LL, et al (2008). An update of prophylactic human papillomavirus L1 virus-like particle vaccine clinical trial results. *Vaccine*, **26**, K53-K61.
- Sherris J, Herdman C, Elias C (2001). Cervical cancer in the developing world. *West J Med*, **175**, 231-3.
- Southwood S, Sidney J, Kondo A, et al (1998). Several common HLA-DR types share largely overlapping peptide binding repertoires. *J Immunol*, **160**, 3363-73.
- Stanley M (2007). Prophylactic HPV vaccines. *J Clin Pathol*, **60**, 961-5.
- Stephens AJ, Burgess-Brown NA, Jiang S (2021). Beyond Just Peptide Antigens: The Complex World of Peptide-Based Cancer Vaccines. *Front Immunol*, **12**.
- Stern LJ, Calvo-Calle JM (2009). HLA-DR: molecular insights and vaccine design. *Curr Pharm Des*, **15**, 3249-61.
- Tavasolian F, Rashidi M, Hatam GR, et al (2021). HLA, Immune Response, and Susceptibility to COVID-19. *Front Immunol*, **11**.
- TopuzoĜullari M, Acar T, Pelit Arayici P, et al (2020). An insight

- into the epitope-based peptide vaccine design strategy and studies against COVID-19. *Turk J Biol TURK J Biol*, **44**, 215-27.
- Torchala M, Bates PA (2014). Predicting the Structure of Protein–Protein Complexes Using the SwarmDock Web Server. *Methods Mol Biol*, **1137**, 181-97
- Torchala M, Moal IH, Chaleil RAG, et al (2013). SwarmDock: a server for flexible protein–protein docking. *Bioinformatics*, **29**, 807-9.
- Tumban E, Peabody J, Tyler M, et al (2012). VLPs displaying a single L2 epitope induce broadly cross-neutralizing antibodies against human papillomavirus. *PLoS One*, **7**, e49751-e.
- Wang D, Li Z, Xiao J, et al (2015). Identification of Broad-Genotype HPV L2 Neutralization Site for Pan-HPV Vaccine Development by a Cross-Neutralizing Antibody. *PLoS One*, **10**, e0123944.
- Wang JW, Roden RBS (2013). L2, the minor capsid protein of papillomavirus. *Virology*, **445**, 175-86.
- Wang P, Sidney J, Kim Y, et al (2010). Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinform*, **11**, 568.
- Wilson KL, Xiang SD, Plebanski M (2016). A Model to Study the Impact of Polymorphism Driven Liver-Stage Immune Evasion by Malaria Parasites, to Help Design Effective Cross-Reactive Vaccines. *Front Microbiol*, **7**.
- Yang A, Jeang J, Cheng K, et al (2016). Current state in the development of candidate therapeutic HPV vaccines. *Expert Rev Vaccines*, **15**, 989-1007.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.