Cloning, Expression and Bioinformatic Analysis of Human Papillomavirus Type 52 L1 Capsid Gene from Indonesian Patient

SONY SUHANDONO[1*], DEWI AYU KENCANA UNGU[1], TATI KRISTIANTI[1],
EDHYANA SAHIRATMADJA[2], AND HERMAN SUSANTO[3]

[1]School of Life Science and Technology, Institut Teknologi Bandung, Jalan Ganeca 10 Bandung 40132, Indonesia;
[2]Department of Biochemistry, Faculty of Medicine, Universitas Padjadjaran, Jalan Eijkman 38, Bandung 40161;
[3]Department of Obstetrics and Gynaecology, Faculty of Medicine Universitas Padjadjaran/
Dr. Hasan Sadikin General Hospital, Jalan Pasir Kaliki 190, Bandung, Indonesia;

Human papillomavirus (HPV) type 52 is the most prevalent type for causing cervical cancer in Indonesian population. Cervical cancer becomes the most common cancer suffered by Indonesian women. Prevention of HPV infection can be achieved using HPV virus-like particle (VLP) vaccine derived from L1 major capsid protein. This study aimed to clone and analyze HPV-52 L1 gene. DNA obtained from biopsy of a cervical cancer patient was amplified using specific primers designed from Asian originated HPV-52 L1 gene available in the GenBank. The isolated HPV-52 L1 gene sequence was submitted to GenBank with accession number KF225497. Expression of HPV-52 L1 gene was performed using pRSET/EmGFP*Escherichia coli* expression vector. We analyzed and compared the HPV-52 L1 gene expressions from recombinant *E.coli* BL21 (DE3) that had been induced for 3 hours with 1 mM IPTG and without induction. The protein was expressed in insoluble form. We performed the following bioinformatic analyses: construction of phlyogenetic tree, T-cell epitopes prediction and 3D proteins structure modelling. We utilized the following softwares: MEGA5 for phylogenetic tree, IEDBann for MHC prediction, CLC DNA Workbench 6.5 for hydrophobicity analysis, and PDB-Viewer Deep for 3D protein structure analysis. The phylogenetic tree which was developed based on KF225497 sequence showed that it shared a branch with Asian countries (Philippines and Thailand). The deduced amino acid sequences of the predicted epitopes that were consistent in all of the programs were 259GTLGDPVPGDLYIQGS274 and 345KKESTYKNE353. This information may be useful to design diagnostic strategies and vaccine suitable for Indonesian population.

Key words: epitope, HPV-L1 52 gene, human papillomavirus L1 protein

Human papillomavirus (HPV) tipe 52 merupakan jenis penyebab kanker serviks yang paling umum pada penduduk Indonesia, terutama kaum perempuan. Pencegahan infeksi HPV dapat dicapai dengan menggunakan vaksin *virus like particle* (VLP) HPV yang berasal dari L1 protein kapsid utama. Penelitian ini bertujuan untuk mengkloning dan menganalisa L1 gen HPV-52. DNA yang diperoleh dari biopsi pasien kanker serviks diamplifikasi menggunakan primer spesifik yang dirancang dari L1 gen HPV-52 berasal dari Asia yang tersedia di GenBank. Urutan L1 gen HPV-52 yang terisolasi didaftarkan ke GenBank dengan nomor akses [KF225497]. Ekspresi L1 gen HPV-52 dilakukan dengan menggunakan vektor ekspresi *Escherichia coli* pRSET/EmGFP. Kami menganalisis dan membandingkan L1 gen HPV-52 dari *E.coli* BL21 (DE3) rekombinan yang telah diinduksi selama 3 jam dengan 1 mM IPTG dan tanpa induksi. Protein diekspresikan dalam bentuk terlarut. Kami melakukan analisis bioinformatika: pembuatan pohon filogenetik, prediksi epitop sel T, dan modeling struktur 3D protein. Kami menggunakan software MEGA5 untuk pohon filogenetik, IEDBann untuk prediksi MHC, CLC DNA Workbench 6.5 untuk analisis hidrofobik dan PDB -Viewer Deep untuk 3D analisis struktur protein . Pohon filogenetik yang dihasilkan berdasarkan pada urutan KF225497 menunjukkan bahwa sekuen L1 gen HPV 52 berbagi cabang dengan negara-negara Asia (Filipina dan Thailand). Urutan asam amino epitop yang diprediksi konsisten dalam semua program adalah 259GTLGDPVPGDLYIQGS274 dan 345KKESTYKNE353. Informasi ini dapat digunakan untuk merancang strategi diagnostik dan vaksin yang cocok untuk penduduk Indonesia.

Kata kunci : epitope, gen HPV-L1 52 , protein L1 human papillomavirus L1

Persistent Human Papillomavirus (HPV) infection is associated with higher risk of cervical cancer development. About 79.14 million Indonesian women aged above 15 have risk of cervical cancer. Based on World Health Organization (WHO) report (2010), 13262 women are diagnosed with cervical cancer and 7493 die from the disease every year. Precaution act using vaccination can suppress the cervical cancer occurrence. Traditionally most prophylactic vaccines are made from live attenuated or inactivated virus

*Corresponding author; Phone: +6222-2511575
Fax +6222-2534107, e-mail: sony@sith.itb.ac.id

(Varsani *et al.* 2003). However, no recent report has shown successful attempt on growing papillomavirus in media culture (Smidkova *et al.* 2010). Major protein L1 and minor protein L2 that composed HPV icosahedral capsid, can be used as vaccine. L1 capsid protein alone has been proven to self assembles into virus-like particle (VLP) (Smidkova *et al.* 2010). HPV VLPs proved to induce high titer of neutralizing antibodies in animal model (Suzich *et al.* 1995 and Bretburd *et al.* 1995). Various expression systems such as animal cells (Zhou *et al.* 1991), insect cells (Rose *et al.* 1993), yeast (Sasagawa *et al.* 1995), plant cells (Biemelt *et al.* 2003 and Warzecha *et al.* 2003) and bacterial cells (Seo *et al.* 2009) have been used to produce VLPs.

Antigenic epitope recognition by T-cell epitope is the key molecular event that triggers the immune response in infected organism (Mohatbakar 2007). Epitope identification was also conducted to assure that the VLPs generated from *Escherichia coli* expression system were suitable for Indonesian patient.

Papillomaviruses have been detected in a various animals as well as in humans. Over 150 types of HPV had been fully characterized (Chen *et al.* 2011). Not all HPV types can infect genital part such as cervix, vagina, vulva, penis and anus. Only thirty HPV types are known as genital types. These genital types can be divided into two types: low and high risk types, according to the virus's association with genital cancer. Low-risk HPV types include types 6, 11, 42, 43,and 44, which usually cause benign genital warts. High-risk HPV types include types 16, 18, 31, 33, 34,35, 39, 45, 51, 52, 56, 58, 59, 66, 68 and 70, which cause genital cancer (Gomez and Santos 2007; De Villiers 2004). Even though HPV types 16 and 18 had been shown to have high prevalence in cervical cancer infection in Indonesia, current study had also shown that HPV-52 also had high prevalence in certain regions of Indonesia (Vet *et al.* 2008).This study suggested incorporation of HPV 52 in Indonesian HPV vaccine design. In this study, we aimed to clone, express and analyze HPV-52 L1 collected from Indonesia.

MATERIALS AND METHODS

HPV-52 L1 Gene Isolation and Cloning. The study protocol was approved by Medical Faculty of Universitas Padjadjaran Ethical Committe no. 190/FKUP-RSHS/KEPK /Kep./EC/2010. Biopsy was performed from one cervical cancer patient infected

with HPV-52 (Panigoro *et al.* 2013). Patient diagnosis was performed by Linear Array HPV genotyping test (Roche). Genomic DNA for PCR template was isolated from a cervical cancer patient with QIAamp DNA mini kit (Qiagen).

Primer was designed based on Asian originated HPV-52 L1 sequence available in GenBank. The HPV-52 L1 gene was amplified using forward primer 5'-CCACTGTGTACCTGCCTCCT3' and reverse primer 5'- ATGCAGGGCGTTTTAGTTTG-3'. PCR carried out with DreamTaq™ Green PCR Master Mix (Fermentas). PCR amplification was conducted using ABI® thermal cycler. Touchdown PCR has been used to amplify HPV L1 gene with the following condition: initial denaturation at 95 °C for 3 min followed by 10 cycle of 95 °C for 30 s, 57 °C for 30 s, 72 °C for 2 min, 10 cycle of 95 °C for 30 s, 57-47 °C (1 °C decreased every cycle) for 30 s, 72 °C for 2 min, 10 cycle of 95 °C for 30 s, 47 °C for 30 s, 72 °C for 2 min. The cycles were followed by final extension at 72 °C for 7 min and holding at 4 °C.

The amplified DNA was extracted from gel electrophoresis using Gel/PCR DNA Fragments Extraction Kit (Geneaid). The PCR product was then cloned to pGEM®-T Easy vector (PROMEGA) and transformed into *E. coli* strain DH5α using heat shock method. Confirmed recombinant plasmid was sequenced using T7 and SP6 primers by Macrogen Inc., South Korea.

Sequence Analysis. Forward and reverse sequences obtained from sequencing by Macrogen Inc. was trimmed and overlapped in order make a DNA contig using CLC DNA Workbench 6.5 software. To check HPV-52 L1 gene homology, the sequence was analyzed using Basic Local Alignment Search Tool (BLAST) available online at http://ncbi.nlm.nih.gov. Protein sequence was obtained using *SIB* (Swiss Institute of Bioinformatics) ExPASy DNA translate tool available online from http://web.expasy.org/translate/(Gasteiger 2003).

Expression Vector Construction. The HPV-52 L1 gene was cloned into pRSET/EmGFP (Invitrogen). Restriction sites were added to HPV L1 gene by PCR using HPV-52 L1 specific primers with restriction site sequence added at the 5' end. The primers were forward primer 5'- *GGATCC*TCCACTGTGTACCTGCCT - 3' and reverse primer 5'-*GAATTC*TTAGATTATGCA GGGCGTTTTAG -3'. Stop codon was added at the 5' end of the reverse primer. Restriction sites used for gene insertion to pRSET-EmGFP expression vector were *BamH*I and *EcoR*I. PCR was carried out with

```
CLUSTAL O(1.2.0) multiple sequence alignment
```

```
1    MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDRWGSSTVYLPPVPVSKVVSTDEYVSRTS
2    ------------------------------TVYLPPVPVSKVVSTDEYVSRTS
                                   ****************

1    IYYYAGSSRLLTVGHPYFSIKNTSSGNGKKVLVPKVSGLQYRVFRIKLPDPNKFGFPDTS
2    IYYYAGSSRLLTVGHPYFSIKNTSSGNGKKVLVPKVSGLQYRVFRIKLPDPNKFGFPDTS
     ************************************************************

1    FYNPETQRLVWACTGLEIGRGQPLGVGISGHPLLNKFDDTETSNKYAGKPGIDNRECLSM
2    FYNPETQRLVWACTGLEIGRGQPLGVGISGHPLLNKFDDTETSNKYAGKPGIDNRECLSM
     ************************************************************

1    DYKQTQLCILGCKPPIGEHWGKGTPCNNNSGNPGDCPPLQLINSVIQDGDMVDTGFGCMD
2    DYKQTQLCILGCKPPIGEHWGKGTPCNNNSGNPGDCPPLQLINSVIQDGDMVDTGFGCMD
     ************************************************************

1    FNTLQASKSDVPIDICSSVCKYPDYLQMASEPYGDSLFFFLRREQMFVRHFFNRAGTLGD
2    FNTLQASKSDVPIDICSSVCKYPDYLQMASEPYGDSLFFFLRREQMFVRHFFNRAGTLGD
     ************************************************************

1    PVPGDLYIQGSNSGNTATVQSSAFFPTPSGSMVTSESQLFNKPYWLQRAQGHNNGICWGN
2    PVPGDLYIQGSNSGNTATVQSSAFFPTPSGSMVTSESQLFNKPYWLQRAQGHNNGICWGN
     ************************************************************

1    QLFVTVVDTTRSTNMTLCAEVKKESTYKNENFKEYLRHGEEFDLQFIFQLCKITLTADVM
2    QLFVTVVDTTRSTNMTLCAEVKKESTYKNENFKEYLRHGEEFDLQFIFQLCKITLTADVM
     ************************************************************

1    TYIHKMDATILEDWQFGLTPPPSASLEDTYRFVTSTAITCQKNTPPKGKEDPLKDYMFWE
2    TYIHKMDATILEDWQFGLTPPPSASLEDTYRFVTSTAITCQKNTPPKGKEDPLKDYMFWE
     ************************************************************

1    VDLKEKFSADLDQFPLGRKFLLQAGLQARPKLKRPA
2    VDLKEKFSADLDQFPLGRKFLLQAGLQARPKLKRPA
     ************************************
```

Fig 1 Aligment result of pRSET/EmGFP-HPV52-L1 amino acid sequence [1] and HPV-52 L1 amino acid sequence [2]. Red: Histag; Yellow: Xpress epitope; Green: Enterokinase binding site.

KAPA2G Fast Hot Start Ready Mix (Kapabiosystem). Touchdown PCR has been used to amplify HPV L1 gene with the same cycling parameters used in cloning. To generate the recombinant plasmid pRSET/EmGFP-HPV52L1, both PCR product and plasmid were digested using *BamH*I and *EcoR*I. Both digested fragment were ligated using T4 DNA ligase overnight at 4 °C. The pRSET/EmGFP-HPV52L1 sequencing was performed in Macrogen, Korea. Restriction analysis and sequencing proved that both recombinant plasmids were constructed correctly.

Recombinant HPV-52 L1 Expression and Inclusion Bodies Isolation. The *E.coli* BL21(DE3) containing recombinant plasmid pRSET/EmGFP-HPV52L1, were grown in 5 mL LB with ampicillin (10^5μg mL^{-1}) overnight. The overnight culture was inoculated to 500 mL LB medium with ampicillin (10^5μg mL^{-1}) and grown at 37 °C, with 200 rpm shaking until exponential phase (OD_{600}= 0.8). The HPV-52 L1 expression was induced by addition of 1 mM IPTG and then incubated for 3 h at 37 °C, with 200 rpm shaking. The induced cell was harvested by centrifugation at 5000 g for 20 min. The cell pellet was resuspended in 6 mL 20 mM Tris-HCl pH 8, sonicated on ice (6x10 s) and centrifuged at 5000 g or 10 minute at 4 °C. Then, the cell pellet was resuspended in 3 mL cold solution containing 2M urea, 20 mMTris-HCl, 0.5 M NaCl, 2% Triton X-100 pH 8 before being sonicated again on
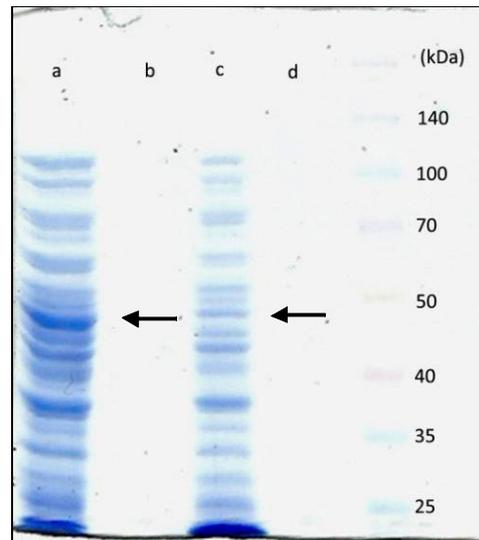

Fig 2 SDS-PAGE pRSET/EmGFP-HPV52-L1. The size of protein of interest is 48 kDa polyacrylamide gel 12%. a:Pellet-Induced ;b: Supernatant-Induced; c: Pellet-Non induced ; d: Supernatant- Non induced

ice (6x10 s) and centrifuged at 5,000 g for 10 minute at 4°C. Afterwards, the cell pellet was washed twice with urea containing-buffer and once with urea lacking-buffer (20 mMTris-HCl). The cell pellet was then stored frozen for later use.

Phylogenetic Tree Construction. Phylogenetic tree was constructed based on the maximum likelihood with MEGA 5.05 software. HPV-52 L1 gene sequences

used to construct phylogenetic tree were obtained from NCBI GenBank Database [http://www.ncbi.nlm.nih. gov]. The sequence origins and accession numbers of HPV-52 L1 genes used in the construction of the phylogenetic tree were Costa Rica [HQ537740], China [GQ472848], Thailand A [HQ537743], Thailand B [Hq537742], Zambia [Hq537736], Rwanda [Hq537735], Panama [U45923], Philippine A [U45922], Philippine B [U45921], Germany [U45920], Bolivia [U45919], and Spain [U45918].

Epitope Prediction. Further analyses had also been conducted using IEDBann software in order to determine the epitope regions. This software is based on artificial neural network (ann) algorithm. IEDBann is available online at http://tools.immuneepitope.org/ analyze/html/mhc_binding.html. This tool was chosen because it mightpredict peptide binding to HLA class I and II molecules (Gustiananda 2011).

Human leukocyte antigen (HLA) class I sequences that had been used for this study were HLA-A*24:07, HLA-A*11:01 and HLA- A*33:03 andHLA-B*15:02 alleles. Human leukocyte antigen (HLA) class II sequences used for this study were HLA-DRB1*07:01, HLA-DRB1*12:02 and HLA-DRB1*15:02. These alleles had been chosen because of their high frequency of occurence in Western Javanese (Indonesia) population (Yuliwulandari 2008).

Hydrophobicity and antigenicity graph was created using CLC DNA Workbench 6.5. L1 protein structure modelling was done using Homology Modelling: Automated Mode which can be accessed through http://swissmodel.expasy.org. L1 protein model then analyzed using SIB:Deep viewer (ExPASy).

RESULTS

HPV-52 L1 Gene Isolation, Cloning and Expression Vector Construction. HPV-52 L1 gene was isolated from a cervical biopsy tissue of an
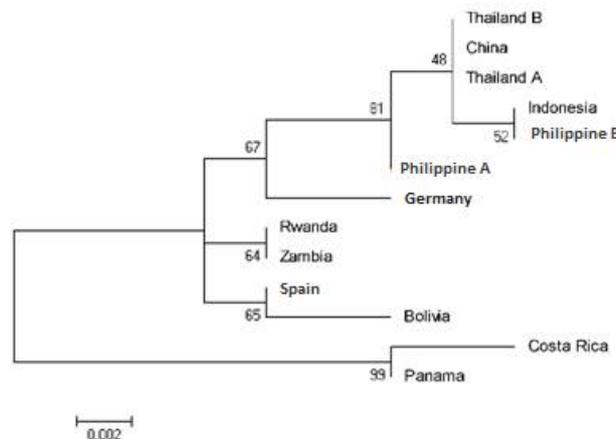

Fig 3 HPV-52 L1 phylogenetic tree. This phylogenetic showed that HPV-52 L1 gene from Asian countries including Indonesia forms its own cluster.

Table 1 T-cell epitopes on sequence of HPV-52 L1 from Indonesia

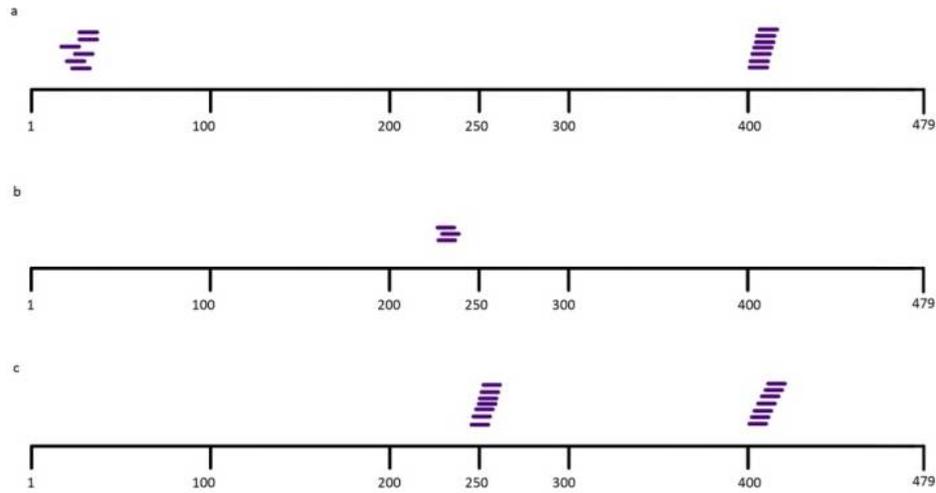| HLA-DRB1*12:02	HLA-DRB1*15:02	HLA-DRB1*07:01
DSLFFFLRREQMFVR	ASLEDTYRFVTSTAI	ASLEDTYRFVTSTAI
SLFFFLRREQMFVRH	SLEDTYRFVTSTAIT	SLEDTYRFVTSTAIT
GDSLFFFLRREQMFV	LEDTYRFVTSTAITC	LEDTYRFVTSTAITC
	EDTYRFVTSTAITCQ	EDTYRFVTSTAITCQ
	DTYRFVTSTAITCQK	DTYRFVTSTAITCQK
	TYRFVTSTAITCQKN	TYRFVTSTAITCQKN
	YRFVTSTAITCQKNT	YRFVTSTAITCQKNT
	LRREQMFVRHFFNRA	TSIYYYAGSSRLLTV
	RREQMFVRHFFNRAG	SRTSIYYYAGSSRLL
	REQMFVRHFFNRAGT	RTSIYYYAGSSRLLT
	EQMFVRHFFNRAGTL	SIYYYAGSSRLLTVG
	QMFVRHFFNRAGTLG	VSRTSIYYYAGSSRL
	MFVRHFFNRAGTLGD	IYYYAGSSRLLTVGH
	FVRHFFNRAGTLGDP	YYYAGSSRLLTVGHP

Fig 4 T-cell epitope candidate site on sequence of Indonesian HPV-52 L1. a: HLA-DRB1*07:01; b: HLA-DRB1*12:02; c: HLA-DRB1*15:02. Purple band represent the peptide sequence of epitope candidate on HPV-52 L1 sequence.
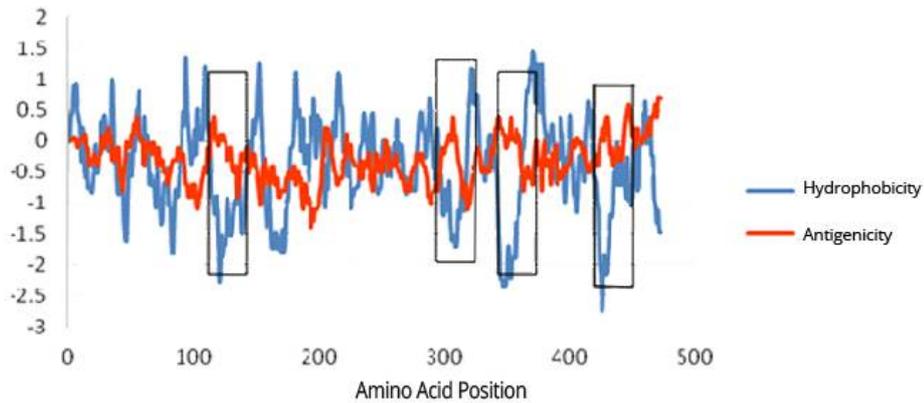

Fig 5 Hydrophobicity and antigenicity plot of HPV-52 L1 protein. This plot showed regions of hydrophobicity and antigenicity on HPV-52 L1 protein sequence. Higher number means higher hydrophobicity and antigenicity property on that particular amino acid position.
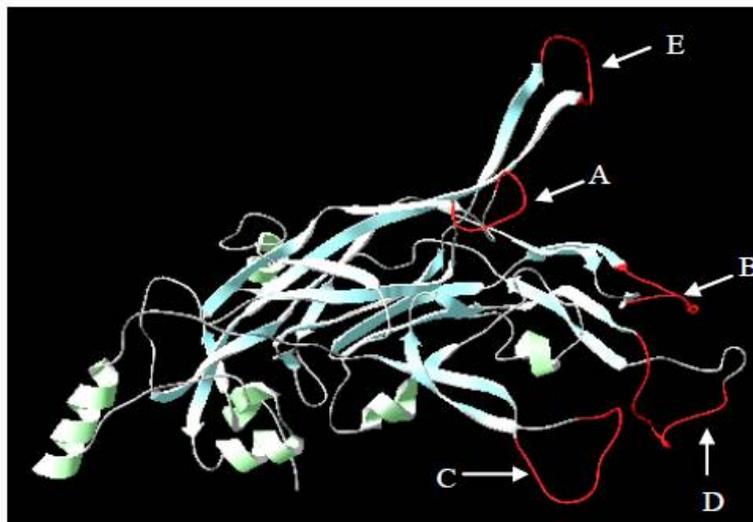

Fig 6 Indonesian HPV-52 L1 3D model structure. Red: Loop; Green: α-helix; Blue: β-strand. A: 46TSSGNGKKVL55, B: 124ETSNKYAGKPGIDNR138, C: 67TPCNNNSGNPGDC179, D: 259GTLGDPVPGDLYIQGS274 and E: 345KKESTYKNE353.

Indonesian patient using specific primers designed for this study. HPV-52 L1 capsid gene was successfully cloned to pGEM®-T Easy cloning vector and pRSET/EmGFP expression vector. Recombinant plasmid sequence analysis showed that the HPV-52 L1 fragment was in-frame with pRSET/EmGFP start codon. To identify the DNA fragment obtained, BLAST analysis was conducted online at http://blast.ncbi.nlm.nih.gov. The result shows 99 % similarity with the sequence in GenBank with accession number JN874433. The cloned HPV-52 L1 sequence did not contain 99 bp of the 5' and 159 bp of the 3' regions of the GenBank sequence JN874433. The HPV-52 L1 gene obtained (1444 bp) was registered to GenBank with accession number Kf225497.

Recombinant HPV-52 L1 Expression and Inclusion Bodies Isolation. The pRSET/EmGFP-HPV52L1 was transformed to *E.coli* BL21(DE3). After the culture reached exponential phase, the expression of HPV-52 L1 protein was induced for 3 hours by 1 mM IPTG for 3 hours at 37 °C, with 200 rpm shaking. Induction started when the transformed *E.coli* BL21(DE3) culture reached exponential phase (OD$_{600}$= 0.8). These expression conditions have been optimized for high HPV-52 L1 protein yield. The HPV-52 L1 was expressed as fusion protein with Histidine tag. SDS-PAGE analysis of the cell lysate showed an approximately 48 kDa protein presents in the cell pellet. However, no significant protein band was found in the supernatant. This result showed that the protein expressed in insoluble form. Inclusion bodies isolation was performed to isolate the HPV-52 L1.

Phylogenetic Tree Construction. Phylogenetic tree was constructed using MEGA 5.05. Maximum likelihood method was used for construction of the phylogenetic tree (Mathura and Kagueane 2009). HPV-52 L1 gene sequences used for constructing phylogenetic tree were obtained from NCBI GenBank Database [http://www.ncbi.nlm.nih.gov]. We used 17 different sequences from different geographic locations to visualize the evolutionary relationship of the Indonesian HPV-52 L1 gene with other known HPV-52 L1 genes. GenBank has limited HPV-52 L1 DNA sequence that has complete origin country data. Unfortunately we were unable to find other HPV-52 L1 originate from Indonesia. We compared the HPV-52 L1 sequence with other HPV-52 L1 sequences from Asia (Philippine, Thailand, and China), Africa (Zambia and Rwanda), America (Panama, Bolivia, and Costa Rica), and Europe (Germany and Spain). The result is presented in Fig. 3.

Epitope Prediction. We combined three analytical methods to determine HPV-52 L1 epitope. The first method that had been used was MHC analysis using IEDBann. Peptides between 8 - 12 mers length were obtained. We restricted the analysis only to peptides with IC50<50nm or netCTLpan score <1% to ensure they bound strongly to the HLA molecules. Peptide with strong affinity to HLA molecule would most likely become T-cell epitope. List of peptide with strong affinity towards HLA molecule is provided in Table 1.

HLA class II alleles that had been used in this study were HLA-DRB1*12:02, HLA-DRB1*15:02 and HLA-DRB1*07:01. These alleles had been chosen because they have high frequency in Western Javanese (Indonesia) population (Yuliwulandari 2008).The alleles HLA-DRB1*15:02 covered 23%, HLA-DRB1*07:01 covered 37.8% and HLA-DRB1*12:02 covered13.1% of western Javanese population (Gustiananda 2011).

We found that the T-cell epitope candidates have different affinity level towards each HLA class we studied. In Figure 3, we described the HPV-52 L1 sequence parts that have high affinity towards HLA molecule. Peptide with strong affinity to HLA molecule would most likely become T-cell epitope. We found out that our T-cell epitope candidates of HPV-52 L1 have strong affinity towards both HLA-DRB1*15:02 and HLA-DRB1*07:01. The T-cell epitope candidates mainly reside around 400aa position.

The second method that has been used is hydrophobicity and antigenicity analysis. In this analysis, we searched for sequences with low hydrophobicity value and high antigenicity value. Hydrophobicity and antigenicity graphs of HPV-52 L1 proteins were obtained from CLC DNA Workbench 6.5. Fig. 5 shows that hydrophobicity peaks are inversed from antigenicity peaks. Regions of antigenicity are 100-120aa, 295-310aa, 340-350aa, and 410-430aa. We suggest that in these areas resides T-cell epitope.

The third and final method is 3D protein structure modelling. We used protein model from Protein Data Bank [PDB id: idzl] as modelling template. This template was chosen because it has deviation value of RMS 0.7Å. RMS deviation values under 1Å indicates good aligment between model and template.

Using SIB:Deep viewer, 5 loop regions in the model of the HPV-52 L1 were determined. The loop regions were determined by the analysis of polar amino

acids and the accessibility of the protein groups. The loop regions obtained are 46TSSGNGKKVL55, 124ETSNKYAGKPGIDNR138, 67TPCNNNSGNPG DC179, 259GTLGDPVPGDLYIQGS274 and 345KKESTYKNE353.We found three structure differences after we compared the resulting HPV-52 L1 model with its template. All difference we discovered located on the loop regions at 23-32aa, 269-279aa, and 349-358aa.

DISCUSSION

Human papillomavirus (HPV) can enter target cells through minor damage of the genital mucosa. If this high risk-HPV infection becomes persistent, it can lead to cervical cancer (Burd 2003). Vaccination can be a feasible solution to prevent cervical cancer development caused by high risk HPV. HPV vaccine can be made using HPV capsid protein called L1 protein. In this study we successfully cloned, expressed and analyzed HPV-52 L1 gene isolated from an Indonesian cervical cancer patient. The HPV-52 L1 obtained was registered to GenBank under accession number KF225497.

E.coli expression vector, pRSET/EmGFP, was used as a vector for HPV-52 L1 expression. Expression of His-HPV52 L1 fusion protein was induced with 1mM IPTG at 37°C when the culture reached exponential phase (OD_{600}=0.8). Induction at mid exponential phase (OD_{600}=0.8) gave the best result compare to induction at late exponential phase (OD_{600}=1.6 or 2.3). His-tag was fused to HPV-52 L1 protein to facilitate protein purification. His-tag binds to nickel or cobalt ions that immobilized on a support matrix such as NTA (nitrilotriacetic acid) (Bornhorst 2000). Ni-NTA affinity chromatography has been proven to produce higher protein yield than GST-tag (Scheich *et al.* 2003). The position of His-tag at a recombinant protein determines the expression level of recombinant protein. His-tag at N-terminus could increase stability of translation initiation region on mRNA structure (Block *et al.* 2009).The expressed His-HPV52 L1 fusion protein was found in insoluble form. This occurred when the protein was expressed at 37°C. Higher solubility might have been achieved if protein had been expressed in cold condition (Seo *et al.* 2011).

The phylogenetic tree showed that the variation of HPV-52 L1 viral DNA occurs geographically. These variations may influence viral infection and development of cervical cancer. Mutation can end up with amino acid substitution that can affect viral assembly, carcinogenic potential and host immune responses. Human immune mechanism against HPV infection was not fully study yet. We cannot be certain that vaccines developed from certain HPV types can protect against infection of other types HPV.

L1 gene variations can influence the persistence of HPV infection to cause cervical cancer (Frati *et al.* 2011). The phylogenetic tree indicates that the HPV-52 L1 Indonesia is located at the same branch with other Asian countries (Thailand and Philippines), while not too close to its counterparts derived from European isolates. According to Chen *et al.* (2011) HPV variants may differ in terms of pathogenicity despite its phylogenetic proximity. HPV type 18, 16 and 52 variants of non-Europeans are commonly found in cancer tissues and cervical injuries (Vet *et al.* 2008 and Chen *et al.* 20). HPV vaccine is commercially available from Cervarix and Gardasil. Both vaccines are made of L1 protein of HPV types 18, 16, 11 and 6,which came from Europe (WHO 2010). Our phylogenetic studies suggest that these vaccines may be less suitable for Indonesian. Both vaccines have been licensed for use in Indonesia since 2009; however there has been no report regarding their degree of efficacy and effectiveness (WHO 2010).

Designing vaccine candidate by epitope prediction *in-silico* can save production cost. In this study we used three methods to determine epitopes of HPV-52. They were MHC analysis, hydrophobicity, antigenicity and protein 3D structure.

MHC class I and II is associated with CD8 + T-cells and CD4 + in cell-mediated immune response (cell mediated immune response / CMI) respectively. Both molecules MHC class I and II have different molecular structures, but both have a peptide binding region. In humans, major histocompability complex (MHC) are also known as human leukocyte antigen (HLA) (Murphy *et al.* 2008). HLA class I and class II genes have the highest rate in the human genome polymorphism. In this study, only HLA genes class II with high allelic frequency in Indonesian population had been used for epitope prediction. The class II HLA-DR,-DQ and -DP are primarily expressed by antigen presenting cells, like dendritic cell, and can present processed antigen to CD4[+] T-cells (Stern 2004). CD4+ cells are important to help B cells to differentiate to antibody secreting plasma cells and memory B-cells (Stanley 2008).

Predicted T-cell epitopes were plotted based on

their position in amino acid sequence obtained. The T cell epitope sites on the sequence of Indonesian L1 protein of HPV-52 are shown in Fig 4. A number of clustered peptide sequences can be seen from this plot. There are two peptide clusters for HLA-DRB1*15:02, one cluster for HLA-DRB1*12:02, and two clusters for HLA-DRB1*07:01. If we align all the HLA alleles' clusters we can see the consensus cluster of the T-cell epitopes is concentrated at end region of the HPV-52 L1 amino acid sequence.

Protein function and interaction were predicted using protein structure analysis. L1 and L2 protein in HPV capsomeres act as a shield to protect the DNA double chain (Burd 2003). L1 protein can form a virus like particle (VLP) that is structurally and immunogenically similar to the infectious HPV. VLP composed of L1 protein has the ability to bind monoclonal antibodies.

Hydrophobicity and antigenicity plot (Fig 5) indicated that the hydrophobicity values are inversed of antigenicity values. This result is supported by the fact that most often antigen determinants (epitopes) are located in polar hydrophilic areas, so it can easily be recognized by antibody (Mohatbakar 2007). Antigen determinants are less likely located in the hydrophobic area because this area tends not to be exposed at the surface.

HPV-52 L1 protein 3D structures (Fig 6) were made using Automated Mode homology modelling that can be accessed online at the site http://swissmodel.expasy.org. This program automatically creates a protein 3D model based on the homology of amino acid sequences (input) with amino acid sequence that has 3D protein structures determined experimentally (template). We used a protein from PDB database as template [PDB id: idzl].

Homology modelling is the most accurate computational method to create reliable 3D protein models (Bordoli et al. 2009). Furthermore, the model protein was analysed using Deep View (ExPASy). Capsid surface exposed loop are determined from this 3D protein model.

According to Stanley et al. (2006) the specific epitopes of the L1 protein is the loop exposed on the surface of capsomeres. Shift of a few angstroms at the surface of the loop or substitution of some residues in the loop may result in changes in the surface antigen determinants. Changes in antigenic determinants (epitopes) on the surface of the loop can cause changes in antibody needed to identify the virus. Epitopes of one HPV type do not necessarily have the same antigenicity for antibodies of other HPV types. Therefore, application of vaccines designed for one type may not be able to prevent HPV infection by other HPV types. It is imperative to add other types of HPV to the vaccine design to achieve broader cases coverage. Addition of HPV type 31, 33, 35, 45, 52 and 58 to the HPV 16/18 vaccine design are estimated to prevent 90% of worldwide cervical cancer burden, which equal to 440,000 cases (Clifford 2006). The incorporation of the data from the MHC analysis, hydrophobicity, antigenicity and protein 3D structure, the epitope regions of the HPV-52 L1 protein isolated from Indonesia used in this study are 259GTLGDPVPGDLYIQGS274 and 345KKESTYKNE353.

REFERENCES

Biemelt S, Sonnewald U, Galmbacher P, Willmitzer L, Muller M. 2003. Production of human papillomavirus type 16 virus-like particle in transgenic plants.J Virol. 77(17): 9211-9220. doi: 10.1128/JVI.77.17.9211-9220.2003.

Block H, Maertens B, Spriestersbach A, Brinker N, Kubicek J, Fabis R, Labahn J, Schäfer F.. 2009. Immobilized-metal affinity chromatography (IMAC): a review. Method Enzymol. 463: 439- 474.

Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. 2009. Protein structure homology modelling using SWISS-MODEL workspace. Nat Protoc. 4: 1-13. doi:10.1038/nprot.2008.197.

Bornhorst JA, Falke JJ. 2000. Purification of protein using polyhistidine affinity tags. Methods Enzymol. 336: 245-254.

Breitburd F, Kirnbauer R, NL. Hubbert, B. Nonnenmacher, C. TrinDinh-Desmarquet, G. Orth, JT. Schiller, DR. Lowy. 1995. Immunization with viruslike particles from cottontail rabbit papillomavirus (CRPV) can protect against experimental CRPV infection. J. Virol. 69:3959–3963.

Burd EM. 2003. Human papillomavirus and cervical cancer.ClinMicrobiol Rev.1: 1-17.

Chen Z, Schiffman M, Herrero R, DeSalle R, Anastos K, Segondy M, Sahasrabuddhe V, Gravitt P, Hsing AW, Burk R. 2011. Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58

and HPV67. PLoS ONE. 6. doi: 10.1371/journal.pone. 0020183.

Clifford G, Franceschi S, Diaz M, Munoz N, Villa LL. 2006. Chapter 3: HPV type-distribution in women with and without cervical neoplastic disease. Vaccine 24S3: 26-34.

De Villiers EM, Fauquet C, Broker TR, Bernard HU, Zur Hausen H. 2004. Classification of papillomavirus. Virology 324(1):17-27.

Frati E, Bianchi S, Colzani D, Zappa A, Orlando G, Tanzi E. 2011. Genetic variability in the major capsid L1 protein of human papillomavirus type 16 (HPV-16) and 18 (HPV-18).Infect Genet Evol. 11: 2119-2124. doi: 10.1016/j.meegid.2011.06.014.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: The Proteomics Server For In Depth Protein Knowledge and Analysis. Nucleic Acis Res.31: 3784-3788. doi: 10.1093/nar/gkg56.

Gomez D, Santoz J. 2007. Human Papillomavirus Infection and Cervical Cancer: Pathogenesis and Epidemiology. Communicating Current Research and Educational Topics and Trends in Applied Microbiology.

Gustiananda M. 2011. Immunoinformatics analysis of H5N1 proteome for designing an epitope-derived vaccine and predicting the prevalence of pre-existing cellular-mediated immunity toward bird flu virus in Indonesian population. Immunome Res. 7: 3.

Mathura VS, Kangueane P. 2009. Bioinformatics : A Concept- Based Introduction. Springer: New York.

Mohatbakar H. 2007. Prediction of epitopes and structural properties of Iranian HPV-16 E6 by bioinformatics methods. Asian Pac J Cancer P, 8: 602-606.

Murphy K, Travers P, Walport M. 2008. Janeway's Immunobiology 7th ed. Garland Science: New York.

Panigoro R, Susanto H, Novel SS, Hartini S, Sahiratmadja E. 2013. HPV genotyping linear assay test comparison in cervical cancer patients: implications for HPV prevalence and molecular epidemiology in a limited-resource area in Bandung, Indonesia.Asian Pac J Cancer Prev. 2013;14(10):5843-7.

Rhodes, G. 2007. Molecular Modelling for Beginner. A c c e s s f r o m : [h t t p : / / s p d b v . v i t a l - it.ch/TheMolecularLevel/SPVTut/index.html].

Rose RC, Bonnez W, Reichman RC, Garcea RL. 1993. Expression of human papillomavirus type 11 L1 protein in insect cells: in vivo and in vitro assembly of virus-like particles. J Virol. 67: 1936-1944.

Sasagawa T, Pushko P, Steers G, Gschmeissner SE, Hajibagheri MA, Finch J, Crawford L, Tommasino M. 1995. Synthesis and assembly of virus-like particles of human papillomaviruses type 6 and type 16 in fission yeast Schizosaccharomycespombe. Virol. 206: 126-135.

Scheich C, Sievert V, Bussow K. 2003. An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinitiychromatography.BMCBiotechnol. 3:12.

Seo P, Heo S, Han EJ, SeoJ,Ghim S, Kim CH. 2009. Bacterial expression and purification of Human papillomavirus type 18 L1. Biotechnol Bioproc Eng. 14:168-174.

Smidkova M, Muller M, Thones N, Piuko K, Angelisova P, Veleminsky J, Angelis KJ. 2010. Transient expression of human papillomavirus type 16 virus-like particles in tobacco and tomato using a tobacco rattle virus expression vector. Biol Plantarum. 54: 451-460.

Stanley M, Lowy DR, Frazer I.2006. Prophylactic HPV vaccines: underlying mechanism. Vaccine 24S3:106-113.

Stanley M. 2008.Immunobiology of HPV and HPV vaccines.GynecolOncol.109:15-21.

SternP. 2004. Immune control of human papillomavirus (HPV) associated anogenital disease and potential for vaccinantion. J of ClinVirol, 32S: 72-81.

Suzich JA, Ghim SJ, Palmer-Hill FJ, White WI, Tamura JK, Bell JA, Newsome JA, Jenson AB, Schlegel R. 1995. Systemic immunization with papillomavirus L1 protein completely prevents the development of viral mucosal papillomas. Proceedings of the National Academy of Science USA.92:11553–11557.

Varsani A, Williamson AL, Rose RC, Jaffer M, Rybicki EP. 2003. Expression of Human papillomavirus type 16 major capsid protein in transgenic *NicotianaTabacum v. Xanthi*.Arch Virol.148: 1771-1786.

Vet JNI, Boer M, Akker BEMW, Siregar B, Lisnawati, BudiningsihS, TyasmorowatiD, Moestikaningsih, Cornain S, Peters AW, FleurenGJ. 2008. Prevalence of human papillomavirus in Indonesia: a population-based study in three regions. Brit J Cancer.99:214 – 218. doi: 10.1038/sj.bjc.6604417.

Warzecha H, Mason HS, Lane C. Tryggvesson A, Rybicki E, Williamson A, Clements JD, Rose RC. 2003. Oral immunigenecity of human papillomavirus-like particle expressed in potato. J Virol.77(16):000.

WHO/ICO Information Centre on HPV and Cervical Cancer (HPV Information Centre).Human papillomavirus and related cancers in Indonesia.*Summary Report 2010*. [10 June 2013]. Available at www.who.int/ hpvcentre.

Yuliwulandari R, Kashiwase K, Nakajima H, Uddin J, Susmiarsih TP, Sofro SM, Tokunaga K. 2008. Polymorphism of HLA genes in western Javanese (Indonesia): close affinities to southeast asianpopulation.TissueAntigens. ISSN 0001-2815.

Zhou J, Sun XY, Stenzel DJ, Frazer I. 1991.Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. Virology 185: 251- 257.