Active sites prediction and binding analysis E1-E2 protein human papillomavirus with biphenylsulfonacetic acid

I Iryani*, F Amelia, I Iswend

Chemistry Department, Faculty of Mathematics and Science, Universitas Negeri Padang, Indonesia

*in.iryani@yahoo.co.id

Abstract. Cervix cancer triggered by Human papillomavirus infection is the second cause to woman death in worldwide. The binding site of E1-E2 protein of HPV 16 is not known from a 3-D structure yet, so in this study we address this issue to study the structure of E1-E2 protein from Human papillomavirus type 16 and to find its potential binding sites using biphenylsulfonacetic acid as inhibitor. Swiss model was used for 3D structure prediction and PDB: 2V9P (E1 protein) and 2NNU (E2 protein) having 52.32% and 100% identity respectively was selected as a template. The 3D model structure developed of E1 and E2 in the core and allowed regions were 99.2% and 99.5%. The ligand binding sites were predicted using online server meta pocket 2.0 and MOE 2009.10 was used for docking. E1-and E2 protein of HPV-16 has three potential binding site that can interact with the inhibitors. The Docking biphenylsulfonacetic acid using these binding sites shows that ligand interact with the protein through hydrogen bonds on Lys 403, Arg 410, His 551 in the first pocket, on Tyr 32, Leu 99 in the second pocket, and Lys 558m Lys 517 in the third pocket.

1. Introduction
Human papillomavirus (HPV) is a double-stranded DNA virus that can induce both benign and malignant tumors on the skin and mucosa epithelium. Nowadays, there are more than 150 types of HPV have been identified and 30 of them can cause genital warts and cancer risk. Various studies have shown that HPV initiate and induce cervical cancer [1].

Based on its ability to induce cervical cancer, HPV is divided into two types of high-risk types (high risk) and low risk (low risk). High risk HPV types are HPV-16, -18, -31, -33, -45 and -56 and low risk types are HPV-4, -11, -13, -27, -34, -44, -65 [2, 3].

Cervical cancer is the second leading cause of death in women with cancer worldwide. Every year, there is an increase of 450,000 cases with 50% of them can lead to death [3]. The higher number of the cervical cancer infection not only in Indonesia which has 11.78% cervical cancer patients but also in United States which has the higher number of women infection of this virus (80%) for woman in the 50 ages [4, 5]. Because of the high impact of HPV infections are harmless, and then the choice of therapy/treatment is needed.

Planning drug use as a target viral replication requires knowledge of the genomic HPV. HPV gene Early (E), Late (L) and Long Control Region (LCR). E1 and E2 genes function as viral DNA replication, gene E4 and E5 genome amplification as well as E6 and E7 are encoprotein. E1 protein has three domains, namely the C-terminal ATPase / helicase domain which serve to release the ends of DNA, DNA-Binding Origin domain (ODB) to recognize viral origin, and the N-terminal domain. E2 protein consists of two domains, N-terminal transactivation domain (TAD) and C-terminal domains. The TAD is involved in transcriptional regulation and are directly related to E1 and C-terminal domains of DNA-binding/ dimerization domain (DBD) [1, 6].
The determination of binding site region is an important factor in an attempt to find a cure cervical cancer. To address this issue, we indicated the amino acid residues of E1-E2 protein HPV which have potential binding to ligand using biphenylsulfonic acid as inhibitor [7].

2. Experimental
The sequences of E1-E2 of human papillomavirus-16 were obtained from the PaVE: the papillomavirus knowledge source online web (http://pave.niaid.nih.gov/). Online Swissmodel (http://swissmodel.expasy.org/) was performed to construct quaternary structure by homology modeling [8]. Selected model validated criteria as assessed by PROCHECK [9]. The binding site prediction of the protein model will be analyzed using metapocket 2.0 online [10, 11] in http://projects.biotec.tu-dresden.de/metapocket. Molecular docking has been done using the MOE 2008.10 software to validate the ligand binding.

3. Results and Discussion
3.1. Structure Analysis
The sequences of E1 and E2 protein HPV-16 has modelled using 2V9P and 2NNU templates respectively (Figure 1). E1 predicted protein has 52.32% similarity to 2V9P while the E2 predicted protein was 100% to 2NNU.

![Figure 1. The predicted homology structure of (a) E1 (b) E2 protein HPV-16.](image)

The predicted analysis of model of E1 and E2 using PROCHECK shows that 88.2 % and 95.7% respectively residues are in the most favorable region and 11.0% and 3.8% in additional allowed region. The percentage of the residues in disallow regions are low, 0.8% and 0.5% in the Ramachandran’s plot (Figure 2). A good quality model would be expected to have over 90% the core and allowed regions[9]. Since 88.2% + 11.0% = 99.2% for E1 and 95.7% + 3.8% = 99.5% for E2, it can be conclude that this model has a good quality of three dimensional structure.
3.2. Binding site protection
Metapocket 2.0 predict the binding site using four methods: LIGSITE (cs), PASS, Q-SiteFinder, and SURFNET to improve the prediction success rate. It calculate both the geometry and energy interaction between the proteins and ligand [10]. The pockets in E1 and E2 protein of HPV-16 were three sites.

Figure 2. Ramachandran plot of the psi/phi distribution of the NS4A model as obtained by PROCHECK (a) E1 (b) E2 protein HPV-16.
Table 1. Protein Residues in Binding Site Analysis using Meta Pocket 2.0.

| Pocket 1 | Protein E1 | Protein E2 |
|----------|------------|------------|
| Polar    | V549 V555 P512 A514 M547 G496 A516 P553 L554 | E74 T78 T81 C140 E141 E142 E80 |
|          | W425 V370 Y428 M420 A415 Y412 A406 Y365 | C109 N84 T107 H161 |
|          | A364 V402 F508 W509 I499 V498 A376 A399 | |
| Non Polar| V549 V555 P512 A514 M547 G496 A516 P553 L554 | L77 A143 Y102 G108 Y138 I139 P106 |
|          | W425 V370 Y428 M420 A415 Y412 A406 Y365 | Y159 |
|          | A364 V402 F508 W509 I499 V498 A376 A399 | |

| Pocket 2 | Protein E1 | Protein E2 |
|----------|------------|------------|
| Polar    | R447 Q441 E603 K483 N597 N599 E591 T565 R462 | S98 K68 Q71 C40 D22 T93 Q95 H18 |
|          | T616 R615 | E100 E39 Q12 S65 Q57 |
| Non Polar| M444 Y448 Y602 F445 L604 I442 L446 F486 L485 | Y19 Y32 W33 M36 A72 L99 I15 L94 |
|          | G482 F592 V601 G598 P600 P593 F594 F456 I473 | V97 Y43 P60 V64 A41 L16 V58 |
|          | L474 L475 F586 F588 L460 F463 L490 L563 G487 | |
|          | M520 I564 A459 V584 V443 F453 F612 L457 W439 | |
|          | V451 | |

| Pocket 3 | Protein E1 | Protein E2 |
|----------|------------|------------|
| Polar    | R447 Q441 E603 K483 N597 N599 E591 T565 R462 | N181 K182 C195 T197 S198 K172 S201 |
|          | T616 R615 | |
| Non Polar| M444 Y448 Y602 F445 L604 I442 L446 F486 L485 | Y154 V183 Y155 W184 P196 V199 |
|          | G482 F592 V601 G598 P600 P593 F594 F456 I473 | F200 |
|          | L474 L475 F586 F588 L460 F463 L490 L563 G487 | |
|          | M520 I564 A459 V584 V443 F453 F612 L457 W439 | |
|          | V451 L464 L620 L622 I467 M422 G466 P561 P560 | |
|          | L557 L513 A516 A514 M420 I518 A415 | |

In an effort to figure out the enzyme-substrate interaction and to determine the key residues responsible for interaction, the model of E1-E2 was docked with biphenylsulfonacetic acid using MOE 2008.10 software. The calculated free energy of binding of protein E1-E2 with biphenylsulfonacetic acid (Figure 3) in pocket 1, 2, and 3 were -15.43137, -14.08253, and -18.11686 kcal/mol respectively. The more negative and lower value of free energy means the more preferable the interaction between protein and ligan [12].

![Figure 3](image-url)  
**Figure 3.** Interaction of E1-E2 Protein HPV-16 with Biphenylsulfonacetic Acid in (a) pocket 1, (b) pocket 2, and (c) pocket 3. Interaction residues with substrate are labeled using arrow.
The Docking biphenylsulfonacetic acid using these binding sites shows that ligand interact with the protein through hydrogen bonds on Lys 403, Arg 410, His 551 in the first pocket, on Tyr 32, Leu 99 in the second pocket, and Lys 558m Lys 517 in the third pocket.

4. Conclusion

The E1-E2 protein HPV-16 has been suggested to be a potential drug target for cervix cancer therapy. The lack of the structural information about this enzyme made us design of the three-dimensional model. The 3D model developed structure of E1 and E2 in the core and allowed regions were 99.2% and 99.5%. They were docked in three different potential binding sites to refine the energy minimization and interaction residues between biphenylsulfonacetic and three pockets. They were Lys 403, Arg 410, His 551 in the first pocket, on Tyr 32, Leu 99 in the second pocket, and Lys 558m Lys 517 in the third pocket.

5. Acknowledgements

This research was founded by the Ministry of Research and Technology and Higher Education of The Republic of Indonesia through Decentralized Competitive Research Fund Scheme (2015) contract no. 205/UN35.2/PG/2015

References

[1] Archambault, C.M.D.A.a.J., Small Molecule Inhibitors of Human Papillomavirus Protein - Protein Interactions. The Open Virology Journal, 2011. 5: p. 80-95.
[2] Woodman, C.B.J., S.I. Collins, and L.S. Young, The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer, 2007. 7(1): p. 11-22.
[3] Kanodia, S.F., Laura M. and W. Martin Kast Mechanisms Used by Human Papillomaviruses to Escape the Host Immune Response. Current Cancer Drug Targets, 2007. 7(1): p. 79-89.
[4] http://www.depkes.go.id/. 28 Mei 2013.
[5] Braaten, K.P. and M.R. Laufer, Human Papillomavirus (HPV), HPV-Related Disease, and the HPV Vaccine. Reviews in Obstetrics and Gynecology, 2008. 1(1): p. 2-10.
[6] Kalvatchev, Z.R., F, Human papilloma virus: realities and perspectives. Biotechnol & Biotechnol 2007. 21(2): p. 137-143.
[7] White, P.W., et al., Biphenylsulfonacetic Acid Inhibitors of the Human Papillomavirus Type 6 E1 Helicase Inhibit ATP Hydrolysis by an Allosteric Mechanism Involving Tyrosine 486. Antimicrobial Agents and Chemotherapy 2005. 49(12): p. 4834-4842.
[8] Biasini, M., et al., SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research, 2014. 42(Web Server issue): p. W252-W258.
[9] Laskowski R A, M.M.W., Moss D, Thornton J M, PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst., 1993. 26: p. 283-291.
[10] Huang, B., metaPocket: a meta approach to improve protein ligand binding site prediction. Omics 2009. 13(4): p. 325-330.
[11] Zengming Zhang, Y.L., Biaoyang Lin, Michael Schroeder and Bingding Huang, Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. Bioinformatics 2011. 27(15): p. 2083-2088.
[12] Kitchen, D.B., et al., Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov, 2004. 3(11): p. 935-949.