14

HPV Bioinformatics: In Silico Detection, Drug Design and Prevention Agent evelopment

Usman Sumo Friend Tambunan and Arli Aditya Parikesit
Department of Chemistry, Faculty of Mathematics and Science, University of Indonesia
Indonesia

1. Introduction

Viral infection is a very serious threat to humanity. It causes malicious diseases, such as HIV/AIDS, dengue, and Avian Influenza, therefore, novel method in virology to combat the viral infection is necessary. Bioinformatics provides outstanding tools for developing vaccines, PCR primers, mutation detection and drugs based design on genetic engineering principles. Those tools are mostly freeware. Algorithm from the computer science has made major contribution to them. Bioinformatics experiment greatly reduces the cost and time in wet laboratory experiment. Our lab has successfully designed PCR primers, vaccine, and mutations prediction. The vaccine design is elaborated for Dengue and HPV. The design has BLAST homology of more than 90%, and RSMD value of 0.1. Those data shown, that the design have identical structure with the native viral protein.

However, their efficacy should be verified in the wet laboratory experiment. The future of medicine will greatly be shaped by advancement in bioinformatics (Tambunan et al, 2010). Moreover, we need to emphasize the needs to understand more about HPV genome and its protein expression.

The human papillomavirus (HPV) is a family of sexually transmitted, double-stranded DNA viruses with over 100 different genotypes identified till date. It is associated with many different types of cancers including cervical, vaginal, head and neck, penile and anal cancer. With approximately 450,000 newly diagnosed cases each year and a 50% mortality rate, cervical cancer is the second most common cause of cancer-related death in women worldwide and it is almost always associated with HPV. Cervical cancer is the most common cancer of women in most developing countries, where it may account for as many as one fourth of female cancers.

HPV genotypes are divided into the low risk and high risk categories based on the spectrum of lesions they induce. The low-risk types induce only benign genital warts and include HPV 6 and 11. The high-risk group containing HPV 16, 18, 31, 33, 45 and 56 is associated with the development of anogenital cancers and can be detected in 99% of cervical cancers , with HPV16 found in 50% of cases.

A consistently effective and safe treatment for HPV infections is currently not yet available. Present therapeutic options are more directed at surgical eradication and/or by destroying

malignant lesions via physical or chemotherapeutical intervention. A majority of these treatments have been developed empirically, few have been thoroughly tested, but none of them are completely satisfactory. In attemps to find additional drugs in the treatment of cervical cancer, inhibitors of the histone deacetylases have received much attention due to their low cytotoxic profiles (Tambunan et al, 2010).

Elaborating a methodological comparison between our computational approaches with the other bioinformatics labs. We will not compare our method directly with wet labs, because our method is developed to supplement it. Our objective is to find the optimal in silico HPV therapeutic design, and implement it in the wet laboratory.

2. In silico HPV related research

2.1 Detection of HPV genome

This section will elaborate more on HPV Genome Annotation. The conserved region of HPV Genome needs to be annotated, in order to design a useful detection tools. Conventional method of detecting cervical cancer is done by carrying out cytological examination, which is more widely known as Pap smear. Due to rapid advancement of molecular biology, molecular based diagnostic and early detection methods on cervical cancer has been highly developed to replace conventional method of detection. Examples include polymerases chain reaction (PCR) and hybrid capture 1, 2 and 3. Hybrid capture 1 is liquid hybridization assaying method designed to detect 14 types of HPV, 9 of which are high risk (type 16, 18,31, 33, 35, 45, 51, 52, 56), while the other 5 are low risk (type 6, 11, 42, 43, and 44). Hybrid capture 2 is a development of hybrid capture 1 which uses microtitre plates instead of tubes and is capable of detecting four additional types of viral oncogenic (type 39, 58, 59, and 68) (Clavel *et al*, 1998). Hybrid capture 3, similar to previous hybrid capture tests, relies on the formation of target HPV DNA-RNA probe heteroduplexes during the hybridization step in specimens containing sufficient HPV DNA. The chemiluminescent detection of these hybrids is by adding an alkaline phosphatase-conjugated monoclonal antibody, specific to the DNA-RNA complexes with dioxetane substrate in a 96-well enzyme-linked immunosorbent assay format (Lorinze and Anthony, 2001). Most of the above mentioned methods are spesific, sensitive, reliable, and easy to perform. Moreover, its routine application has been very much improved by the use of non-radioactive enzyme immunoassay detection procedure (Clavel *et al*, 1998). Modification of the Hybrid Capture method is expected to be achieved through the use of a customized oligonucleotide probe able to detect multiple high-risk HPV infection.

There are 10 genes encoded in HPV genome. These gene may be classified into two groups, namely Gene E (early) which encodes regulatory proteins and Gene L (late) which encodes structural proteins. The region of L1 and L2 of Gene L is responsible for encoding capsid protein to be used as DNA envelope in HPV. The capsid protein will serves as protection system to HPV genetic materials. It is generally assumed that there are many nucleotides sequence in the region of L1 and L2 that is conserved throughout the evolution process of HPV (Dahlgren, 2005).

2.1.1 Other labs approach

Mendez-Tenorio and his group were using DNA fingerprinting. Identification of microorganisms by whole genome DNA fingerprinting was tested “in silico”. 94 HPV

genome sequences were submitted to virtual hybridization analysis on a DNA chip with 342 probes. This Universal Fingerprinting Chip (UFC) constitutes a representative set of probes of all the possible 8-mer sequences having at least two internal and non-contiguous sequence differences between all them. A virtual hybridization analysis was performed in order to find the fingerprinting pattern that represents the signals produced for the hybridization of the probes allowing at most a single mismatch. All the fingerprints for each virus were compared against each other in order to obtain all the pairwise distances measures. A match-extension strategy was applied to identify only the shared signals corresponding to the hybridization of the probes with homologous sequences between two HPV genomes. A phylogenetic tree was constructed from the fingerprint distances using the Neighbor-Joining algorithm implemented in the program Phylip 3.61. This tree was compared with that produced from the alignment of whole genome HPV sequences calculated with the program Clustal_X 1.83. The similarities between both trees are suggesting that the UFC-8 is able to discriminate accurately between viral genomes. A fingerprint comparative analysis suggests that the UFC-8 can differentiate between HPV types and subtypes (Méndez-Tenorio *et al,* 2006)

Kaladhar and his group are working on annotation HPV-92 genome. Most of the biologists focus to explore innovations of their research in faster rate using developments in Information technology. The gene identification, characterization and modeling of the proteins in HPV 92 is done using bioinformatics tools. A complete genome of HPV-92 with NCBI's accession number NC_004500 was submitted to FGENES V0, a viral gene prediction server, predicts six genes. These six genes are characterized as E6 oncoprotein, E7 oncoprotein, E1 Replication protein, E2 Regulatory protein, L1 major capsid protein and L2 minor capsid protein. Isoelectric points and Molecular weights of all the six proteins vary largely and the modeled structures are shown. The research can provide characterization and modeling of genome which can further implemented in drug designing methods using bioinformatics tools (Kaladhar et al, 2010).

Eom and his group has interesting approach towards mapping HPV genome. They are using genetic mining algorithm. Classifying the type of HPV is very important to the treat of cervical cancer. The machine learning approach to mine the structure of HPV DNA sequence for effective classification of the HPV risk types has been introduced. The most informative subsequence segment sets and its weights with genetic algorithm to classify the risk types of each HPV has been determined and learnt. To resolve the problem of computational complexity of genetic algorithm, distributed intelligent data engineering platform based on active grid concept called IDEA@Home was used. The proposed genetic mining method, with the described platform, shows about 85.6% classification accuracy with relatively fast mining speed (Eom et al, 2004).

Lee and his groups are using in silico DNA microarry for detecting HPV genome. DNA microarrays are widely used techniques in molecular biology and DNA computing area. It consists of the DNA sequences called probes, which are DNA complementaries to the genes of interest, on solid surfaces. And its reliability seriously depends on the quality of the probe sequences. Therefore, one must carefully choose the probe sets in target sequences. The probe design for DNA microarrays was formulated as the multi-objective optimization problem. Multi-objective evolutionary approach was proposed, which is known to be suitable for this kind of optimization problem. Since a multi-objective evolutionary

algorithm can find multiple solutions at a time, thermodynamic criteria was used to choose the most suitable one. For the experiments, the probe set generated by the proposed method is compared to the sequences used in commercial microarrays, which detects a set of Human Papillomavirus (HPV). The comparison result supports that the approach can be useful to optimize probe sequences (Lee et al, 2004).

2.1.2 Our approach

The aim of our study was to determine the conserved regions of late genes L1 and L2 from 74 sequenced and published HPV genome (Icenogle, 1995). The result was used to predict candidate template for oligonucleotide probes that are specific on types of HPV, which cause cervical cancer. Nevertheless, the spesific purpose of this study is to design primer that is able to detach on the open reading frame region and also to develop a new assay for the detection of high risk HPV DNA.

This study was carried out to determine the conserved regions of late genes from sequenced HPV types. HPV genome sequences were collected from the Los Alamos National Laboratory *papillomavirus* database. There are 74 types of HPV in the database, which have completely documented genome sequences as well as their translation product. Specific types of HPV, which may cause cervical cancer, are grouped into high risk or low risk, according to their risk potential. This classification may differ from one research methodology to another. In order to access a representative classification, three sets of classification were studied for this research. HPV type 16 and 18 are consistently grouped as high risk, while other types of HPV varied randomly. Sequence alignment was taken and the result shows 62 conserved regions as a primer template for L1 and L2 genes. These conserved regions were then subjected to BLASTn operation in order to search the conserved region with least similarity to low risk HPV and human genome.

Finally, 7 selected conserved regions were examined for secondary structures using NetPrimer program. From this operation, only region 52 (5'-ACAGGCTATGGTGCTATGGA-3') met the criteria to be used as an oligonucleotide primer (Tambunan et al, 2007).

Oligonucleotide primers are considered based upon certain properties, namely: they must not have potential secondary structures such as hairpins or dimmers have a GC content of 45-60%; have a Tm between 52-58%; their 5' ends stability has to be greater than the stability of their 3' ends; be 17-25 nucleotides in length. From NetPrimer analysis results of seven regions, only region 52 meet with the above mentioned criteria. with a NetPrimer rating of 100 (maxium) (Tambunan et al, 2007). The result is shown in table 1. Based on sequence similarity, 62 conserved regions were found. Out of the 62, 7 regions were then used as templates for primers used in detection of high and low risk HPV. From the 7 template candidates, only one met the criteria to be used as an oligonucleotide primer, namely region 52. From the study, region 52 is predicted to be selective to be used in the detection of oncogenic Human *papillomavirus.*

2.2 HPV vaccine design and its post translation studies

A new paradigm of vaccine design is now emerging, following essential discoveries in immunology and the development of bioinformatics tools for T-cell epitope prediction from

Nucleotide sequence	Alignment	Region	BLASTn result and Position on Genome according to HPV type	
	Type	1	Type	16 (5734-5751)
TATCATGCA	16;18;31;45		Type	31 (5647-5660)
			Type	33 (1920- 1929)
			Type	35 (5703-5719)
			Type	45 (5703-5726)
			Type	51 (5616-5639)
			Type	52 (5742-5761)
			Type	53 (7424-7434)
			Type	56 (5693-5715)
			Type	59 (5714-5724)
			Type	66 (5743-5759)
	Type	21	Type	11 (6350-6372)
ATATGGTTG	11;16;18;3T		Type	16 (6227-6246)
	35;68		Type	31 (6143-6165)
			Type	68 (2205-2229)
	Type	31	Type	16 (4907-5003)
	11;16;18;31;		Type	33 (4960-4976)
	35;68		Type	35 (4950-4974)
			Type	50 (4994-5004)
			Type	59 (6143-6165)
	Type	43	Type	16 (5933-5945)
GTTTGGGCCT	11;16;18;		Type	31(5046-5061)
	31;35;68		Type	35 (5009-5906)
			Type	39 (7001-7014)
			Type	45 (2951-2961)
			Type	52 (7096-7003)
			Type	68 (4997-4907)
	Type	45	Type	06 (6063-6001)
	11;16;18		Type	18 (5096-5915)
	35;68		Type	31 (5040-5057)
			Type	39 (5926-5944)
			Type	45 (2951-2961)
			Type	52 (7069-7004)
	Type	46	Type	16 (5979-5909)
TAGTGGCCAT-3	11;16;18		Type	31 (2303-2393)
	31; 35;68		Type	42 (6176-6193)
	Type	52	Type	16 (6240-6264)
	16;18;52		Type	61 (6373-6303)
			Type	68 (2223-2242)
			Type	70 (6190-6209)

Table 1. Conserved regions selected as templates

primary protein sequences. One rationale for this new paradigm is that following exposure to a pathogen, epitope-specific memory T-cell clones are established. These clones respond rapidly and efficiently upon any subsequent infection, elaborating cytokines, killing infected host cells, and marshalling humoral and cellular defences against the pathogen. The most efficient immune response to some pathogens is derived from a number of different T cells that respond to an ensemble of pathogen-derived short peptides called epitopes. Whether an immune response is directed against a single immunodominant epitope or against many epitopes, the generation of a protective immune response does not require the development of T-cell memory to every possible peptide in the entire pathogen. T-cell response to the ensemble of epitopes, not the whole pathogen, is the source from which a protective immune response is derived. Similarly, if an individual is previously exposed to a language, upon hearing just a few words of that language he/she will usually recognize, for example, that French or English is being spoken. Complete mastery of the language is not required for this recognition. Using this analogy to describe epitopes, one could say that they are pathogen-specific 'words' that alert the immune system to the presence of a pathogen. It is now possible to envisage the design of vaccines based on an ensemble of epitopes (a string of words, a few sentences, a paragraph, or a chapter) derived from the genome of a pathogen, using tools that have been developed in the field of immuno-informatics (De Groot et al, 2002). Knowledge about immunology is crucial in designing vaccine. Immunoinformatics, which is a branch of Bioinformatics, is a flourishing field.

The significant breakthrough in HPV vaccine research was accomplished, when capsid protein L1 and L2 was found to be able to assemble themselves to be Virus Like Particles (VLP) during cell expression. VLP is very similar to the native HPV particle, and it includes the conformation epitope that induce the viral neutralization antibody. This is very crucial for the immune system, in order to detect VLP as viral infection, and giving the proper response. Because VLP is coreless and didn't contain the viral DNA, then it is expected that it won't create infection. The produced VLPs are type 6,11, 16, 18, 31, 33, 35, 39, 45,and 58). One VLP Chimeric (cVLP) model has sucessfully induced Mice Cytotoxic T Cell, by joining E6 and E7 capsid protein in the VLP. Some scientist believes, that cVLP has huge potentials to be utilized as infection prevention (Kolls et al, 2000).

Protein must be folded like its native conformation, in order to be activated as mature protein. The protein modification into its native conformation is called as post-translation modification. The polypeptide chain, which consisted of more than 200 residues, is usually folded into two or more globular domain. Most of the domain has 100 until 200 amino acid residues, and having diameter of ~25 Å (Voet *et al*, 1995).

DNA vaccine is designed using choice of a suitable expression vector, ensuring optimal expression by codon optimization, engineering CpG motifs for enhancing immune responses and providing additional sequence signals for efficient translation. DNA vaccines have been one of the latest developments in vaccine technology. DNA vaccines are essentially plasmids capable of expressing an antigenicpeptide in the host. These expressed proteins are recognized as foreign in the cells of the body. They are processed by the host cells and displayed on their surface to alert the immune system and trigger body's immune responses. DNA vaccines have become anattractive alternative to conventional methods due to the fact that it can elicit sustained cell-mediated as well as humoral immun responses, which is very much important in combating pathogenic organisms, especially intracellular

pathogens. Vaccine efficacy can be assessed by correlating the vaccine's immunogenicity such as its ability to induce CD8+ or CD4+ T cells to the HPV oncoproteins with its ability to protect vaccinated animals against formation of tumors or to cause clearance of already established tumors. Recently several techniques like optimizing codons, CpG optimization and promoter and resistance gene insertion have been tried to enhance the immunogenicity of DNA vaccines (Gupta et al, 2009).

2.2.1 Other labs approach

Gupta and his group are working with DNA vaccine. There is a need to develop a new prophylactic DNA vaccine, which can work against different strains of HPVs and may lead to protection of cervical cancer against new pandemicviruses. Potential prophylactic DNA vaccine has been designed by using all the consensus epitopic sequences of HPVs L2 capsid protein and performed in silico cloning of multiepitopic antigenic DNA sequence in pVAX-1 vector. Immunogenicity of vaccine has been enhanced by techniques like codon optimization, engineering CpG motifs, introducing promoters and co-injection with plasmids expressing immune-stimulatory molecules (Gupta et al, 2009).

2.2.2 Our approach

Unlike others, our laboratory is using chimeric protein for designing the vaccine. The cVLP HPV-16 ANN1, ANN2, HMM1, and HMM2 in silico vaccine design were discovered. The BLAST test towards them was generating 96% identity with native L1 HPV-16 protein. Therefore, it is expected that the vaccines could cause same level of immunogenicity with the native protein. The Ramacandran Plot of them showed that the disallowed region plot of non-glycine residue was less than 15%. Henceforth, the quality of the vaccine could be structuraly good. The VAST test toward them showed the RMSD of 0,1 Å, which shows that they have a high structural similarity.

Based upon in silico predition, it was found that post-translational modification could occur at cVLP. During the formation of cVLP, the possible occurred post-translational modification is N-Glicocylation. It is because this modification has N-Xaa-S/T motif which found at our in silico detection method. Although it is expected to happen, it is predicted that this modification wouldn't affect the tability of the cVLP, because its epitope did not affected.

Chimeric virus like particles (cVLP) has been developed as vaccine candidate for preventing cervical cancer. cVLPs are improvement of Virus Like Particles (VLP) by substituting the epitope of L1 HPV -18 and -52 protein to L1 HPV -16 protein. They are ANN1, ANN2, HMM1, and HMM2. The impact of post translation modification will be determined. Based on In Silico study, the dominant post translation modification is glycosylation (Tambunan et al 2007).

However, the next step is to develop in silico plasmid vector for expressing cVLP at the euchariotic host cell. The necessary step is to conduct in vitro experimentation to construct the cVLP HPV L1 at the proper host cell. After the cVLP has sucessfully produced, we could conduct in vivo research to determine the immunogenicity of cVLP at the animal testing, for example at rabbits or mice. Figure 1 shows the amino acid sequence of our vaccine design, while figure 2 shows its 3D visualization.

>SequenceANN1:

KVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIHLPDP
NKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASAYAAN
AGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTQVAVQPGDCPPLELINTVIQDG
DMVDTGFGAMDFTTLQANKLFLRNVNVFSICKYPDYIKMVSEPYGDSLFFYLRREQMFVR
HLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFNKPYWLQRAQ
GHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRFYILVIFYYIFQLC
KITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLDTYRFVTSQAIACQKHTPPAPKEDPL
KKYTFWEVNLKEKFSADLDQFPLGRKFLLQLGL

Fig. 1. The sequence of Our ANN1 cVLP L1 HPV Vaccine design.

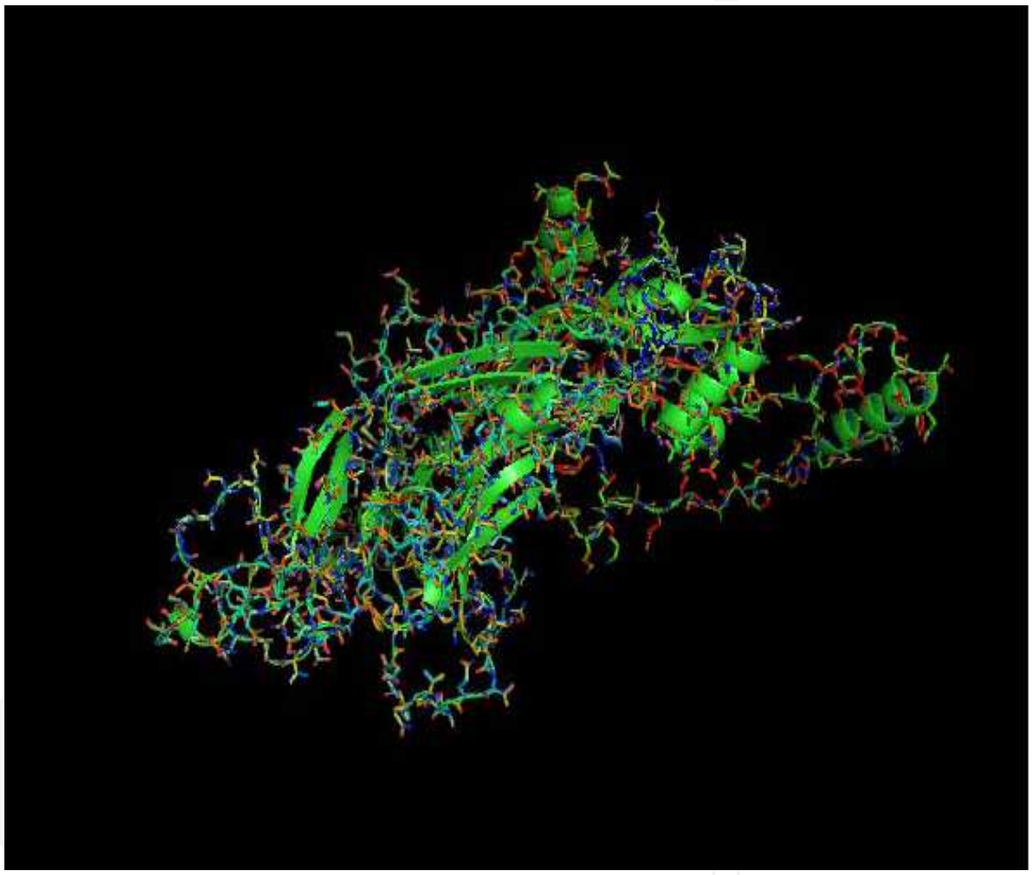


Fig. 2. Our ANN1 cVLP L1 HPV Vaccine Design. The vaccine was visualized by using MacPymol application. The ribbons inside the chains are the vaccine backbone (Tambunan et al, 2010).

2.3 HPV drugs design

This section will elaborate more on Drugs design. Knowledge about protein receptor-inhibitor interaction is very important for designing drugs. Computational Chemistry is the major supporting science in it. The structural modification of histones is playing important roles in the knowledge of HPV drug design.

However, we are going to explain more about drug design biochemistry background. The structural modification of histones is regulated mainly by acetylation/deacetylation of the N-terminal tail and is crucial in modulating gene expression, because it affects the interaction of DNA with transcription-regulatory non-nucleosomal protein complexes. The balance between the acetylated/deacetylated states of histones is mediated by two different sets of enzymes: histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs prefentially acetylate specific lysine substrates among other non-histones protein sub- strates and transcription factors, affecting DNA-binding properties and, in turn, altering gene transcription. HDACs restore the positive charge on lysine residues by removing acetyl groups and thus are involved primarily in the repression of gene transcription by compacting chromatin structure. Therefore, open lysine residues attach firmly to the phosphate backbone of the DNA, preventing transcription. In this tight conformation, transcription factors, regulatory complexes, and RNA polymerases cannot bind to DNA Acetylation relaxes the DNA conformation, making it accessible to transcription machinery. High levels of acetylation of core histones are seen in chromatin-containing genes, which are highly transcribed genes; genes that are silent are associated with low levels of acetylation. Inappropriate silencing of critical genes can result in one or both hits of tumor suppressor gene inactivation in cancer. Members of the classical HDAC family fall into two different phylogenetic classes, namely class I and class II. The class I HDACs (HDAC1, HDAC2, HDAC3, and HDAC8) are most closely related to the yeast (*Saccharomyces cerevisiae*). Class II HDACs (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, and HDAC10) share domains with similarity to HDA1, another deacetylase found in yeast (de Ruitter et al, 2003).

The inhibition of HDAC activity by a specific inhibitor induces growth arrest, differentiation, and apoptosis of transformed cells as well as several cancer cells (Subha et al, 2008). Recent studies were directed to investigate the molecular effects of HDAC inhibition on cervical carcinoma cells as well as on primary human foreskin keratino-cytes, separately immortalized with amphotropic retroviruses that carry the open reading frames of HPV 16 E6, E7 or E6/E7. In these experiments one could show that E6/E7 oncogene function of human papillomavirus can be completely bypassed by HDAC inhibition. Both malignant and immortalized HPV 16/18-positive cells became blocked in G1/S transition despite ongoing viral gene expression. G1 arrest was accompanied by a down-regulation of cyclin D and cyclin A and a concomitant up regulation of the cyclin kinase inhibitors (CKI) p21 and p27. Binding of both CKIs led to a complete block of the cyclin-dependent kinase (cdk2) activity and in turn prevented binding of E7. This was intriguing with respect to the reversibility of HPV transformation process, since it is thought that the abrogation of the growth inhibitory function of p21 and p27 through E7 represents a key event in HPV-induced carcinogenesis. HDAC inhibitors also trigger pRb degradation, while E2F expression remained unaffected. pRb degradation is an E7-specific phenomenon, since in E6-positive cells pRb only became hypophosphorylated. The presence of E2F under cell cycle arrest led to a classical “conflict situation” which finally induced apoptosis (Finzer et al, 2001; Finzer et al, 2002; Finzer et al, 2004). Hence, the knowledge how the transforming potential of HPV can be bypassed without switching off viral transcription could open new therapeutical perspectives for the treatment of cervical cancer (Acharya et al, 2006). The drug design on HPV are mainly tampering with the reactivity of E(early) protein.

2.3.1 Other labs approach

Rehmi and his group are mainly working on E2 proteins. The E2 protein from HPV 16 was selected as a molecular target and its known structures were exploited for broad scope of "hits" to be identified in the screening process. They compared both structure-based and ligand-based design approaches for virtual screening. Databases enriched in natural compounds were used for virtual screening based on molecular docking. In this study, they identified novel classes of HPV inhibitors by means of a structure-based drug-design protocol involving Pharmacophore based virtual screening with molecular docking simulation (Rehmi et al, 2009).

Baleja and his group have different approach, because they are using E6 protein as vaccine template. The E6 protein from the high-risk HPV types represents an attractive target for intervention because of its roles in viral propagation and cellular transformation. E6 functions in part by interaction with human cellular proteins, several of which possess a helical E6-binding motif. The role for each amino acid in this motif for binding E6 has been tested through structure determination and site-directed mutagenesis. These structural and molecular biological approaches defined the spatial geometry of functional groups necessary for binding to E6. This E6-binding information (the E6-binding pharmacophore) was transferred into a three-dimensional query format suitable for computational screening of large chemical databases. Compounds were identified and tested using in vitro and cell culture-based assays. Several compounds selectively inhibited E6 interaction with the E6-binding protein E6AP and interfered with the ability of E6 to promote p53 degradation. Such compounds provide leads for the development of new pharmacologic agents to treat papillomavirus infections and their associated cancers (Baleja et al, 2006).

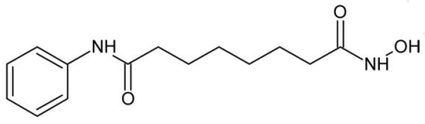
2.3.2 Our approach

Our approach was done by Identification of a better *Homo sapiens* Class II HDAC inhibitor. The aim of this work is to analyze the interaction of *Homo sapiens* class II HDACs with SAHA and TSA that are already in the phase I/II clinical trials based on their binding affinity and pharmacological properties. Since, no theoretical works have been carried out in identifying the properties and specificity, we intend to identify the group that could act as potential binding inhibitors.

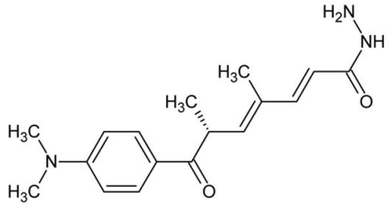
In this paper, we present homology models of six *Homo sapiens* Class II HDACs (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, and HDAC10) that are validated by comparison with the X-ray structure of HDAC4 and HDAC7, which became available during the course of our study. Two HDAC inhibitors (SAHA and TSA) are docked to the six homology models. The pharmacological properties of SAHA and TSA were identified using Molinspiration, Osiris Property, Tox- Boxes, and Toxmatch-v1.06. Therefore, the molecular binding interactions between the histone deacetylases with SAHA and TSA were analyzed to provide some insights into the molecular interactions and designing new HDAC inhibitors.

The certain types of HPV are involved in the development of cervical cancer. In attempts to find additional drugs in the treatment of cervical cancer, inhibitors of the histone deacetylases (HDAC) have received much attention due to their low cytotoxic profiles and the E6/E7 oncogene function of human papilomavirus can be completely by passed by HDAC inhibition. The histone deacetylase inhibitors can induce growth arrest,

differentiation and apoptosis of cancer cells. HDAC class I and class II are considered the main targets for cancer. Therefore, the six HDACs class II was modelled and about two inhibitors (SAHA and TSA) were docked using AutoDock 4.2, to each of the inhibitor in order to identify the pharmacological properties. Based on the results of docking, SAHA and TSA were able to bind with zinc ion in HDACs models as a drug target. SAHA was satisfied almost all the properties i.e., binding affinity, the Drug-Likeness value and Drug Score with 70% oral bioavailability and the carbonyl group of these compound fits well into the active site of the target where the zinc is present. Hence, SAHA could be developed as potential inhibitors of class II HDACs and valuable cervical cancer drug candidate (Tambunan et al, 2010).



Suberoyl Anilide Hydroxamic Acid (SAHA)



Trichostatin A (TSA)

Fig. 3. 2D Structure of SAHA and TSA. SAHA and TSA are hydroxamic acid derivatives that can be HDAV inhibitors.

Each ligand shows different affinity with class II HDAC, for example SAHA compound show best affinity with HDAC 5 based on Autodock calculation, and HDAC 10 based on APBS calculation. Whereas the same compound was found to be rank 2 with HDAC7 (-7.42 kcal/mol) based on AutoDock calculation and HDAC6 (-213.60 kJ/mol) based on APBS calculation, and rank 3 (-6.72 kcal/mol) with HDAC10 (AutoDock) and HDAC7 (-203.21 kJ/mol, APBS). Local free binding energy obtained from AutoDock of *Homo sapiens* class II HDACs complexed with an inhibitor showed that SAHA is a weaker inhibitor of HDACs than TSA. But, global binding energy of *Homo sapiens* class II HDACs and inhibitors obtained from APBS, showed that TSA to be a weaker inhibitor of HDACs than SAHA. There are differences in binding energy calculated with AutoDock
and APBS; this is because AutoDock does not calculate columbic contribution from all of atoms in protein like APBS.

The further descriptor analysis and the toxicity prediction helped in the identification of the better inhibitor. Drug Score and the Drug-Likeness are the two properties that are important for considering a compound to become a successful drug. TSA had a drug score of 0.37 and drug likeness property score of 1.24, which is higher than those for SAHA with respective scores of 0.35 and -8.87. The molecular weight of SAHA was 264.32 g/mol and that of TSA was 302.37 g/mol (Table 1), between the preferred range of molecular weight for drug

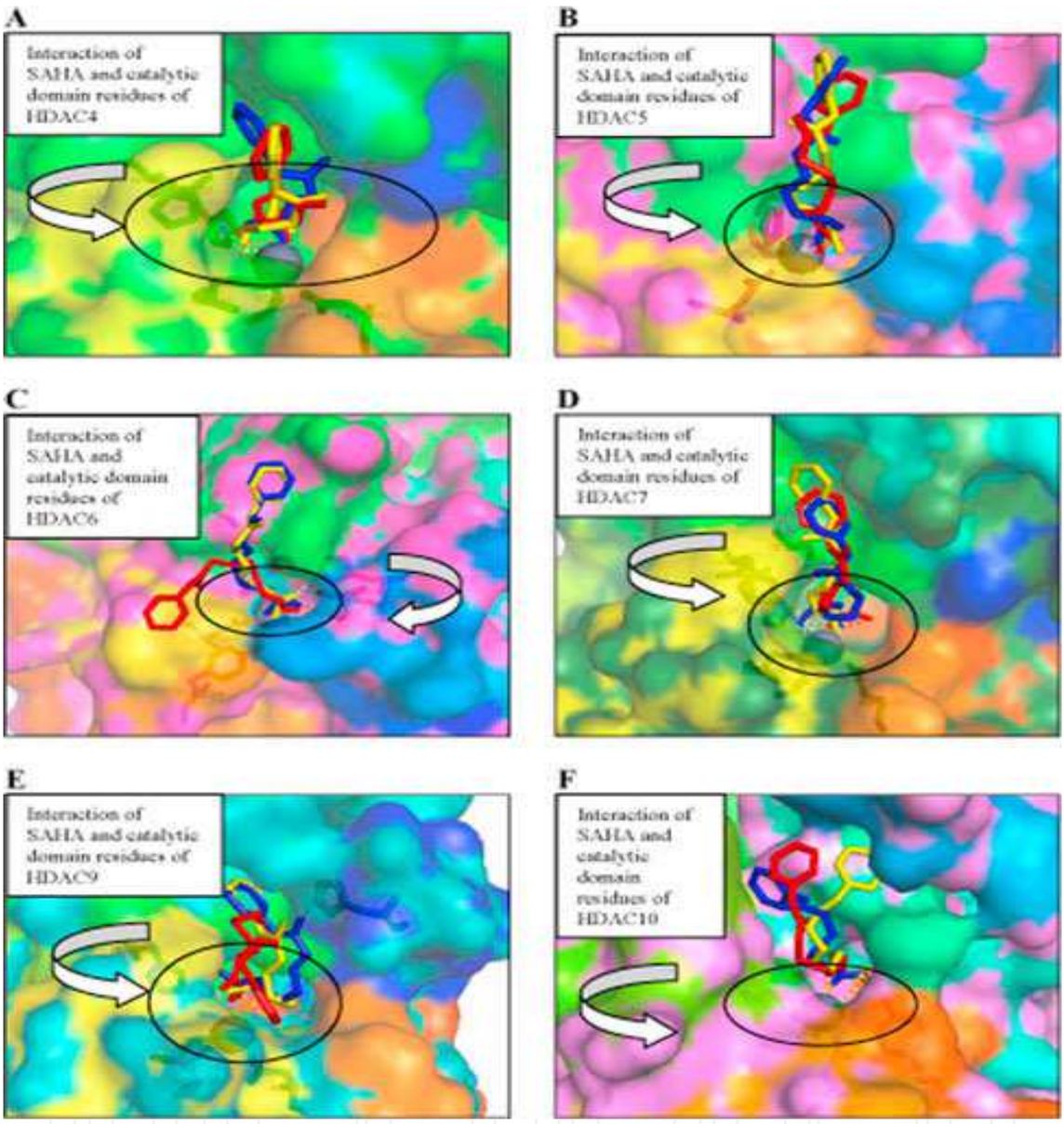


Fig. 4. Structures of docked SAHA with *Homo sapiens* Class II HDACs. Three Conformations of Structures of docked SAHA with (A)HDAC4, (B) HDAC5, (C) HDAC6, (D) HDAC7, (E) HDAC9 and (F) HDAC10. A surface representation of catalytic domain of Homo sapiens Class II HDACs bound to SAHA. The zinc ion is shown as gray sphere. SAHA are shown as stick models colored as per docked type: red, blue, and yellow. Amino acids coordinating the zinc and forming the trihedrally coordinates are shown as sticks. Some catalytic domain of Homo sapiens Class II HDACs residues interacting with the docked SAHA are shown as stick models. In HDAC4-SAHA, HDAC5-SAHA, HDAC6-SAHA and HDAC7-SAHA complexes, SAHA binds the catalytic zinc ion in a bidentate fashion, with its carbonyl and hydroxyl bound to catalytic zinc ion. Whereas, in HDAC9-SAHA and HDAC10-SAHA complexes bind the catalytic zinc ion in a monodentate fashion, with its carbonyl bound to catalytic zinc ion.

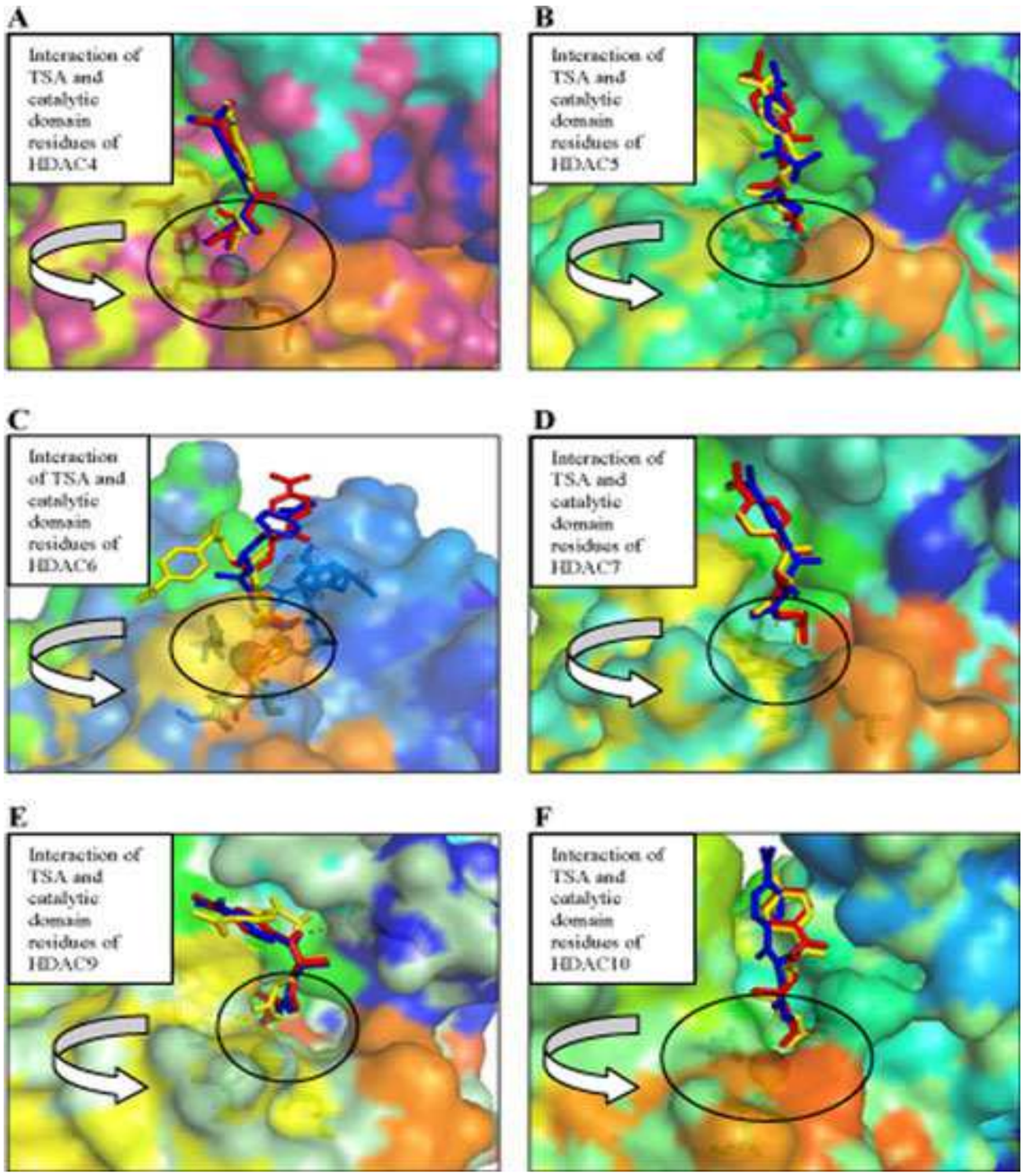


Fig. 5. Structures of docked TSA with HDAC Class II *Homo sapiens*. Three Conformations of Structures of docked TSA with (A) HDAC4, (B) HDAC5, (C) HDAC6, (D) HDAC7, (E) HDAC9 and (F) HDAC10. A surface representation of catalytic domain of *Homo sapiens* Class II HDACs bound to TSA. The zinc ion is shown as gray sphere. TSA are shown as stick models colored as per docked type: red, blue, and yellow. Amino acids coordinating the zinc and forming the trihedrally coordinates are shown as sticks. Some catalytic domain of *Homo sapiens* Class II HDACs residues interacting with the docked TSA are shown as stick models. Except HDAC5-TSA complex, all of *Homo sapiens* class II HDACs-TSA complexes bind the catalytic zinc ion in a monodentate fashion.

likeness property (160-480 g/mol). TSA was having tumorigenic property. SAHA was the compound that had the acceptable range for toxicity risk. These values were also taken into account to decide the best inhibitor. Thus, SAHA was the best drug candidate than TSA and also found to possess better global binding affinity score.

The ADME-TOX box results showed that the SAHA has an oral bioavailability of more than 70% i.e., good solubility and stability. It acts as a non-substrate and non-inhibitor of P-gp. SAHA does not undergo significant first-pass metabolism.

The three-dimensional models for six class II histone deacetylases (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, and HDAC10), which were built using homology modeling and validated by bioinformatics techniques and by comparison to an X-ray structures, were docked to SAHA and TSA. Our studies provide the structural view of the catalytic domain of a class II HDAC and reveal for this subclass specific features: (i) novel zinc binding motif that is likely to be involved in substrate binding and/or protein-protein interactions and may provide a site for modulation of activity, and (ii) a unique active site topology in catalytic activity. SAHA and TSA predicted to inhibit the class II HDACs are effective to all forms HDACs. SAHA was satisfied almost all the properties i.e., binding affinity scores of SAHA in the six class II HDAC enzymes was -156.94 kJ/mol, -171.77 kJ/mol, -213.60 kJ/mol, -203.21 kJ/mol, -179.29 kJ/mol & -263.15 kJ/mol respectively, the Drug Likeness value (1.24) and drug score (0.37) with 70% oral bioavailability and the carbonyl group of these com- pounds fits well into the active site of the target where the zinc is present. Hence, SAHA could be developed as potential inhibitors of class II HDACs and valuable anti cancer agents.

3. Conclusion

The IT industry has provided strong and robust computing power, with low cost expediture. Nowadays, a powerful low cost multiprocessor computers are available,which made the modeling of complicated proteins and sophisticated drug design possible. The major computer operating system, such as MacOSX, Linux, and Windows are already supporting open source bioinformatics software. They could do the functionalities of the commercial software, with the same robustness. Nowadays, the field of bioinformatics is growing. The In Silico (Bioinformatics) experiment will be considered as important as wet experiment by biologist and/or biotechnologist. The In silico approach did not designed to replace wet experiment, but it's in order to supplement it. Open source implementation will help bioinformaticians to solve viral threat in efficient and effective manner. There will be more robust bioinformatics tools available in the future for solving crucial virology related problems. Our laboratory has successfully designing primer and vaccine for therapeutics. However, the efficacy of the design must be proven in the wet labs experiment. Synthesizing them by using latest molecular biology instrument is crucial for progressing towards clinical trial. Conducting it will require us to form strong cooperation with faculty of medicine in our university. We already have cooperation with them, and will verify our design in the future.

HPV Bioinformatics is a growing and developing field. Our labs has sucessfully developed HPV Genome detection, Vaccine, and drugs design. We found 7 conserve region candidates for PCR design, HPV L1 cVLP for vaccine design, and SAHA/TSA drug design. In silico PCR Primer, Vaccine, and drugs design are possible with the newest development in

algorithm and programming. Our labs and others has successfully developed them. The next challenges would be implementing them in the wet laboratory research.

4. Aknowledgement

This publication is supported by Universitas Indonesia. The authors are grateful to to Dr. Ridla Bakri, Chairman of Department of Chemistry, Faculty of Mathematics and Natural Science, University of Indonesia for his support.

5. References

Acharya, MR; Sparreboom, A; Sausville, EA; Conley, BA; Doroshow, JH; Venitz, J & Figg, WD. (2006). Interspecies differences in plasma protein binding of MS-275, a novel histone deacetylase inhibitor, *Cancer Chemotherapy and Pharmacology*, 57:275-81. ISSN: 1432-0843

Alfonso, M; Perla, F; Armando, G; Hueman, J; Emma, R; Arcadio, M; Mercedes, E; Rogelio, M & Loren, BK. (2006). In silico evaluation of a novel DNA chip based fingerprinting technology for viral identification, *Revista latinoamericana de microbiología*, 48 (2): 56-65

Baleja, JD; Cherry, JJ; Liu, Z; Gao, Hua; Nicklaus, MC; Voigt, JH; Chen, JJ & Androphy, EJ. (2006). Identification of inhibitors to papillomavirus type 16 E6 protein based on three-dimensional structures of interacting proteins, *Antiviral Research*, Volume 72, Issue 1, 2006, Pp 49-59, ISSN 0166-3542,

Clavel, C; Masure, M & Putaut, I. (1998). Hybrid Capture II, a new sensitive test for human *papillomavirus* detection. Comparison with hybrid capture I and PCR results in cervical lesions. *Journal of Clinical Pathology*, 51:737-740. ISSN: 1472-4146

Dahlgren, L. (2005). *Studies on the Presence and Influence of human papillomavirus (HPV) in Head and Neck Tumors*. ISBN: 91-7140-289-6. Karolinska University of Stockholms

De Groot, AS;, Sbai, H;, Aubin, CS; McMurry, J & Martin, W. (2002). Immuno-informatics: Mining genomes for vaccine components. *Immunology and Cell Biology. Nature Publishing Group*. Volume 80 issue 3. pp 255- 269. ISSN: 1440-1711

De Ruijter, AJ; van Gennip, AH; Caron, HN; Kemp, S & van Kuilenburg, AB. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family, *Journal of Biochem* , 370:737-49. ISSN: 1470-8728

Eom, J; Park, S & Zhang, B. (2004). Genetic Mining of DNA Sequence Structures for Effective Classification of the Risk Types of Human Papillomavirus (HPV). *Neural Information Processing: Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. pp 1334-1343
Volume: 3316. ISSN: 0302-9743

Finzer, P; Krueger, A; Stöhr, M; Brenner, D; Soto, U; Kuntzen, C; Krammer, PH & Rösl, F. (2004). HDAC inhibitors trigger apoptosis in HPV-positive cells by inducing the E2F-p73 pathway. *Oncogene*, 23:4807-17.ISSN: 09509232

Finzer, P; Ventz, R; Kuntzen, C; Seibert, N; Soto, U & Rösl, F. (2002) Growth arrest of HPV-positive cells after histone deacetylase inhibition is independent of E6/E7 oncogene expression, *Journal of Virology*, 304:265-73. ISSN: 1096-0341

Finzer, P; Kuntzen, C; Soto, U; zur Hausen, H & Rösl, F. (2001). Inhibitors of histone deacetylase arrest cell cycle and induce apoptosis in cervical carcinoma cells

circumventing human papillomavirus oncogene expression,*Oncogene* , 20:4768-76. ISSN: 09509232

Gupta, SK; Singh, A; Srivastava, M; Gupta, SK & Akhoon, BA. (2009) In silico DNA vaccine designing against human papillomavirus (HPV) causing cervical cancer, *Vaccine*, Volume 28, Issue 1, Pages 120-131, ISSN 0264-410X

Icenogle, J. (1995). *Analysis of the Sequences of the L1 and L2 Capsid Proteins of Papillomavirus*. Centers for Disease Control, Atlanta, Georgia.

Kaladhar, DSVGK; Devi, TU & Nageswara, RPV. (2010). An In Silico Genome Wide Identification, Characterization and Modeling of Human Papilloma Virus Strain 92. *International Journal of Engineering Science and Technology.*Vol. 2(9), 4288-4291. ISSN: 0975–5462.

Kolls, A & Sherris, J. (2000). *HPV Vaccines: Promoises and Challenges*. PATH Seattle.

Lee, I; Kim, S & Zhang, B. Multi-objective Evolutionary Probe Design Based on Thermodynamic Criteria for HPV Detection. PRICAI 2004: Trends in Artificial Intelligence. Lecture Notes in Computer Science. 2004. Springer Berlin / Heidelberg. pp: 742-750.Volume: 3157

Lorinez, A. & J. Anthony. (2001). Hybrid Capture method for detection of human papillomavirus DNA in clinical specimens: a tool for clinical management of equivocal Pap smears and for population screening, J*ournal of obstetrics and gynaecology*, 12:145-154. ISSN: 1341-8076

Reshmi, G & Pillai, MR. (2009). In Silico Screening of Novel Inhibitors for HPV: A Rational Structure Based Approach (Docking Versus Pharmacophore Model Generation), *Letters in Drug Design & Discovery,* Volume 6 Issue 7.pp.494-501. ISSN: 1570-1808.

Subha, K & Kumar, GR. (2008). Assessment for the identification of better HDAC inhibitor class through binding energy calculations and descriptor analysis., *Bioinformation,* 3:218-222. ISSN: 0973-2063

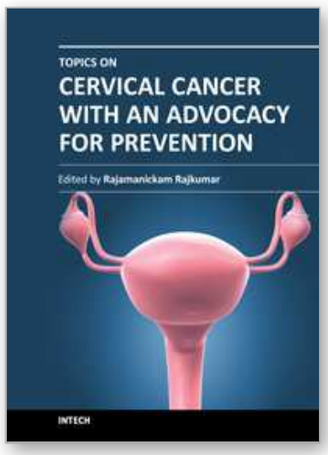
Tambunan, USF; & Wulandari, EK. (2010).Identification of a better *Homo sapiens* Class II HDAC inhibitor through binding energy calculations and descriptor analysis, *BMC Bioinformatics* , 11(Suppl 7):S16. ISSN:1471-2105

Tambunan, USF & Parikesit, AA. (2010). Cracking The Genetic Code of Human Virus by using Open Source Bioinformatics tools, J*ournal of Fundamental Science*, 6(1).ISSN: 1823-626X

Tambunan, USF; Parikesit, AA;, Tochary, TA & Sugiono, D. (2007). Computational Study of Post Translation Modification in Chimeric Virus Like Particles Vaccine of Human Papilloma Virus with Virion Capsid L1, *Makara Sains Journal,* 11(2): 56-62. ISSN: 1693-6671

Tambunan, USF; Butar-Butar ,HW; Umbas, R & Hidayah, Z. (2007). Conserved Region Analysis of Oncogenic Human Papillomavirus Genome., *Biotechnology,* (6)1: 93-96. ISSN: 16822978

Voet, D., Voet J.G. (1995). *Biochemistry,* John Willey and Sons Inc,. ISBN-13: 978-0471586517, New York



Topics on Cervical Cancer With an Advocacy for Prevention

Edited by Dr. R. Rajamanickam

ISBN 978-953-51-0183-3
Hard cover, 284 pages
Publisher InTech
Published online 02, March, 2012
Published in print edition March, 2012

Cervical Cancer is one of the leading cancers among women, especially in developing countries. Prevention and control are the most important public health strategies. Empowerment of women, education, "earlier" screening by affordable technologies like visual inspection, and treatment of precancers by cryotherapy/ LEEP are the most promising interventions to reduce the burden of cervical cancer.Dr Rajamanickam Rajkumar had the privilege of establishing a rural population based cancer registry in South India in 1996, as well as planning and implementing a large scale screening program for cervical cancer in 2000. The program was able to show a reduction in the incidence rate of cervical cancer by 25%, and reduction in mortality rate by 35%. This was the greatest inspiration for him to work on cerrvical cancer prevention, and he edited this book to inspire others to initiate such programs in developing countries. InTech - Open Access Publisher plays a major role in this crusade against cancer, and the authors have contributed to it very well.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen