Research Article **Open Access**

Investigation of potential antiviral natural products with an effect on HPV18 E6 protein by molecular docking method

Mehdi Fazeli[1], Hosna Sarvazad[2], Nasrin Rahnejat[3], Rezvan Rostampour[2], Mahtab Ghanbari rad[4], Narges Eskandari Roozbahani[2*]

[1]Department of Basic Sciences, School of Veterinary Medicine, Shiraz University, Shiraz, Iran; [2]Clinical Research Development Center, Imam Reza Hospital, Kermanshah University of Medical Sciences, Kermanshah, Iran; [3]Kermanshah University of Medical Sciences, Faculty of Nursing and Midwifery, Kermanshah, Iran; [4]Department of Physiology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

[*]**Corresponding Author:** Dr. Narges Eskandari Roozbahani, Clinical Research Development Center, Imam Reza Hospital, Kermanshah University of Medical Sciences, Kermanshah, Iran

Submission Date: September 1st, 2021; **Acceptance Date:** November 8th, 2021; **Publication Date:** November 12th, 2021

Please cite this article as: Fazeli M., Sarvazad H., Rahnejat N., Rostampour R., Ghanbari rad M., Roozbahani N.E. Investigation of potential antiviral natural products with an effect on HPV18 E6 protein by molecular docking method. *Functional Foods in Health and Disease* 2021; 11(11): 586-603. DOI: https://www.doi.org/10.31989/ffhd.v11i11.840

ABSTRACT

Background: Infection with the Human Papillomavirus (HPV) causes cellular dysplasia, which leads to cervical cancers in women and penile or rectal cancers in men.

Objective: This *in silico* study identified the plant compounds with potential therapeutic effects against HPV 18 oncogenic virus using the molecular docking method.

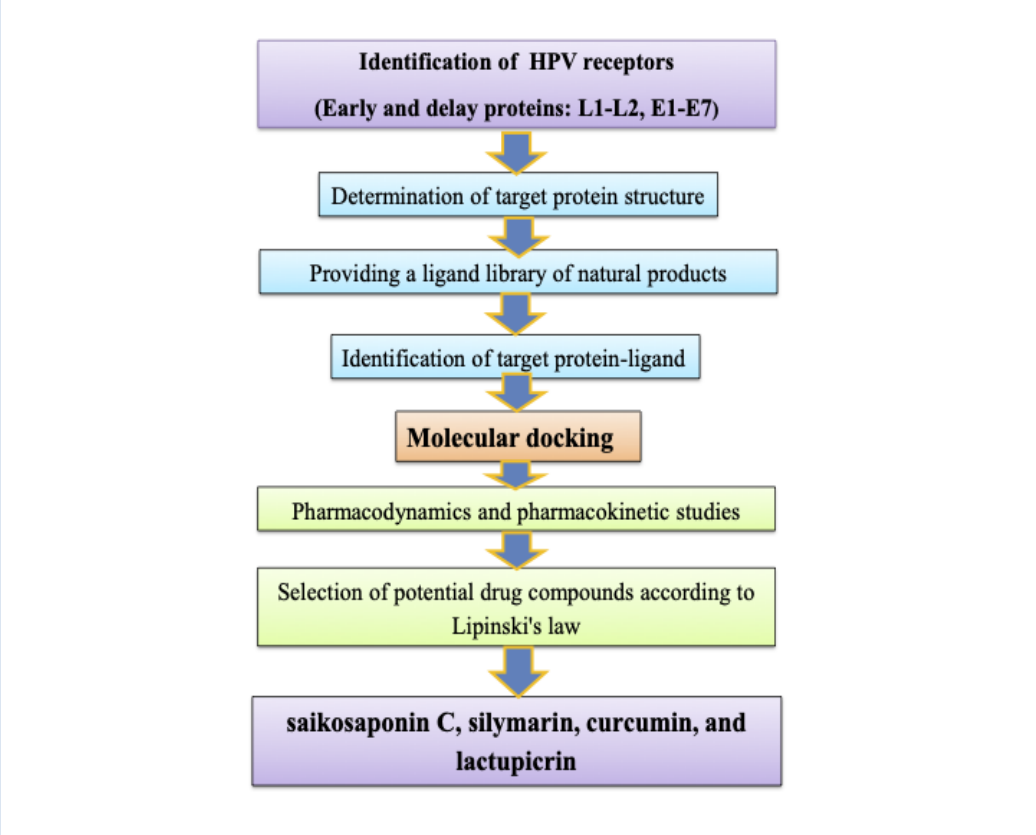
Methods: The three-dimensional (3D) structure of HPV18 E6 protein, as the target protein, and the 3D structure of plant compounds with potential therapeutic effect against viruses, as ligands, was obtained from the protein databases (RCSB) and PubChem, respectively. Both structures of ligands and target protein were subjected to AutoDock tools-1.5.6, *ver*.4 separately. The structure with the most negative affinity was docked to reconsider its connection location. The results were analyzed more based on pharmacodynamic and pharmacokinetic parameters.

Results: The docking of HPV18 E6 protein with 19 selected ligands resulted in four compounds, curcumin, silymarin, saikosaponin c, and lactupicrin, showing the best docking scores; they had better binding free energies with HPV E6 protein. Among four compounds against HPV18 E6, silymarin and curcumin were less dangerous than other compounds

due to the lack of inhibition of the human Ether-à-go-go-Related Gene (hERG). Of these two compounds, silymarin had lower oral absorption, lactopicrin had less skin absorption, lactopicrin is the substrate of P-gp, and saikosaponin c crosses the blood-brain barrier.

Conclusion: Among potential antiviral plants against HPV18E6, four compounds were found to be effective. According to these findings, it is recommended that *in vitro* and *in vivo* examinations be conducted to determine the effectiveness of these compounds against HPV18

Keywords: Biological products, Antiviral agents, HPV18, Molecular docking, Computational biology**,** E6 protein



INTRODUCTION

Human Papilloma Virus (HPV) is the most common sexually transmitted disease, affecting 13% of the world's population [1]. According to previous reports, 34,800 new HPV-related cancers can be expected to occur each year if HPV is not prevented [2].

More than 100 types of HPV have been identified, with 40 strains affecting the anogenic region. Some HPV strains cause cell dysplasia leading to cervical cancer in women and penile or anal cancer in men. Among them, HPV types 16 and 18 are high-risk subtypes for malignancy [1].

The HPV genome expresses two classes of proteins through alternative splicing; the first is regulatory or nonstructural proteins, which include E1 to E7, and the second category is delayed or structural proteins, which include L1 and L2. Among the primary proteins, E1 and E2 are regulatory proteins while E6 and E7 are oncogenic proteins. Depending on their oncogenic potential, HPVs

types are divided into high-risk and low-risk types, the first of which is the cause of many benign warts, and the latter is associated with malignant diseases [3]. There are approximately 15 types of high-risk HPV, the most important of which are HPV16 and HPV18, which are responsible for more than 70% of cervical cancers in women [4]. In cervical cancer, the integration of the virus genome leads to abnormal expression of E6, E7, and loss of expression of E1, E2, E4, E5, as well as capsid proteins of L1, L2. E6 is a very small protein with about 150 amino acids and contains two zinc-binding domains, E6C and E6N. Due to E6 associated protein (E6AP), E6 protein can bind to P53 and inhibit P53 function. Ubiquitination of P53 after binding to E6 exposes it to proteasomes. Disruption of the E6 dimer has been shown to increase E6 solubility and decrease P53 degradation due to mutations in the E6N domain. Each E6 molecule of E6 dimer binds to an E6AP-related ubiquitin ligase to promote the p53 polyubiquitination, and degradation of the P53. Only HPV oncogenic variants, HPV16 and HPV18, contain dimeric E6 proteins, which give them the ability to polyubiquitination of the P53 [5]. The LxxLL motif of E6AP is a short, leucine-rich sequence located in the second ubiquitin ligase E6AP that binds to E6. In high-risk HPV variants, p53 degradation occurs by the E6 dimerization model through its N-terminus domain. Each E6 molecule of the dimer is attached to an E6AP ubiquitin ligase molecule and p53 molecule by the E6/E6AP complex to amplify p53 polyubiquitination. Modulation of p53 containing cells leads to degradation, destruction of the cell cycle, and subsequent abnormal cell proliferation [3]. The molecular treatment for cervical cancer can also be possible by an inhibitory ligand that disrupts the binding site interactions of the E6AP and the E6 HPV16 protein [5].

Identifying the HPV interactions of E6 and E7 oncoproteins with their multiple cellular targets could motivate targeted drug therapy for cervical cancer. In particular, *in silico* studies show that herbal compounds can be used to treat cervical cancer because they can block the binding of HPV18 E6 to p53 and thus prevent its degradation [6]. In addition, competitive antagonist ligands, benzopyranone derivatives, which disrupt the E6-E6AP interaction, are formulated to prevent the formation of complexes containing E6, E6AP, and p53. These promising anticancer agents were used in binding and functional assays [7].

In vitro and *ex vivo* methods are time-consuming and not economical compared to the in silico as a predictable method. Molecular docking is a method that can design possible reactions between target protein complexes and ligands and select appropriate treatment candidates from the available library before performing *in vitro* and *in vitro* tests [8].

This study aimed to investigate the potential plant compounds effective against HPV16 by acting on the active site of E6 protein using the molecular docking method. The results of this study could accelerate the discovery of more effective anti-HPV compounds.

METHODS

This docking study investigates the interactions between the HPV18 E6 protein and potentially effective plant compounds. The target protein (E6) was selected by examining the structure of HPV18 in viral zone media at https://viralzone.expasy.org/ and by studying various articles that examined the structure of the virus.

(i) Suitable structure of the target protein and energy minimization: The suitable structure of the target protein was obtained by the x-ray method (resolution of 1.34%) through the protein database (https://www.rcsb.org). This synthetic structure expressed in Escherichia coli BL21 (DE3), was examined in SPDBV software. The active site of the protein was determined by examining the structure of the virus in various articles. The energy minimizing of the structure was performed 10 times and the final format was saved (Figure 1).

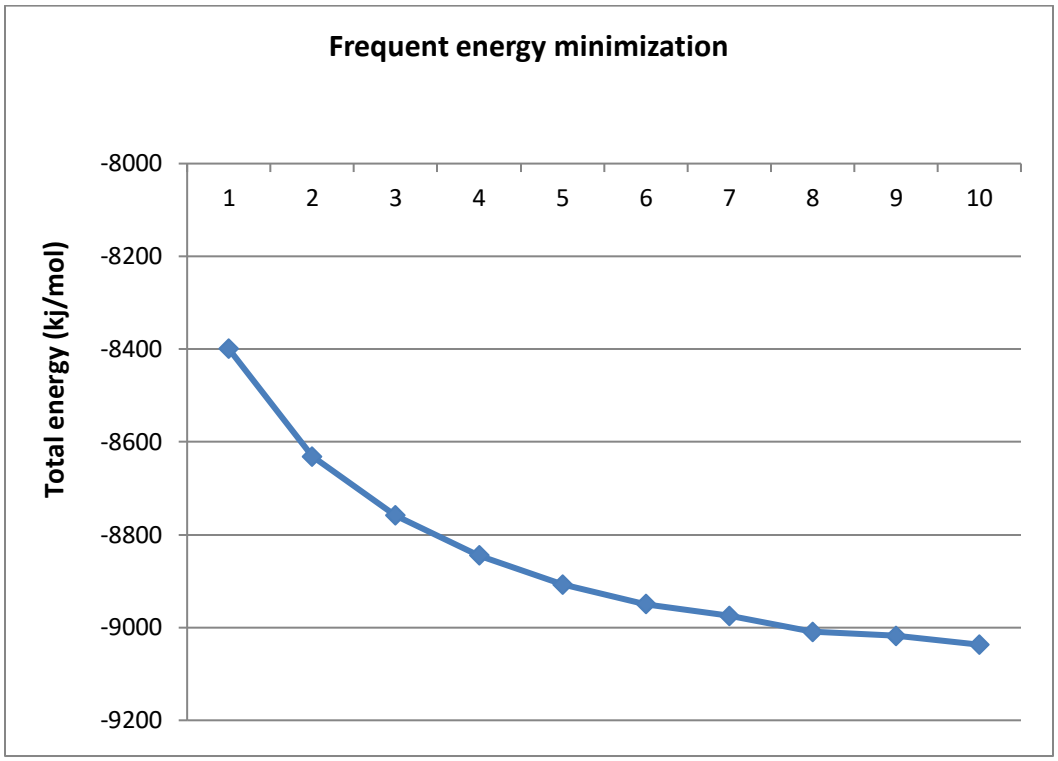


Figure 1. Energy minimization diagram after ten times energy minimization of target E6 protein

(***ii) Ligand selection:*** An overview of articles that mentioned the effective herbal compounds against viruses was performed. Several herbal compounds that had therapeutic potential against other viruses and HPV were selected as ligands and a ligand library was established. The formula of the compounds and the access code to them in PubChem and the three-dimensional (3D) structure of these compounds were collected from the same database at https://pubchem.ncbi.nlm.nih.gov. The Open Babel software was used to convert the SDF file into a PDB file.

(iii) Ligand and target protein interaction: The interaction of the ligand and the target protein (active site of the target protein), as a receptor, was investigated to determine the best binding state between the two molecules. To achieve this goal, docking them was modeled. At this stage, the academic software Auto Dock Tools-1.5.6 *ver.4* was used. Parameters used to perform molecular docking in this software for all compounds was fixed and include standard docking type, genetic algorithm, spacing zone diameter 0.375 A°, number of interactions 70, coordinate size center-x = 19.486, center-y = 18.653, center-z = 20.552, size-x = 86, size- y = 44, size-z = 74. The assay of hydrogen, electrostatic, and van der Waals interactions in the whole of the active site was done as well.

(iv) Pharmacodynamics parameters: To run the docking algorithm, the Cygwin Terminal Ink software was used. The run that had the lowest mean binding energy (ΔG) was selected, and information such as free binding energy, inhibition constant (Ki), or Michaelis-Menten constant for the inhibitor, Intramolecular energy, van der Waals energy, hydrogen energy, electrostatic energy, total internal energy, and free rotating energy were recorded in the table prepared for this purpose. The selected run was analyzed by UCSF Chimera 1.11

software and the number of hydrogen bonds (if any) and the size of the hydrogen bond were also recorded. To check the above features in Chimera software, an image file of the desired ligand was created using the grep, cut, cat commands in Cygwin software, and the format created in Chimera was retrieved.

(v)Pharmacokinetic parameters: Physicochemical properties such as water solubility, plasma protein binding, bioavailability, human intestinal absorption (HIA), CaCo2 cell permeability, mutagenicity (Ames test), carcinogenic effects on the murine model, and cardiac lethal effects (inhibiting the human Ether-à-go-go-Related Gene (hERG)) were predicted using PREADMET servers (https://preadmet.bmdrc.kr/) and ADMETSAR (http://lmmd.ecust.edu.cn/admetsar2/). In the mentioned servers, the information of chemical molecules in the form of *Mol* or *Smile* file provides the possibility of predicting physicochemical properties.

RESULTS

(i) Suitable structure of the target protein and energy minimization: The different structures of the E6 target protein in the UniProt (code: UniProtKB-P06463 (VE6_HPV 18)) were obtained, and the selected structure was created by the X-ray method was obtained through the protein database at https://www.rcsb.org/ with access code 4joR (Figure 2A). By reviewing previous reports, the possible active site of the protein in the C-terminal was identified (Figure 2B).

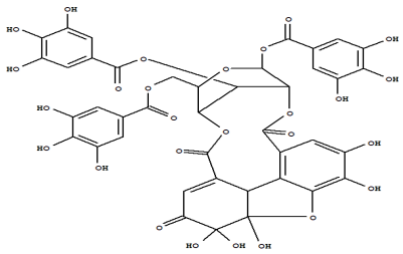
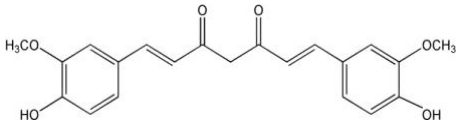
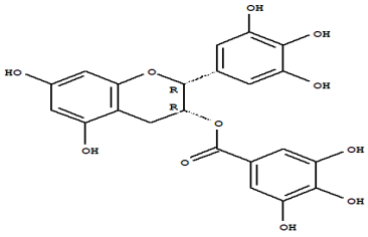
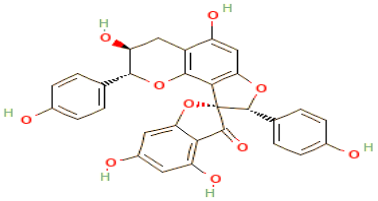
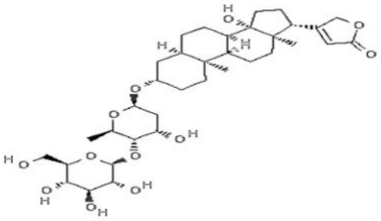
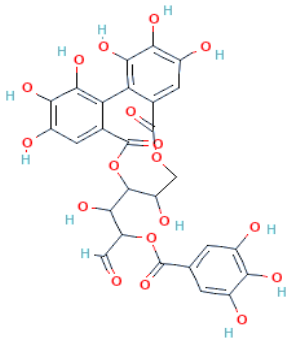
In SPDBV software, energy minimization was performed with the command select → all (Ctrl + A) and then tools → Energy minimization (Ctrl + N), and each time the total energy (KJ/mol) was calculated and recorded in excel software. After ten times recording the total energy minimization, the final format was saved and the diagram was drawn (Figure 1).

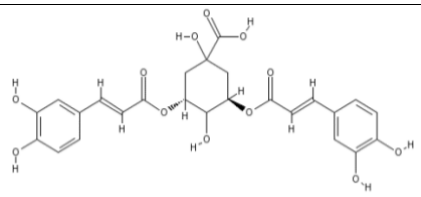
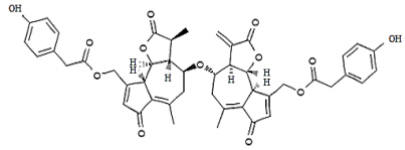
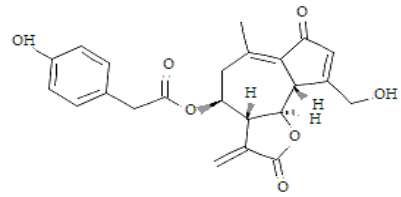
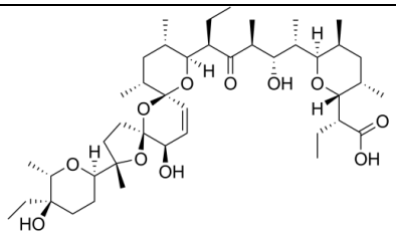
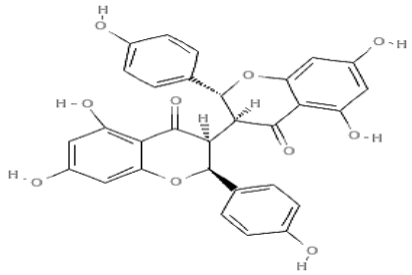
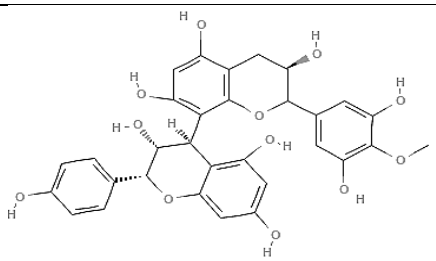
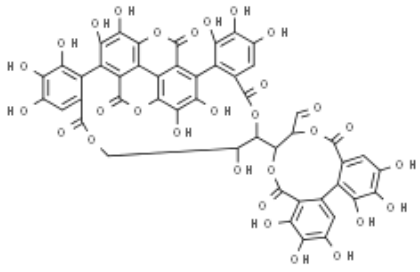
(ii) Ligand selection: After reviewing the articles in which plant compounds were used as antivirus, a total of 50 compounds were obtained. Among the initial compounds, 19 compounds with lower binding energies (based on the defined cut-off) were isolated and collected in a table along with the chemical formula and access code in PubChem (Table 1) [9].

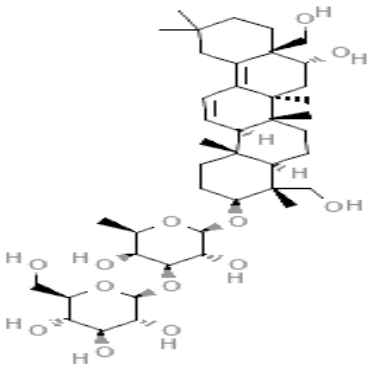
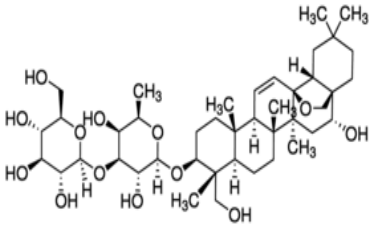
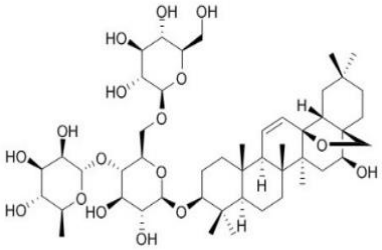
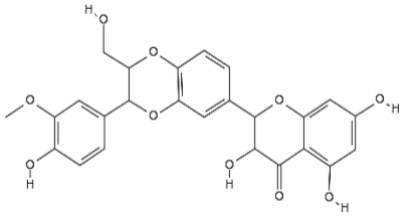
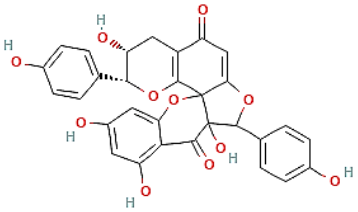
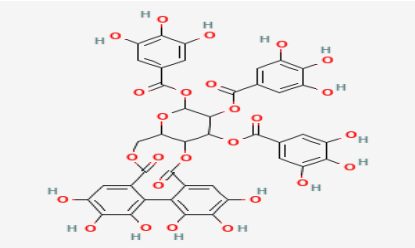
(iii) Docking results based on pharmacodynamics parameters: The pharmacodynamics parameters of the binding of ligands to the target protein are presented in Table 2. By evaluating these parameters and based on the binding energies obtained from different ligands, the selected compounds that had optimal binding energy were separated to conduct further studies. Regarding the average binding energies of ligands in the range of -7.3 to -15, the cut-off point was determined (Table 3).

(iv) Docking results based on pharmacokinetic parameters: The pharmacokinetic parameters of the compounds that were considered to be optimal based on pharmacodynamics parameters were listed in Table 3. To investigate the toxicity of the selected ligands, the information obtained from the ADMETsar server was given in Table 4. Finally, the compounds that were considered potential drugs according to Lipinski's law were listed in Table 5. How the selected ligands bind to the target protein and the length of the hydrogen bond were depicted in Figure 3.

Table 1. Herbal ingredient with drug potential against viral infections with PubChem access code (https://pubchem.ncbi.nlm.nih.gov)

	Ligand Name	PubChem CID	Chemical formula	Chemical structure	*Source*
1	**Chebulagic acid**	442674	C41H30O27		*Terminalia chebula*
2	**Curcumin**	969516	C21H20O6		*Turmeric yellow*
3	**Epigallocatechin Gallate**	65064	C22H18O11		*Green tea*
4	**Genkuwanol A**	394845	C30H22O10		*Daphne genkwa*
5	**Glucoevatromonoside**	15137997	C35H54O12		*Digitalis lanata*
6	**Hippomannina**	191266	C27H22O18		*Hippomane Mancinelli*

	Ligand Name	PubChem CID	Chemical formula	Chemical structure	*Source*
7	**Isochlorogenic acid A**	6474310	C25H24O12		*Suaeda glauca*
8	**Lactucain C**	6918760	C46H44O13		*Lactuca indica*
9	**Lactupicrin; Lactucopicrin;**	174880	C23H22O7		*Lactuca virosa (wild lettuce)*
10	**Narasin**	65452	C43H72O11		*Monteban*
11	**Neochamaejasmin B**	21636084	C30H22O10		*Stellera chamaejasme*
12	**Proanthocyanidin**	108065	C31H28O12		*Grapeseed*
13	**Punicalagin**	44584733	C48H28O30		*Pomegranate*

	Ligand Name	PubChem CID	Chemical formula	Chemical structure	*Source*
14	**Saikosaponin B2**	21637642	C42H68O13		*Bupleurum spp*
15	**Saikosaponin D**	107793	C42H68O13		*Bupleurum spp*
16	**Saikosaponin C**	131801344	C48H78O18		*Bupleurum spp*
17	**Silymarin**	5213	C25H22O10		*Flavobin Spofa*
18	**Stelleranol**	131676072	C30H22O11		*Radix Wikstroemiae*
19	**Tellimagrandin II**	151590	C41H30O26		*Filipendula palmata*

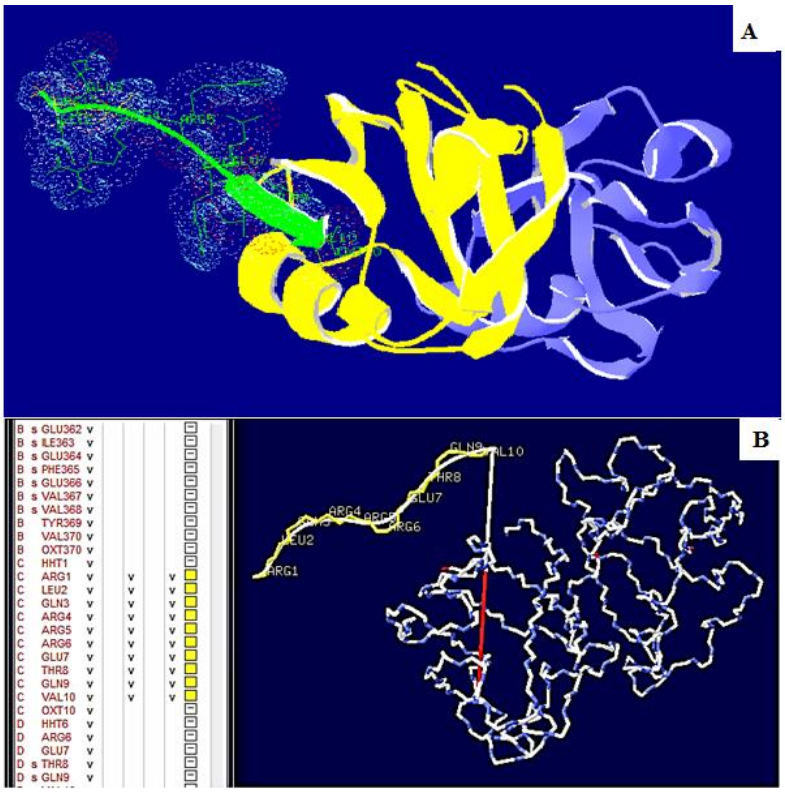


Figure 2. (A) X-ray structure of HPV 18 E6 protein (4joR) from RCSB database (ribbon format SPDBV software), (B) The active site of the E6 target protein and the amino acid sequence **(ARG1, LEU2, GLN3, ARG4, ARG5, ARG6, GLU7, THR8, GLN9, VAL10)** specified in the figure in SPDBV software

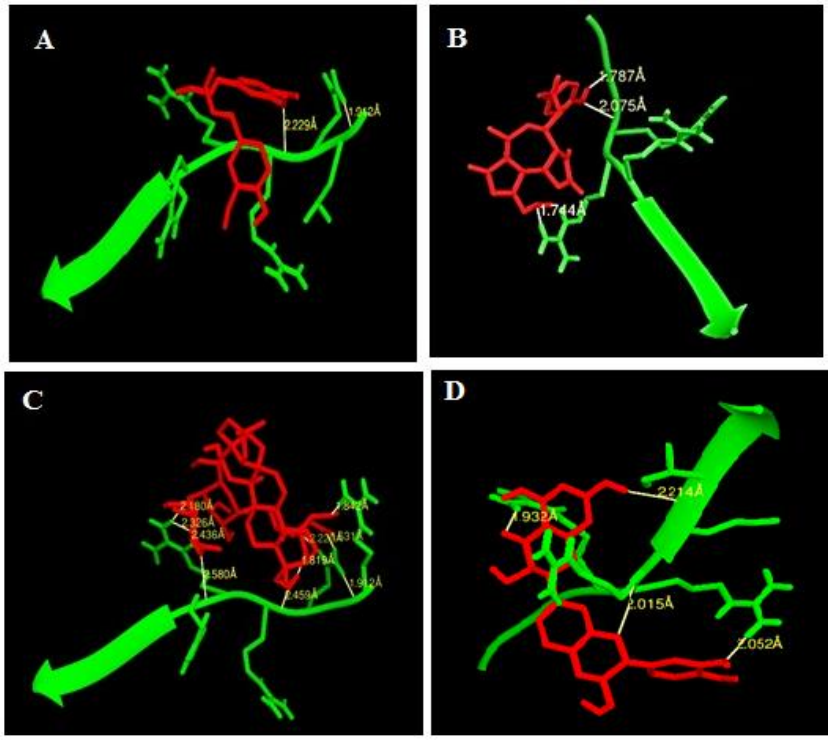


Figure 3. The hydrogen bonds distance (A°) between the ligand and the target protein (Chimera software); HPV 18 E6 target protein (green) with (A) curcumin; (B) Lactupicrin; (C) Saikosaponin C, and (D) Silymarin ligands (red)

Table 2. Pharmacodynamics parameters of selected ligands based on defined cut-off point of binding energy

	Ligand Name	Energy Binding (Kcal/mol)	Ki	Final Intermolecular Energy (Kcal/mol)	Vdw + H-bond + desolv energy (Kcal/mol)	Electrostatic Energy (Kcal/mol)	Final Total Internal Energy (Kcal/mol)	Torsional Free Energy (Kcal/mol)	Unbound System's Energy (Kcal/mol)	Num Of H bond
1	**Tellimagrandin II**	-14.77	14.96 pM	-10.8	-10.37	-0.43	-16.14	7.16	-5.01	7
2	**Hippomannina**	-13.61	105.24 pM	-6.89	-6.63	-0.26	-14.33	5.07	-2.54	4
3	**Chebulagic acid**	-12.97	309.87 pM	-9.81	-8.34	-1.46	-11.51	5.37	-2.97	4
4	**Punicalagin**	-12.83	393.99 pM	-9.71	-9.44	-0.27	-12.54	5.07	-4.35	1
5	**Saikosaponin C**	-12.21	1.12 nM	-9.55	-9.27	-0.28	-11.25	5.97	-2.62	7
6	**Saikosaponin B2**	-11.86	2.01 nM	-9.47	-9.31	-0.17	-9	4.77	-1.83	3
7	**Lactuca C**	-11.34	4.85 nM	-10.32	-10.07	-0.26	-7.47	4.18	-2.27	3
8	**Narasin**	-10.84	11.37nM	-11.13	-9.91	-1.23	-6.03	4.77	-1.55	5
9	**Epigallocatechin gallate**	-10.58	17.6nM	-9.81	-9.33	-0.48	-6.03	3.58	-1.69	5
10	**Saikosaponin D**	-10.13	37.80 nM	-8.32	-8.18	-0.13	-7.77	4.18	-1.79	1
11	**Isochlorogenic acid A**	-9.76	70.46 nM	-9.23	-7.82	-1.41	-7.33	4.77	-2.03	3
12	**Proanthocyanidin**	-9.76	70.31 nM	-10.06	-9.61	-0.45	-5.31	3.88	-1.74	4
13	**Glucoevatromonoside**	-9.16	192.74nM	-9.48	-9.07	-0.42	-5.27	3.58	-2.02	5
14	**Neochamaejasmin B**	-7.93	1.54 uM	-9.06	-8.87	-0.19	-3.78	2.68	-2.23	0
15	**Silymarin**	-7.88	1.67 uM	-7.96	-7.67	-0.29	-3.7	2.68	-1.1	4
16	**Stelleranol**	-7.69	2.30 uM	-9.01	-8.66	-0.35	-2.9	2.39	-1.83	6
17	**Curcumin**	-7.52	3.07 uM	-8.72	-8.53	-0.2	-2.53	2.98	-0.75	3
18	**Genkuwanola**	-7.52	3.06 uM	-8.67	-8.57	-0.11	-2.42	2.39	-1.18	3
19	**Lactupicrin; Lactucopicrin;**	-7.48	3.30 uM	-7.93	-7.68	-0.24	-2.32	2.09	-0.68	4

VdW: Van der Waals; Ki: inhibitor constant

Table 3. Pharmacokinetic parameters of selected ligands

Ligand Name	Bioavailability	BBB permeant	Drug likeness	TPSA Å²	P-gp substrate	Log S pH=7	Log Kp cm/s	Log p	HIA
Chebulagic acid	0.11	No	violated	447.09	Yes	-5.92 (moderately)	-11.87	-0.92	95.58327 (low)
Curcumin	0.55	No	suitable	93.06	No	-3.94 (moderately)	-6.28	3.03	94.40339 (high)
Epigallocatechin gallate	0.17	No	violated	197.37	No	-3.56 (soluble)	-8.27	0.95	20.7125 (low)
Genkuwanola	0.7429	No	violated	166.14	No	-6.01 (poorly)	-6.80	2.56	80.08263 (low)
Glucoevatromonoside	0.6571	No	violated	184.60	Yes	-4.01 (moderately)	-9.88	1.33	52.64456 (low)
Hippomannina	0.6571	No	violated	318.50	Yes	-3.88 (poorly)	-9.94	-0.55	0.632311 (low)
Isochlorogenic acid A	0.7	No	violated	211.28	Yes	-3.65 (moderately)	-8.37	0.79	23.12314 (low)
Lactucain C	0.5857	No	violated	189.03	Yes	-6.26 (poorly)	-8.87	4.09	97.48107 (low)
Lactupicrin; Lactucopicrin;	0.6	No	suitable	110.13	Yes	-2.90 (moderately)	-8.02	1.92	94.27315 (high)
Narasin	0.8143	No	violated	161.21	No	-7.67 (poorly)	-6.60	5.20	91.8641 (low)
Neochamaejasmin B	0.6	No	violated	173.98	No	-6.36 (poorly)	-6.33	2.88	78.60049 (low)
Proanthocyanidin	0.8143	No	violated	209.76	No	-5.36 (moderately)	8.00	1.85	37.48425 (low)
Punicalagin	0.6	No	violated	518.76	Yes	-8.05 (poorly)	-11.67	0.02	0 (low)
Saikosaponin B2	0.8143	No	violated	218.99	Yes	-5.81 (moderately)	-9.27	1.91	29.67371 (low)
Saikosaponin D	0.7714	No	violated	207.99	Yes	-5.87 (moderately)	-9.27	2.10	42.0265 (low)
Saikosaponin C	46.3	Yes	violated	29.10	No	-1.85 (soluble)	-6.37	1.80	12.73696 (high)
Silymarin	0.7571	No	suitable	155.14	No	-4.14 (moderately)	-7.89	1.59	78.55065 (low)
Stelleranol	0.7143	No	violated	183.21	No	-4.55 (moderately)	-8.51	1.34	75.75563 (low)
Tellimagrandin II	0.5429	No	violated	444.18	Yes	-6.91 (poorly)	-10.32	0.24	0 (low)

BBB: blood-brain barrier, P-gp: P-glycoprotein, HIA: Human intestinal absorption, TPSA: Topological polar surface area, log Kp: logarithmic skin permeation coefficient, LogP: octanol-water partition coefficient

Table 4 Investigation of toxicity parameters of candidate ligands

Ligand Name	Energy Binding (Kcal/mol)	Cytochrome inhibitor	Plasma protein binding %	Caco2 cell permeability	hERG_inhibition	Carcino – Mouse	Ames test
Chebulagic acid	-12.97	No	100	17.696	ambiguous	positive	non-mutagen
Curcumin	-7.52	CYP2C9, CYP3A4	88.030378	20.0731	medium_risk	negative	non-mutagen
Epigallocatechin gallate	-10.58	No	100	12.0421	high_risk	negative	non-mutagen
Genkuwanola	-7.52	CYP2C9, CYP3A4	100	10.388	high_risk	negative	non-mutagen
Glucoevatromonoside	-9.16	No	71.909222	20.4775	ambiguous	positive	non-mutagen
Hippomannin A	-13.61	No	100	15.4439	ambiguous	negative	non-mutagen
Isochlorogenic acid A	-9.76	No	86.055118	19.321	high_risk	positive	mutagen
Lactucain C	-11.34	CYP2C9	89.928982	19.7665	low_risk	positive	non-mutagen
Lactupicrin;	-7.48	No	78.861696	20.6973	high_risk	positive	non-mutagen
Narasin	-10.84	CYP3A4	90.546537	39.9512	ambiguous	positive	mutagen
Neochamaejasmin B	-7.93	CYP2C9, CYP3A4	100	12.4974	medium_risk	negative	non-mutagen
Proanthocyanidin	-9.76	CYP3A4	100	13.9636	high_risk	negative	non-mutagen
Punicalagin	-12.83	No	100	16.0334	ambiguous	positive	non-mutagen
Saikosaponin B2	-11.86	No	60.927339	20.5797	ambiguous	positive	non-mutagen
Saikosaponin D	-10.13	No	59.692509	20.3021	ambiguous	positive	non-mutagen
Saikosaponin C	-12.21	CYP1A2	46.362348	19.6592	ambiguous	positive	non-mutagen
Silymarin	-7.88	CYP3A4	87.754608	4.84461	medium_risk	negative	mutagen
Stelleranol	-7.69	CYP2C9	100	9.91683	high_risk	positive	non-mutagen
Tellimagrandin II	-14.77	No	100	15.751	ambiguous	positive	non-mutagen

hERG: human Ether-à-go-go-Related Gene,

Table 5 Details of selected ligands according to Lipinski's law

Ligand Name	Drug likeness	Lipinski low
Chebulagic acid	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Curcumin	suitable	Yes; 0 violation
Epigallocatechin Gallate	violated	No; 2 violations: Nor O>10, NH or OH>5
Genkuwanola	violated	No; 2 violations: MW>500, NH or OH>5
Glucoevatromonoside	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Hippomannina	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Isochlorogenic acid A	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Lactucain C	violated	No; 2 violations: MW>500, Nor O>10
Lactupicrin; Lactucopicrin	suitable	Yes; 0 violation
Narasin	violated	No; 2 violations: MW>500, Nor O>10
Neochamaejasmin B	violated	No; 2 violations: MW>500, NH or OH>5
Proanthocyanidin	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Punicalagin	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Saikosaponin B2	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Saikosaponin D	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Saikosaponin C	violated	Yes;0 violation MW<200
Silymarin	suitable	Yes; 0 violation
Stelleranol	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5
Tellimagrandin II	violated	No; 3 violations: MW>500, Nor O>10, NH or OH>5

MW: molecular weight.

DISCUSSION

Here, we identified some ligands of plant origin against HPV18E6 protein, which have been defined in previous studies as a potential target protein. After ligand library formation and molecular docking, the compounds that had the most negative binding energy for the active site of the target protein were isolated. Following pharmacodynamics and pharmacokinetic studies, four compounds including silymarin, saikosaponin c (SSC), curcumin, and lactupicrin were isolated in terms of toxicity and drug potential according to Lipinski's law.

Attempts to discover new antiviral agents in both synthetic and herbal medicinal products have been made. Natural products have metabolites that can contribute to antiviral activity by interfering with the interactions of viruses and host receptors. Apart from the low cost of herbal medicinal products, the native plants of each region have special properties that can act with different antiviral activities than existing antiviral drugs. Today, the management of HPV infections is more focused on active infections, and most treatment strategies are along with recurrence [10]. Various studies have been performed on the effectiveness of herbal medicines and extracts from herbal mixtures such as *Pinus densiflora Sieb* [11] *Echinacea purpurea, Epigallocatechin gallate,* and *Carrageenan* [12].

The *In silico* research can be used to obtain basic information about natural products that have treatment effects on HPV [13]. The molecular docking method is one of the most well-known molecular tools for discovering new drugs. Finding the lowest energy level or optimizing is the most important docking technique, according to which the ligand can find a place for connection with the least amount of energy. Docking also shows the best orientation of the ligand to the position of the active site of the protein [14].

Lipinski's Rule of Five (ROF) was developed to set guidelines for the potency of drugability of compounds. In the drug discovery process, this rule predicts the adsorption or permeability behavior of compounds. When the considered agent has more than 5 H-bond donors, 10 H-bond acceptors, and a molecular weight of more than 500 kDa, and has a calculated log P (Clog P) of more than 5, it is very poor in terms of absorption [15]. In the SWISS ADME server, the rule of 5 and some properties such as Clog P, which indicates the lipophilicity of the substance (preferably from 1 to 5), Log S, which indicates the solubility of the substance (more than zero, high solubility, and less than 10 insoluble), TPSA, area of polar parts or polarity (more than 140 angstroms is without proper permeability) are available.

In table 5, the selected compounds that have high binding energy are separated according to Lipinski's law in terms of drug potential. Accordingly, only four compounds of silymarin, SSC, curcumin, and lactupicrin were included in this classification. To investigate these compounds' toxicity and other adverse effects, the information obtained from the ADMETSAR server was reviewed.

The result of pharmacodynamics and pharmacokinetic factors must be considered for the effectiveness of a drug orally, by injection, or by skin absorption. A set of pharmacodynamic and pharmacokinetic factors of selected ligands are given in tables 2 and 3.

The binding energy of the selected compounds in the most negative binding energies is SSC (-12.21 kcal/mol), silymarin (-7.88 kcal/mol), curcumin (-7.52 kcal/mol), and lactupicrin (-7.48 kcal/mol), respectively. Among them, only Silymarin has mutagenic properties in confirming the Ames test and the rest are non-mutagenic. However, this compound along with curcumin has no carcinogenic effects in mice. In terms of hERG inhibition, lactupicrin is a high-risk compound, while curcumin and silymarin are low risks.

Except for SSC, the other compounds do not pass through the blood-brain barrier (BBB), and all compounds except Silymarin have high digestive absorption and can be easily consumed orally. All compounds have relative solubility.

According to the logarithmic skin permeation coefficient (log Kp), the skin absorption of lactupicrin is more negative than the others (-8.02 cm/s) and curcumin (-6.28 cm/s) is the most. Therefore, curcumin has both oral and skin absorption. lactopicrin is a p-glycoprotein (p-gp) substrate and has the potential to be reversed at the site of action.

Based on topological polar surface area (TPSA), which determines the degree of polarity of the compounds and the range of 20-130 Å² is desirable, all our selected compounds except Silymarin (155.14 Å²) were in the desired range (Table 3). The lowest TPSA was observed in SSC (29.10 Å²), indicating that the compound passes easily through the membrane and blood brain barrier. This compound is not known as a drug in some cases outside of Lipinski's law due to its molecular weight of < 200 KDa.

The bioavailability of all selected compounds is higher than 50% (Table 3), so they can be taken orally, but the percentage of binding of these compounds to plasma proteins, except for SSC (46.3%), are more than 70%, and this causes the compounds to stay longer in the blood and have a longer half-life and the compounds do not reach the target drug site.

Curcumin has been shown to have a cytotoxic effect on HPV-infected cells in a concentration- and time-dependent manner. This compound reduces the regulation of both serine kinase AKT/nuclear factor κB pathways by sensitizing cancer cells [16]. Curcumin also inhibits telomerase activity, the RAS and extracellular signal-regulated kinases (ERKs) signaling pathways, cyclin D1, cyclooxygenase 2, and nitric oxide synthase activity, and inhibits cell growth by reducing the expression of HPV oncoproteins [17-19].

The therapeutic effects and possible mechanisms of antiviral effects of the four compounds mentioned above have been evaluated in the previous studies.

Yu et al. showed that Silymarin dramatically suppresses cancer cell survival and expressed that the possible mechanisms are modulating B-cell lymphoma 2 (Bcl-2) family proteins, activating caspase-3, inhibiting Akt phosphorylation by increasing phosphatase and tensin homolog expression, inhibition of virus migration to the wound, and inhibition of metalloproteinase-9 matrix [20]. Also, it has been found that silymarin binds through the formation of hydrogen bonds to Tyr32 and Arg55 and is a suitable candidate for further study [21].

Saikosaponins (SSs) are triterpenoid saponins and are divided into seven types according to different aglycones, which are SSA, SSD, and SSC epoxy ether saikosaponins (type I). Although SSC has the same basic structure as SSA and SSD, its pharmacological activity has been reported to be much weaker. However, the anti-apoptotic effects of SSC have been reported by suppressing caspase-3 activation and then degrading focal adhesion kinase and other cell adhesion signals. On the other hand, the antiviral activity of SSC against hepatitis B virus (HBV) has been shown by inhibiting DNA replication [22].

Lactucopicrin is a bitter substance that has central sedative and analgesic effects [23]. It is a sesquiterpene lactone and is part of the lactocarium derived from plants such as *Lactuca virosa* (wild lettuce), *Cichorium intybus*, and *dandelion coffee* [24]. Among the traditional treatments of lactopicrin, apart from its analgesic effect, its antimalarial effect has been studied [25]. Previous research has shown that this compound can act as an acetylcholinesterase inhibitor [26]. However, so far no reliable report has been provided on the antiviral effect of this substance.

One of the limitations of the studies of *in silico* in the discovery and development of drugs is the use of the rules that determine the drugability of various compounds. Although there are several approaches in the discovery and development of new drugs, lipinski's rule of five or ROF is the simplest and most widespread. The ROF-Score is lying between '0' and '4'; molecules with ROF-Scores greater than one are considered for further development. One of the problems with this approach is that it is very strict in scoring compounds, and a slight difference from the values set out in this law causes many compounds to be ignored for further research and development. Besides, this rule is applied only to permeation by passive diffusion of drugs through cell membranes; drugs that are actively transported through cell membranes by transporter proteins are exceptions to this rule. However, this rule should be used with a lot of caution and other sophisticated metrics should be utilized where appropriate [27]

CONCLUSION

Our study revealed that natural compounds curcumin, silymarin, saikosaponin c, and lactupicrin had better binding free energies with HPV E6 protein. Silymarin and curcumin were less dangerous than other compounds due to the lack of inhibition of the hERG. Of these compounds, silymarin had lower oral absorption, lactopicrin had less skin absorption, lactopicrin is the substrate of P-gp, and saikosaponin c crosses the blood-brain barrier.

By using the *in-silico* method in the discovery of effective biological compounds against pathogens the costs of research, *in vivo* and *in vitro* will be reduced. Molecular docking identifies effective therapeutic agents and target proteins by evaluating the interactions between receptors and ligands and provides more effective compounds to the next phase of research. On the other hand, identifying the mechanism of action of active biological compounds used in alternative medicine with computational biology can bring a new horizon for researchers around the world that do not have access to native plants in each region. According to these findings, it is recommended that *in vitro* and *in vivo* examinations be conducted to determine the effectiveness of these compounds against HPV18.

Abbreviation list: HPV: Human Papilloma Virus, E6AP: E6 associated protein, 3D: three-dimensional, SSC: Saikosaponin C, Clog P: calculated log P, BBB: blood-brain barrier, P-gp: P-glycoprotein, SS: Saikosaponin, hERG: human Ether-à-go-go-Related Gene, HIA: Human intestinal absorption, TPSA: Topological polar surface area, log Kp: logarithmic skin permeation coefficient, Bcl-2: B-cell lymphoma 2, HBV: hepatitis B virus, vdW: Van der Waals.

Conflicts of interest/Competing interests: None

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article

Acknowledgments: The authors would like to thank the *Clinical Research Development Center of Imam Reza Hospital* for their kindly assist and advice.

REFERENCES

1. Pennycook KB, McCready TA: Condyloma Acuminata. In: StatPearls. Treasure Island (FL): StatPearls Publishing, StatPearls Publishing LLC.; 2021.
2. Senkomago V, Henley SJ, Thomas CC, Mix JM, Markowitz LE, Saraiya M: Human Papillomavirus-Attributable Cancers - the United States, 2012-2016. MMWR Morbidity and mortality weekly report 2019, 68(33):724-728. https://doi.org/10.15585/mmwr.mm6833a3
3. Pappa KI, Kontostathi G, Lygirou V, Zoidakis J, Anagnou NP: Novel structural approaches concerning HPV proteins: Insight into targeted therapies for cervical cancer. Oncology reports 2018, 39(4):1547-1554. https://doi.org/10.3892/or.2018.6257

4. Chapter 5: Human Papillomavirus, Manual for the surveillance of vaccine-preventable diseases [https://www.cdc.gov/vaccines/pubs/surv-manual/chpt05-hpv.html] Retrieved September 1, 2021.
5. Pal A, Kundu R: Human Papillomavirus E6 and E7: The Cervical Cancer Hallmarks and Targets for Therapy. Frontiers in microbiology 2020, 10:3116-3116. https://doi.org/10.3389/fmicb.2019.03116
6. Kumar S, Jena L, Sahoo M, Kakde M, Daf S, Varma AK: In silico docking to explicate interface between plant-originated inhibitors and E6 oncogenic protein of highly threatening human papillomavirus 18. Genomics and informatics 2015, 13(2):60. https://doi.org/10.5808/GI.2015.13.2.60
7. Rietz A, Petrov DP, Bartolowits M, DeSmet M, Davisson VJ, Androphy EJ: Molecular probing of the HPV-16 E6 protein alpha helix binding groove with small molecule inhibitors. PLoS One 2016, 11(2):e0149845. https://doi.org/10.1371/journal.pone.0149845
8. Sethi A, Joshi K, Sasikala K, Alvala M: Molecular docking in modern drug discovery: principles and recent applications. Drug Discovery and Development-New Advances 2019:1-21. https://doi.org/10.5772/intechopen.85991
9. National Libraray of Medicine National Center for Biotechnology Information [https://pubchem.ncbi.nlm.nih.gov] Retrieved September 1, 2021.
10. Garbuglia AR, Lapa D, Sias C, Capobianchi MR, Del Porto P: The Use of Both Therapeutic and Prophylactic Vaccines in the Therapy of Papillomavirus Disease. Frontiers in Immunology 2020, 11(188). https://doi.org/10.3389/fimmu.2020.00188
11. Lee H-J, Park M, Choi H, Nowakowska A, Moon C, Kwak JH, Kim YB: Pine Needle Extract Applicable to Topical Treatment for the Prevention of Human Papillomavirus Infection. J Microbiol Biotechnol 2021, 31(1):137-143. https://doi.org/10.4014/jmb.2010.10055
12. Yusuf L, Bala J, Aliyu I, Kabir I, Abdulkadir S, Doro A, Kumurya A: Phytotherapy as an alternative for the treatment of human papillomavirus infections in Nigeria: a review. African Journal of Clinical and Experimental Microbiology 2020, 21(3):175-184.
13. Kotadiya R, Johnson A, Georrge JJ: Natural products as inhibitors of E6 and E7 proteins of Human papillomavirus (HPV): an in silico approach. In: Proceedings of the National Conference on Innovations in Biological Sciences (NCIBS): 2020. https://doi.org/10.2139/ssrn.3574478
14. Pinzi L, Rastelli G: Molecular Docking: Shifting Paradigms in Drug Discovery. Int J Mol Sci 2019, 20(18):4331. https://doi.org/10.3390/ijms20184331
15. Benet LZ, Hosey CM, Ursu O, Oprea TI: BDDCS, the Rule of 5 and drugability. Advanced drug delivery reviews 2016, 101:89-98. https://doi.org/10.1016/j.addr.2016.05.007
16. Zhai K, Brockmüller A, Kubatka P, Shakibaei M, Büsselberg D: Curcumin's beneficial effects on neuroblastoma: Mechanisms, challenges, and potential solutions. Biomolecules 2020, 10(11):1469. https://doi.org/10.3390/biom10111469
17. Lee T-Y, Tseng Y-H: The Potential of Phytochemicals in Oral Cancer Prevention and Therapy: A Review of the Evidence. Biomolecules 2020, 10(8):1150. https://doi.org/10.3390/biom10081150
18. Lindsay C, Kostiuk M, Conrad D, O'Connell DA, Harris J, Seikaly H, Biron VL: Antitumour effects of metformin and curcumin in human papillomavirus positive and negative head and neck cancer cells. Molecular Carcinogenesis 2019, 58(11):1946-1959. https://doi.org/10.1002/mc.23087
19. MOUTINHO MSS, ARAGÃO S, CARMO D, CASACA F, SILVA S, RIBEIRO J, SOUSA H, PIRES I, QUEIROGA F, COLAÇO B, et al: Curcumin and Rutin Down-regulate Cyclooxygenase-2 and Reduce Tumor-associated Inflammation in HPV16-Transgenic Mice. Anticancer Research 2018, 38(3):1461-1466. https://doi.org/10.21873/anticanres.12371
20. Yu H-C, Chen L-J, Cheng K-C, Li Y-X, Yeh C-H, Cheng J-T: Silymarin Inhibits Cervical Cancer Cell Through an Increase of Phosphatase and Tensin Homolog. Phytotherapy Research 2012, 26(5):709-715. https://doi.org/10.1002/ptr.3618
21. Małecka M, Skoczyńska A, Goodman DM, Hartinger CG, Budzisz E: Biological properties of ruthenium(II)/(III) complexes with flavonoids as ligands. Coordination Chemistry Reviews 2021, 436:213849. https://doi.org/10.1016/j.ccr.2021.213849
22. Yuan B, Yang R, Ma Y, Zhou S, Zhang X, Liu Y: A systematic review of the active saikosaponins and extracts isolated from Radix Bupleuri and their applications. Pharmaceutical Biology 2017, 55(1):620-635. https://doi.org/10.1080/13880209.2016.1262433
23. Abdul-Jalil TZ. Lactuca serriola: Short Review of its Phytochemical and Pharmacological Profiles. International Journal of Drug Delivery Technology. 2020;10(3):505-508.
24. Sessa RA, Bennett MH, Lewis MJ, Mansfield JW, Beale MH: Metabolite profiling of sesquiterpene lactones from Lactuca species. Major latex components are novel oxalate and sulfate conjugates of lactucin and its derivatives. The Journal of biological chemistry 2000, 275(35):26877-26884. https://doi.org/10.1016/S0021-9258(19)61456-0
25. Channotiya J, Tiwari A, Taj G, Verma AK, Dubey A: In-silico and Molecular Docking Studies on Germacrene A Synthase enzyme and sesquiterpene lactone (Lactucin) involved in the antimalarial activity of Cichorium intybus. Network Modeling Analysis in Health

Informatics and Bioinformatics 2021, 10(1):24. https://doi.org/10.1007/s13721-021-00288-5

26. Zhang L, Rocchetti G, Zengin G, Ak G, Yıldıztugay E, Mahomoodally MF, Picot-Allain MCN, Lucini L: Profiling of polyphenols and sesquiterpenoids using different extraction methods in Muscari turcicum, an endemic plant from Turkey. Industrial Crops and Products 2020, 154:112626. https://doi.org/10.1016/j.indcrop.2020.112626
27. Shultz MD: Two Decades under the Influence of the Rule of Five and the Changing Properties of Approved Oral Drugs. Journal of medicinal chemistry 2019, 62(4):1701-1714. https://doi.org/10.1021/acs.jmedchem.8b00686