

Molecular Modeling and Docking Studies on Phyto-compounds against Caspase-3, BRCA1, and Rb

Asita Elengoe^{1,*} , Vishalani Loganathan¹

¹ Department of Biotechnology, Faculty of Science, Lincoln University College, 47301 Petaling Jaya, Selangor, Malaysia; asitaelengoe@yahoo.com (A.E.); vishalaniloganathan94@gmail.com (V.L.)

* Correspondence: asitaelengoe@yahoo.com (A.E.)

Scopus Author ID 56118646500

Received: 01.11.2021; Revised: 20.11.2021; Accepted: 21.11.2021; Published: 25.11.2021

Abstract: Breast cancer is one of the well-known diseases analyzed in women compared to men worldwide. There are few studies about plant compounds that have been identified to have anticancer properties. Consequently, phyto-compounds have the capability of evolving new drugs. In this research, the three-dimensional (3D) structure of breast cancer cell line proteins, caspase-3, breast cancer susceptibility type 1 (BRCA1), and retinoblastoma (Rb) were generated, and docking with plant compounds (ferulic acid and quercetin, respectively) was studied. Swiss model was used to build the 3D structure of protein models. Then, the protein models were assessed using the validation tools (PROCHECK, ProQ, ERRAT, and Verify 3D programs). Lastly, the protein was docked successfully with ferulic acid (PubChem ID: 445858) and quercetin (PubChem ID: 5280343), respectively, using the SwissDock server and visualized with Discovery Studio (DS) 4.0 software. The results show that the protein models were stable after the validation process. The binding energy of the protein-phyto-compound complexes (Rb-Ferulic acid and Rb-Quercetin) were -6.6 and -7.8 kcal/mol, respectively. These proteins had a stable bond with their phyto-compounds. The toxicity prediction analysis revealed that ferulic acid (PubChem ID: 445858) is safe to use as a drug. This current study of the protein-phyto-compound-complex interaction will help in designing new clinical medications.

Keywords: breast cancer susceptibility type 1 (BRCA-1); caspase-3; docking; modelling; retinoblastoma (Rb)

© 2021 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Breast cancer is the most common cancer among women, which has an impact, with an estimated 2.1 million new cases reported each year [1]. According to the Section of Cancer Surveillance, World Health Organization (WHO) (2018), it is estimated that 627,000 women died from breast cancer that is approximately 15% of all cancer deaths among women. It is also in ranking as the second-largest cancer worldwide, contributing 12.3% of the total number of cases diagnosed in 2018 (2,088,849), followed by lung cancer [2]. Breast cancer is higher among women in more developed regions (75.2) than less developed regions (32.8) which is a total of 46.3 worldwide [2].

According to the Ministry of Health, the Malaysian Study on Cancer Survival (MySCan) 2018 report demonstrated that cancer was the fourth common cause of death in Malaysia. Malaysian National Cancer Registry (MNCR), under the Ministry's National Cancer Institute, published a report on the results of five-year relative survival for fifteen of the most common cancers, which include breast cancer too. Based on the database obtained from WHO

(2018), breast cancer is in the first rank for both sexes, with an estimation of 7593 (17.3%) in Malaysia [3]. The risk factors for breast cancer are age, family history, medical history, weight, physical activity, smoking, alcohol consumption, and unhealthy diet [4,5].

Breast cancer is a genetic illness, meaning it is brought on by changes in DNA. These mutations can be passed down through the generations or acquired after birth. Several genetic alterations had already accumulated in the tumor cells at the initial identification of clinical malignancy. BRCA1, BRCA2, Caspase-3 and Rb gene mutations are all important in breast tumorigenesis [6,7].

Researchers worldwide are attempting to improve the quality of life of patients and survivors by finding better ways to prevent, detect, and treat breast cancer [1]. There are few studies about plant compounds that have been identified to have anticancer properties [8]. Consequently, phyto-compounds have the capability of evolving new drugs. Plant compounds are safe and efficient drugs for breast cancer treatment compared to conventional methods such as chemotherapy, radiotherapy, and surgery. The conventional methods cause different side effects such as liver, heart, and kidney failure, damage to normal cells, etc.

Stigmasterol is one of the chemical constituents found in the leaves of *Clinacanthus nutans*. It is a phytosterol. Phytosterol is defined as a steroid derived from plants. It plays an important role in lowering cholesterol absorption in the intestines and act as an anticancer agent [9]. Ferulic acid is a phenolic phytochemical which is found in the cell walls of plants. It can be absorbed by the small intestine and excreted through the urine. Research studies have been demonstrated that ferulic acid shows not only positive results for cancer patients but also other diseases such as diabetes, hypertension, etc. Ferulic acid plays an essential role in autophagy which triggers apoptosis [10]. Quercetin is a flavonoid present in many plants and fruits (apples, honey, lemon, orange, tomato, raspberries, cranberries, onions, broccoli, red grapes, and green leafy vegetables). Several epidemiological studies have been reported that a positive correlation between the dietary consumption of flavonoids and decreased incidence and mortality from cardiovascular disease and cancer [11-14].

Bioinformatics tools such as molecular modeling and docking aid in developing substrate-based drugs (SBD). It also helps in understanding the protein-protein interaction between cancer cell line protein (target protein) and plant compound (ligand) which plays key role in cellular signaling, apoptosis, cell proliferation, etc. These protein-protein interactions will use to create a protein-protein network that aid in understanding the cancer pathways better. These bioinformatics tools are inexpensive, save time and energy; flexible, and easy compared to the tedious experimental lab works. Based on Gurung *et al.* (2021) study, it has been demonstrated that β -Bourbonene from *Ficus carica* plant extract had the best docking score with the three cancer target proteins (topoisomerase-I, topoisomerase-II, and VEGFR-2) [15]. They obtained these results from molecular dynamics simulation and molecular docking approaches. They found that these phytochemicals could be developed into attractive multi-target therapeutic candidates that suppress cancer cell proliferation while also inducing apoptosis. According to Singh and Bast (2015) study, it has been reported that epigallocatechin gallate (EGCG) (bioactive compound) had the best binding affinity with IGF1R (PDB ID: 1K3A) and VEGFIIR (PDB ID: 2OH4) [16]. They studied the interaction between the cancer target protein and plant compound through the molecular docking approach (GLIDE (Grid-based Ligand Docking with Energetics)). In this research, the 3D structure of breast cancer cell line proteins, caspase-3, breast cancer susceptibility type 1 (BRCA1), and retinoblastoma (Rb)

were generated, and docking with plant compounds (stigmasterol, ferulic acid, and quercetin, respectively) was studied.

2. Materials and Methods

2.1. Target protein sequence.

The complete amino acid sequence of caspase-3 (GI:16516817), breast cancer type 1 susceptibility protein (BRCA1) (GI:1698399), and retinoblastoma (Rb) (GI:292421) were obtained from the National Center for Biotechnology Information (NCBI) [17]. Caspase-3, BRCA1, and retinoblastoma contain 277, 1863, and 3418 amino acids, respectively.

2.2. Homology modeling.

The 3D models of caspase-3, BRCA1, and retinoblastoma were generated using their relative amino acid sequence in the SWISS Model [18]. The 3D models were then visualized in Discovery Studio (DS) 4.0 [19].

2.3. Physicochemical characterization.

The physicochemical characterizations of the target proteins were determined using ExPASy's ProtParam Proteomics server [20,21].

The salt bridges in the protein models were discovered by using the salt bridges program [22]. Salt bridges program analyses the salt bridges in a protein structure as a negative oxygen atom in Asp and Glu residues and a positive nitrogen atom in Arg, Lys, or His residues. The interatomic distance must be < 7.0 Angstrom.

The Cys_Rec program was used to calculate the number of disulfide bonds present in the protein models. The program finds out the positions of Cys, the sum of Cys present, and identifies the most probable disulfide bond pattern of pairs in the protein sequence [23,24].

2.4. Prediction of secondary structures.

The Self-Optimized Prediction Method from Alignment (SOPMA) server was used in the current study to obtain the secondary structures of caspase-3, BRCA1, and retinoblastoma. SOPMA predicts amino acids for a three-state description of the secondary structure (alpha-helix, beta-sheet, and coil) in a protein model [25].

2.5. Validation tools.

The 3D structures of the protein models were validated using tools such as PROCHECK [26], ProQ [27], ERRAT [28], and Verify 3D [29].

2.6. Target proteins' active sites identification.

After validation was done, the protein models caspase-3, BRCA1, and retinoblastoma were submitted to SCFBio server to predict the binding sites of each protein, respectively. Supercomputing Facility Bioinformatics (SCFBio) is a tool to interpret the language of Genomic DNA from a new physicochemical perspective (Chemgenome). It is also able to address the challenge or problem of protein tertiary structure prediction [30].

2.7. Preparation of ligand models.

The retrieved plant compounds (stigmasterol, ferulic acid, and quercetin) in sdf format from the PubChem database; were prepared using the DS 4.0 'Prepare 'ligand' technique, which deleted duplicates, counted tautomers/isomers, inserted hydrogen bonds, and minimized energy using the CHARMM force field (Chemistry at Harvard Macromolecular Mechanics) [31]. 'Lipinski's Rule of Five and 'Vebers' protocol (Ro5 & VP) were used to filter the produced ligands, which establishes criteria for drug-like qualities and focuses on medication bioavailability [32-34]. The compounds were screened using Ro5 and VP based on molecular weight ($MW \leq 500$ daltons), the number of hydrogen bond donors ($HBD \leq 5$) and hydrogen bond acceptors ($HBA \leq 10$) in each molecule, the number of rotatable bonds ($RB \leq 10$) in each molecule, logP value (≤ 5) and polar surface area ($PSA \leq 140 \text{ \AA}^2$) [35,36]. The ligands that had been screened were then sent to be molecularly docked with target proteins.

2.8. Docking tool.

The docking of the target protein with its relevant phyto-component was performed using SwissDock [37]. The model of the target protein-phyto-component complex was viewed using DS 4.0 [38]. The binding energy, number of hydrogen bonds, and hydrogen bond distance between the target protein and phyto-component were recorded [39-41].

2.9. Evaluation of pharmacokinetics.

The DS 4.0 *in silico* tool 'ADMET 'descriptors' can aid in the evaluation of pharmacokinetic parameters and the assessment of a 'molecule's quality in terms of absorption, distribution, metabolism, excretion, and toxicity following human consumption [42]. This method lowers the expense of new medication development as well as the likelihood of clinical failure. Human intestinal absorption, aqueous solubility, carcinogenicity, human Etherà-go-go-Related Gene (hERG) toxicity, AMES toxicity, hepatotoxicity, fish toxicity, *Tetrahymena pyriformis* toxicity, honeybee toxicity, CYP2D6 inhibition, lethal dose LD₅₀, and plasma protein binding (PPB) were among the metrics calculated by this descriptor [43].

3. Results and Discussion

3.1. Physicochemical characterization.

The total number of amino acids for caspase-3, BRCA1, and Rb are 277, 1863, and 928, respectively. The molecular weight of caspase-3, BRCA1, and Rb proteins are 31641.92, 207720.85, and 106159.41 Daltons, respectively. Next, the isoelectric point (pI) of caspase-3 and Rb is more than 7, indicating alkaline characteristics. However, the isoelectric point (pI) of BRCA1 is less than 7, indicating acidic characteristics. Fundamentally, the isoelectric point (pI) of a protein is the essential characteristic that influences its overall electrostatic behavior [44]. Further, the amount of negatively charged residues (Asp+Glu) and positively charged residues (Arg+Lys) for caspase-3 is 40 and 36, respectively. Meanwhile, BRCA1 has 281 negatively and 213 positively charged residues, whereas Rb has 188 negatively and 122 positively charged residues. Additionally, the instability index of caspase-3, BRCA1, and Rb are computed to be 40.58, 54.68, and 47.85, respectively. Substantially, all three proteins are classified as unstable.

Next, the Cys_Rec tool was used to predict the position of a cysteine residue, the total number of cysteines present, and the pattern if pairs are present in the protein sequence as output. Caspase-3, BRCA1, and Rb proteins had a total of 8, 43, and 15 disulfide bonds, respectively which were calculated using the Cys_Rec tool (Table 1).

Further, salt bridges are the electrostatic interactions that provide the overall proteins' stability. These interactions play an important part in the nucleation process of the hierarchical protein folding model. Based on this study, the number of salt bridges of caspase-3, BRCA1, and Rb obtained from the salt bridges program were 23, 11, and 61, respectively.

Table 1. Cys_Rec result on prediction of disulfide bonds.

Protein	Cys_Rec	Score
Caspase-3	Cys_47	-67.8
	Cys_116	-60.0
	Cys_148	-66.7
	Cys_163	-68.7
	Cys_170	-44.6
	Cys_184	-57.9
	Cys_220	-58.3
	Cys_164	-59.4
BRCA 1	Cys_24	-33.2
	Cys_27	-32.2
	Cys_39	44.2
	Cys_44	41.6
	Cys_47	-21.0
	Cys_61	-14.7
	Cys_64	-4.4
	Cys_91	-67.3
	Cys_197	-11.9
	Cys_226	-11.7
	Cys_274	-16.6
	Cys_305	-3.0
	Cys_328	11.6
	Cys_348	-15.9
	Cys_360	-9.9
	Cys_442	-16.4
	Cys_636	-1.4
	Cys_644	-2.1
	Cys_712	-16.7
	Cys_801	-11.7
	Cys_805	-9.0
	Cys_818	-13.3
	Cys_882	-11.7
	Cys_903	2.7
	Cys_944	-3.5
	Cys_953	4.3
	Cys_994	-7.4
	Cys_1103	-15.7
	Cys_1146	-6.1
	Cys_1225	-27.5
	Cys_1251	-5.5
	Cys_1270	-24.4
	Cys_1291	8.7
Cys_1300	-3.0	
Cys_1372	3.4	
Cys_1382	2.6	
Cys_1501	-7.3	
Cys_1513	3.7	
Cys_1697	-26.9	
Cys_1767	-16.1	
Cys_1768	-8.3	
Cys_1787	-22.4	
Cys_1828	-28.4	
Cys_1847	-17.2	
Rb	Cys_61	-16.7

Protein	Cys_Rec	Score
	Cys_102	-36.7
	Cys_169	-18.9
	Cys_221	-30.1
	Cys_278	-21.0
	Cys_283	-35.3
	Cys_407	-55.5
	Cys_438	-49.9
	Cys_489	-73.1
	Cys_553	-60.9
	Cys_590	-14.9
	Cys_666	-60.3
	Cys_706	-51.3
	Cys_712	-45.3
	Cys_853	-17.1

3.2. Prediction of secondary structures.

SOPMA tool used in this study demonstrated the different regions of secondary structures found in the entire protein sequence, which consists of alpha-helix and beta sheets. The protein models consist of many random coils and a low of beta turns in their structures. Retinoblastoma had the longest α -helix (51 residues). However, caspase-3 consists of 19 residues of α -helix, which was the shortest α -helix. The number of alpha-helix is very important in the protein structure because it maintains the stability of the target protein. Moreover, it develops a strong interaction between target protein (cancer cell line protein) and ligand (compound).

3.3. Validation tools.

PROCHECK was used to determine the stereochemistry and the residues of all three protein structures (caspase-3, BRCA1, and retinoblastoma). Figure 1 shows the analysis of the Ramachandran plot based on the selected protein. According to the PROCHECK results, it was reported that the residues of the protein models were in the most favorable region (> 80%). Validation of all three proteins was indicated in Table 2, which indicates the evaluation of the stability of the selected proteins.

Then, ProQ was carried out to determine the proteins' quality based on the Levitt-Gerstein (LG) score and maximum subarray (MaxSub). According to the results obtained for the LG score, all three proteins were considered extremely good models. MaxSub score was a very good model as indicated in ProQ based on the ranges given to predict results (Table 2).

Validation of proteins was done in ERRAT, which is also an analysis tool for assessing the target protein models evaluated by x-ray crystallography. The value of protein models in ERRAT depends on the statistics of nonbonded atomic interactions in the 3-D protein structure. Proteins are evaluated based on the quality factor, which should be more than 50%. From the gained results in this study, the overall quality factor of caspase-3, BRCA1, and retinoblastoma are 89.565%, 96.954%, and 86.494%, respectively. Thus, all the 3 proteins are confirmed to be of good quality. Figure 2 shows the ERRAT results of caspase-3, BRCA1, and retinoblastoma.

Final validation of protein was done using Verify 3D, which evaluates the quality of protein based on the number of residues available in each protein. According to Verify 3D tool, the proteins with a residue number of more than 80% are predicted to be good protein. As per obtained results of caspase-3, BRCA1 and retinoblastoma were 80.32%, 88.79%, and 89.99%, respectively shown in Figure 3.

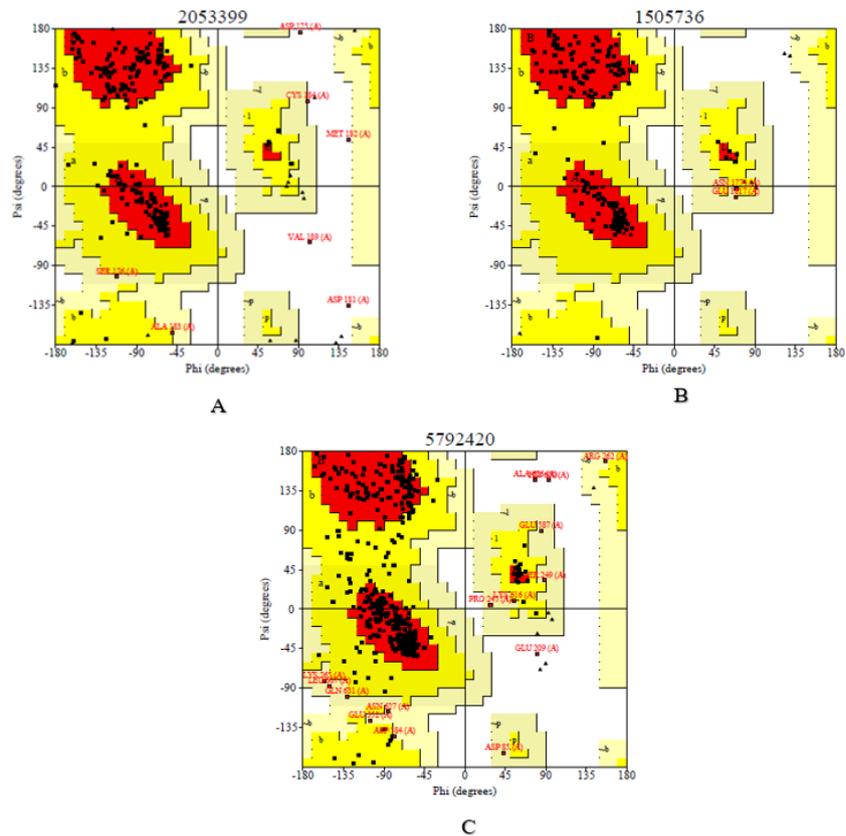
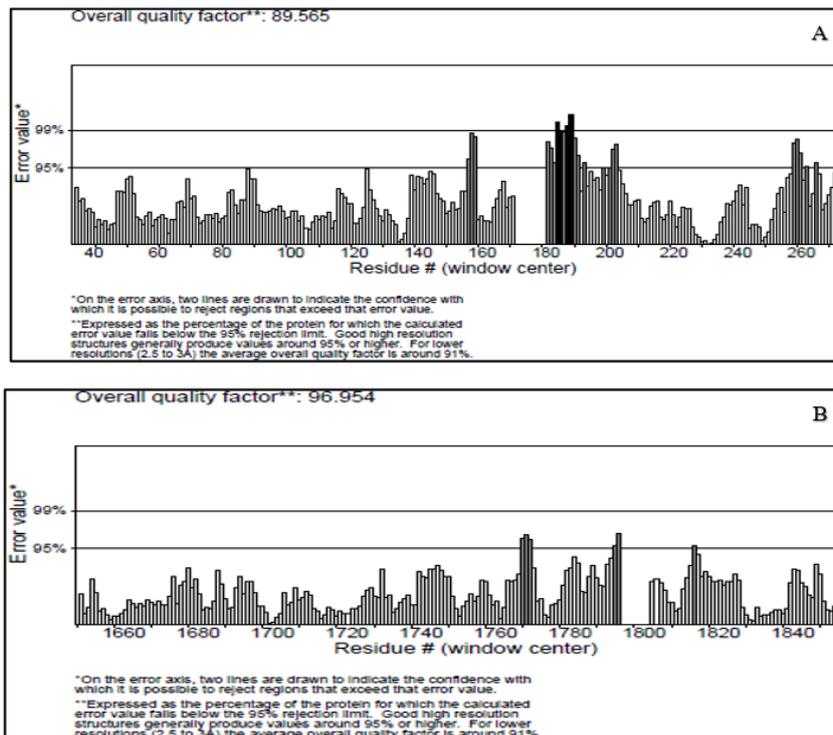


Figure 1. Ramachandran plots for (A) Caspase-3, (B) BRCA1, and (C) Rb.

Table 2. Validation of caspase-3, BRCA1, and Rb protein using PROCHECK program; and LG score and MaxSub using ProQ tool.

Structure	Ramachandran plot statistics (%)				ProQ	
	Most favored	Additionally allowed	Generously allowed	Disallowed	LG score	Max Sub
Caspase-3	87.7	9.8	1.3	1.8	4.386	0.422
BRCA1	91.6	7.4	1.1	0.0	6.538	0.557
Rb	87.0	10.9	1.6	0.4	6.054	0.502



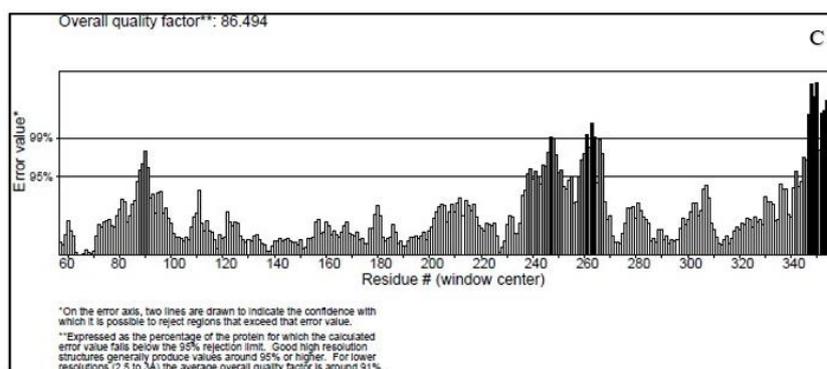


Figure 2. ERRAT result for (A) Caspase-3, (B) BRCA1, and (C) Rb.

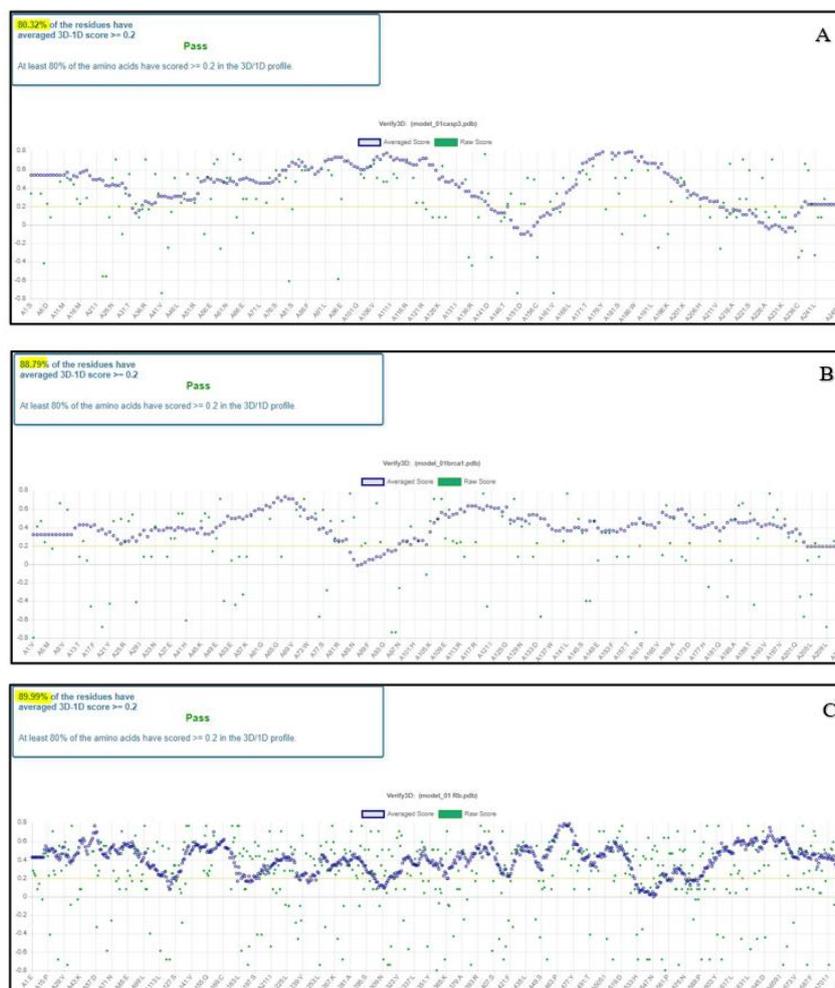


Figure 3. Verify 3D result of (A) Caspase-3, (B) BRCA1, and (C) Rb.

3.4. Target proteins' active sites identification.

The active site of the proteins was identified using the SCFBio server. Based on the results obtained, the protein volume of caspase-3, BRCA1, and Rb are 1422, 1095, and 2158 Å³, respectively. Table 3 shows the identified active sites of Caspase-3, BRCA1, and Rb.

3.5. Screening of plant compounds.

The plant compounds (stigmasterol, ferulic acid, and quercetin) were used in the *in silico* investigation (Table 4).

Table 3. The predicted active site of Caspase-3, BRCA1, and Rb.

Protein	Volume (Å ³)	Pocket Forming Residues
Caspase-3	1556	PRO127, LEU129, ILE132, ALA134, ARG136, GLU139, ASP141, ILE144, ASP147, SER148, GLY149, VAL150, ASP151, ASP152, ASP153, MET154, ALA155, CYS156, HIE157, TYR169, SER170, ALA172, PRO173, GLY174, TYR175, TRP178, TRP186, PHE187, LEU191, MET194, GLN197, TYR198, LEU202, GLU203, PHE204, MET205, HIE206, ILE207, LEU208, THR209, ARG210, ASN212, ARG213, LYS214, VAL215, ALA216, THR217, GLU218, PHE219, LYS232, GLN233, ILE234, PRO235, CYS236, ILE237, VAL238, SER239, MET240, LEU241, THR242, LYS243, GLU244, LEU245, PHE247, PHE50, VAL87, THR9, GLU96
BRCA1	1350	SER10, GLY11, LEU12, PHE127, THR128, ASN129, THR13, MET130, PRO131, ASP133, GLN134, TRP137, PRO14, GLU16, GLN166, PHE17, ARG190, GLU191, LEU31, THR32, ASN33, LEU34, ILE35, THR36, THR39, LYS45, THR46, ASP47, CYS52, GLU53, ARG54, THR55, LEU56, LYS57, PHE59, LEU60, VAL9, VAL95, VAL96
Rb	2158	MET405, GLU406, SER407, MET408, LEU409, SER411, GLU412, GLU413, GLU414, ARG415, LEU416, SER417, ILE418, ASN420, PHE421, SER422, LYS423, LEU424, ASN426, ASP427, ILE429, PHE430, HIE431, LEU434, LEU471, LYS472, PHE474, ASP475, TYR477, LYS478, VAL479, ILE480, GLU481, SER482, ILE484, LYS485, ALA486, GLU487, GLY488, LEU490, ILE495, GLU499, GLU502, MET506, PRO516, LEU520, PRO543, GLN545, ASN546, ASN547, HIE548, THR549, ALA550, ALA551, ASP552, MET553, TYR554, SER560, PRO561, LYS562, LYS563

After preparation, all ligands were exposed to Ro5 and VP filtration, and ferulic acid and quercetin followed the rules (Table 5). However, stigmaterol violated the rules. It had a high value (6.95) for lipophilicity (>5) and a low value for gastrointestinal absorption. This shows that the stigmaterol had minimal absorption and permeability across cell membranes.

Table 4. The list of identified phyto-compounds.

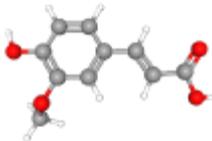
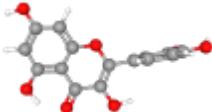
Compound name	PubChem ID	Chemical formula	3D structure
Stigmaterol	5280794	C ₂₉ H ₄₈ O	
Ferulic acid	445858	C ₁₀ H ₁₀ O ₄	
Quercetin	5280343	C ₁₅ H ₁₀ O ₇	

Table 5. The list of pharmacokinetics properties includes physicochemical properties, bioactivity, polar surface area, synthetic accessibility (SA), gastrointestinal (GI) absorption, and Lipinski 5' Rules of all plant compounds.

Compound name	PubChem CID	MW (<500 Daltons)	HBD (≤5)	HBA (≤10)	RB (≤10)	logP (≤5)	PSA (<140Å ²)	SA	GI	LR
Stigmaterol	5280794	412.702	1	1	5	6.95	20.23	Moderate	Low	NO
Ferulic acid	445858	194.18	2	4	3	1.62	66.76	Easy	High	YES
Quercetin	5280343	302.24	5	7	1	1.63	127.45	Easy	High	YES

3.6. Docking tool.

In this study, the SwissDock server was used to dock each protein with its relevant phyto-component. As a result, the lowest binding energy was chosen for each protein-ligand complex because the target protein had the most stable interaction with the plant compound (ligand). Rb-ferulic acid and Rb-quercetin had the lowest negative value for binding energy (-

6.6 and -7.8 kcal/mol, respectively) (Table 6). The 3D structures of the target protein and phyto-component complexes are shown in Figure 4. Ferulic acid interacted with Rb at the residues THR502, SER501, and THR497 through hydrogen bond lengths 2.29 Å, 2.72 Å, and 3.18 Å, respectively (Table 7). Quercetin had three hydrogen bonds with Rb at the residues GLN702, GLN738, and ASN505.

Table 6. The binding affinities of the identified plant compound-target protein interaction. The top, binding affinities are highlighted.

Compound name	Pubchem CID	Caspase-3	BRCA1	Rb
Ferulic acid	445858	-5.7	-5.1	-6.6
Quercetin	5280343	-7.4	-6.5	-7.8

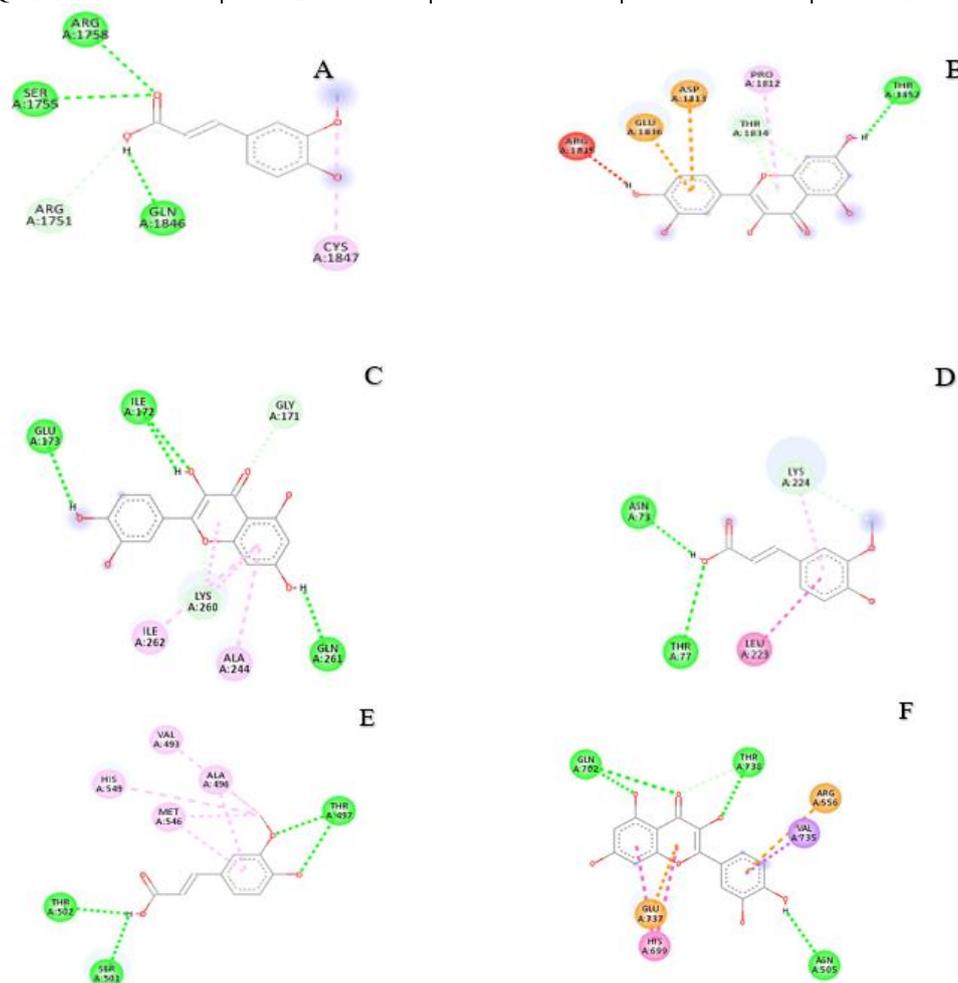


Figure 4. The interaction of BRCA1 with (A) Ferulic acid and (B) Quercetin; interaction of CASPASE 3 with (C) Ferulic acid and (D) Quercetin; Interaction of Rb with (E) Ferulic acid and (F) Quercetin.

Table 7. List of hydrogen bond interactions between BRCA1, Caspase-3, and Rb.

Target protein	Ligand(ID)	Residues	Distance (Å)	Bond	Bond type
BRCA1	Ferulic acid (445858)	ARG1758	3.06	Hydrogen	Conventional hydrogen
		SER1755	3.06	Hydrogen	Conventional hydrogen
		GLN1846	2.14	Hydrogen	Conventional hydrogen
CASPASE 3	Quercetin (5280343)	THR1852	2.07	Hydrogen	Conventional hydrogen
		ILE172	2.95	Hydrogen	Conventional hydrogen
		ILE172	2.44	Hydrogen	Conventional hydrogen
	Ferulic acid (445858)	GLU173	2.55	Hydrogen	Conventional hydrogen
		GLN261	2.44	Hydrogen	Conventional hydrogen
		THR77	3.34	Hydrogen	Conventional hydrogen
Rb	Ferulic acid(445858)	ASN73	2.12	Hydrogen	Conventional hydrogen
		THR502	2.29	Hydrogen	Conventional hydrogen
		SER501	2.72	Hydrogen	Conventional hydrogen

Target protein	Ligand(ID)	Residues	Distance (Å)	Bond	Bond type
		THR497	3.18	Hydrogen	Conventional hydrogen
		THR497	3.06	Hydrogen	Conventional hydrogen
	Quercetin (5280343)	GLN702	3.17	Hydrogen	Conventional hydrogen
		GLN702	3.06	Hydrogen	Conventional hydrogen
		THR738	2.98	Hydrogen	Conventional hydrogen
		ASN505	2.67	Hydrogen	Conventional hydrogen

3.7. Evaluation of pharmacokinetics.

Using the admetSAR 2.0 web server, and *in silico* toxicity test was carried out to discover the negative effects of two plant compounds. Table 8 shows drug-induced hERG toxicity, AMES toxicity, carcinogenicity, P-glycoprotein inhibitor (PGI), fish toxicity, *Tetrahymena pyriformis* (TP) toxicity, honeybee (HB) toxicity, hepatotoxicity, plasma protein binding (PPB), and Rat lethal dose (LD₅₀) discovered by the server. In this study, the toxicity analysis revealed that quercetin (ID: 5280343) had hepatotoxicity. Some flavonoid molecules, such as phloroglucinol, kaempferol, isobutyl isothiocyanate, taurine, and apigenin, are hepatotoxic. Many investigations on the therapeutic properties of these phyto-compounds have already been published. These chemicals could be used to create medicine with a low dosage that has a lower harmful effect on the liver. Plasma protein binding (PPB) of ferulic acid and quercetin was found to be good. This shows that the two plant compounds are pharmacologically active and quickly detach themselves from plasma protein [45].

Table 8. Toxicity test of the plant compounds.

Plant compound (ID)	Ferulic acid (445858)	Quercetin (5280343)
hERG toxicity	No	No
AMES toxicity	No	Yes
Carcinogenicity	No	No
PGI	No	No
Fish toxicity	Yes	Yes
<i>Tetrahymena pyriformis</i> toxicity	Yes	Yes
Honey bee toxicity	Yes	Yes
Hepatotoxicity	No	Yes
Plasma protein binding	0.925	1.175
RAT (LD ₅₀)	1.407	2.559

Currently, the development of advanced computational biology tools is increasing; thus, substrate-based drug design (SBDD) is becoming an important approach in developing target-based therapies. Computer-aided-drug design tools such as molecular dynamics simulation, molecular modeling, molecular docking, etc. approach will help generate the 3-D structure of the protein, analyze the active sites of the protein models and determine the protein-ligand complex interaction. Therefore, these approaches aid in understanding the mechanisms of cancer target proteins and modulate their functions to decrease or stop cancer activities in humans. Scientists found that medicinal plants were the best solution for cancer treatment. They isolated and purified the novel compounds from the plant extracts. Plant compounds potential as anticancer agents. Hence, plant compounds are used as ligands in the SBDD approach. The plant compounds enter the system biology era through drug design tools.

Based on Kasilingam and Elengoe (2018) study, p53, caspase-3, and MADCAM 1 (target proteins) developed a strong interaction with the apigenin (plant compound) due to the lowest binding energy. p53, caspase-3 and MADCAM 1 successfully bound with apigenin at -4.611, -5.750 and -5.307 kcal/mol respectively [46]. The interaction between the target protein

and phyto-compound was made through the hydrogen bonds. Understanding the interactions will aid in developing novel and effective structure-based drugs (SBD) for cancer patients.

Maruthalia and her colleagues (2019) found that myricetin, quercetin, apigenin, luteolin, and baicalein (plant compound) successfully docked with human estrogen receptor ligand-binding domain (hERLBD) at the binding affinity of -10.78, -9.48, -8.92, -8.87, and -8.82 kcal mol⁻¹ respectively [47]. The best interaction was determined based on the highest glide score. The phyto-compounds were anti-estrogens. Schrödinger's (Maestro 9.5) software was used for the molecular docking approach. Luteolin and Baicalein were proven that they be the most promising anti-breast agents among all through laboratory experiments. They showed positive results against the MCF-7 cell line using MTT assay [47].

According to Suhaibun *et al.* (2020) study, the target proteins (p53, caspase-3, and Rb1) were docked successfully with plant compounds (garcinone E, triterpenoid, and gallic acid). The p53-garcinone E, caspase-3- triterpenoid, and Rb1-gallic acid complexes had their docking scores of 3.873, 4.321, and 3.051, respectively [48]. The plant compounds could be potential as an effective anticancer agent.

Mutazah and her colleagues (2020) found that entadamide C and clinamide D had the best binding affinity with caspase-3 at -4.28 kcal/mol and -4.84 kcal/mol, respectively. Entadamide C and clinamide D were the phyto-components derived from methanol extract of *Clinacanthus nutans* leaves. Moreover, these two plant compounds showed positive results for anticancer activity against MDA-MB 231 and MCF-7 cells. The cytotoxicity analysis was carried out using an MTT assay [49]. Therefore, these phyto-components can be further analyzed in *in vivo* study.

Zubair and his co-researchers (2016) performed *in silico* molecular docking between 62 plant compounds and EGFR-TK (target protein). The 62 plant compounds were extracted from nine *Begonia* species. Cyanidin 3-(6''-(Z)-p-coumarylsophoroside) (phyto-compound) bound successfully with the binding site of EGFR-TK. It has the best docking score of -120.2330 among all the plant compounds [50].

From the above study, each protein (caspase-3, BRCA1, retinoblastoma) was successfully docked with its related phyto-component. The protein models had a strong interaction with the selected phyto-component which resulted in good binding energy. According to the lowest binding energy, the Rb-quercetin complex had the most stable binding affinity (-7.8 kcal/mol) among all the protein-phyto-compound complexes (caspase-3-quercetin, caspase-3-ferulic acid, BRCA1-ferulic acid, BRCA1-quercetin, Rb-ferulic acid). These potential drug candidates can then be tested in the lab to ensure that they work properly against the breast cancer cells. Therefore, based on the results obtained can be concluded that the current study can be used to design and develop a more powerful structure-based drug.

4. Conclusions

In conclusion, the functions of caspase -3, BRCA1, and Rb (breast cancer cell target proteins) could be modified successfully with ferulic acid and quercetin (plant compounds) respectively through the docking approach. The plant compounds had a strong interaction with the target proteins based on their lowest docking score. This interaction will help enhance or decrease the particular activity of target proteins. This *in silico* study of the interaction between cancer cell protein and plant compound will aid in developing a new and effective drug for breast cancer treatment.

Funding

This research received no external funding.

Acknowledgments

This work was supported by the Biotechnology Department, Faculty of Science, Lincoln University College, Malaysia.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Supramaniam, G.; Elengoe, A. *In silico* molecular docking of glycyrrhizin and breast cancer cell line proteins. In: *Plant-derived Bioactives*. Springer **2020**; pp. 575-589.
2. GLOBOCAN. Estimated age standardized incidence rates of all cancers, Worldwide 2018. **2019**. Available: <https://gco.iarc.fr/today/home> (accessed on 16 April 2021).
3. World Health Organization. WHO, Cancer. **2018**. Available: <https://www.who.int/news-room/fact-sheets/detail/cancer> (accessed on 16 April 2021).
4. Nordqvist, C. What you need to know about breast cancer. **2017**. Available: <https://www.medicalnewstoday.com/articles/37136.php> (accessed on 20 April 2021).
5. Verma, R.; Bowen, R.L.; Slater, S.E.; Mihaimed, F.; Jones, J.L. Pathological and epidemiological factors associated with advanced stage at diagnosis of breast cancer. *British medical bulletin* **2012**, *103*, 129-145, <https://doi.org/10.1093/bmb/lds018>.
6. Dunning, A.M.; Healey, C.S.; Pharoah, P.D.; Teare, M.D.; Ponder, B.A.; Easton, D.F. A systematic review of genetic polymorphisms and breast cancer risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **1999**, *8*, 843-854.
7. Cavalieri, E.; Chakravarti, D.; Guttenplan, J.; Hart, E.; Ingle, J.; Jankowiak, R.; Muti, P.; Rogan, E.; Russo, J.; Santen, R.; Sutter, T. Catechol estrogen quinones as initiators of breast and other human cancers: implications for biomarkers of susceptibility and cancer prevention. *Biochimica et biophysica acta* **2006**, *1766*, 63-78, <https://doi.org/10.1016/j.bbcan.2006.03.001>.
8. Ikwu, F.A.; Shallangwa, G.A.; Mamza, P.A.; Uzairu, A. In silico studies of piperazine derivatives as potent anti-proliferative agents against PC-3 prostate cancer cell lines. *Heliyon* **2020**, *6*, e03273, <https://doi.org/10.1016/j.heliyon.2020.e03273>.
9. Khoo, L.W.; Audrey Kow, S.; Lee, M.T.; Tan, C.P.; Shaari, K.; Tham, C.L.; Abas, F. A Comprehensive Review on Phytochemistry and Pharmacological Activities of *Clinacanthus nutans* (Burm.f.) Lindau. *Evidence-Based Complementary and Alternative Medicine* **2018**, *2018*, 9276260, <https://doi.org/10.1155/2018/9276260>.
10. Maiuri, M.C.; Zalckvar, E.; Kimchi, A.; Kroemer, G. Self-eating and self-killing: crosstalk between autophagy and apoptosis. *Nature Reviews Molecular Cell Biology* **2007**, *8*, 741-752, <https://doi.org/10.1038/nrm2239>
11. Rodríguez-García, C.; Sánchez-Quesada, C.; Gaforio, J.J. Dietary Flavonoids as Cancer Chemopreventive Agents: An Updated Review of Human Studies. *Antioxidants (Basel, Switzerland)* **2019**, *8*, <https://doi.org/10.3390/antiox8050137>.
12. Bondonno, N.P.; Dalgaard, F.; Kyrø, C.; Murray, K.; Bondonno, C.P.; Lewis, J.R.; Croft, K.D.; Gislason, G.; Scalbert, A.; Cassidy, A.; Tjønneland, A.; Overvad, K.; Hodgson, J.M. Flavonoid intake is associated with lower mortality in the Danish Diet Cancer and Health Cohort. *Nature Communications* **2019**, *10*, <https://doi.org/10.1038/s41467-019-11622-x>
13. Arts, I.C. A review of the epidemiological evidence on tea, flavonoids, and lung cancer. *The Journal of nutrition* **2008**, *138*, 1561s-1566s, <https://doi.org/10.1093/jn/138.8.1561S>.
14. Ciumărnean, L.; Milaciu, M.V.; Runcan, O.; Vesa, Ș.C.; Răchișan, A.L.; Negrean, V.; Perné, M.-G.; Donca, V.I.; Alexescu, T.-G.; Para, I.; Dogaru, G. The Effects of Flavonoids in Cardiovascular Diseases. *Molecules (Basel, Switzerland)* **2020**, *25*, <https://doi.org/10.3390/molecules25184320>.
15. Gurung, A.B.; Ali, M.A.; Lee, J.; Farah, M.A.; Al-Anazi, K.M. Molecular docking and dynamics simulation study of bioactive compounds from *Ficus carica* L. with important anticancer drug targets. *PLoS One* **2021**, *16*, <https://doi.org/10.1371/journal.pone.0254035>.

16. Singh, P.; Bast, F. Screening of multi-targeted natural compounds for receptor tyrosine kinases inhibitors and biological evaluation on cancer cell lines, in silico and in vitro. *Medical oncology (Northwood, London, England)* **2015**, *32*, 1-18, <https://doi.org/10.1007/s12032-015-0678-8>.
17. National Center for Biotechnology Information (NCBI). **2021**. Available: <https://www.ncbi.nlm.nih.gov/> (accessed on 21 April 2021).
18. Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Gallo Cassarino, T.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research* **2014**, *42*, W252-258, <https://doi.org/10.1093/nar/gku340>.
19. BIOVIA, Dassault Systèmes, Discovery Studio, 4.0, San Diego: Dassault Systèmes, **2021**.
20. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In: *The Proteomics Protocols Handbook*. **2005**; pp. 571-607, <https://doi.org/10.1385/1-59259-890-0:571>.
21. Ahmed, E.Y.; Abdel Latif, N.A.; El-Mansy, M.F.; Elserwy, W.S.; Abdelhafez, O.M. VEGFR-2 inhibiting effect and molecular modeling of newly synthesized coumarin derivatives as anti-breast cancer agents. *Bioorganic & Medicinal Chemistry* **2020**, *28*, <https://doi.org/10.1016/j.bmc.2020.115328>.
22. Costantini, S.; Colonna, G.; Facchiano, A.M. ESBRI: a web server for evaluating salt bridges in proteins. *Bioinformatics* **2008**, *3*, 137-138, <https://doi.org/10.6026/97320630003137>.
23. Roy, S.; Maheshwari, N.; Chauhan, R.; Sen, N.K.; Sharma, A. Structure prediction and functional characterization of secondary metabolite proteins of Ocimum. *Bioinformatics* **2011**, *6*, 315-319, <https://doi.org/10.6026/97320630006315>.
24. Elengoe, A.; Sebastian, E. In silico molecular modelling and docking of allicin, epigallocatechin-3-gallate and gingerol against colon cancer cell proteins. *Asia Pacific Journal of Molecular Biology and Biotechnology* **2020**, 51-67, <https://doi.org/10.35118/apjmbb.2020.028.4.05>.
25. Geourjon, C.; Deléage, G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer applications in the biosciences : CABIOS* **1995**, *11*, 681-684, <https://doi.org/10.1093/bioinformatics/11.6.681>.
26. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **1993**, *26*, 283-291, <https://doi.org/10.1107/S0021889892009944>.
27. Wallner, B.; Elofsson, A. Can correct protein models be identified? *Protein Sci* **2003**, *12*, 1073-1086, <https://doi.org/10.1110/ps.0236803>.
28. Colovos, C.; Yeates, T.O. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* **1993**, *2*, 1511-1519, <https://doi.org/10.1002/pro.5560020916>.
29. Eisenberg, D.; Lüthy, R.; Bowie, J.U. VERIFY3D: assessment of protein models with three-dimensional profiles. In: *Methods in enzymology*. Academic Press **1997**.
30. Singh, T.; Biswas, D.; Jayaram, B. AADS - An Automated Active Site Identification, Docking, and Scoring Protocol for Protein Targets Based on Physicochemical Descriptors. *Journal of Chemical Information and Modeling* **2011**, *51*, 2515-2527.
31. Brooks, B.R.; Brooks, C.L., 3rd; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A.R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.W.; Post, C.B.; Pu, J.Z.; Schaefer, M.; Tidor, B.; Venable, R.M.; Woodcock, H.L.; Wu, X.; Yang, W.; York, D.M.; Karplus, M. CHARMM: the biomolecular simulation program. *Journal of computational chemistry* **2009**, *30*, 1545-1614, <https://doi.org/10.1002/jcc.21287>.
32. Lipinski, C.A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies* **2004**, *1*, 337-341, <https://doi.org/10.1016/j.ddtec.2004.11.007>.
33. Veber, D.F.; Johnson, S.R.; Cheng, H.-Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *Journal of Medicinal Chemistry* **2002**, *45*, 2615-2623, <https://doi.org/10.1021/jm020017n>.
34. Jagtap, N.; Yadav, A.; Mohite, S. Synthesis, molecular docking studies and anticancer activity of 1, 3, 4-oxadiazole-3 (2h)-thione derivatives. *Journal of University of Shanghai for Science and Technology* **2020**, *22*, 535-550.
35. Rodrigues, J.; Hullatti, K.K.; Jalalpure, S.; Khanal, P. In-vitro Cytotoxicity and in silico Molecular Docking of Alkaloids from *Tiliacora acuminata*. *Indian Journal of Pharmaceutical Education and Research* **2020**, *54*, s295-s300, <https://doi.org/10.5530/ijper.54.2s.86>.
36. Tantawy, E.; Amer, A.; Mohamed, E.; Alla, M.; Nafie, M. Synthesis, characterization of some pyrazine derivatives as anticancer agents: In vitro and in Silico approaches. *Journal of Molecular Structure* **2020**, *1210*, <https://doi.org/10.1016/j.molstruc.2020.128013>.
37. Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic acids research* **2011**, *39*, W270-277, <https://doi.org/10.1093/nar/gkr366>.

38. Abd El-Sattar, N.E.A.; El-Adl, K.; El-Hashash, M.A.; Salama, S.A.; Elhady, M.M. Design, synthesis, molecular docking and in silico ADMET profile of pyrano[2,3-d]pyrimidine derivatives as antimicrobial and anticancer agents. *Bioorganic Chemistry* **2021**, *115*, <https://doi.org/10.1016/j.bioorg.2021.105186>.
39. Parmar, D.R.; Soni, J.Y.; Guduru, R.; Rayani, R.H.; Kusurkar, R.V.; Vala, A.G.; Talukdar, S.N.; Eissa, I.H.; Metwaly, A.M.; Khalil, A.; Zunjar, V.; Battula, S. Discovery of new anticancer thiourea-azetidine hybrids: design, synthesis, in vitro antiproliferative, SAR, in silico molecular docking against VEGFR-2, ADMET, toxicity, and DFT studies. *Bioorganic Chemistry* **2021**, *115*, <https://doi.org/10.1016/j.bioorg.2021.105206>.
40. Alam, S.; Nasreen, S.; Ahmad, A.; Darokar, M.P.; Khan, F. Detection of Natural Inhibitors against Human Liver Cancer Cell Lines through QSAR, Molecular Docking and ADMET Studies. *Curr Top Med Chem* **2021**, *21*, 686-695, <https://doi.org/10.2174/1568026620666201204155830>.
41. Ikwu, F.; Shallangwa, G.; Mamza, P. QSAR, QSTR, and molecular docking studies of the anti-proliferative activity of phenylpiperazine derivatives against DU145 prostate cancer cell lines. *Beni-Suef University Journal of Basic and Applied Sciences* **2020**, *9*, 1-12, <https://doi.org/10.1186/s43088-020-00054-y>.
42. Tian, S.; Wang, J.; Li, Y.; Li, D.; Xu, L.; Hou, T. The application of in silico drug-likeness predictions in pharmaceutical research. *Advanced Drug Delivery Reviews* **2015**, *86*, 2-10, <https://doi.org/10.1016/j.addr.2015.01.009>.
43. Usha, T.; Goyal, A.K.; Lubna, S.; Prashanth, H.; Mohan, T.M.; Pande, V.; Middha, S.K. Identification of anticancer targets of eco-friendly waste Punica granatum peel by dual reverse virtual screening and binding analysis. *Asian Pacific journal of cancer prevention* **2014**, *15*, 10345-10350.
44. Łapińska, U.; Saar, K.L.; Yates, E.V.; Herling, T.W.; Müller, T.; Challa, P.K.; Dobson, C.M.; Knowles, T.P.J. Gradient-free determination of isoelectric points of proteins on chip. *Physical Chemistry Chemical Physics* **2017**, *19*, 23060-23067, <https://doi.org/10.1039/C7CP01503H>.
45. Smith, D.A.; Di, L.; Kerns, E.H. The effect of plasma protein binding on in vivo efficacy: misconceptions in drug discovery. *Nature Reviews Drug Discovery* **2010**, *9*, 929-939, <https://doi.org/10.1038/nrd3287>.
46. Kasilingam, T.; Elengoe, A. In silico molecular modeling and docking of apigenin against the lung cancer cell proteins. *Asian Journal of Pharmaceutical and Clinical Research* **2018**, *11*, 246-252, <https://doi.org/10.22159/ajpcr.2018.v11i9.26649>.
47. Maruthanila, V.L.; Elancheran, R.; Roy, N.K.; Bhattacharya, A.; Kunnumakkara, A.B.; Kabilan, S.; Kotoky, J. In silico Molecular Modelling of Selected Natural Ligands and their Binding Features with Estrogen Receptor Alpha. *Current computer-aided drug design* **2019**, *15*, 89-96, <https://doi.org/10.2174/1573409914666181008165356>.
48. Suhaibun, S.; Elengoe, A.; Poddar, R. Technology Advance in Drug Design Using Computational Biology Tool. *Malaysian Journal of Medicine and Health Sciences* **2020**, *16*, 2636-9346.
49. Mutazah, R.; Hamid, H.A.; Mazila Ramli, A.N.; Fasihi Mohd Aluwi, M.F.; Yusoff, M.M. In vitro cytotoxicity of Clinacanthus nutans fractions on breast cancer cells and molecular docking study of sulphur containing compounds against caspase-3. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* **2020**, *135*, <https://doi.org/10.1016/j.fct.2019.110869>.
50. Zubair, M.S.; Anam, S.; Khumaidi, A.; Susanto, Y.; Hidayat, M.; Ridhay, A. Molecular docking approach to identify potential anticancer compounds from Begonia (Begonia sp). *AIP Conference Proceedings* **2016**, *1755*, <https://doi.org/10.1063/1.4958513>.