ORIGINAL ARTICLE

Technology Advance in Drug Design Using Computational Biology Tool

Sherman Ryner Suhaibun[1], Asita Elengoe[1], Ruma Poddar[2]

[1] Department of Biotechnology, Faculty of Science, Lincoln University College, No. 2, Jalan Stadium, SS 7/15, Kelana Jaya,47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia
[2] Faculty of Science, Lincoln University College, No. 2, Jalan Stadium, SS 7/15, Kelana Jaya, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia

ABSTRACT

Introduction: Worldwide, breast cancer is the most life-threatening disease among women. There is always high search to find a cure for cancer. Plant compounds have been identified that they have anti-cancer properties. Therefore, phyto-compounds can be potential for the development of new drugs. In this research, three-dimensional (3-D) structure of breast cancer cell line proteins, tumor suppressor gene (p53), caspase-3 and retinoblastoma-1 were generated and docking with plant compounds (garcinone E, triterpenoid and gallic acid respectively) was studied. **Methods:** The three-dimensional models of proteins were built using SWISS model. Then, the physical and chemical characters of the protein models were determined using ExPASy - ProtParam tool. Next, the proteins were assessed using validation tools such as PROCHECK, ProQ, ERRAT and Verify 3D programs. **Results:** The results show that the proteins were stable. Lastly, the protein models were docked successfully with garcinone E, triterpenoid and gallic acid respectively using BSP-slim server. The docking scores of the protein-phyto-compound complexes (p53-garcinone E, caspase-3- triterpenoid and Rb1-gallic acid) were 3.873, 4.321 and 3.051 respectively. The proteins had a stable bond with phyto-compounds. **Conclusion:** The study of the protein-phyto-compound complex interaction will aid in designing new clinical drugs.

Keywords: Caspase-3, Docking, Phyto-compounds, p53, Retinoblastoma-1

Corresponding Author:
Asita Elengoe, PhD
Email: asitaelengoe@yahoo.com

INTRODUCTION

Globally, breast cancer is the most leading cancer among women. World Health Organization (WHO) reported about 2.09 million new cases in 2018. It is the most prevalent cancer in women in both highly and least developed regions where less developed regions are estimated to have significantly more cases compared to highly developed regions due to low awareness about breast cancer. It is responsible for the killing of 627, 000 people. It listed at the fifth rank for cancer death from overall cancer. Based on the Section of Cancer Surveillance, World Health Organization, breast cancer is the most common cause for 324,000 deaths (14.3%) among women in the least developed region. It also causes 198,000 women deaths (15.4%) in more developed regions. An estimated global death from breast cancer by country development level report was done and described by GLOBOCAN, 1990, 2010 and 2030. The report shows the difference between the estimate of death between a less developed country and more developed nation (1).

The lack of equipment for diagnosis may be the possible reason why death cases in less developed countries are higher compared to cases in more developed countries because breast cancer is hard to detect especially in early stages. Even in Malaysia alone, based on the database obtained from World Health Organization (2015) (2), a total of 19,301 cases of cancer reported and 28% (5,410) from the total cases are breast cancer.

According to a newsletter by Christian Nordqvist published in Medical News Today (3), the exact cause of breast cancer remains unknown. However, some risk factors might be the risk for breast cancer development such as age, genetics, previous history of breast cancer, body weight, alcohol consumption, exposure to radiation, hormone treatments and occupational hazards.

Thus, there is an urgency to develop a new treatment for breast cancer. Medicinal plants are used effectively in the cancer treatment as they naturally derive anti-cancer properties. Moreover, the approach of using medicinal plants is safe and efficient for cancer treatment. This is due to the usage of drugs in chemotherapy and radiation has a lot of side effects such as killing of normal and healthy cells altogether with the cancer cells.

Phytochemicals are chemical compounds that are produced by plants via primary or secondary metabolism (4). Garcinone E is found in fruits and it is a constituent of the fruit Garcinia mangostana (Mangosteen). According to Jung et al. (2006) study, garcinone E can induce cell cycle arrest at G1 phase (5). Gallic acid is an organic acid that is also known as 3,4,5-trihydroxy benzoic acid and found in Amoora rohituka plant. Gallic acid was studied for its ability to induce cell cycle arrest by p53 mediated mechanism (6). Triterpenoids are a byproduct of squalene 2,3-oxide decomposition. It is found in God's Crown (Phaleria macrocarpa) plant. Squalene is a precursor to the biosynthesis of sterol. Triterpenoid is derived biosynthetically from units of isoprene. Based on Rabi et al. (2013) study, it has been reported that triterpenic acid can induce apoptosis in breast cancer through up-regulation of caspase-3 (7).

Computational biology tools such as homology modelling and molecular docking help in designing substrate-based drugs and study the interaction between the target protein (cancer cell proteins) and ligand (plant compound). This tool cuts cost, energy, time, simple and easy. In this study, the docking between breast cancer cell line proteins (tumor suppressor gene p53, caspase-3, retinoblastoma-1 (Rb-1)) and phyto-compounds (garcinone E, triterpenoid and gallic acid) were determined.

MATERIALS AND METHODS

Target protein sequence

The complete amino acid sequence of p53 (GI: 23491729), caspase-3 (GI: 16516817) and retinoblastoma-1 (GI: 26252120) was obtained from the National Center for Biotechnology (NCBI) (https://www.ncbi.nlm.nih.gov/). p53, caspase-3 and retinoblastoma-1 contain 227, 285 and 928 amino acids respectively.

Homology modelling

The three-dimensional (3-D) models of p53, caspase-3 and retinoblastoma-1 were generated using the SWISS Model (8). At SWISS-MODEL Interactive Workspace, the 3D models were generated and saved in PDB format.

Physiochemical characterisation

Physical and chemical characters of the protein structures were determined using ExPASy - ProtParam tool (9). The discovery of salt bridges between the protein chains was performed using the ESBRI program (10). The disulphide bonds between the Cys-Cys residues were evaluated by Cys_Rec Program (11).

Prediction of secondary structures

The secondary structures of protein (alpha helix, extended strand, beta turn and random coil) were predicted using the Self-Optimized Prediction Method with Alignment (SOPMA) program (12).

Validation tools

The 3D protein structures were verified through Structural Analysis and Verification Server (SAVES v5.0) (13, 14, 16) and ProQ (15).

Target proteins' active sites identification

To explore the active sites of p53, caspase-3 and Rb-1, these protein models submitted into Active-Site Prediction Server (SCFBio) (17).

Preparation of ligand models

The complete sequence of the ligand model was retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and upon completion of the ligand model preparation; the models were downloaded into the PDB file format.

Docking tool

The docking of the target protein and phyto-components were done using the BSP-Slim server. A blind molecular docking approach was carried out (18).

RESULTS

Physiochemical characterisation

Computed isoelectric (pI) value for p53 and retinoblastoma-1 was more than 7. Therefore, they indicate alkaline characteristics while caspase-3 computed isoelectric was less than 7 which indicates acidic characteristics. The molecular weight of p53, caspase-3 and retinoblastoma-1 protein were 25583.93, 32672.02, and 106159.41 Daltons respectively. Moreover, three of these proteins had the different sum of negatively (ASP+GLU) and positively (ARG+LYS) charged residues. For p53, (ASP+GLU) was 21 and (ARG+LYS) was 27. Caspase-3 had 40 (ASP+GLU) and 36 (ARG+LYS) while retinoblastoma-1 had 118 (ASP+GLU) and 122 (ARG+LYS). Based on data obtained from Expasy's ProtParam instability index, p53 and retinoblastoma-1 were classified as an unstable protein with an instability index of 69.83 for p53 and 47.85 for retinoblastoma-1. However, caspase-3 protein was classified as a stable protein with an instability index of 39.53.

Salt bridges also play an important role in the model and stability of a protein. Interruption of a salt bridge will reduce the stability of a protein (19). In this study, salt bridges of p53, caspase-3 and Rb-1 had a total of 71, 30, and 30 salt bridges respectively which were obtained from ESBRI analysis. Moreover, p53, caspase-3 and Rb-1 had a sum of 10, 8 and 15 disulphide bonds respectively which were predicted using Cys_Rec analysis (Table I).

Prediction of secondary structures

The protein models had the highest number of random coils while they had the lowest number of beta turns at different positions in proteins. The composition of α-helix in p53, caspase-3 and retinoblastoma-1 were

Table I: Cys_Rec result on prediction of disulfide bonds

Protein	Cys_Rec	Score
p53	Cys_124	-10.0
	Cys_135	0.3
	Cys_141	-9.7
	Cys_176	-12.1
	Cys_182	-46.5
	Cys_229	-55.4
	Cys_238	-1.2
	Cys_242	-9.5
	Cys_275	-35.1
Caspase-3	Cys_277	-50.8
	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_264	-59.4
Retinoblastoma-1	Cys_61	-16.1
	Cys_102	-36.7
	Cys_169	-18.9
	Cys_221	-30.1
	Cys_278	-21.0
	Cys_283	-35.3
	Cys_407	-55.5

22.39, 30.69 and 48.71 respectively. Retinoblastoma-1had the longest α-helix (51 residues) but p53 consists of 17 residues of α-helix which was the shortest α-helix. The alpha helix is important because it forms the hydrogen bond between the target protein and plant compound.

Validation tools

Chemistry and good-quality stereoisomers of protein models were then validated by Ramachandran plot calculation through PROCHECK software. From the analysis that was done, it was reported that all the protein models were within the acceptable range (>80% as shown in Ramachandran plots statistic) (Fig. 1). Table II indicates that these proteins were evaluated to be great and stable.

ProQ was then used to evaluate the quality of the proteins using Levitt-Gerstein (LG) score and maximum subarray (MaxSub). Results obtained from ProQ stated that proteins with LG score higher than 1.5 and MaxSub score higher than 0.1. This result shows that the proteins were good and stable in condition. LG and MaxSub score for p53, caspase-3 and retinoblastoma-1 were shown in Table II.

Another validation tool used was ERRAT which is an analysis for evaluating protein models determined by

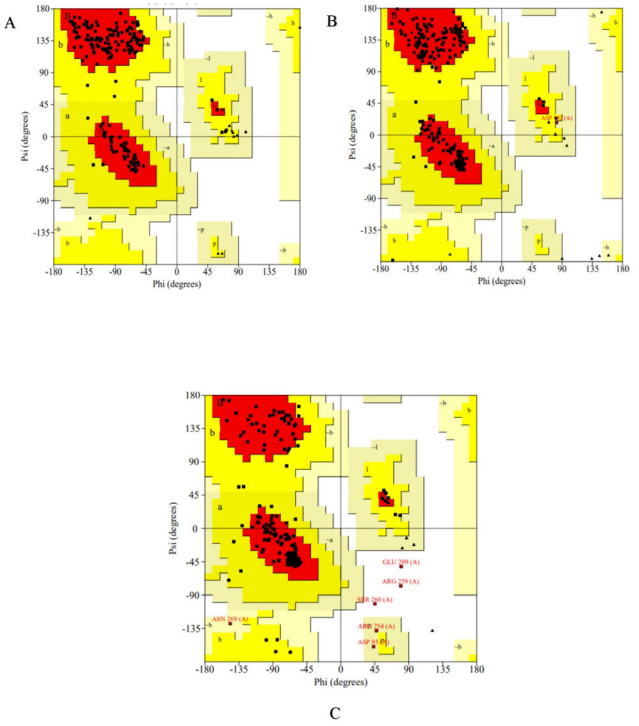

Figure 1: Ramachandran plots for (A) p53, (B) Caspase-3, and (C) Retinoblastoma-1

Table II: Validation of p53, caspase-3 and retinoblastoma-1 protein using PROCHECK program; and LG score and MaxSub using ProQ tool

Structure	Ramachandran plot statistics (%)				ProQ	
	Most favoured	Additionally allowed	Generously allowed	Disallowed	LG score	Max Sub
p53	92.9	7.1	0.0	0.0	3.988	0.356
Caspase-3	92.5	7.0	0.4	0.0	4.903	0.403
Retinoblastoma-1	90.8	7.0	1.1	1.1	6.272	0.473

x-ray crystallography. ERRAT value depends on the statistics of non-bonded atomic interactions in the three-dimensional protein structure. The quality factor of more than 50% is accepted as a good quality protein. The quality factor for p53, caspase-3 and retinoblastoma-1 were 81.250%, 84.188% and 89.985% respectively (Fig. 2).

Upon completion of protein validation using ERRAT, Verify 3D was then used to validate the protein models thus showing the percentage of residues. p53, caspase-3 and retinoblastoma-1 had 93.47%, 84.77%, and 90.44% of the residues respectively and an average 3D-1D score of more than 0.2 which indicates that all the sequence was compatible with its protein model (Fig. 3). When all of the protein models have been validated and passed the requirement needed for protein evaluation, proteins were then sent for docking analysis with phyto-components.

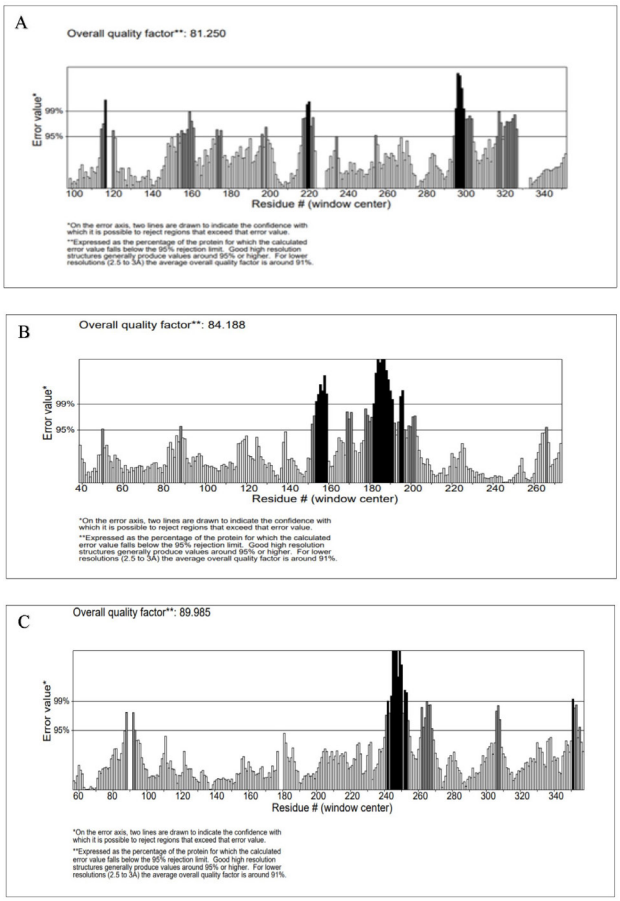

Figure 2: ERRAT result for (A) p53, (B) Caspase-3, and (C) Retinoblastoma-1

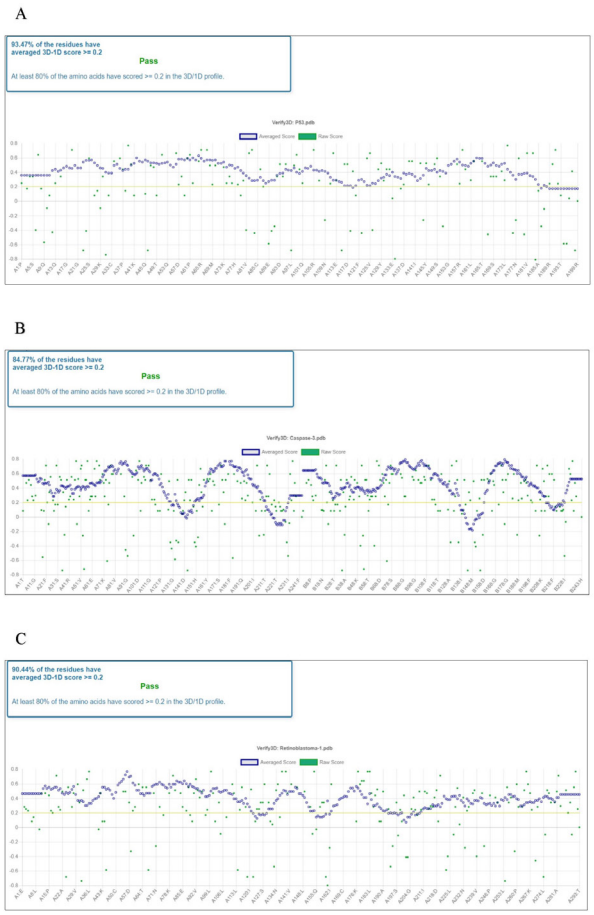

Figure 3: Verify 3D result of (A) p53, (B) Caspase-3, and (C) Retinoblastoma-1

Table III: Predicted active site of p53, Caspase-3 and Retinoblastoma-1

Protein	Volume	Pocket Forming Residues
p53	272	VAL56, GLY17, GLN13, ARG19, TRP55, PHE18, TYR12, THR11, LYS10, LEU54, ASN177, LEU20, GLN53, GLY21, ARG176, PHE22, HIE24, ILE163, GLN9, SER178, TYR143, LEU23, TYR35, PRO37, PRO51, PHE179, ASN40, LEU161, ALA38, TYR145, MET42, LEU39, LYS73
Caspase-3	178	ILE20, MET72, LEU108, ILE131, PHE114, CYS88, ILE111, LYS109, ILE132, PHE130, THR112, PHE130, ASN113, SER76, VAL87, LEU129, PHE86, PHE115, ARG116, LYS128, SER85, CYS120, GLY117, PRO127, LYS126, THR124, TYR9
Retinoblastoma-1	229	LEU137, ALA136, LEU139, HIE287, GLU114, LYS140, VAL262, LEU285, ASP286, SER142, PHE110, GLU263, VAL141, VAL262, TRP143, PHE284, THR145, ILE144, ASP288, SER266, LEU283, ARG282, LEU291, LYS267, LYS70, LEU148, LEU147, GLU270, ASP280, LEU106, ALA281, THR66, TYR269, LYS150, GLU270, TYR273, ILE272, TYR269

Target proteins' active sites identification

The active site of the target proteins was obtained using the SCFBio server. Results obtained show that the protein volume for p53, caspase-3 and Rb1 were 272, 178 and 229 A3 respectively. The predicted active sites were shown in Table III.

Docking tool

In this study, target proteins were docked successfully with phyto-components through the BSP-Slim Server tool. Based on the lowest docking score, the best docking orientation was chosen. The docking scores of the protein-phyto-compound complexes (p53-garcinone E,

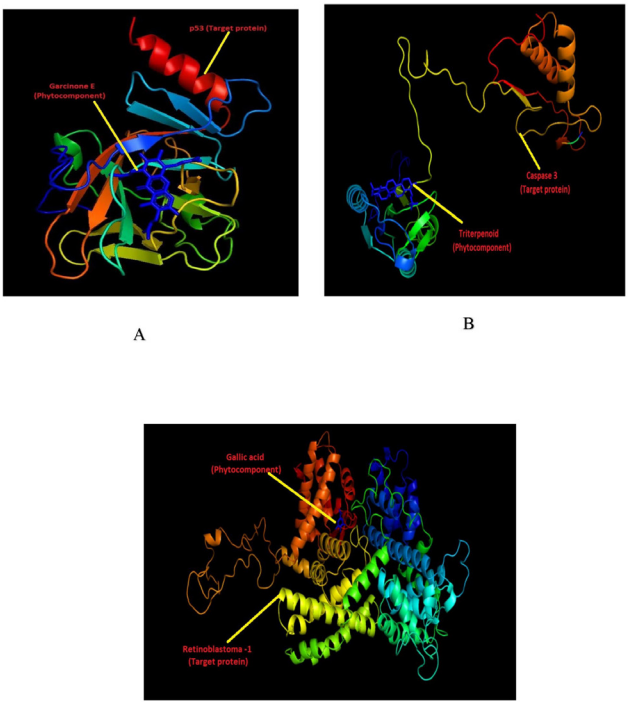

Figure 4: (A) Docking of p53 (target protein) and Garcinone E (phyto-component); (B) Docking of Caspase-3 (target protein) with Triterpenoid (phyto-component); (C) Docking of Rb1 (target protein) with Gallic Acid (phyto-component)

caspase-3- triterpenoid and rb1-gallic acid) were 3.873, 4.321 and 3.051 respectively. The 3D structures of the target protein and phyto-component complexes were shown in Fig. 4.

DISCUSSION

Cancer remains as one of the leading causes of morbidity and mortality until today and therefore researchers and healthcare communities are working on significant advances in cancer research and therapy (20). With our recent advancements and developments in biotechnology especially in the study of genome, proteomes and transcriptomes, researchers can develop computational methods for data analysis and thus improving our understanding of cancer. Currently, computational approaches are used in the treatment of cancer by identifying the cancer genes and pathways (21).

A research experiment was conducted to explore the therapeutic ability of fenugreek against breast cancer by employing molecular docking which according to this study; fenugreek is a type of spice that is being used around the world (22) and it also possesses anti-cancer properties that are known to induce apoptosis (23). One of the breast cancer causes is the hereditary factor. It is due to a mutation of BRCA1 and BRCA2 genes. The previous studies have been reported that the fenugreek seed has anti-cancer activity towards breast cancer. It also revealed that it controls the proliferation of MDA-MB 231-induced mammary hyperplasia (24). However, the most effective compound has not yet been delineated and therefore current investigations are being done by employing computational techniques such as molecular modelling and docking. The interaction between the target protein (cancer cell protein) and ligand (drug or plant component) will be studied through a docking approach (25).

Recently, medicinal plants are the potential to be promising anti-cancer agents. The in silico study of cytotoxic activity of isoquinoline alkaloid from Leguminosae (Erythrina poeppigiana) against breast cancer cell line (MCF-7) was done. The local name for this plant is “dadap belendung” which known in Indonesia. The leaves were traditionally used for the treatment of infection, inflammation and fever (26). The plant constitutes was extracted and isolated to determine the cytotoxicity activity against MCF-7 cell line through lab experiment. The interaction between the plant component and epidermal growth factor receptor 2 (EGFR 2) was studied using molecular docking approach. The in silico technique helps to find out how the phyto-component blocks the activity of EGFR.

According to Acharya et al. (2019) study, the three-dimensional models of estrogen receptor alpha (ERα), progesterone receptor (PR), EGFR and mechanistic target of rapamycin(mTOR) were attained from the PDB and docked with 23 furanocoumarin compounds which were obtained from PubChem using FlexX docking approach (27). The study has been reported that xanthotoxol had the best binding affinity with breast cancer protein when compared with other plant compounds such as bergapten, angelicin, psoralen and isoimperatorin.
Mutazah and his colleagues (2019) found that entadamide C and clinamide D plant compounds from Clinacanthus nutans could bind successfully to the caspase 3 (breast cancer cell protein) (28). Entadamide C and clinamide D have a docking score of -4.28 kcal/mol and -4.84 kcal/mol, respectively.

Suganya and her co-workers performed in silico molecular docking study on flavonoids against breast cancer cell protein (estrogen receptor). The human ER was obtained from PDB. They found that two plant compounds (chrysin and equol) which were retrieved from PubChem; docked successfully with human ER at a binding energy of -11.0189 kcal/mol and -10.8354 kcal/mol respectively (29).

In this study, p53, caspase-3 and rb1 protein models were docked successfully with garcinone E, triterpenoid and gallic acid respectively using the BSP-Slim server. Based on the lowest binding energy, Rb1- gallic acid complex had the strongest binding affinity and the most stable protein-phyto-component compelx among the three protein-plant compound complexes (p53-

garcinone E, caspase-3- triterpenoid and retinoblastoma 1-gallic acid). With the various types of genomic information on cancer cell lines, the target of drugs and pharmacological information, the discovery of a new drug combination may improve future cancer therapy (30).

CONCLUSION

In conclusion, p53, Caspase-3 and Rb1 (breast cancer cell proteins) were successfully docked with garcinone E, triterpenoid and gallic acid (plant compounds) respectively. The protein models had a strong interaction between phyto-components due to their lowest binding score. This in silico investigation of the target protein and phyto-component complex will support in creating a novel, effective and more potent drug.

ACKNOWLEDGEMENT

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

REFERENCES

1. GLOBOCAN. Estimated age standardized incidence rates of all cancers, Worldwide 2018. 2019. Available: https://gco.iarc.fr/today/home.
2. WHO. Breast Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012. 2015. Available: http://globocan.iarc.fr/old/FactSheets/cancers/breast-new.asp.
3. Nordqvist C. What you need to know about breast cancer. 2017. Available: https://www.medicalnewstoday.com/articles/37136.php.
4. Molyneux RJ, Lee ST, Gardner DR, Panter KE, James LF. Phytochemicals: the good, the bad and the ugly?. Phytochemistry. 2007; 68(22-24): 2973-2985.
5. Jung HA, Su BN, Keller WJ, Mehta RG, Kinghorn AD. Antioxidant xanthones from the pericarp of Garcinia mangostana (Mangosteen). Journal of agricultural and food chemistry. 2006; 54(6): 2077-2082.
6. Tandrasasmita OM, Lee JS, Baek SH, Tjandrawinata RR. Induction of cellular apoptosis in human breast cancer by DLBS1425, a Phaleria macrocarpa compound extract, via down-regulation of PI3-kinase/AKT pathway. Cancer biology & therapy. 2010; 10(8): 814-823.
7. Rabi T, Ramachandran C, Fonseca HB, Nair RP, Alamo A, Melnick SJ, Escalon E. Novel drug amooranin induces apoptosis through caspase activity in human breast carcinoma cell lines. Breast cancer research and treatment. 2003; 80(3): 321-330.
8. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Bordoli L. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic acids research. 2014; 42(W1): W252-W258.
9. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server The proteomics protocols handbook, Springer, 2005, pp. 571-607.
10. Costantini S, Colonna G, Facchiano AM. ESBRI: a web server for evaluating salt bridges in proteins. Bioinformation. 2008; 3(3): 137.
11. Roy S, Maheshwari N, Chauhan R, Sen NK, Sharma A. Structure prediction and functional characterization of secondary metabolite proteins of Ocimum. Bioinformation. 2011; 6(8): 315.
12. Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics. 1995; 11(6): 681-684.
13. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of applied crystallography. 1993; 26(2): 283-291.
14. Colovos C, Yeates TO. Verification of protein structures: patterns of non-bonded atomic interactions. Protein science. 1993; 2(9): 1511-1519.
15. Wallner B, Elofsson A. Can correct protein models be identified? Protein science. 2003; 12(5): 1073-1086.
16. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods in enzymology. 1993; 277: 396-404.
17. 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 modelling. 2011; 51(10): 2515-2527.
18. Lee HS, Zhang Y. BSP-SLIM: a blind low-resolution ligand-protein docking approach using predicted protein structures. Proteins. 2012; 80(1): 93-110. doi:10.1002/prot.23165.
19. Zhao WS, Sun MY, Sun LF, Liu Y, Yang Y, Huang LD, Fan YZ, Cheng XY, Cao P, Hu YM, Li L. A highly conserved salt bridge stabilizes the kinked conformation of β2, 3-sheet essential for channel function of P2X4 receptors. Journal of Biological Chemistry. 2016; 291(15): 7990-8003.
20. Zhang M, Hua L, Zhai W, Zhao Y. Novel computational approaches and applications in cancer research. BioMed research international. 2017; 2017.
21. Dimitrakopoulos CM, Beerenwinkel N. Computational approaches for the identification of cancer genes and pathways. Wiley Interdisciplinary Reviews: Systems Biology and Medicine. 2017;

9(1): e1364.
22. Rampogu S, Parameswaran S, Lemuel MR, Lee KW. Exploring the therapeutic ability of fenugreek against type 2 diabetes and breast cancer employing molecular docking and molecular dynamics simulations. Evidence-Based Complementary and Alternative Medicine. 2018; 2018.
23. Khoja KK, Shaf G, Hasan TN, Syed NA, Al-Khalifa AS, Al-Assaf AH, Alshatwi AA. Fenugreek, a naturally occurring edible spice, kills MCF-7 human breast cancer cells via an apoptotic pathway. Asian Pacific Journal of Cancer Prevention. 2011; 12(12): 3299-3304.
24. Amin A, Alkaabi A, Al-Falasi S, Daoud SA. Chemopreventive activities of Trigonella foenum graecum (Fenugreek) against breast cancer. Cell biology International. 2005; 29(8): 687-94.
25. Wang G, Zhu W. Molecular docking for drug discovery and development: a widely used approach but far from perfect. 2016.
26. Herlina T, Mardianingrum R, Gaffar S, Supratman U. Isoquinoline alkaloids from Erythrinapoeppigiana (Leguminosae) and cytotoxic activity against breast cancer cells line MCF-7 in silico. In Journal of Physics: Conference Series 2017 Feb (Vol. 812, No. 1, p. 012091). IOP Publishing.
27. Acharya R, Chacko S, Bose P, Lapenna A, Pattanayak SP. Structure based multitargeted molecular docking analysis of selected furanocoumarins against breast cancer. Scientific reports. 2019; 9(1):1-3.
28. Mutazah R, Hamid HA, Ramli AN, Aluwi MF, Yusoff MM. 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. 2020; 135: 110869.
29. Suganya J, Radha M, Naorem DL, Nishandhini M. In silico docking studies of selected flavonoids-natural healing agents against breast cancer. Asian Pacific Journal of Cancer Prevention. 2014; 15(19): 8155-9.
30. Jeon M, Kim S, Park S, Lee H, Kang J. In silico drug combination discovery for personalized cancer therapy. BMC systems biology. 2018; 12(2): 16.