

# In silico-based virtual drug screening and molecular docking analysis of phyto chemical derived compounds and FDA approved drugs against BRCA1 Receptor

## Abstract

**Background:** Molecular docking is used to identify natural photochemical and commercial drugs that are equivalent to treat BRCA1-based breast cancers. BRCA1 type breast cancers have high mortality and recurrence rate. BRCA1 gene which is also known as Breast Cancer Susceptibility gene 1, BRCA1 gene has high probability of mutation or duplication which further leads to breast cancer.

**Materials and Methods:** There are numerous photochemical and chemical drugs that are administered against breast cancer. These compounds are taken from the literature survey and they are meant to drug screening process against BRCA1 receptor (PDBID-4Y2G) using iGEMDOCKv2.1 and further the binding site of the Receptor is determined and the top 5 molecules that were attained from both phytochemicals and commercial drugs are proceeded to Active Site docking using Autodock vina software.

**Results:** Based on their Drug Screening scores and Binding affinity scores, one best phytochemical Epi Gallo Catechin Gallate-EGCG (Green tea extract), and one commercial drug, Doxorubicin Hydrochloride with best energies from docking are obtained.

**Keywords:** molecular docking, BRCA1, drug screening, EGCG, doxorubicin hydrochloride

Volume 8 Issue 2 - 2017

Rahul Ravichandran,<sup>1</sup> Raji Sundararajan<sup>2</sup><sup>1</sup>SASTRA University, India<sup>2</sup>Purdue University, USA

**Correspondence:** Rahul Ravichandran, School of Chemical and Biotechnology, SASTRA University, Tamil Nadu, India, Email rsundara@purdue.edu

**Received:** April 28, 2017 | **Published:** June 12, 2017

## Introduction

Computer-Aided-Drug-Discovery approaches have been extensively used in drug investigation to increase the efficiency of the drug discovery and development pipeline, subject to the purpose and systems of interest<sup>1</sup>. Docking is a software technique that fits a molecule into target binding sites. It is used for binding modes of protein with ligands and inhibitors<sup>2</sup>. It is used to predict the structure of the intermolecular complex formed between two or more molecules. Thus, it is used for drug discovery and matching molecules, with overall minimum energy. In this research, the equivalent natural phytochemical and commercial drugs are identified using molecular docking for treating BRCA1 breast cancer.

The BRCA1 gene is situated on the long arm of Chromosome 17 at 17q21 and BRCA1 also contains 24 coding exons range over 80 kb<sup>3</sup>. BRCA1 was first mapped and was consequently cloned in the years 1990 and 1994.<sup>4,5</sup> BRCA1 interacts with quite a lot of proteins that are involved in cell cycle progression, cellular pathways, ubiquitination, gene transcription regulation and DNA damage response.<sup>6</sup> BRCA1 is a pleiotropic DNA Damage Response protein (DDR) that functions in both DNA activation and checkpoint repair.<sup>7</sup> BRCA1 can bind to RAD51, p53, RAD50-MRE11-NBS1 (R-M-N) complex and leading to homologous recombination or Non-Homologous End-Joining (NHEJ) which is having excessive importance in DNA damage and its repair.<sup>8</sup> Deficiency in BRCA1 as a result causes abnormalities in the centrosome duplication, the G2/M checkpoint, the spindle checkpoint and S-phase checkpoint.<sup>9</sup> One of the most important features of BRCA1 associated breast cancer is that it comprises a complex unit of aneuploidy than tumours without Breast Cancer Susceptibility gene 1 mutations.<sup>10</sup> HCC1937 cell line derived from homozygous BRCA1

deficient breast tumour shows a high number of chromosomal gains and losses, heterozygosity loss at numerous loci identified to be involved in the development of breast cancer.<sup>11</sup>

Deficient cells in BRCA1 revealed centrosome amplification, abnormalities in all most important cell cycle checkpoints and impaired DNA damage repair.<sup>1,12</sup> Molecular docking has turned out to be an ever more important tool in the case of drug discovery. The main objective of molecular docking is to contribute a prediction of the ligand-receptor complex structure by means of computational approaches.<sup>13</sup> Docking can be used to execute virtual screening on enormous libraries of compounds, rank the results based on the scores and put forward structural hypothetical theories on how the ligands inhibit the target receptor, which is vital in lead optimization.<sup>14</sup> Factors affecting docking are: intramolecular (bond width, bond angle and dihedral angle) and intermolecular (electrostatic, dipolar, H-bonding, hydrophobicity and van der Waals) forces.

## Materials and methods

### Primary sequence retrieval

The amino acid Fasta sequence of BRCA1 gene has been obtained from Uniprot database with ID (UniProtKB - P38398 (BRCA1\_HUMAN)).<sup>15</sup> This fasta sequence was used for further investigation.

### Protein structure retrieval

The fasta sequence of BRCA1 was subjected to sequence similarity search using BLASTP<sup>16</sup> and the one with 100% sequence identity structure was retrieved from Brookhaven RCSB PDB database with (PDBID-4Y2G).<sup>17</sup> This PDB structure was used for further analysis.

## Ligand archive research

Literature survey lead to the identification of 18 natural compounds that are belonging to Flavonoids, Isoflavones, Catechine, polyphenolic and terpenophenolic families<sup>18</sup> and 26 Chemical drugs approved by the Food and Drug Administration (FDA) for Breast Cancer treatment.<sup>19</sup> The 3D structures of these natural compounds and chemical drugs are downloaded from the PubChem database<sup>20</sup> in SDF format and they are converted to the PDB format using Open Babel version 2.3.2 software.<sup>21</sup>

## Drug screening using iGEMDOCK

All the 17 phytochemicals and the 26 chemical compounds are subjected to Drug Screening process against BRCA1 receptor (PDBID-4Y2G) using iGEMDOCKv2.1<sup>22</sup> a standalone virtual screening software The Drug Screening module is set with the Default parameters such as Population Size = 200, Generations = 70, Number of solutions =3. Out of three solutions the one with the least energy values are considered.

## Active site prediction

Active Site region of the BRCA1 gene is predicted using CASTp webserver<sup>23</sup> which gives the residues that are participating in the active site of the macromolecule including the active site's area and volume. The active site residues are further mapped on to the structure to check exactly where the residues are present.

## Molecular docking studies

Molecular Docking is a technique used to analyze the position and the inhibition interaction between the protein and the small molecule. Docking was performed using Auto Dock Vina<sup>24</sup> which is standalone software. As a preprocessing step, The PDB format of Macromolecule and SDF format of small molecules are converted to Atudock's PDBQT format. From the Auto Dock Vina the macromolecule was added with a Gasteiger Partial charges and further Hydrogens were added. The docking of small molecule to the macromolecule was focused on the specified binding site. The total number of rotatable bonds of the ligands is calculated. The Grid was defined to the Binding site of the protein structure with the configurations of x/y/z coordinates was set to size\_x=62, size\_y=84, size\_z=60 centre of the grid box was set to center\_x=-1.141, center\_y=2.034, center\_z=-24.991 in X,Y,Z dimensions, in which the grid was covered to the binding site of the macromolecule. The Binding energies between receptor and the ligands are attained in terms of Kcal/mol.

## Protein-ligand interactions

The interactions like Hydrogen Bonding and other non-bonded terms between the top docked Phytochemicals, Commercial drug and the receptor are seen using Accelrys Discovery Studio Visualizer software<sup>25</sup>. Figure 1 shows the methodology.

## Results and discussion

### Protein structure visualization

The protein structure retrieved from RCSB PDB with (PDBID-4Y2G) was visualized using UCSF chimera which is molecular visualization software.<sup>26</sup> The bound ligands are already removed and the structure without bound ligand has been taken for the *in silico* analysis. Figure 2 shows BRCA1 susceptibility protein structure visualisation using Chimera 1.10.2 software. BRCA1 structure has double chain such as chain-A represented in blue colour, chain-B

represented in white with various atom representing various colours and its N-Terminal indicating starting point of the structure and C-Terminal indicating the end point of the structure.

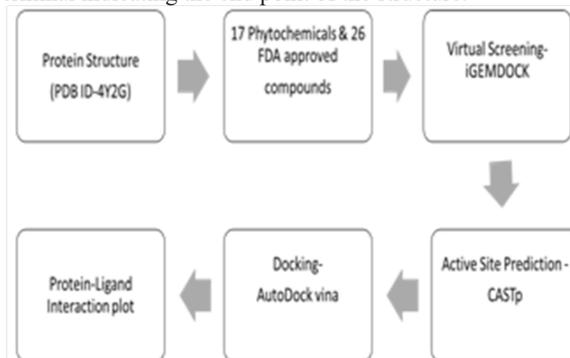


Figure 1 Methodology of Docking Study.

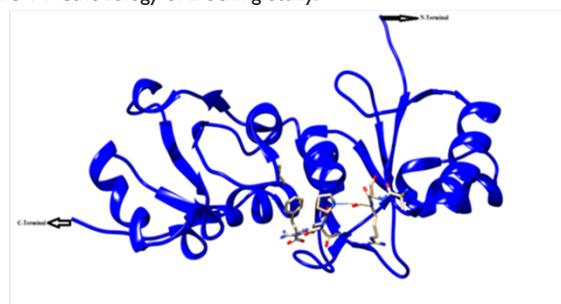


Figure 2 BRCA1 susceptibility protein structure visualisation using Chimera 1.10.2 software.

### Residues involved in active sites

The active site of the macromolecule BRCA1 has been determined using CASTp webserver. The determination of residues that are involved in the active sites are taken from Pocket number27 with an Area of 443Angstrom<sup>2</sup>, Volume of 701.4 Angstrom<sup>3</sup> and the active site residues are shaded in Green colour as shown in Figure 3.

Chain	Residue	Color	
Chain A	1646- V N K R M S M I V S G L T P E E N L V Y K F A R K H H I T L N L I T E E T T H V V M R D A E F	Green	
	1696- V C I N T L K Y F L G I A G G K W V V S Y F W V T Q S I K E R K M L N E H D F E V R O D V V N G R N	Green	
	1746- H Q G P K R A R E S Q D R K I F R O L E I C C Y G P F T N M P T D Q L E W M V Q L C G A S V V K E L	Green	
	1796- S S F T L G T G V H P I V V V Q P D A W T E D N G F A I G O M C E A P V V T R E W L V S V A L Y	Green	
	1846- Q C Q E L D T Y L I P Q I P	Green	
	Chain B	403- Y S R - P T F	Green

Figure 3 Shaded Regions representing active site residues using CASTp. Various residues from both the chains-A,B are involved. These residues are further mapped onto the structure to know where exactly the active site is present in the molecule.

### Drug screening using iGEMDOCKv2.1

As the phytochemicals and commercial drugs are subjected to drug screening, the top five small molecules of both Phytochemicals and commercial drugs were short-listed based upon their scores and they are subjected to further molecular docking studies. The features of the top five drug Screened compounds are summarized in Table 2. The virtual drug screening scores for these top five phytochemicals and commercial drugs were obtained from iGEMDOCK. The highly negative energy score represents the stronger binding affinity between the receptor and the small molecule, whereas positive score represents to poor or no effective binding. Curcumin (turmeric extract), Lapatinib Ditosylate (commercial drug) attained the highest score.

**Table 1** The 18 compounds of phyto chemicals and the 26 compounds from commercial drug list that were studied

Phytochemicals	Pubchem ID	Commercial Drugs	Pubchem ID
Apigenin	Pubchem CID-5280443	5-Fluoro Uracil	Pubchem CID-3385
Biochainin A	Pubchem CID-5280373	Abitrexate	Pubchem CID-126941
Carotenoids	Pubchem CID-16061280	Abiraxane	Pubchem CID-36314
CBD	Pubchem CID-26346	Anastrozole	Pubchem CID-2187
Curcumin	Pubchem CID-969516	Aredia	Pubchem CID-73351
Cyanidin	Pubchem CID- 68247	Aromasin	Pubchem CID-60918
Daidzein	Pubchem CID-5281708	Capecitabine	Pubchem CID-60953
EGCG	Pubchem CID- 65064	Clafen	Pubchem CID-2907
Gabridin	Pubchem CID-124052	Dofetaxel	Pubchem CID-148124
Genistein	Pubchem CID-5281377	Doxorubicin Hydrochloride	Pubchem CID-443939
Hesperitin	Pubchem CID-72281	Ellence	Pubchem CID-41867
Ipriflavone	Pubchem CID-3747	Eribulin Mesylate	Pubchem CID-17755248
Kaempferol	Pubchem CID-5280863	Everolimus	Pubchem CID-6442177
Lignans	Pubchem CID-261166	Fareston	Pubchem CID-3005573
Luteolin	Pubchem CID-5280445	Faslodex	Pubchem CID-104741
Naringenin	Pubchem CID-439246	Femara	Pubchem CID-3902
Quercetin	Pubchem CID-5280343	Gemicitabine Hydrochloride	Pubchem CID-60749
		Goseralin Acetate	Pubchem CID-16052011
		Ibrance	Pubchem CID-5330286
		Ixabepilone	Pubchem CID-6445540
		Lapatinib Ditosylate	Pubchem CID-9941095
Resveratrol	Pubchem CID-445154	Megestrol Acetate	Pubchem CID-11683
		Nolvadex	Pubchem CID-2733526
		Thiotepa	Pubchem CID-5453
		Trastuzumab	Pubchem CID-3001322
		Velban	Pubchem CID-46783244

**Table 2** Summary of Top-5 Phytochemicals and Commercial Drugs Virtual Drug Screening Scores obtained from iGEMDOCK

Name of receptor	Natural compound	iGEMDOCK energy (phytochemicals)	Commercial drug	iGEMDOCK energy (commercial drugs)
BRCA1 (4Y2G)	Curcumin	-101.5	Lapatinib Ditosylate	-115.6
BRCA1 (4Y2G)	EGCG	-99.9	Doxorubicin Hydrochloride	-113.1
BRCA1 (4Y2G)	Apigenin	-95	Ellence	-109.2
BRCA1 (4Y2G)	Quercetin	-93.9	Abitrexate	-105.5
BRCA1 (4Y2G)	Kaempferol	-93.6	Trastuzumab	-105.1

**Table 3** Summary of Top-5 Phytochemicals and Commercial Drugs Binding Afiinity scores obtained from Auto Dockvina

Name of the receptor	natural compound	Autodock energy Kcal/mol	Commercial compound	Autodock energy Kcal/mol
BRCA1 (4Y2G)	Curcumin	-6.1	Lapatinib Ditosylate	-7.5
BRCA1 (4Y2G)	EGCG	-7.9	Doxorubicin Hydrochloride	-8.1
BRCA1 (4Y2G)	Apigenin	-7.1	Ellence	-8
BRCA1 (4Y2G)	Quercetin	-6.6	Abitrexate	-6.1
BRCA1 (4Y2G)	Kaempferol	-6.7	Trastuzumab	-7.6

## Molecular docking

Molecular Docking was done taking these top five small molecules from both phytochemicals and commercial drugs and further the molecules were docked to the active site of the protein. Various binding affinities are obtained for all the ligands in terms of (Kcal/mol). The various features and the binding affinities of the docked complexes of are summarized in Table 3, which shows the summary of Top five Phytochemicals and Commercial Drugs binding affinity scores obtained from AutoDockvina interms of kcal/mol. The highly negative energy score represents the stronger binding affinity between the receptor whereas EGCG (green tea extract) and Doxorubicin

Hydrochloride attained the highest binding affinities from the top five screened compounds. Table 4 shows the chemical structure of the top 2 Phytochemicals and Commercial Drugs Binding Afiinity scores obtained from iGEMDOCK and AutoDockvina.

### BRCA1-EGCG complex

Figure 4 shows the docked complex of BRCA1 receptor. Dark blue represents chain-A and white represents chain-B whereas light blue represents EGCG compound in Sticks model highlighted in the box. This docked confirmation of EGCG with BRCA1 has obtained the highest binding affinity of -7.9 Kcal/mol using Autodock vina.

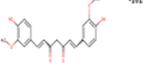
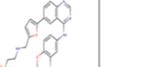
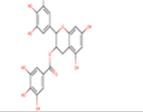
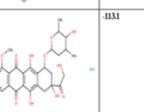
Phytochemical	Structure	Igemdock Score	Autodock Score	Commercial Drug	Structure	Igemdock score	Autodock Score
Curcumin		-10.5	-6.1	Lipistatib Dipeptate		115.6	-7.5
EGCG		-9.9	-7.9	Doxorubicin Hydrochloride		113.1	-8.1

Figure 4 Summary of Top-2 Phytochemicals and Commercial Drugs Binding Affinity scores obtained from iGEMDOCK and AutoDockvina.

### Ligand plot for residual wise interaction

Figure 5 shows BRCA1 Protein and EGCG Receptor interaction plot visualised using Discovery studio visualizer. It represents various residues corresponding to different types of interactions of BRCA1 with EGCG. A conventional hydrogen bond is seen between Leucine:1657 (A-chain) of BRCA1 and the EGCG molecule and other non-bonded interacting residues can be seen.

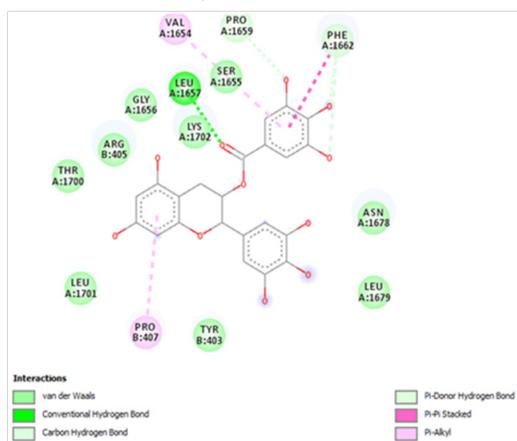


Figure 5 BRCA1 Protein and EGCG Receptor interaction plot.

### 4Y2G doxorubicin hydrochloride complex

Figure 6 shows the docked complex of BRCA1 receptor. Here, dark blue represents chain-A, white represents chain-B and light blue represents Doxorubicin Hydrochloride in Sticks model highlighted in the box. This docked confirmation of Doxorubicin Hydrochloride with BRCA1 has been obtained with the highest binding affinity of -8.1Kcal/mol using Autodock vina.

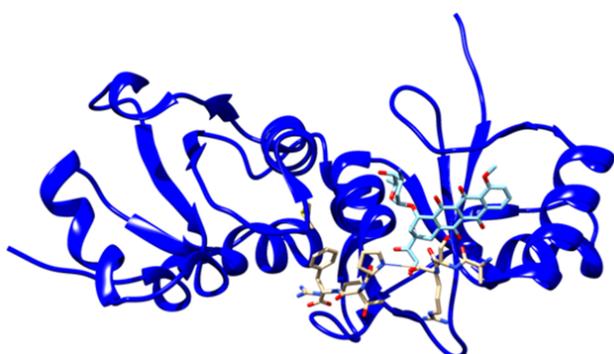


Figure 6 Docked complex of BRCA1 receptor with EGCG phytochemical.

### Ligand plot for residual wise interaction

Figure 7 shows BRCA1 Protein and Doxorubicin Hydrochloride drug interaction plot visualised using Discovery studio visualizer. It represents various residues corresponding to different types of interactions of BRCA1 with Doxorubicin Hydrochloride. A conventional hydrogen bond is seen between Proline 407 (B-chain) of BRCA1 and the Doxorubicin Hydrochloride molecule and other non-bonded interacting residues can be seen.

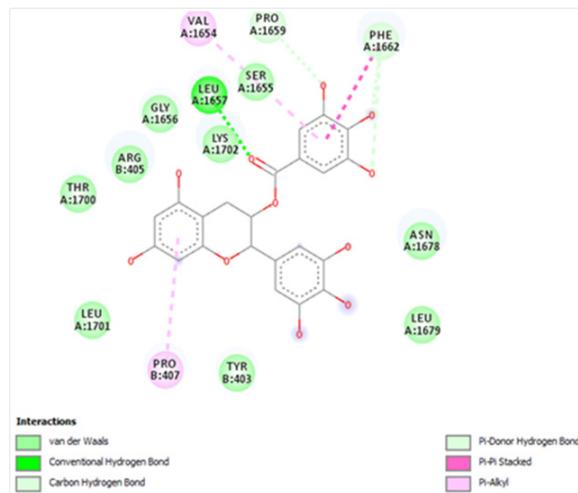


Figure 7 BRCA1 Protein and Doxorubicin Hydrochloride drug interaction plot.

### Conclusion

It takes years and millions of dollars to discover a drug and bring to real world use. To reduce the cost and time, it is of practical interest to study the drug molecules and their characteristics using commercial/ industrial software. Drugs interact with receptors in a highly specific and complimentary manner. Towards this, in this study, we compared the naturally occurring phytochemical compounds and commercial drugs which are similar in characteristics for BRCA1 type breast cancers. Breast cancer was considered as it is the most common type of cancer in women worldwide. Each year over 1 million new cases occur and over 500,000 die. For the patients on chemo therapy, side effects and drug cost add to the physical suffering due to the cancer. Hence, it is critical that we identify natural compounds which have least or no side effects and are also affordable.

The current proposed hypothesis deals with the Virtual Screening & molecular docking procedures for different small molecules against BRCA1 protein structure for checking its binding affinities and inhibitory activity of small molecules against the receptor. The screened compounds such as EGCG (Phytochemical) & Doxorubicin Hydrochloride (commercial drug) has shown the best binding affinity towards the BRCA1 receptor and EGCG's Energy is -7.9kcal/mol whereas Doxorubicin Hydrochloride's energy is -8.1kcal/mol as both the compounds have very slight difference in their binding affinity scores can be hypothesized that the Phytochemicals can also be considered as an inhibitor against Breast cancer susceptibility gene 1 for breast cancer treatment apart from these commercial drugs. Further experimental studies are needed to understand the mechanism of action of phytochemicals against Breast cancer.

### Acknowledgements

None.

## Conflicts of interest

The authors declare that there are no conflicts of interest.

## Funding

None.

## References

1. Zhang S. Computer-aided drug discovery and development. *Methods Mol Biol.* 2011;716:23–38.
2. Meng XY, Zhang HX, Mezei M, et al. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr Comput Aided Drug Des.* 2011;7(2):146–157.
3. Miki Y, Swensen J, Eidens SD, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science.* 1994;266(5182):66–71.
4. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science.* 1990;250(4988):1684–1689.
5. Miki T, Bottaro DP, Fleming TP, et al. Determination of ligand-binding specificity by alternative splicing: two distinct growth factor receptors encoded by a single gene. *Proc Natl Acad Sci USA.* 1992;89(1):246–250.
6. Rosen EM, Fan S, Ma X. BRCA1 regulation of transcription. *Cancer Lett.* 2006;236(2):175–185.
7. Roy R, Chun J, Powell SN. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nature Reviews Cancer.* 2011;12(1):68–78.
8. Teng LS, Zheng Y, Wang HH. BRCA1/2 associated hereditary breast cancer. *J Zhejiang Univ Sci B.* 2008;9(2):85–89.
9. Deng CX. BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic Acids Res.* 2006;34(5):1416–1426.
10. Tirkkonen M, Johannsson O, Agnarsson BA, et al. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. *Cancer Res.* 1997;57(7):1222–1227.
11. Tomlinson GE, Chen TT, Stastny VA, et al. Characterization of a breast cancer cell line derived from a germ-line BRCA1 mutation carrier. *Cancer Res.* 1998;58(15):3237–3242.
12. Deng CX. Tumor formation in Brca1 conditional mutant mice. *Environ Mol Mutagen.* 2002;39(2–3):171–177.
13. Deng CX, Scott F. Role of the tumor suppressor gene Brca1 in genetic stability and mammary gland tumor formation. *Oncogene.* 2000;19(18):1059–1064.
14. Morris GM, Wilby ML. Molecular Docking. *Methods Mol Biol.* 2008;443:365–382.
15. UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* 2015;43:D204–D212.
16. Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–410.
17. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank Nucleic Acids Research. 20a0;28(1):235–242.
18. Siddiqui JA, Singh A, Chagtoo M, et al. Phytochemicals for Breast Cancer Therapy: Current Status and Future Implications. *Curr Cancer Drug Targets.* 2016;15(2):116135.
19. <https://www.cancer.gov/about-cancer/treatment/drugs/breast#2>.
20. Kim S, Thiessen PA, Bolton EE, et al. PubChem Substance and Compound databases. *Nucleic Acids Res.* 2016;44(D1):D1202–D1213.
21. Boyle ONM, Banck M, James CA, et al. Open Babel: An open chemical toolbox. *J Cheminform.* 2011;3:33.
22. Hsu KC, Chen YF, Lin SR, et al. iGEMDOCK: A Graphical Environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. *BMC Bioinformatics.* 2011;12(Suppl 1):S33.
23. Dundas J, Ouyang Z, Tseng J, et al. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* 2006;34:W116–W118.
24. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem.* 2010;455–461.
25. Accelrys. Discovery Studio Visualizer Software. Version 16.1.0.15350. 2016.
26. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem.* 2014;25(13):1605–1612.