QSAR Pharmacophore-based Virtual Screening, CoMFA and CoMSIA Modeling and Molecular Docking towards Identifying Lead Compounds for Breast Cancer Protease Inhibitors

Lan Huang[1], Xuan R. Zhang[1*], Pei H. Luo[1], Lun Yuan[2], Xang Z. Zhou[2], X. Gao[1] and Ling S. Li[2]

[1]*School of Chemical Engineering, Sichuan University, Chengdu 610065, China.*
[2]*Department of Pharmaceutical Engineering, Sichuan University, Chengdu 610041, China.*

Authors' contributions

This work was carried out in collaboration between all authors. Authors LH, XRZ and XG designed the study, wrote protocol and wrote the first draft of the manuscript. Authors LH, PHL and LY performed the computational modeling and statistical analysis. Authors XZZ and LSL managed the analysis of the study and the literature searches. All authors read and approved the final manuscript.

Original Research Article

ABSTRACT

Aim: This study used QSAR Pharmacophore-based virtual screening and molecular docking to identify lead compounds and determine structural requirements for breast cancer inhibitor development. CoMFA and CoMSIA modeling was employed to design more potential inhibitors.
Materials and Methods: 3D-QSAR pharmacophore models were developed using HypoGen Module and validated by Fischer's model and decoy test. The best pharmacophore model was employed to screen ZINC chemical library to obtain reasonable hits. Following ADMET filtering, 18 hits were subjected to further filter through docking. CoMFA and CoMSIA models were built by partial least squares on phenylindole-3-carbaldehydes derivatives.
Results: 19 random runs from Fischer's validation and decoy test which led to an enrichment

_____

*Corresponding author: E-mail: icddpharmscu@hotmail.com;

factor of 48.23 and Guner-Henry factor of 0.774 show that the identified pharmacophore model is highly predictive. Top three hits (IC_{50}=0.01~0.05 μM, fitness =52~62) were identified as potential inhibitory candidates from virtual screening and docking, and three new lead compounds were designed with predicted inhibiting potencies by pIC_{50} value of 8.55 from CoMFA and CoMSIA modeling and fitness value of ~59 from docking.

Conclusion: Validation results and decoy test indicate that the developed pharmacophore model is highly predictive. Residue Sep6 and Cys 5 were observed as important active sites for ligand-protein binding. Top three hits were identified as more potential inhibitors, and the designed compounds show more inhibiting potencies. The QSAR and docking results obtained from this work should be useful in determining structural requirements for inhibitor development as well as in designing more potential inhibitors.

1. INTRODUCTION

Breast cancers are threatening to women's health and life and the population of women developing breast cancer is gradually increasing every year in the world. Breast cancers arise from the epithelial cells, which display tumor heterogeneity. Usually, breast cancer tumor development may depend on HER2, progesterone receptors and estrogen receptors (ER). Especially, ER tumors dominate nearly two thirds of breast cancer women, predominantly infecting postmenopausal women, while minority of women developing breast cancers has triple-negative breast cancer tumors. Therefore hormonal and endocrine therapies are primary strategies against breast cancer ER tumors [1,2]. Estrogen receptors are subdivided into two types: alpha hERR and beta hER (ERα and ERβ). ERα is found to be more expressed in breast cancer tumors. Thus ERα is identified as a key target for developing new anti-breast cancer drugs with endocrine therapy [3,4].

Tamoxifen-selective ER modulator and aromatase inhibitors including exemestane - steroidal and letrozole-nonsteroidal compounds are currently recommended for clinical treatment of ER breast cancer tumors. Tamoxifen is often used for treatment of ER-positive or metastatic cancers. However, it has been observed that tamoxifen may lead to endometrial enlargement or cancer and also can cause thrombosis as it may induce hormone activity. As compared to tamoxifen, aromatase drugs are recommended for the first choice for the treatment of ER breast cancers since they indicated better benefits for patient survival and greatly decrease recurrence of breast cancers. However, aromatase drugs may cause muscular and joint syndrome, such as lower back pain and feet pain as well as neuropathy and myalgia symptoms [3-5]. These side effects lead to discontinue aromatase inhibitors and tamoxifen, and patients are more probably to stop treatment. Therefore, researchers are currently more likely to develop novel scaffolds or inhibitory candidates *in vivo* or *in vitro* to avoid these side effects and clinically adverse treatment. Foudah et al. [5] investigated anti-breast cancer proliferative bioactivities of natural products-terpenoids with new sipholenol scaffolds and used 3D QSAR pharmacophore modeling to correlate structures and anti-breast cancer bioactivities of these compounds. Busnena et al. [6] assessed the inhibitory activity and anti-proliferative of olive secoiridoids and bioisostere analogues against highly metastatic human breast cancer cell line. Others such as thiosemicarbazone and alkylindole derivatives with nonsteroidal analogs were reported [7,8]. However, all these inhibitory activities were submicromolar. Several different factors interposing in antiproliferative activity against breast tumor cells may not be observed *in vivo* or *in vitro*. Thus 3D QSAR pharmacophore modeling and molecular docking may not only discover lead compounds with novel scaffolds but also provide important structural information for future structure-based drug design. Furthermore, comparative molecular field analysis (CoMFA) and comparative molecular similarity analysis (CoMSIA) may also be useful in designing more potential inhibitors [9-14].

2. MATERIALS AND METHODS

2.1 3D-QSAR Pharmacophore Modeling and Virtual Screening

3D-QSAR pharmacophore models were developed using HypoGen mode in program Discovery Studio 2.5. 46 known breast cancer

inhibitors with IC_{50} values were collected from previous publications, which were subdivided into training data set (TDS) (Fig. 1) and testing data set [5,15-18]. All molecular conformations were produced using "Best conformation method", and Uncertain value was given as 2.0 with other values kept as defaults. The test data set, Fischer's randomization and decoy test were used to validate pharmacophore models [9-11]. In the decoy test, the database contains 1224 compounds, including 19 known inhibitors. The well-validated pharmacophore model was employed to identify ligands and aligns ligands to the query throughout ZINC databases. All hits were further screened using ADMET descriptors. A penalty function which introduces each parameter to fall into suitable region was used to evaluate completion of ADMET.

2.2 Molecular Docking

Docking ligands into active sites of breast cancer proteins was performed by Gold 4.1.2 program. Breast cancer proteases (PDB code: 4IGK and 4IFI) were prepared for molecular docking as given previously [9,10]. Protease atoms were comprised within the 3D spatial region of 7 Å surrounding the reference compound. Active site of docking was confined to the spatial region of 8.0 Å using the coordinates of crystal structure of reference ligand. For studied systems, the bins of receptor and ligand was fixed to 0.28-0.55 Å, outline bins was fixed to 0.22-0.44 Å, and the distance tolerant range for aligning a ligand and a receptor sphere was fixed to 1.3-1.55 Å. The structures of breast cancer protein outside of active sites remained unchanged, while active sites were kept partially flexible upon docking. The binding competence of hit compounds was evaluated by a docking-score method [11,19-21].

2.3 CoMFA and CoMSIA Modeling

CoMFA and CoMSIA models were built on 20 phenylindole-3-carbaldehydes derivatives by SYBYL-X 1.3 program [8]. Reasonable molecular conformations were obtained from docking into crystal structure of breast cancer protein, which were used in molecular alignment. Biological activities of compounds were determined using partial least squares (PLS) analysis. Inhibitory bioactivity values and CoMFA descriptors were employed as dependent and independent variables for PLS analysis, respectively. 30 kcal/mol cut-off energy was maintained to remove unfavorable electrostatic and steric energies and a filtering value of 1.8 kcal/mol was

kept to improve modeling signal [22,23]. In CoMSIA modeling, five similarity fields, i.e. hydrogen bond acceptor, hydrogen bond donor, hydrophobic, steric and electrostatic were calculated. The most important contributions to ligand-protein binding were characterized by these fields. Probe atom with charge +1 around radius 1 Å, hydrogen bond accepting +1, hydrogen bond donating +1 and hydrophobicity +1 was used to compute the corresponding fields. A test set of five compounds was used to perform external validation of CoMFA and CoMSIA models [23].

3. RESULTS AND DISCUSSION

3.1 3D-QSAR Pharmacophore Model and Virtual Screening

The identified pharmacophore model from ten hypotheses is shown in Fig. 2, which includes four elements, i.e. Hydrophobic feature (HP), aromatic ring (AR), hydrogen bond acceptor (HBA) and hydrogen bond acceptor-lipid (HBAL). Calculated statistical results indicate that cost difference between null and total cost is 168.8984 (>60), the difference between null cost and the fixed cost is 177.2944 (>80), and statistic errors from bioactivity prediction range from -2 to +2 (Table 1) [11,12]. All 19 random runs from Fischer's validation give lower correlation coefficients and higher cost values compared with the identified model. Calculations on the decoy test indicate that the enrichment factor is 48.23 and Guner-Henry factor is 0.774 [10]. Fig. 2 shows that higher bioactive compound **13** was mapped well with all four elements, while lower bioactive compound **5** or **2** was mapped only with three or two elements. All these results confirm that the identified model is highly predictive [9-12].

The pharmacophore model was employed to discover novel and inhibitory lead compounds by matching pharmacophore features and spatial requirements from ZINC library. 48 hits with estimated IC_{50} values in the range of 0.02~2.80 µM were obtained through virtual screening of about 2.0 million compounds, which were subsequently subjected to filter through ADMET descriptors. Finally, 18 of those hits met key parameters for lead compounds, such as BBB penetration, toxicity, distribution coefficient, aqueous solubility, and percent human oral absorption. The final hits were further visually analyzed based on functional and geometrical space necessities.

Table 1. Estimated from the pharmacophore model and experimental bioactivity (IC$_{50}$ in μM) for known breast cancer inhibitors in the TrDS

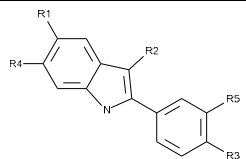| Comp. no	Bioactive	Predicted	Errors	Fitted value
1	130	93.0933	-1.3964	4.2588
2	120	65.2257	-1.8397	4.4133
3	11	14.2543	1.3296	5.0738
4	8.23	23.837	1.5037	4.1349
5	8.9	2.3749	-3.7474	5.8521
6	80	81.8047	1.0225	4.315
7	68	46.6102	-1.4589	4.5593
8	12.2	7.1431	-1.7079	5.3739
9	5.9	5.5676	-1.0597	5.4821
10	24.5	50.2526	2.0511	4.5266
11	0.26	0.734	2.8231	6.3621
12	0.28	0.7299	2.6069	6.3645
13	0.06	0.1059	1.7658	7.2026
14	0.552	0.5004	-1.1029	6.5283
15	1.015	0.4929	-2.0588	6.5349
16	1.99	0.7228	-2.7528	6.3687
17	0.0055	0.006626	1.2047	8.4065
18	0.0074	0.008434	1.1398	8.3017
19	40	88.1022	2.2025	4.2828

3.2 Docking Modeling

18 hits were docked into the cavities of active sites of breast cancer proteins (PDB code: 4IGK and 4IFI). Each compound with different conformations was matched to the crystal structures of the two proteins to achieve the satisfied docking results. Fitness values for 18 hits range from 41 to 62. Finally, based on higher fitness values and lower IC$_{50}$ values predicted from pharmacophore screening, top three hits with novel scaffolds: ZINC67688504 (IC$_{50}$=0.01962 μM, fitness =52.23), ZINC32752710 (0.1277, 61.13), and ZINC37577736 (0.050478, 55.41), were identified as potential lead compounds or inhibitory candidates for breast cancer as shown in Fig. 3.

The compound ZINC32752710 is involved in four hydrogen bonds with Ser-P6, Ile1, Lys1 and Cys5 of breast cancer protein as shown in Fig. 4. Carboxyl group on the main chain interacts with amino group of Lys1 and methylene of Ser-P6 to form two hydrogen bonds with 2.582 and 2.247Å distance respectively. Oxygen atom of oxadiazole interacts with amino group of Cys5 to engage in hydrogen bond, and oxygen atom of C$_5$O$_2$ ring also interacts with NH group of Ile1 to form the other HB. $\pi - \pi$ stacking is evident between oxadiazole ring and the methyl group of Sep6. Furthermore, methyl groups make hydrophobic interactions with Asn1, Ile1, and Lys1, which stabilizes the interaction of the

compound and the protein. ZINC67688504 was predicted to make three hydrogen bonds with Lys1-Ys17, and Sep6 of breast cancer protein as shown in Fig. 4. Oxygen atoms of thiophene ring accept hydrogen atom of amino group of Sep6-Ep6 to form a hydrogen bond. Carboxyl group on the main chain also accepts hydrogen atom of amino group of Lys1 and hydrogen atom of Sep6 to form two hydrogen bonds respectively. The compound also makes short contacts with Ser-P6, Glu-1698, Gly 1656 and Leu 1657 to form hydrophobic interaction and Van der waals bond, which further stabilizes its binding. Interestingly, residue Sep6 and Cys5 were observed as important active sites for ligand-protein binding.

3.3 CoMFA and CoMSIA Modeling

We note that the identified pharmacophore model above may predict lead compounds or inhibitory candidates with novel scaffolds, but sometimes may not predict more potential inhibitory candidates or lead compounds since the features of the pharmacophore model were determined on the basis of experimentally known inhibitors and their structures and thus other important features or factors may not be included in the pharmacophore model. Considering the limitation of 3D-QSAR model, we employed CoMFA and CoMSIA models built on phenylindole-3-carbaldehydes derivatives [8] (Table 2) to predict more potential lead compounds as a supplement.

Table 2. Structures and activities of breast cancer inhibitors (MCF-7) used for CoMFA and CoMSIA modeling

No.	R1	R2	R3	R4	R5	Observed activity pIC_{50}	CoMFA	CoMSIA
1	Ome	COH	OMe	H	H	6.58	6.77	6.64
2	H	COH	OMe	OMe	H	7.45	7.17	7.18
3	H	COH	OMe	F	H	7.22	7.03	7.12
4	F	COH	OMe	H	H	6.26	6.81	6.56
5	Me	COH	OMe	H	H	7.06	6.98	7.14
6	n-Bu	COH	OMe	H	H	8.17	7.84	7.67
7	t-Bu	COH	OMe	H	H	6.55	6.50	6.63
8	n-Pent	COH	OMe	H	H	8.25	8.34	8.27
9	n-Hex	COH	OMe	H	H	8.13	8.35	8.46
10	H	COH	H	OMe	OMe	5.99	6.03	5.91
11	H	COH	OMe	OMe	OH	6.09	6.13	6.55
12	H	COH	Me	OMe	H	7.51	7.66	7.61
13	H	COH	Me	Cl	H	8.11	7.58	8.13
14	n-Bu	COH	Et	H	H	7.57	7.61	7.51
15	n-Bu	COH	F	H	H	6.45	6.99	6.59
16	n-Hex	COH	CF3	H	H	7.36	7.15	7.40
17	H	CH=NMe	OMe	OMe	H	7.46	7.50	7.58
18	n-Pent	CH=NMe	OMe	H	H	8.22	8.34	8.47
19	n-Bu	CH=NOH	OMe	H	H	7.39	7.65	7.63
20	n-Bu	CH=NOH	CF3	H	H	6.30	6.21	6.41
21*	Me	COH	OMe	Cl	H	7.59	7.78	-
22*	H	COH	OMe	OMe	OMe	6.56	5.88	-
23*	n-Bu	CH=NOMe	OMe	H	H	8.21	7.84	-
24*	n-Bu	COH	Me	H	H	7.46	7.40	-
25*	Et	COH	n-Bu	H	H	6.52	7.39	-

*test set for validation of CoMFA and CoMSIA models

Since reliable CoMFA and CoMSIA models depend on two important factors: conformation generation and compound alignment, in this case flexible docking was employed to achieve compound alignment. Statistic results show that cross-validation coefficient q^2 is greater than 0.419 for both CoMFA and CoMSIA models (Tables 2, 3). The calculation on test cross-validation produced the cross validated coefficient q^2 of ~ 0.65 for both CoMFA and CoMSIA models. To achieve a best validation, a test set of five compounds with similar activities and structures as those in TDS was used to perform external validation of CoMFA and CoMSIA models using the criteria of validation proposed by Golbraikh [24]. The identified CoMFA and CoMSIA models indicate the best R^2 (>0.78) and q^2 (>0.59), satisfying the conditions of validation [24,25]. The predicted active values were given in Table 2. These analyses confirm that the built CoMFA and CoMSIA models are stable and reliable.

CoMFA contour maps of compound 6 (Table 2) are shown in Fig. 5 (a) and (b). Methyl groups at both side ends occupy sterically favor green region, locating at the corresponding AR region in the final pharmacophore model; Phenyl group occupies the yellow regions, lying towards corresponding H space. Methyl group connecting to oxygen atom is also situated at positive charge favor blue region, lying towards HBD space; indole ring occupies electronegative favor red region, lying towards HBA space. n-Bu and OCH_3 groups occupy sterically favored green regions at both side sites leading to potent bioactivity of $pIC_{50}=8.17$. This trend is also observed for other phenylindole-3-carbaldehydes derivatives at the positions by large groups (compounds 6, 8, 9, 18, 23 with n-pent and n-hex

substituents). However, the two green contours are unoccupied by compounds 10 (pIC$_{50}$=5.99) and 15 (pIC$_{50}$=6.45) due to lack of steric bulk group, reflecting the importance of steric bulk substituents in these regions. Occupancy of a yellow steric unfavorable areas around phenyl ring would lead to lower activity (for example,

OMe substituent in compounds 10 and 22 (pIC$_{50}$=6.56)). Localization of n-Bu and OCH$_3$ groups within hydrophobic favored region in white in CoMSIA field (Fig. 5e)) and steric favored region in green in CoMFA field sugeests that hydrophobic and steric bulk groups may play an important role at the positions [22].

Table 3. Statistic and cross-validation parameters for the best CoMFA and CoMSIA models

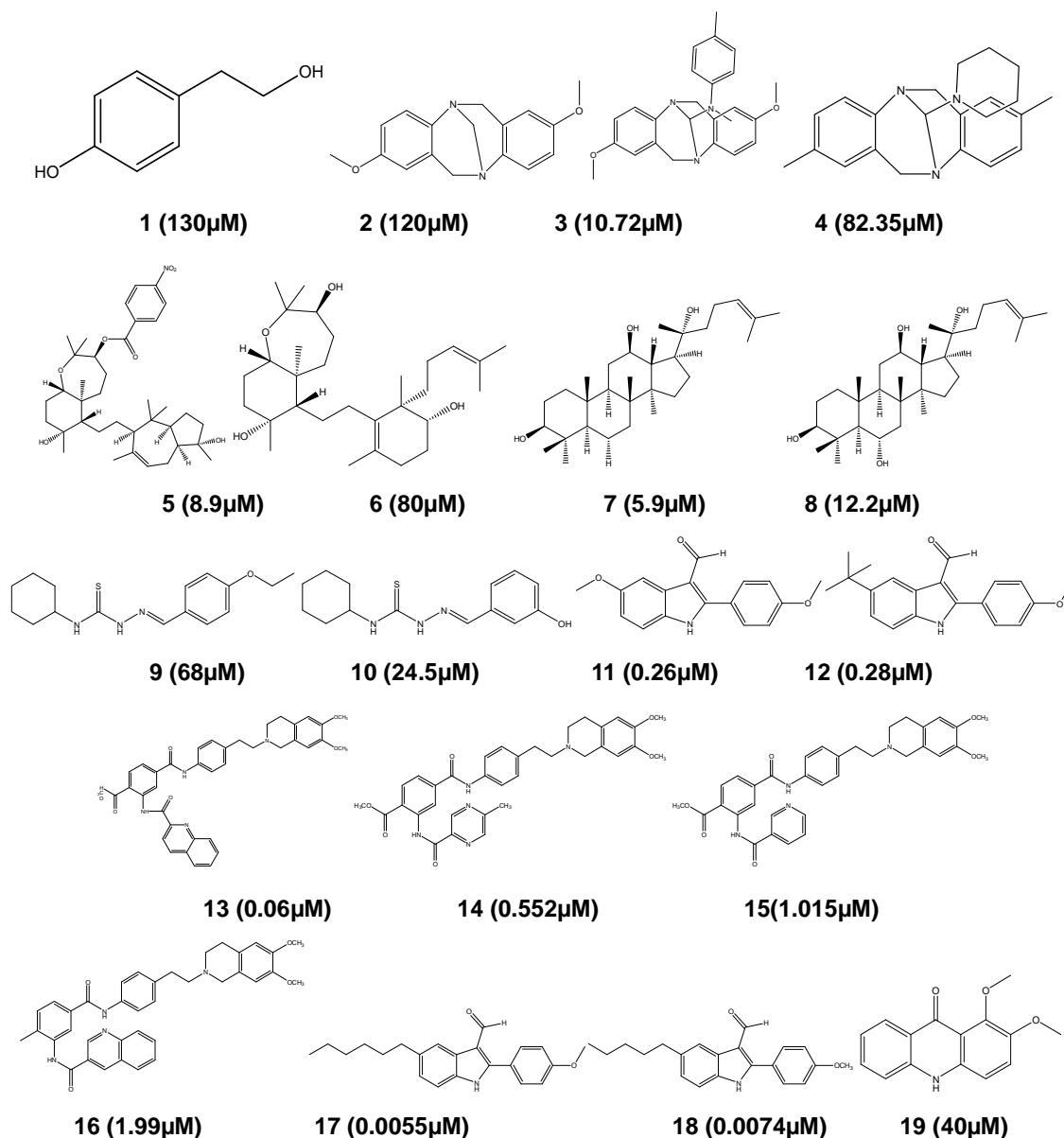| Model	N	q^2	r^2	R^2	SEE	F
CoMFA	4	0.49	0.88	-	0.212	-
CoMSIA	6	0.55	0.85	0.975	0.147	82.97

Fig. 1. 2D structures of 19 breast cancer inhibitors in the training data set for pharmacophore modeling. The values in parentheses represent IC$_{50}$ (µM) obtained experimentally
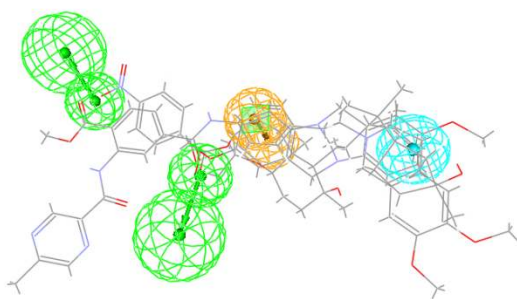
Fig. 2. The identified pharmacophore model. HP is shown by cyan sphere, AR is shown by orange spheres, HBA and HBAL are shown by green spheres. The space distances between the features are shown in Å. The inhibitors 2 (120 µM), 5 (8.9 µM) and 15 (1.015 µM) in TDS are mapped with the model

Carboxyl group (electron withdrawing) on phenyl ring, occupying favored positive atomic charge region in blue (Fig. 5b), makes the moiety electronic density more positive and thus leads to compounds 8 ($pIC_{50}=8.25$) and 23 ($pIC_{50}=8.21$) with more potential inhibitory activity. The least active compound 10 contains O-CH_3 substituent at this position which makes the moiety electronic density more negative. Negative charge favorable area in red is mainly occupied by negative partial charge atoms such as oxygen and nitrogen around indole ring, indicating more potential inhibitory activity in compounds 6, 8, 18, 23. In contrast to these compounds, compound 17 with CH=NMe group (electron donating) at indole ring indicates lower inhibitory activity, reflecting the importance of negative partial charge atoms at this position.
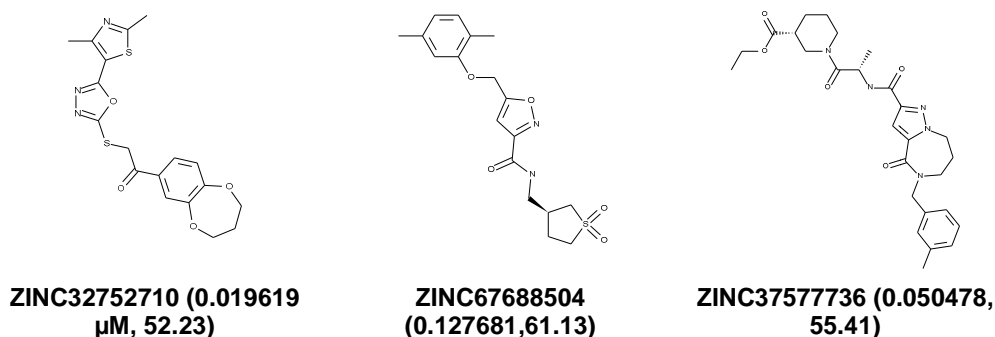

ZINC32752710 (0.019619 µM, 52.23)

ZINC67688504 (0.127681,61.13)

ZINC37577736 (0.050478, 55.41)

Fig. 3. 2D structures of top three hits discovered by docking and virtual screening. IC_{50} in µM and fitness values are listed in the parentheses, respectively

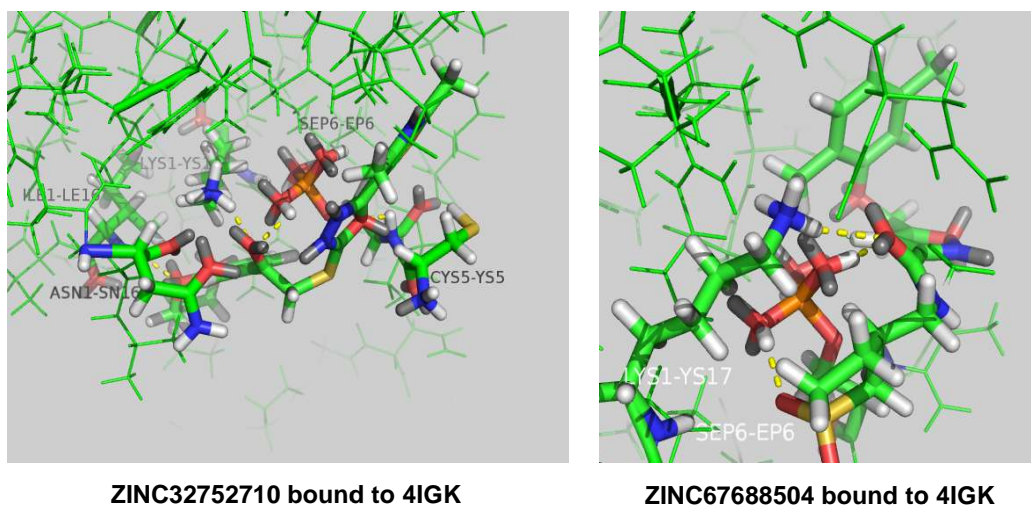**ZINC32752710 bound to 4IGK**

ZINC67688504 bound to 4IGK

Fig. 4. The docking modes of ZINC32752710 and ZINC67688504 to 4IGK protein. The hydrogen bonds are shown in dotted yellow lines. Important residues are shown and labeled

Electrostatic and steric contour maps in CoMSIA are similar to those of CoMFA field and the colors in both fields provide the same significance. Fig. 5 c and d show HB acceptor and donor contour maps of CoMSIA models. Carboxyl substituent at indole ring occupies magenta favored HB acceptor area, leading to potent inhibitory bioactivity in compounds 6, 9 (pIC_{50}=8.13), 18 (pIC_{50}=8.22). Red unfavorable HB acceptor areas are observed around substituents CH=NMe and CH=NOMe at indole ring exhibiting potent activity in compounds 18 and 23, which

agrees well with CoMFA maps. Cyan favored HB donor contours are occupied by phenyl and indole rings in compounds 6, 9, 23 with higher activity, reflecting favored hydrogen atom contributions. Purple unfavorable HB donor contours are occupied by amide and hydroxyl groups at these positions leading to the decrement in activity in compounds 1,10, 20 (pIC_{50}<6.30). In addition, compounds 17and 19 with N-H containing substituent's at indole ring wrapped by the purple contour show less active.
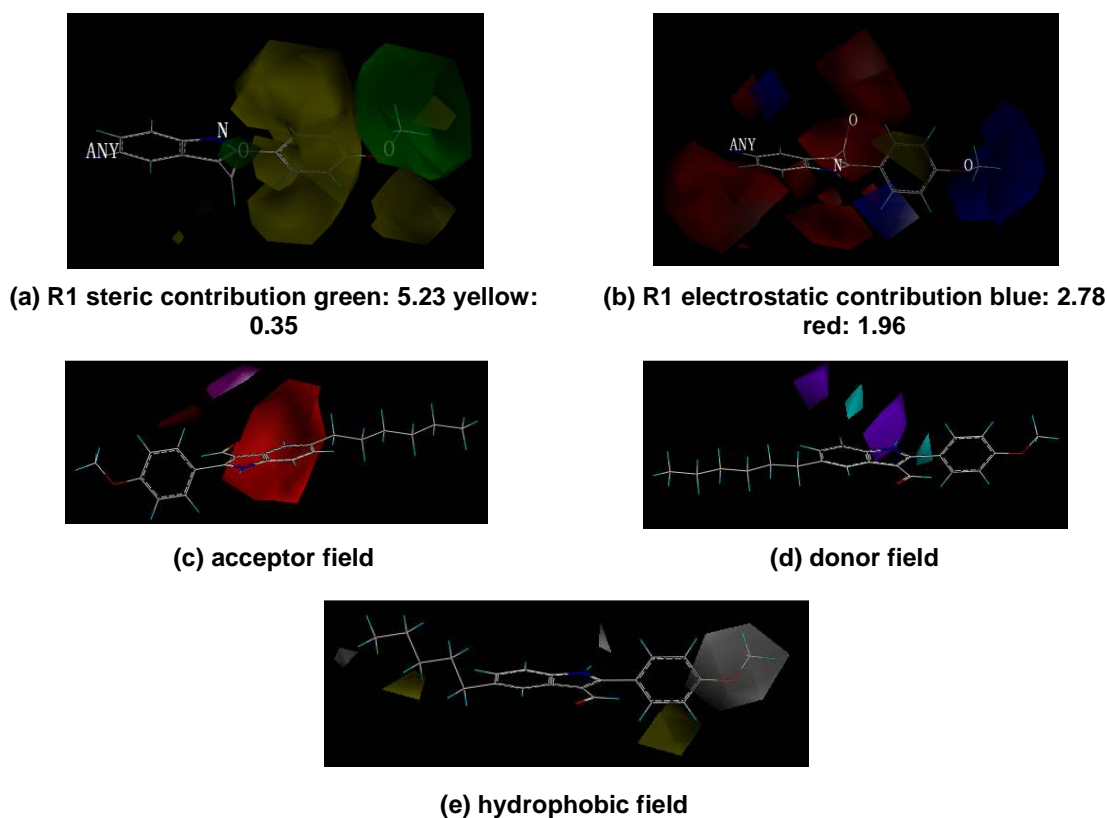

(a) R1 steric contribution green: 5.23 yellow: 0.35

(b) R1 electrostatic contribution blue: 2.78 red: 1.96

(c) acceptor field

(d) donor field

(e) hydrophobic field

Fig. 5. Contour maps of CoMFA model (a),(b) and CoMSIA model (c), (d), (e) for compound 6 (Table 2)

1 (8.50,53.33)   **2 (8.53,65.20)**   **3 (8.56,53.43)**
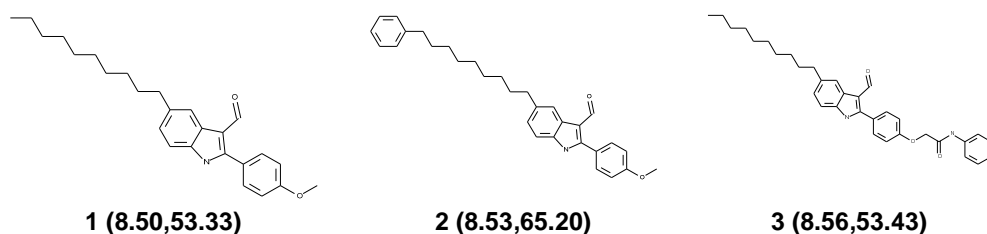
Fig. 6. 2D structures and predicted bio-activities of designed compounds based on CoMFA and CoMSIA modeling. The values in parentheses represent pIC_{50} and fitness from docking, respectively

8

Considering these structure requirements for the template compound, three new compounds were designed as inhibitor as shown in Fig. 6 (compounds **1**, **2**, **3**). To further evaluate their activity, molecular docking of the threecompounds to breast cancer protease and ADMET filtering were performed respectively. Fitness values from docking are 53.33 (**1**), 65.20 (**2**) and 53.43 (**3**) respectively. Compound **1** is involved in four hydrogen bonds with carboxyl group of Asn1678, OH group of Sep6 and amide group of Cys5 and Sep6, respectively. Compound **2** was predicted to make five hydrogen bonds with carboxyl group of Asn1678, amide group of Gln1779, Ala4 and Cys5, and OH of Sep6, respectively. Compound **3** was also predicted to make three hydrogen bonds with carboxyl group of Asn1678, OH of Sep6, and amide group of Sep6, respectively. ADMET analysis also shows that these compounds met key parameters for drug use. These results indicate promising potential as breast cancer inhibitors for the three compounds. It is implied from CoMFA, CoMSIA modeling and docking that substituent at the side site of compound 6 (OCH_3 group) such as phenyl-$CONH_2$ group and ethane group at the other side site may improve bioactivity by pIC_{50} value of 8.55 and fitness value of ~59, respectively.

4. CONCLUSION

Validation results and decoy test indicate that the developed 3D pharmacophore model is highly predictive. Residue Sep6 and Cys 5 were observed as important active sites for ligand-protein binding. Top three hits with novel scaffolds: ZINC67688504 (IC_{50}=0.01962 μM, fitness =52.23), ZINC32752710 (0.1277, 61.13), and ZINC37577736 (0.050478, 55.41), were identified as potential lead compounds or inhibitory candidates for breast cancer. Three new lead compounds were designed with predicted inhibiting potencies. It is implied from CoMFA and CoMSIA modeling that substituent at the side site of compound 6 (OCH_3 group) such as phenyl-$CONH_2$ group and ethane group at the other side site may improve bioactivity by pIC_{50} value of 8.55 and fitness value of ~59, respectively. The QSAR and docking results obtained from this work should be useful in determining structural requirements for inhibitor development as well as in designing more potential inhibitors.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Salgado TM, Davis EJ, Farris KB, Fawaz PN, Henry NL. Identifying socio-demographic and clinical characteristics associated with medication beliefs about aromatase inhibitors among postmenopausal women with breast cancer. Brea Canc Res Treat. 2017;163(2):311-19.

2. Sugimotoa Y, Sawant DB, Fisk HA, Mao L, Li C, Chettiar S, Li P, Darbya MV, Brueggemeier RW. Novel pyrrolopyrimidines as Mps1/TTK kinase inhibitors for breast cancer. Bioorg Med Chem. 2017;25(7):2156-66.

3. Lukong KE. Understanding breast cancer - the long and winding road. BBA Clinical. 2017;7:64-67.

4. Roberts K, Rickett K, Greerb R, Woodward N. Management of aromatase inhibitor induced musculoskeletal symptoms in postmenopausal early Breast cancer: A systematic review and meta-analysis. Criti Rev in Oncology/Hemato. 2017;111:66-80.

5. Foudah AI, Sallam AA, Akl MR, Sayed KA. Optimization, pharmacophore modeling and 3D-QSAR studies of sipholanes as breast cancer migration and proliferation inhibitors. Europ J Med Chem. 2014;73: 310-24.

6. Busnena BA, Foudah AI, Melancon T, Sayed KA. Olive secoiridoids and semisynthetic bioisostere analogues for the control of metastatic breast cancer. Bioorg Med Chem. 2013;21:2117-27.

7. Bhat MA, Dhfyan AA, Harbi N, Manogaran PS, Alanazi AM, Fun H, Omar MA. Targeting HER-2 over expressed breast cancer cells with 2-cyclohexyl-N- [(Z)-(substitutedphenyl /furan-2-yl/thiophene-2-yl)methylidene] hydrazinecarbothioamide. Bioorg Med Chem Lett. 2015;25:83-87.

8. Kaufmann D, Pojarova M, Vogel S, Liebl R, Gastpar R. Antimitotic activities of 2-phenylindole-3-carbaldehydes in human

breast cancer cells. Bioorg Med Chem. 2007;15:5122-36.

9.   Bharatham N, Bharatham K, Lee KW. Pharmacophore identification and virtual screening for methionyl-tRNA synthetase inhibitors. J Mol Graph Model. 2007;25: 813–23.

10. Clement OO, Freeman CM, Hartmann RW. Three dimensional pharmacophore modeling of human CYP17 inhibitors. Potential agents for prostate cancer therapy. J Med Chem. 2003;46:2345-51.

11. Leach AR, Gillet VJ. An introduction to chemoinformatics. 2nd Ed. Springer: Kluwer Dordrecht; 2011.

12. Golbraikh A, Bernard P, Chretien JR. Validation of protein based alignment in 3D quantitative structure–activity relationships with CoMFA models. Eur J Med Chem. 2000;35:123–36.

13. Shoichet BK. Virtual screening of chemical libraries. Nature. 2004;432:862–65.

14. Sastry GM, Dixon SL, Sherman W. Rapid shape-based ligand alignment and virtual screening method based on atom/feature-pair similarities and volume overlap scoring. J Chem Inf Model. 2011;51:2455-66.

15. Norinder U, Bergstrom CAS. Prediction of ADMET properties. Chem Med Chem. 2006;1:920-37.

16. Kwak JH, Park JY, Lee D, Kwaka JY, Park EH. Inhibitory effects of ginseng sapogenins on the proliferation of triple negative breast cancer MDA-MB-231 cells. Bioorg Med Chem Lett. 2014;24(23):5409-12.

17. Boumendjel A, Macalou S, Belkacem A, Blanca M, Pietro AD. Acridone derivatives: Design, synthesis, and inhibition of breast cancer resistance protein ABCG2. Bioorg Med Chem. 2007;15:2892-97.

18. Manda BR, Alla M, Ganji RJ, Addlagatta A. Discovery of Troger's base analogues as selective inhibitors against human breast cancer cell line: Design, synthesis and cytotoxic evaluation. Euro J Med Chem. 2014;86:39-47.

19. Kuhnle M, Egger M, Muller C, Mahringer A, Bernhardt G. Potent and selective inhibitors of breast cancer resistance protein (ABCG2) derived from the p-glycoprotein (ABCB1) modulator tariquidar. J Med Chem. 2009;52:1190-97.

20. Juvale K, Stefan K, Wiese M. Synthesis and biological evaluation of flavones and benzoflavones as inhibitors of BCRP/ABCG2. Euro J Med Chem. 2013;67:115-26.

21. Srisawat T, Sukpondma Y, Chimplee S, Kanokwiroon K. Extracts from vatica diospyroides type SS fruit show low dose activity against MDA-MB-468 breast cancer cell-line via apoptotic action. Bio Med Res Inter. 2014;11:1-8. ID: 479602.

22. Golbraikh A, Bernard P, Chretien JR. Validation of protein-based alignment in 3D quantitative structure-activity relationships with CoMFA models. Eur J Med Chem. 2000;35(1):123-26.

23. Gupte A, Buolamwini JK. CoMFA and CoMSIA 3D-QSAR studies on S6-(4-nitrobenzyl) mercaptopurine riboside (NBMPR) analogs as inhibitors of human equilibrative nucleoside transporter 1 (hENT1). Bioorg Med Chem Lett. 2009;19:314–18.

24. Golbraikh A, Tropsha A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. J Comput Aided Mol Des. 2002;16:357-69.

25. Alexander G, Alexander T. Beware of Q2. J Mol Graph Model. 2002;20:269–76.