IOSR Journal Of Pharmacy And Biological Sciences (IOSR-JPBS)
e-ISSN:2278-3008, p-ISSN:2319-7676. Volume 14, Issue 6 Ser. I (Nov –Dec 2019), PP 55-68
www.Iosrjournals.Org

In Silico Structure Activity Relationship Analysis on a Set Of ERA Inhibitors

Jakkanaboina Tilakvijay[1*] Kandimalla Vivek Babu [2] Addepally Uma [1]
1. Centre for Biotechnology, Institute Of Science & Technology, JNTUH Hyderabad, Telangana
2. Biological E. Limited Shameerpet HyderabadTelangana
Corresponding Author: Jakkanaboina Tilakvijay

Abstract: *Breast cancer is the most common cancer among women and majority of diagnosed cancers are estrogen receptor (ER)-positive. Estrogen facilitates its effects by binding to its receptors, estrogen receptor (ER)-α and ER-β. In a search to identify novel ERα inhibitors, a multivariate regression analysis was carried out on a set of 80 compounds known to inhibit ERα in vitro belonging to benzofurans, diphenyl amine analogs, sulfoximine-based acyclic triaryl olefins, isoxazole, thiazolidinone derivatives, tamoxifen mimics, pyrazolo(1,5-a)pyrimidines and chromen-2-one derivatives, respectively, and all molecules are known to inhibit the estrogen receptor in MCF-7 cancer cell lines. Nearly three new QSAR models were built by dividing the complete data set into 63 molecule training set and a 6 molecule validation set, after excluding outlying data based on Relative Error and Standardized Residuals. Predictive ability of new models (7, 4 and 6 variables) were evaluated and observed that all statistical values are within limits. R^2_{cvext} was found to be 0.99 for all the three model equations. Therefore, to define the statistical quality of activity prediction, FIT Kubinyi function was used, where the 7-variable model was chosen as best model as it possessed high FIT value than others. This model displayed good internal predictivity with q^2 value of 0.99 and was able to explain 91 % variance of inhibitory activities. The model was further validated by applying the y-randomization test and the low R^2 and Q^2 values indicated that the results obtained in the QSAR model are not due to chance correlation. The model indicated an increase in HOMO, H-bond acceptors, donors and LogP with reduction in LUMO and lipophilic character would enhance ERα inhibition.*

KeyWords: *Estrogen, QSAR, FIT Kubinyi, correlation coefficient, H-bond acceptors, H-bond donors, HOMO, LUMO*

Date of Submission: 02-11-2019 Date of Acceptance: 18-11-2019

I. Introduction

Breast cancer is the most common cancer among women in all parts of the world[1]. Majority of breast cancers diagnosed today are estrogen receptor (ER)-positive. However, in certain cases, progesterone receptor-positive (PR-positive) is also responsible for breast cancer. Hence, during hormone therapy, it is more important to diagnose the received signal from Estrogen receptor (ER) or progesterone receptor (PR)[2]. When circulating estrogen binds ER, it stimulates cell division and tumor growth [3]. Tumors that are ER/PR-positive are much more likely to respond to hormone therapy than tumors that are ER/PR-negative[4]. Breast cancer starts when cells in the breast begin to grow out of control. These cells usually form a tumor that can often be seen on an x-ray or felt as a lump. The tumor is malignant (cancer) if the cells can grow into (invade) surrounding tissues or spread (metastasize) to distant areas of the body[5]. Estrogen, a sex steroid hormone is produced by the ovaries and affects growth, differentiation, and function of the mammary gland. Estrogen facilitates its effects by binding to its receptors, estrogen receptor (ER)-α and ER-β [6]. In premenopausal women, estrogen production is high in ovaries. In these women, surgical, radiation, and pharmacologic ablation of the ovaries can be employed to decrease estrogen production. In postmenopausal women, relatively small amounts of estrogensare produced in peripheral tissues by conversion of androgens produced by the adrenal glands. These low levels of estrogens can be inhibited either by blocking the estrogen receptor, or by inhibiting the peripheral conversion of androgens to estrogens [7]. Among the pharmacologic endocrine therapies for breast cancer are treatment with antiestrogens (including selective estrogen receptor modulator (SERMs)), luteinizing hormone-releasing hormone (LHRH) analogs, aromatase inhibitors, estrogens, progestins, and androgens. Since years, several scaffolds have been developed as potential agents against breast cancer. Tamoxifen is the most extensively used and studied antiestrogen and its role in the management of patients with breast cancer is well established. However, extensive evaluation of tamoxifen treatment revealed significant side effects such as endometrial cancer, blood clots and the development of acquired resistance.Hence, there is a pressing need for the improvement and/or development of new antiestrogens for the prevention and treatment of breast cancer. Herein, we report the

computer-aided QSAR (Quantitative Structure Activity Relationship) analysis on a diverse set of ERα inhibitors to extract the important characteristic features of ligands responsible for bioactivity against ERα.

II. Materials and Methods

Dataset

ERα antagonists directly block the active site of ERα to prevent any estrogen from binding to it as well as to stop the function of hormone. In breast cancer cell, estrogen activates ERα by binding to its active site, which induces conformational changes that allow co-activators to attach on the complex. Several ligands have been put forward by many researchers as antagonists of ERα such as benzofurans[8], diphenyl amine analogs[9], sulfoximine-based acyclic triaryl olefins[10], isoxazole derivatives[11], thiazolidinone derivatives[12], tamoxifen mimics[13], pyrazolo(1,5-a)pyrimidine conjugates [14], chromen-2-one derivatives [15] etc. Many of those compounds are serving as anticancer agents [16], antifungal agents [17], and antiinflammatory agents etc. [18]

Molecular Descriptors

In our study, forty various physico-chemical, topological and electrostatic descriptors were evaluated in terms of their efficacy to predict the activities of the investigated inhibitors. Molecular descriptors chosen in the study: topological, shape and connectivity indices, total dipole and lipole, molecular weight, h-bond donors, h-bond acceptors, logP and rotatable bond counts. A semi-empirical molecular orbital package was used to calculate thermodynamic property like heat of formation and electrostatic properties like HOMO (Highest Occupied Molecular Orbital), LUMO (Lowest Unoccupied Molecular Orbital) components.

Multivariate Regression Analysis

QSAR models were constructed on complete and training sets, respectively. Validation was done internally using leave-one-out (LOO) technique and externally by predicting the activities of validation set. The relationship between dependent variable ($\log 1/IC_{50}$) and independent variables was established by linear multiple regression analysis. Significant descriptors were chosen based on the statistical data of analysis. The generated QSAR equation was judged based on the parameters like correlation coefficient (r), standard error of estimate (s), F-value, cross-validation r^2 (q^2).

Predictive Ability of QSAR model

Predictive ability of the generated model was estimated externally by predicting the activities of validation set. This criterion may not be sufficient for a QSAR model to be truly predictive [19]. An additional condition for high predictive ability of QSAR model is based on external set cross-validation r^2, ($R^2_{cv,ext}$) and the regression of observed activities against predicted activities and vice versa for validation set, if the following conditions are satisfied [20]

$$R^2_{cv,ext} > 0.5 \quad (1)$$

$$R^2 > 0.6 \quad (2)$$

$$(R^2 - R_0^2) / R^2 < 0.1 \text{ or } (R^2 - R_0'^2) / R^2 < 0.1 \quad (3)$$

$$0.85 \leq k \leq 1.15 \text{ or } 0.85 \leq k' \leq 1.15 \quad (4)$$

Calculations relating to $R^2_{cv,ext}$, R_0^2 and the slopes, *k* and *k'* are based on regression of observed values against predicted values and vice versa.

Y-randomization

This test confirms the sturdiness of a QSAR model [21] and to evaluate the multiple linear regression models obtained by variables. In y-randomization test, the dependent variable is shuffled randomly and a new model is builtwith X-data intact. The new models are expected to have low R2 and Q2 values, which determine the statistical significance of the original model.

III. Results and Discussion

A multivariate regression analysis was carried out on a set of 80 compounds which are known to inhibit ERα *in vitro* as all molecules are known to inhibit the estrogen receptor in MCF-7 cancer cell lines with better inhibitory concentrations. A multiple linear regression analysis has been initiated on these compounds to delineate the important features of these set of compounds. The biological activity of few molecules was reported as IC_{50} (M) while few others as Relative Binding Affinity (RBA %) against ERα in MCF-7 cancer cell lines. Therefore, the RBA activity data was converted to corresponding IC_{50} values.

The 80 molecule dataset was subjected to QSAR analysis to identify the influential parameters responsible for biological activity. The activity data of is given in Table 1.

Table 1: Biological activity data of 80 compounds dataset (IC_{50} (μM)).

S. No.	Molecule No.	Simplified molecular-input line-entry system (SMILES)	Experimental Activity (IC50μM)	1/IC50	log (1/IC50)
1	11_6a.mol	Fc1ccc(cc1)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccccc1	1.79	0.558659	-0.25285
2	11_6b.mol	Clc1ccc(cc1Cl)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccccc1	5.3	0.188679	-0.72428
3	11_6c.mol	COc1ccc(cc1)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccccc1	2.16	0.462963	-0.33445
4	11_6d.mol	COc1ccc(cc1OC)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccccc1	6.12	0.163399	-0.78675
5	11_6e.mol	COc1cc(cc(c1OC)OC)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccccc1	6.93	0.1443	-0.84073
6	11_6f.mol	Fc1ccc(cc1)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(cc1)F)c1ccccc1	3.34	0.299401	-0.52375
7	11_6g.mol	Fc1ccc(cc1)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(c(c1)Cl)Cl)c1ccccc1	4.73	0.211416	-0.67486
8	11_6h.mol	COc1ccc(cc1)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccc(cc1)F	4.97	0.201207	-0.69636
9	11_6i.mol	COc1ccc(cc1OC)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccc(cc1)F	7.02	0.14245	-0.84634
10	11_6j.mol	COc1cc(cc(c1OC)OC)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccc(cc1)F	6.28	0.159236	-0.79796
11	11_6k.mol	COc1ccc(cc1)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(cc1)F)c1ccccc1	8.08	0.123762	-0.90741
12	11_6l.mol	COc1ccc(cc1)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(c(c1)Cl)Cl)c1ccccc1	6.12	0.163399	-0.78675
13	11_6m.mol	COc1ccc(cc1)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(cc1)OC)c1ccccc1	2.93	0.341297	-0.46687
14	11_6n.mol	COc1ccc(cc1)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(c(c1)OC)OC)c1ccccc1	2.85	0.350877	-0.45484
15	11_6o.mol	COc1ccc(cc1)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1cc(c(c(c1)OC)OC)OC)c1ccccc1	3.02	0.331126	-0.48001
16	11_6p.mol	COc1ccc(cc1OC)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(cc1)F)c1ccccc1	2.69	0.371747	-0.42975
17	11_6q.mol	COc1ccc(cc1OC)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(c(c1)Cl)Cl)c1ccccc1	5.95	0.168067	-0.77452
18	11_6r.mol	COc1ccc(cc1)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1ccc(c(c1)OC)OC	2.53	0.395257	-0.40312
19	11_6s.mol	COc1ccc(cc1OC)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(c(c1)OC)OC)c1ccccc1	2.61	0.383142	-0.41664
20	11_6t.mol	COc1ccc(cc1OC)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1cc(c(c(c1)OC)OC)OC)c1ccccc1	4.24	0.235849	-0.62737
21	11_6u.mol	COc1cc(cc(c1OC)OC)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(cc1)F)c1ccccc1	7.59	0.131752	-0.88024
22	11_6v.mol	COc1cc(cc(c1OC)OC)Nc1ncccc1C(=O)N1CCN(CC1)C(=O)C1=Nc2cc(nn2C(=C1)c1ccc(c(c1)Cl)Cl)c1ccccc1	5.38	0.185874	-0.73078
23	11_6w.mol	COc1ccc(cc1)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1cc(c(c(c1)OC)OC)OC	3.83	0.261097	-0.5832
24	11_6x.mol	COc1ccc(cc1OC)C1=CC(=Nc2cc(nn21)c1ccccc1)C(=O)N1CCN(CC1)C(=O)c1cccnc1Nc1cc(c(c(c1)OC)OC)OC	7.34	0.13624	-0.8657
25	6_10.mol	Oc1ccc(cc1)N1(C@@H)(SCC1=O)c1ccc(cc1)Cl	5	0.2	-0.69897
26	6_12.mol	Cc1ccc(cc1)N1(C@@H)(SCC1=O)c1ccc(cc1)Cl	0.81	1.234568	0.091515
27	6_13.mol	O=C1CS(C@H)(N1c1ccccc1)c1ccccc1	0.25	4	0.60206
28	6_14.mol	Cc1ccc(cc1)N1(C@@H)(SCC1=O)c1cccc2ccccc21	0.23	4.347826	0.638272
29	6_9.mol	COc1ccc(cc1OC)(C@@H)1SCC(=O)N1c1ccc(cc1)O	2.58	0.387597	-0.41162
30	8_11a.mol	O=C1OC2(C=CC(=O)C=C2)C(=C1c1sc2ccccc2c1)c1ccccc1	5.9	0.169492	-0.77085
31	8_11b.mol	Fc1ccc(cc1)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	6.32	0.158228	-0.80072
32	8_11c.mol	O=C1OC2(C=CC(=O)C=C2)C(=C1c1oc2ccccc2c1)c1ccccc1	5.8	0.172414	-0.76343
33	8_11d.m	Oc1ccc(cc1)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	6	0.166	-

	ol			667	0.77815
34	8_11e.mol	(O-)(N+)(=O)c1ccc(cc1)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	5.99	0.166945	-0.77743
35	8_11f.mol	COc1ccc(cc1)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	5.86	0.170648	-0.7679
36	8_11g.mol	OCc1ccc(cc1)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	5.76	0.173611	-0.76042
37	8_11h.mol	COc1cc(cc(c1OC)OC)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	5.79	0.172712	-0.76268
38	8_11i.mol	COc1ccc(cc1C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2)F	5.73	0.17452	-0.75815
39	8_11j.mol	COc1c(cc(cc1C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2)C)Br	5.85	0.17094	-0.76716
40	8_11k.mol	O=C1C=CC2(OC(=O)C(=C2c2ccccc2)c2ccc3c(c2)OCO3)C=C1	5.81	0.172117	-0.76418
41	8_11l.mol	O=C1C=CC2(OC(=O)C(=C2c2ccccc2)c2ccc(cc2)C#N)C=C1	19	0.052632	-1.27875
42	8_11m.mol	OC(=O)c1ccc(cc1)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	5.95	0.168067	-0.77452
43	8_11n.mol	O=Cc1sc(cc1)C1=C(c2ccccc2)C2(OC1=O)C=CC(=O)C=C2	6.23	0.160514	-0.79449
44	1_4d.mol	Oc1ccc(cc1)N(CC1CC1)c1ccc(cc1)O	7.41	0.135	-0.86967
45	1_4e.mol	CC(C)CCN(c1ccc(cc1)O)c1ccc(cc1)O	3.10	0.3225	-0.49147
46	1_4g.mol	Oc1ccc(cc1)N(CC1CCCCC1)c1ccc(cc1)O	0.45	2.225	0.34733
47	1_4h.mol	Oc1ccc(cc1)N(CCC1CCCCC1)c1ccc(cc1)O	0.11	9.1075	0.959399
48	1_4i.mol	Oc1ccc(cc1)N(C(C@@)12C(C@@H)3C(C@@H)(C(C@@)(Br)(C3)C1)C2)c1ccc(cc1)O	0.73	1.3775	0.139092
49	1_4j.mol	Oc1ccc(cc1)N(Cc1ccccc1)c1ccc(cc1)O	0.61	1.6375	0.214181
50	1_4k.mol	Oc1ccc(cc1)CN(c1ccc(cc1)O)c1ccc(cc1)O	0.09	10.65	1.02735
51	1_4l.mol	Oc1ccc(cc1)N(CC1CCCCC1)c1ccc(cc1)O	0.35	2.825	0.451018
52	1_4m.mol	Oc1ccc(cc1)N(c1ccccc1)c1ccc(cc1)O	0.31	3.2375	0.51021
53	14_15a.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN(C)C)c2ccccc2OC1=O	1.14	0.875	-0.05799
54	14_15b.mol	CCN(CC)CCOc1ccc(cc1)NC1=C(C(=O)Oc2ccccc21)c1ccc(cc1)OC	2.35	0.425	-0.37161
55	14_15c.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCCC2)c2ccccc2OC1=O	1.05	0.95	-0.02228
56	14_15d.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCCCC2)c2ccccc2OC1=O	0.77	1.3	0.113943
57	14_15f.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCOCC2)c2ccccc2OC1=O	4.60	0.2175	-0.66254
58	14_16a.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN(C)C)c2ccc(cc2OC1=O)OC	0.56	1.8	0.255273
59	14_16b.mol	CCN(CC)CCOc1ccc(cc1)NC1=C(C(=O)Oc2cc(ccc21)OC)c1ccc(cc1)OC	0.87	1.15	0.060698
60	14_16c.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCCC2)c2ccc(cc2OC1=O)OC	0.91	1.1	0.041393
61	14_16d.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCCCC2)c2ccc(cc2OC1=O)OC	0.53	1.875	0.273001
62	14_16f.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCN(C)CC2)c2ccc(cc2OC1=O)OC	3.33	0.3	-0.52288
63	14_18a.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN(C)C)c2ccc(cc2OC1=O)O	0.20	5.05	0.703291
64	14_18b.mol	CCN(CC)CCOc1ccc(cc1)NC1=C(C(=O)Oc2cc(ccc21)O)c1ccc(cc1)OC	0.22	4.65	0.667453
65	14_18c.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCCC2)c2ccc(cc2OC1=O)O	0.23	4.3	0.633468
66	14_18d.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCCCC2)c2ccc(cc2OC1=O)O	0.14	7.075	0.849726
67	14_18e.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCOCC2)c2ccc(cc2OC1=O)O	7.55	0.1325	-0.87778
68	14_18f.mol	COc1ccc(cc1)C1=C(Nc2ccc(cc2)OCCN2CCN(C)CC2)c2ccc(cc2OC1=O)O	0.83	1.2	0.079181
69	4_4a.mol	Oc1ccc(cc1)c1onc(c1/C=C/c1ccccc1)c1ccc(cc1)O	0.11	8.75	0.94200

					8
70	4_4c.mol	Cc1ccc(cc1)\C=C\c1c(onc1c1ccc(cc1)O)c1ccc(cc1)O	0.67	1.5	0.176091
71	4_4d.mol	Oc1ccc(cc1)c1onc(c1/C=C/c1ccc(cc1)F)c1ccc(cc1)O	0.13	7.75	0.889302
72	4_4e.mol	Oc1ccc(cc1)c1onc(c1/C=C/c1ccc(cc1)Cl)c1ccc(cc1)O	0.50	2	0.30103
73	4_4f.mol	Oc1ccc(cc1)c1onc(c1/C=C/c1ccc(cc1)C(F)(F)F)c1ccc(cc1)O	0.08	13.25	1.122216
74	4_4g.mol	Oc1ccc(cc1)c1onc(c1/C=C/c1cccc(c1)C(F)(F)F)c1ccc(cc1)O	0.27	3.75	0.574031
75	4_4h.mol	Oc1ccc(cc1)\C=C\c1c(onc1c1ccc(cc1)O)c1ccc(cc1)O	0.02	41	1.612784
76	4_4i.mol	Oc1ccc(cc1)c1onc(c1/C=C/c1cccc(c1)O)c1ccc(cc1)O	0.05	21.75	1.337459
77	4_4j.mol	CCCCCCC\C=C\c1c(onc1c1ccc(cc1)O)c1ccc(cc1)O	0.22	4.5	0.653213
78	4_4k.mol	CCCCCCCC\C=C\c1c(onc1c1ccc(cc1)O)c1ccc(cc1)O	0.16	6.25	0.79588
79	4_4l.mol	CCCCCCCCC\C=C\c1c(onc1c1ccc(cc1)O)c1ccc(cc1)O	0.31	3.25	0.511883
80	4_4m.mol	CC(C)(C)\C=C\c1c(onc1c1ccc(cc1)O)c1ccc(cc1)O	0.09	10.75	1.031408

Dataset (80 compounds - IC_{50} (μM) activity)

The inhibitory activities of dataset reported in terms of IC_{50} in μM transformed into their respective logarithmic values in order to overcome overlapping data. Therefore, to assure linear distribution of data, the enzyme inhibition data converted to negative logarithmic values and then subjected to QSAR analysis.

The equation 1 given below represents the linear QSAR model from a set of 80 ERα inhibitors.

Complete Data set:

$\log(1/IC_{50})$ = + 0.071831*Total Dipole
- 0.033945*Total Lipole
- 4.4167678e-005*Weiner Index
+ 0.3128*H-bond Donors
+ 0.18529*LogP
- 1.1803 (Eq 1)

r =0.818, r^2 = 0.669, q^2 =0.540, F = 29.961, n = 80, s = 0.406

From the above equation, it is evidenced that the parameters of r, r2 values are within limits and the properties that appeared in the equation included total dipole and lipole on the molecules where a reduced lipole and increase in dipole characteristics would favour better values or in other words ERα inhibitory features of ligands. Similarly, decrease in Weiner index with an increase in H-bond donors and logP on ligands favour ERα inhibition.

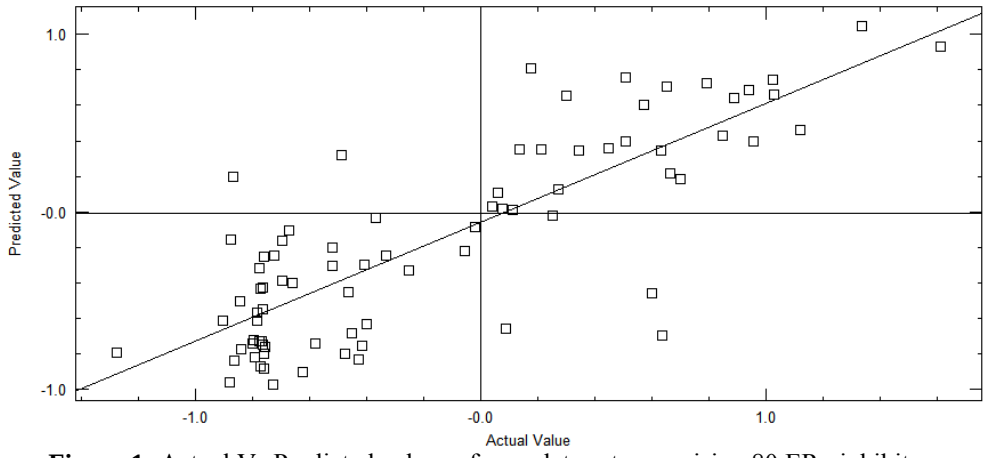


Figure 1: Actual Vs Predicted values of complete set comprising 80 ERα inhibitors.

A plot showing observed values versus predicted values is given in Figure 1 where it was observed that few data points are away from the regression line. Such points refer to noise and are represented as outlying data. Hence, outlier detection was carried out by two methods.

Outlier Detection

The criterion for removing outliers is based on Relative Error calculation and Standardized Residuals.

Relative Error Calculation

This method was employed to calculate the relative error (Eq. 2) of all compounds in the data set. Ten compounds with high relative error (more than 100 relative error %) are highlighted and the data represented were 1, 2, 5, 11, 13, 24, 25, 26, 68 and 71 respectively (Table 2). Moreover, it should be noted that the QSAR model was good, however, the model prediction led to a high relative error for compounds and hence they should be excluded from the study as they influence the outcome in a significant manner.

Relative Error = Residual Value / Actual Value (Eq 2)

Table 2. Outlier calculation on complete data set. Compounds 1, 2, 5, 11, 13, 24, 25, 26, 68 and 71 regarded as outliers based on relative error %. Standardized residual data reported 1, 2, 24, 25, 26 and 76 compounds.

S. No.	Actual Value	Predicted Value	Residual Value	Standardized residuals	relative error	error %
1	-0.86967	0.208898	-1.07856	-2.74214	1.2402	124.02
2	-0.49147	0.332473	-0.82394	-2.0948	1.676489	167.6489
3	0.34733	0.355744	-0.00841	-0.02139	-0.02423	-2.42253
4	0.959399	0.408035	0.551364	1.401793	0.574697	57.46973
5	0.139092	0.361172	-0.22208	-0.56462	-1.59665	-159.665
6	0.214181	0.364884	-0.1507	-0.38315	-0.70362	-70.362
7	1.02735	0.75344	0.27391	0.696391	0.266618	26.6618
8	0.451018	0.368172	0.082847	0.21063	0.183688	18.36882
9	0.51021	0.408451	0.101758	0.258711	0.199443	19.94434
10	0.889302	0.651965	0.237337	0.603408	0.26688	26.68801
11	0.30103	0.666116	-0.36509	-0.9282	-1.21279	-121.279
12	0.942008	0.700609	0.241399	0.613735	0.25626	25.626
13	0.176091	0.822442	-0.64635	-1.64329	-3.67055	-367.055
14	1.12222	0.473535	0.648681	1.649213	0.578034	57.80337
15	0.574031	0.612531	-0.0385	-0.09788	-0.06707	-6.70697
16	1.61278	0.940882	0.671902	1.70825	0.416611	41.66111
17	1.33746	1.05765	0.279811	0.711394	0.209211	20.92107
18	0.653212	0.716121	-0.06291	-0.15994	-0.09631	-9.63063
19	0.79588	0.738725	0.057155	0.14531	0.071813	7.181309
20	0.511883	0.765816	-0.25393	-0.6456	-0.49608	-49.6076
21	1.03141	0.669747	0.361661	0.91949	0.350647	35.06472
22	-0.41162	-0.29189	-0.11973	-0.3044	0.290875	29.08751
23	-0.69897	-0.38265	-0.31632	-0.80422	0.452553	45.2553
24	0.091515	-0.6541	0.745616	1.895661	8.147473	814.7473
25	0.60206	-0.45422	1.05628	2.685496	1.754443	175.4443
26	0.638272	-0.69108	1.32936	3.379778	2.082748	208.2748
27	-0.77815	-0.31369	-0.46446	-1.18085	0.596876	59.68764
28	-0.77743	-0.72067	-0.05675	-0.14429	0.073001	7.300119
29	-0.77085	-0.72305	-0.0478	-0.12154	0.062014	6.20145
30	-0.80072	-0.73389	-0.06683	-0.1699	0.083459	8.345932
31	-0.76343	-0.79368	0.030256	0.076922	-0.03963	-3.96314
32	-0.7679	-0.54645	-0.22145	-0.56302	0.288386	28.8386
33	-0.76042	-0.24966	-0.51076	-1.29856	0.671681	67.1681
34	-0.76268	-0.87597	0.113294	0.28804	-0.14855	-14.8547
35	-0.75816	-0.75497	-0.00318	-0.00809	0.004196	0.419559
36	-0.76716	-0.42186	-0.34529	-0.87787	0.450092	45.00923
37	-0.76418	-0.73961	-0.02457	-0.06246	0.032148	3.214822
38	-1.27875	-0.79016	-0.4886	-1.24221	0.38209	38.20895
39	-0.77452	-0.42623	-0.34829	-0.8855	0.449687	44.96867
40	-0.79449	-0.81449	0.020002	0.050854	-0.02518	-2.51761
41	-0.78675	-0.56272	-0.22403	-0.56958	0.284757	28.47572
42	-0.84073	-0.76978	-0.07095	-0.18038	0.08439	8.439029
43	-0.25285	-0.32353	0.070677	0.179689	-0.27952	-27.9517
44	-0.72428	-0.23868	-0.48559	-1.23458	0.670454	67.04544
45	-0.33445	-0.24165	-0.0928	-0.23593	0.277466	27.74663
46	-0.52375	-0.19619	-0.32756	-0.83278	0.62541	62.541

47	-0.67486	-0.10245	-0.57241	-1.4553	0.848188	84.8188
48	-0.69636	-0.15781	-0.53855	-1.36922	0.773386	77.3386
49	-0.84634	-0.50084	-0.3455	-0.87841	0.408232	40.82322
50	-0.79796	-0.7138	-0.08416	-0.21398	0.105474	10.54736
51	-0.90741	-0.60487	-0.30254	-0.76918	0.333411	33.34112
52	-0.78675	-0.60529	-0.18147	-0.46136	0.230652	23.06524
53	-0.46687	-0.44579	-0.02108	-0.05359	0.045144	4.514445
54	-0.45485	-0.67813	0.223281	0.567672	-0.49089	-49.0895
55	-0.48001	-0.7916	0.311594	0.7922	-0.64914	-64.9145
56	-0.42975	-0.82431	0.39456	1.003133	-0.91811	-91.8111
57	-0.77452	-0.86483	0.090314	0.229615	-0.11661	-11.6607
58	-0.40312	-0.62382	0.220704	0.56112	-0.54749	-54.7488
59	-0.41664	-0.75031	0.333674	0.848336	-0.80087	-80.0867
60	-0.62737	-0.89435	0.26698	0.678773	-0.42556	-42.5557
61	-0.88024	-0.95381	0.073568	0.187039	-0.08358	-8.35766
62	-0.73078	-0.96475	0.23397	0.594848	-0.32016	-32.0164
63	-0.5832	-0.73342	0.150219	0.381918	-0.25758	-25.7578
64	-0.8657	-0.83031	-0.03538	-0.08996	0.040874	4.08737
65	0.113943	0.02225	0.091693	0.233122	0.804729	80.47287
66	-0.05799	-0.21436	0.156371	0.397559	-2.69643	-269.643
67	-0.37161	-0.028	-0.34361	-0.87359	0.924639	92.46389
68	-0.02228	-0.0775	0.055225	0.140404	-2.47906	-247.906
69	-0.66254	-0.39231	-0.27024	-0.68705	0.407878	40.78782
70	0.273001	0.135151	0.13785	0.350471	0.504943	50.49432
71	0.255273	-0.01813	0.273407	0.695113	1.071038	107.1038
72	0.060698	0.117054	-0.05636	-0.14328	-0.92848	-92.8477
73	0.041393	0.041643	-0.00025	-0.00064	-0.00604	-0.60419
74	-0.52288	-0.29579	-0.22709	-0.57736	0.434309	43.43089
75	0.849726	0.440946	0.40878	1.039286	0.481073	48.10727
76	-0.87778	-0.15149	-0.7263	-1.84655	0.827422	82.74222
77	0.703291	0.193346	0.509945	1.296489	0.725084	72.50839
78	0.667453	0.230621	0.436832	1.110606	0.654476	65.4476
79	0.633468	0.35461	0.278858	0.708971	0.440209	44.02085
80	0.079181	0.029929	0.049252	0.125219	0.622018	62.20176
		stdev	0.393328			

Standardized Residuals

The data set was analyzed for the presence of outliers, by calculating the standard residuals. Standardized residuals greater than 2 and less than -2 are usually considered large. Outliers should be removed in order to obtain the best statistical result.

From table 1, Compounds 1, 2, 24, 25, 26 and 76 have high standardized residuals and are safely removed from the dataset.

A new QSAR model was built (Eq 3) with n=69, after excluding eleven outlying data (1, 2, 5, 11, 13, 24, 25, 26, 68, 71 and 76) (based on relative error % and standardized residuals) and the graph plotted based on actual versus predicted values of complete set is given in Figure 2, where better predictive nature was displayed when compared with Eq 3 data, given in Figure 1.

Complete Data set after removing Outliers:

$\log (1/IC_{50})$ = - 0.027521*Total Lipole
- 0.00014816*Weiner index
+ 0.19378*H-bond Acceptors
+ 0.43732* H-bond Donors
+ 0.20667*LogP
- 0.25956*LUMO
+ 0.55245*HOMO
+ 2.9308 (Eq 3)

r =0.958, r^2 = 0.918, q^2 =0.844, F = 97.392, n = 69, s = 0.214

Equation 3 displays better correlation coefficient values than Eq 1, which can be attributed to the removal of outlying points from the data set. Further, new properties entered the regression equation such as H-bond Acceptors, HOMO and LUMO, where, the former properties needs to be enhanced on molecules with a concomitant decrease in LUMO parameter in order to attain better inhibition.

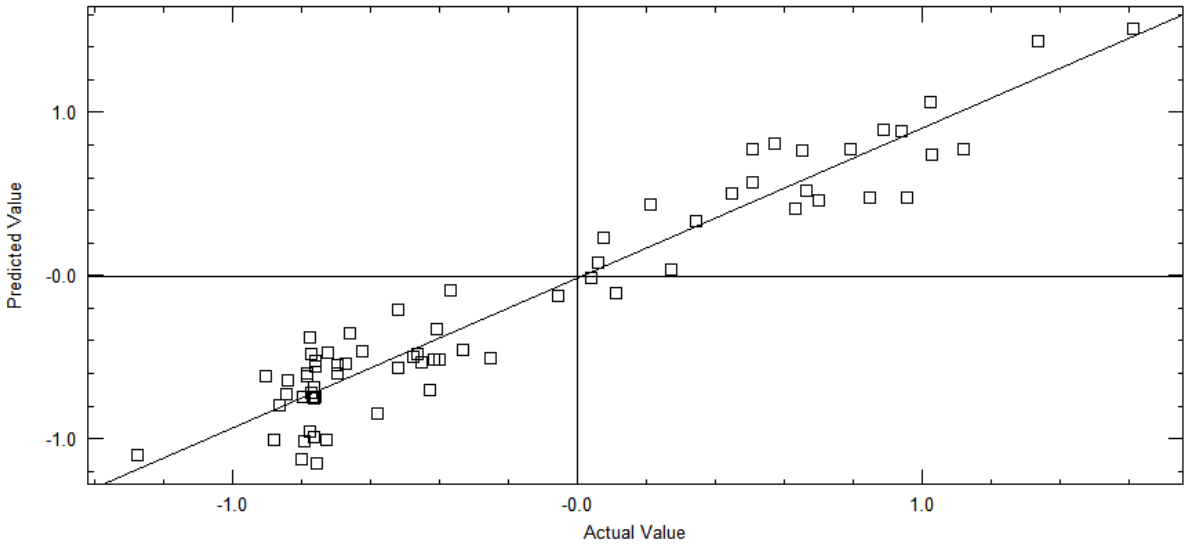


Figure 2: Observed and predicted values of molecules of 69 compound dataset after removing outliers.

New QSAR Models

Around three new QSAR models were attempted by dividing the 69 compound ERα inhibitor data set as a 63 molecule training set and a 6 molecule validation set (Table 3) based on visual inspection after rejection of outliers from the data set. More specifically, the selection of molecules in the training set was made according to the IC_{50} values; so that representatives of a wide range of structures with different substituents and activity were included. The distribution of activity values for the validation set follows the similar distribution of the activity values for the training set. The results obtained from the multiple linear regression procedure with varied number of descriptors are shown in Table 3 with their statistics.

Given below are a set of 3 different models obtained and are statistically significant (Table 3).

Table 3. Descriptor data and statistical values of three model equations.

Descriptor	**Coefficient**		
	Model-1	**Model-2**	**Model-3**
Total Dipole	-	-	+0.03416
Total Lipole	-0.02735	-0.02198	-0.02786
Weiner index	-0.00014	-8.8e-005	-8.8e-005
H-bond acceptors	+0.19400	-	+0.12182
H-bond donors	+0.43721	+0.66001	+0.56131
LogP	+0.20483	-	+0.16957
LUMO	-0.24939	-	-
HOMO	+0.54437	-	-
KAlpha3 index	-	+0.15236	-
Constant	+2.8706	-1.1104	-1.6867
Statistics			
r	0.954	0.928	0.938
r^2	0.910	0.861	0.879
q^2	0.822	0.761	0.764
F	79.47	89.68	68.41
n	63	63	63
s	0.22	0.27	0.25
No. of Descriptors	7	4	6
Equation No.	**4**	**5**	**6**

Test set data:
Different compounds were selected as test/validation set. They are:

Set-1: 38, 41, 1, 8, 14, 62
Set-2: 23, 39, 58, 5, 7, 4
Set-3: 12, 26, 38, 7, 62, 68

Test set data for equations 4, 5 and 6 given below and the predictive ability of test sets were given in Table 4.

Test set data-1 (Equation 4):
Test set graphs were plotted for calculations (Figure 3). Values corresponding to k (actual Vs predicted) and k' (predicted Vs actual) and $\mathbf{R^2\text{-}R_0^2/R^2}$ are given below. It was observed that these parameters are within the limits.

Actual Vs Predicted – Test set-1

Actual Values	Predicted Values
0.34733	0.346533
0.942008	0.88496
0.79588	0.780828
-0.52375	-0.56807
-0.84634	-0.71774
0.060698	0.087232

k = 0.9534

R^2 = 0.9942
R_0^2 = 0.9937

$R^2\text{-}R_0^2/R^2$ = 0.0005

Predicted Vs Actual – Test set-1

Actual Values	Predicted Values
0.346533	0.34733
0.88496	0.942008
0.780828	0.79588
-0.56807	-0.52375
-0.71774	-0.84634
0.087232	0.060698

k' = 1.0488

R^2 = 0.9942
R_0^2 = 0.9938

$R^2\text{-}R_0'^2/R^2$ = 0.0004

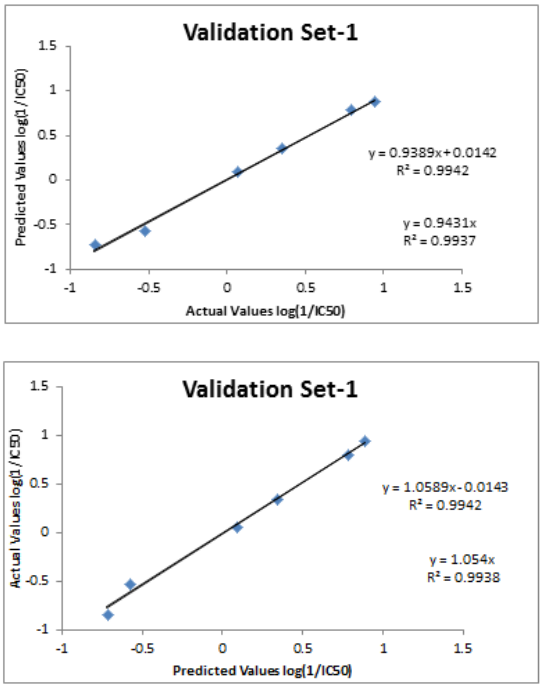


Figure 3: Graph showing validation set-1 data comprising 6 compounds. A.) Actual versus Predicted data B.) Predicted versus Actual values data

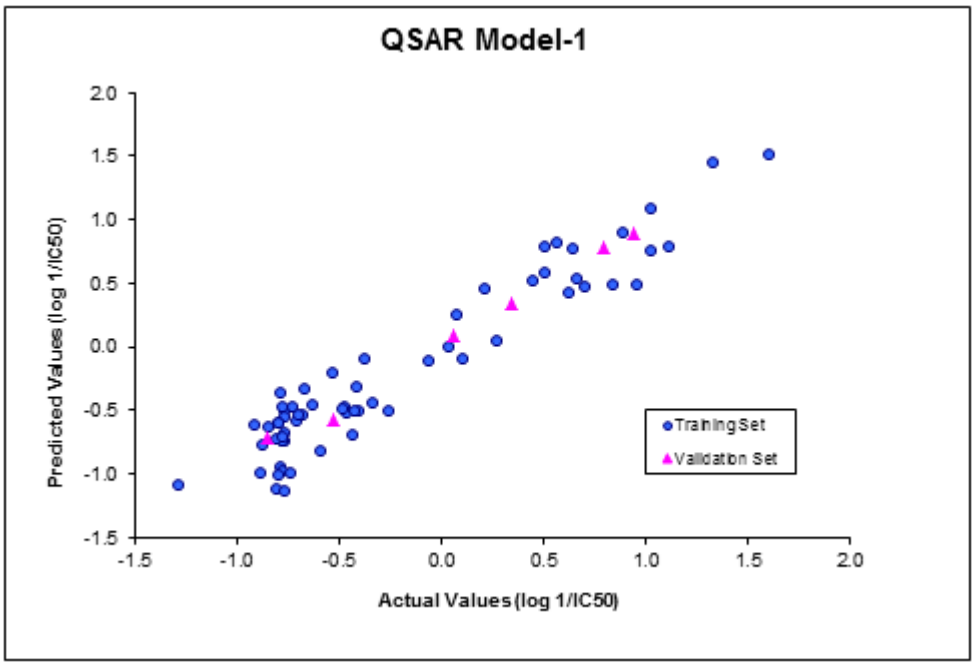


Figure 4: Regression plot showing training set (blue spheres) data and test set (triangles) of QSAR Model-1

Test set data-2 (Equation 5):

Test set data-2 graphs were plotted, given in Figures 5 and 6. Values corresponding to k (actual Vs predicted) and k' (predicted Vs actual) and $\mathbf{R^2\text{-}R_0^2/R^2}$ are presented here, where it was observed that these parameters are within the limits.

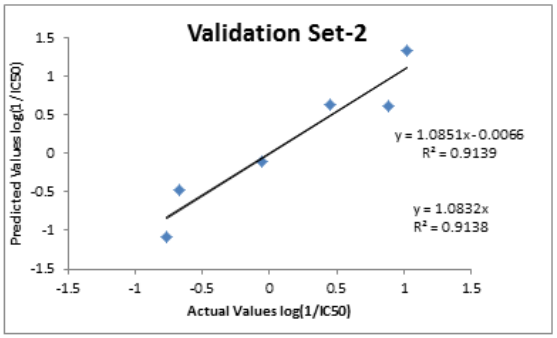
Actual Vs Predicted – Test set-2

Actual Values	Predicted Values
1.02735	1.33183
0.451018	0.634692
0.889302	0.607083
-0.76343	-1.08321
-0.67486	-0.47565
-0.05799	-0.10878

k = 0.9618

R^2 = 0.9139
R_0^2 = 0.9138

$R^2\text{-}R_0^2/R^2$ = 0.0001



Predicted Vs Actual – Test set-2

Actual Values	Predicted Values
1.33183	1.02735
0.634692	0.451018
0.607083	0.889302
-1.08321	-0.76343
-0.47565	-0.67486
-0.10878	-0.05799

k' = 1.0396

R^2 = 0.9139
R_0^2 = 0.9132

$R^2\text{-}R_0'^2/R^2$ = 0.0007

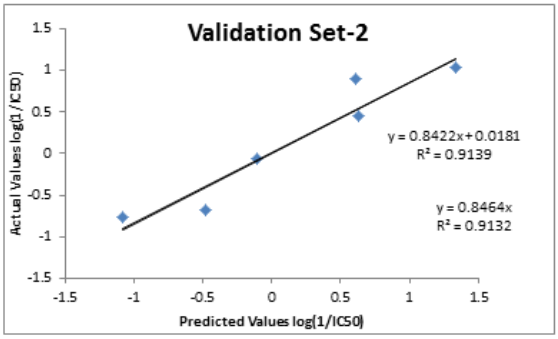


Figure 5: Graph showing validation set-2 data comprising 6 compounds. A.) Actual versus Predicted data B.) Predicted versus Actual values data

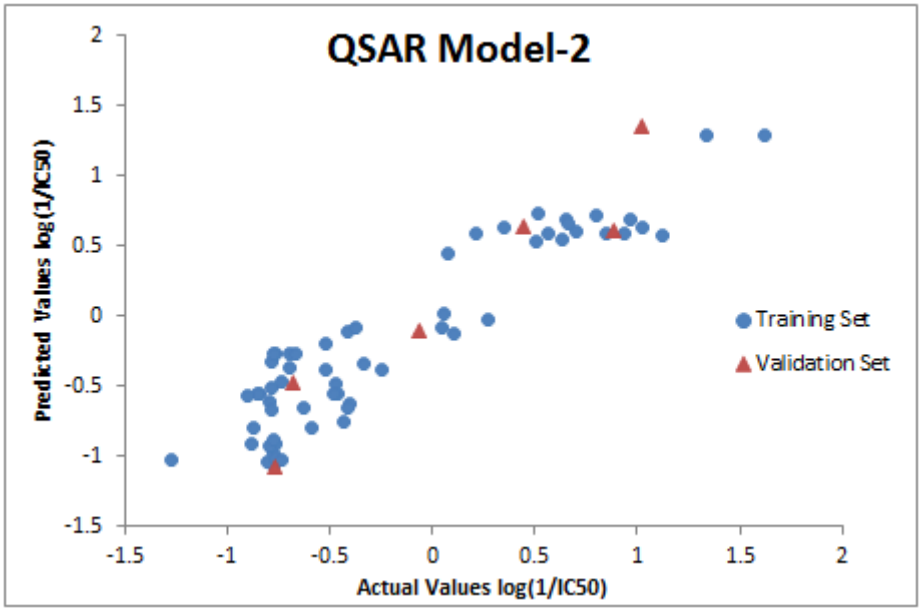


Figure 6: Regression plot showing training set (blue spheres) data and test set (triangles) of QSAR Model-2.

Test set data-3 (Equation 6):

Test set data-3 graphs were given in Figures 7 and 8. Values corresponding to k (actual Vs predicted) and k' (predicted Vs actual) and $\mathbf{R^2\text{-}R_0^2/R^2}$ are presented here, where it was observed that these parameters are within the limits.

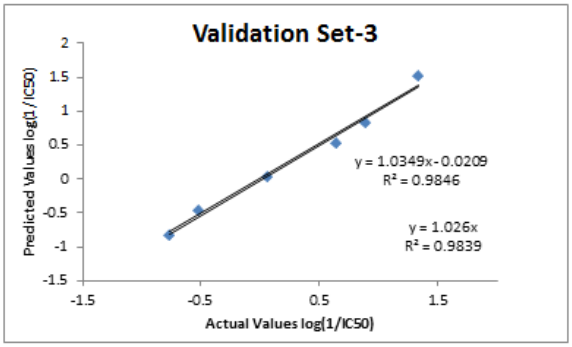
Actual Vs Predicted – Test set-3

Actual Values	Predicted Values
0.889302	0.808132
1.33746	1.50851
-0.76268	-0.83829
-0.52375	-0.45109
0.060698	0.022322
0.633468	0.516812

k = 1.0434

R^2 = 0.9846
R_0^2 = 0.9839

$R^2\text{-}R_0^2/R^2$ = 0.0007



Predicted Vs Actual – Test set-3

Actual Values	Predicted Values
0.808132	0.889302
1.50851	1.33746
-0.83829	-0.76268
-0.45109	-0.52375
0.022322	0.060698
0.516812	0.633468

k' = 0.9583

R^2 = 0.9846
R_0^2 = 0.9836

$R^2\text{-}R_0'^2/R^2$ = 0.001

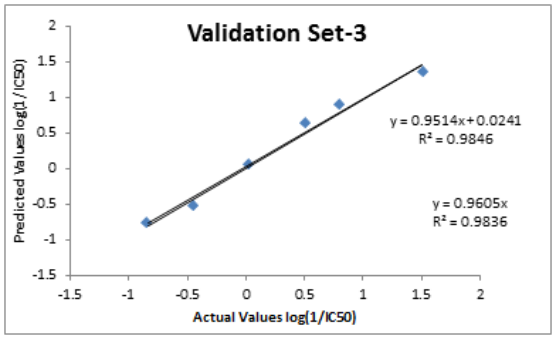


Figure 7: Graph showing validation set-3 data comprising 6 compounds. A.) Actual versus Predicted data B.) Predicted versus Actual values data

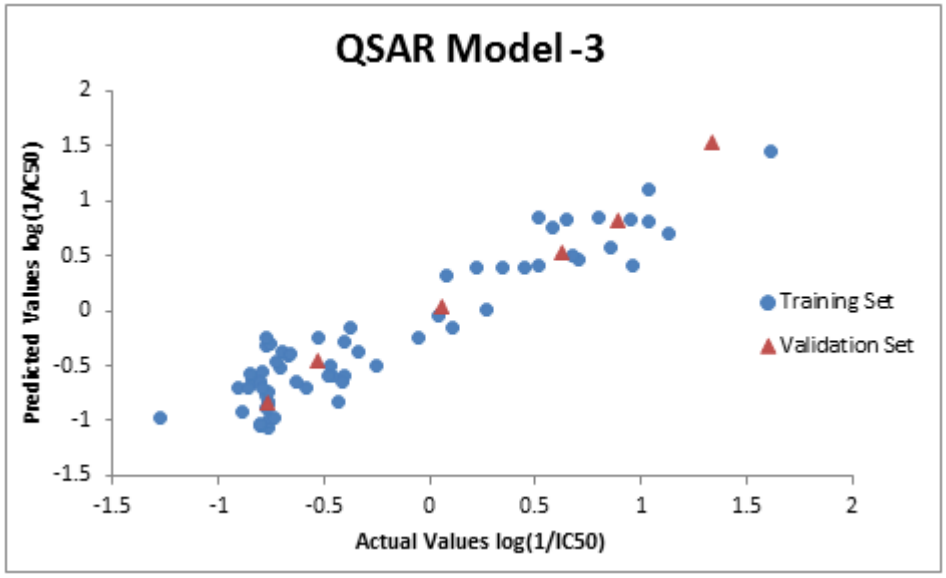


Figure 8: Regression plot showing training set (blue spheres) data and test set (triangles) of QSAR Model-3.

Table 4 represents the predictive ability of all newly generated models. R^2_{cvext} which is an external set cross validation was found to be 0.99 for all the three model equations.

Table 4. Predictive ability of validation sets for all 3 equations obtained as models.

Var[a]	$R^2_{cv,ext}$ (q^2)	R^2	*k*	*k'*	Eq[b]	Eq[c]
7	0.999	0.910	0.953	1.048	0.0005	0.0004
4	0.999	0.861	0.961	1.039	0.0001	0.0007
6	0.999	0.879	1.043	0.958	0.0007	0.001

[a]number of significant variables
[b] $(R^2 - R_0^2) / R^2$
[c] $(R^2 - R_0'^2) / R^2$

FIT Kubinyi function

To define the statistical quality of activity prediction, the number of variables that enter in to a QSAR model are compared by using FIT Kubinyi function (Eq. 7), a criteria closely related to F value was proven to be useful. The best model will be the one that possess a high value of this function.

$$FIT = R^2 (n - k - 1) / (n + k^2) (1 - R^2) \quad \text{(Eq 7)}$$

Where *n* is the number of compounds in training set and *k* is the number of variables in the QSAR equation.

Table 5. FIT Kubinyi data obtained for all QSAR models.

Eq No.	r^2	*k*	n	FIT
4	**0.910**	**7**	**63**	**4.96**
5	0.861	4	63	3.04
6	0.879	6	63	3.57

According to the statistical values of the models reported in Table 5, we choose the model with seven variables (Eq. 4) since this showed high FIT than others. The observed, calculated and predicted values of the statistically significant seven parameter QSAR model (Eq. 4) is presented in Table 3.

Equation 4 accounts for the significant correlation of descriptors with biological activity and displayed good internal predictivity as shown by q^2 value of 0.99 and was able to explain 91 % variance of inhibitory activities of derivatives. Observed verses predicted values of molecules in training and validation set are shown graphically in Figures 3 and 4. The proposed QSAR model Eq. 4 illustrated the predictive ability of the model.

The model was further validated by applying the y-randomization test. Random shuffles of the dependent as well as independent variables were performed and the results presented in Table 6. The low R^2 and Q^2 values indicate that the results obtained in the QSAR model (Eq. 4) are not due to chance correlation.

Table 6. R^2 and Q^2 values after several y-randomization tests

Iteration	R^2	Q^2
1	0.22	0.34
2	0.25	0.11
3	0.36	0.12
4	0.27	0.15
5	0.30	0.24
6	0.41	0.17
7	0.33	0.12
8	0.29	0.18
9	0.43	0.29
10	0.18	0.22

The generated QSAR model (Eq. 4) indicates that a high value of LUMO energy contributes negatively to the activity whereas an increase in HOMO generates positive inhibition. Molecular Orbital (MO) surfaces denote several constant electronic distributions of a chemical compound or ligand. As per Frontier Orbital theory, the highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) are critical in forecasting the reactivity of a species. HOMO is the outermost orbital comprising the electron and LUMO is the first orbital that does not encompass an electron. The electron donating nature of a compound is measured by HOMO and the energy of the LUMO measures electron accepting property (Hall LH, *et al.,*1991). The lower the LUMO value, the stronger is the electrophilicity.

Electron-withdrawing substituents (for example, halogens) decrease the LUMO energy on the molecule. Molecules with low-lying LUMOs have greater tendencies to accept electrons than those with high-energy LUMOs. As LUMOincreases the molecule becomes less reactive (Hall LH1991). Thus, designing analogs with electron-withdrawing substituents would improve ERα inhibitory activity.

From equation 4 it can be observed that an increase in H-bond acceptors, donors and LogP would enhance ERα inhibition. Proper spatial orientation of H bond donor and acceptor groups of ligand is important to interact with the acceptor and donor atoms of amino acid residues in the active site region of ERα. On the other hand, reduction of lipophilic character on the compounds would increase bioactivity.

IV. Conclusion

One of the most important contributions of the MCF-7 cell line to breast cancer research has been its utility for the study of the estrogen receptor (ER) alpha, as this cell line is one of a very few to express substantial levels of ER. It was reported that anti-estrogens inhibited growth of MCF-7 cells. In search to identify novel ERα inhibitors, QSAR analysis was carried out on a set of 80 ERα inhibitors reported in literature to identify the influential parameters responsible for biological activity. About 11 data points as outliers were removed from analysis based on Relative Error calculation and Standardized Residuals and three new QSAR models were constructed with 7, 4 and 6-variables, wherein the model with 7 variables was found to be best model based on FIT Kubinyi function. This model suggested an increase in HOMO, H-bond acceptors, donors and LogP with reduction in LUMO and lipophilic character would enhance ERα inhibition. Further, virtual screening of novel analogs with these associated properties is under investigation using molecular docking techniques.

References

[1]. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 2014;64(1):9–29.
[2]. Marco M Hefti et al. Estrogen receptor negative/progesterone receptor positive breast cancer is not a reproducible subtype. Breast Cancer Research 2013 15:R68.
[3]. Osborne CK, Ravdin PM: Adjuvant systemic therapy of primary breast cancer. In Diseases of the Breast. Edited by: Harris JR, Lippman ME, Morrow M, Osborne CK. 2000, Philadelphia: Lippincott Williams & Wilkins, 599-632. 2.
[4]. Hammond ME, Hayes DF, Dowsett M, Allred DC et al: American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J ClinOncol.2010, 28: 2784-2795.
[5]. Sotiriou C, Pusztai L: Gene-expression signatures in breast cancer. N Engl J Med. 2009, 360: 790-800.
[6]. Roger P, Sahla ME, Mäkelä S, Gustafsson JA, Baldet P, Rochefort H. Decreased expression of estrogen receptor beta protein in proliferative preinvasive mammary tumors. Cancer Res. 2001 Mar 15; 61(6):2537-41
[7]. Hayes DF, Henderson IC, Shapiro CL: Treatment of metastatic breast cancer: Present and future prospects. SeminOncol 22:5-19, 1995 (suppl 5).
[8]. Xiao-Yan Li et al. 3-Acyl-5-hydroxybenzofuran derivatives as potential anti-estrogen breast cancer agents: A combined experimental and theoretical investigation. Bioorganic & Medicinal Chemistry Letters 23 (2013) 4617–4621,
[9]. KiminoriOhta et al. Promising core structure for nuclear receptor ligands: Design and synthesis of novel estrogen receptor ligands based on diphenylamine skeleton. Bioorganic & Medicinal Chemistry Letters 18 (2008) 5050–5053
[10]. Xiao Yun Chen et al. Syntheses and biological activities of sulfoximine-based acyclic triaryl olefins. Bioorganic & Medicinal Chemistry Letters 22 (2012) 4307–4309
[11]. Terra Haddad et al. Synthesis and evaluation of 4-(substituted styryl/alkenyl)-3,5-bis(4-hydroxyphenyl)-isoxazoles as ligands for the estrogen receptor. Bioorganic & Medicinal Chemistry Letters 22 (2012) 5999–6003

[12]. Marina Sala et al. Synthesis and cytotoxic activity evaluation of 2,3-thiazolidin-4-one derivatives on human breast cancer cell lines. Bioorganic & Medicinal Chemistry Letters 23 (2013) 4990-5
[13]. SivakumarArchana et al. Development of constrained tamoxifen mimics and their antiproliferative properties against breast cancer cells. Bioorganic & Medicinal Chemistry Letters 25 (2014) 680–4
[14]. Ahmed Kamal et al. Synthesis and mechanistic aspects of 2-anilinonicotinyl-pyrazolo (1,5-a)pyrimidine conjugates that regulate cell proliferation in MCF-7 cells via estrogensignaling. Bioorganic & Medicinal Chemistry Letters 26 (2016) 2077–83
[15]. GuoshunLuo et al. Design, synthesis, biological evaluation and molecular docking studies of novel 3-aryl-4-anilino-2H-chromen-2-one derivatives targeting ERα as anti-breast cancer agents. Bioorganic & Medicinal Chemistry Letters 27 (2017) 2668–73
[16]. Rohaiza, S.; Yaacob, W.; Din, L.; Nazlina, I. Afr. J. Pharm. Pharmacol. 2011, 5, 1272
[17]. Ryu, C. K.; Song, A. L.; Lee, J. Y.; Hong, J. A.; Yoon, J. H.; Kim, A. Bioorg. Med. Chem. Lett. 2010, 20, 6777
[18]. Hu, Z. F.; Chen, L. L.; Qi, J.; Wang, Y. H.; Zhang, H.; Yu, B. Y. Fitoterapia 2011, 82, 190
[19]. Golbraikh A, Tropsha A. Beware of q2!. J. Mol. Graph. Model. 2002, 20, 269-276; Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J, Igglessi-Markopoulou O. A novel QSAR model for predicting induction of apoptosis by 4-aryl-4H-chromenes.Bioorg. Med. Chem. 14, 2006, 6686-6694.
[20]. Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J, Igglessi-Markopoulou O. A novel QSAR model for predicting induction of apoptosis by 4-aryl-4H-chromenes.Bioorg. Med. Chem. 14, 2006, 6686-6694
[21]. Kim D, Hong S, Lee D. The Quantitative Structure-Mutagenicity Relationship of Polycylic Aromatic Hydrocarbon Metabolites. Int. J. Mol. Sci. 7, 2006, 556-570

IOSR Journal of Pharmacy and Biological Sciences (IOSR-JPBS) is UGC approved Journal with Sl. No. 5012, Journal no. 49063.

Jakkanaboina Tilakvijay " In Silico Structure Activity Relationship Analysis on a Set Of ERA Inhibitors" IOSR Journal of Pharmacy and Biological Sciences (IOSR-JPBS) 14.6 (2019): 55-68