In Silico Screening Studies on Methanesulfonamide Derivatives as Dual Hsp27 and Tubulin Inhibitors Using QSAR and Molecular Docking

A. Mostoufi[1,2], H. Eucefifar[1,2], F. Beygi[1], M. Fereidoonnezhad[*1]

[1] *Department of Medicinal Chemistry, School of Pharmacy, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Islamic Republic of Iran*
[2] *Student Research Committee, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Islamic Republic of Iran*

Abstract

The expression of heat shock protein 27 (Hsp27) as a chaperone protein, is increased in response to various stress stimuli such as anticancer chemotherapy. This phenomenon can lead to survive of the cells and causes drug resistance. In this study, a series of methanesulfonamide derivatives as dual Hsp27 and tubulin inhibitors in the treatment of cancer were applied to quantitative structure–activity relationship (QSAR) analysis. A collection of chemometrics methods such as MLR, FA-MLR, PCR, and GA-PLS was applied to make relations between structural characteristics and anti-proliferative activity of them against SKBR-3 breast cancer cell line. The best multiple linear regression equation was obtained from GA-PLS. Concerning this model, new potent lead compounds were designed based on new structural patterns using in silico-screening study. To obtain their binding mode, binding site and types of their interactions to both tubulin and HSP27, molecular docking studies were also conducted on these compounds. The validity of docking protocol was also explored. The results obtained from this docking study indicate that the important amino acids inside the active site cavity that are in charge of essential interactions with HSP27 are Arg140, Thr139, Phe138, Cys137, Arg136, Phe104, His103, Val101, and Asp100. And this important amino acids in essential interactions with tubulin are Asn258, Val238, Cys241, Asn350, Leu255, Met259, Val315, Thr353, Arg221, Thr179 and Ser178.

Keywords: QSAR, Molecular Docking; Heat shock protein 27 (Hsp27) inhibitors; Tubulin inhibitors.

Introduction

The expression of heat shock proteins (Hsp) is increased in response to various stress stimuli such as anticancer chemotherapy [1-3]. Heat shock protein 27 (Hsp27) as a small heat shock protein, shows a very close relationship with anticancer drug resistance [3, 4]. The Hsp27 has strong antiapoptotic properties which come from its molecular chaperone functions at

multiple steps of the apoptotic signaling pathways [5].

Large oligomers of Hsp27 which protect cells from oxidative stress, are required for the chaperone activity [6, 7]. It has been well investigated that inhibiting the formation of Hsp27 large oligomers can promote programmed cell death and decrease the chaperone activity [8]. Hsp27 causes a great limitation in the efficacy of cancer therapy due to its overexpression in a wide range of tumors cells and tissues [1, 2, 9-13]. Hence, targeting Hsp27 has a great importance in cancer chemotherapy, and it's a good anticancer target [5].

One of the great limitation in developing small molecule Hsp27 inhibitors is the absence of endogenous ligand for Hsp27 and hence lack of information about its binding site [5]. Recently, Methanesulfonamide derivatives show promising anticancer activities in preclinical models against a wide range of cancer cell line with high expression levels of Hsp27 [9, 14, 15]. These potent anticancer agents are dual ligands binding to Hsp27 and tubulin [16, 17]. These compounds interfere directly with the tubulin system and can inhibit tubulin polymerization with high potency [5].

In this work, for a series of Methanesulfonamide derivatives as dual Hsp27 and tubulin inhibitors, two different drug design methodologies have been applied: QSAR and molecular docking studies. In a plenary study, to describe the physicochemical properties of the molecules, we used a very large descriptor set such as geometrical, functional, charge, constitutional, topological, atom-centered fragments, 2D auto-correlation, quantum, electrostatic, aromaticity indices, empirical and chemical. Also, to model the relationship between the structural characteristics and anti-proliferative activity of the studied compounds against SKBR-3 breast cancer cell line, the different chemometrics methods were used: 1) multiple linear regression (MLR), 2) factor analysis-based multiple linear regression (FA-MLR), 3) principal component regression (PCR) and 4) partial least squared combined with genetic algorithm for variable selection (GA-PLS). A validated molecular docking simulation technique was also applied, on all compounds of the dataset and those designed from in silico screening, to achieve the detailed molecular binding site of them in interacting with major amino acids in the active site of Hsp27 and tubulin.

Materials and Methods

Data set

A data set consisting of 49 methanesulfonamide derivatives as a series of dual Hsp27 and tubulin

Table 1. Chemical structure of the compounds used in this study, experimental and cross-validated predicted activity as well as their docking binding energy on HSP27 and tubulin.
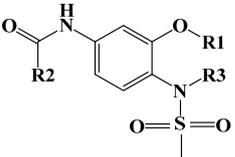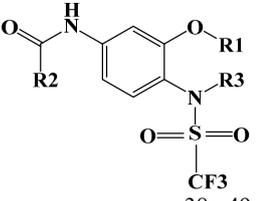
Name	R_1	R_2	R_3	Exp.pIC_{50}	Pred.pIC_{50}	ΔE^1 (kcal/mol) Hsp27	ΔE^1 (kcal/mol) Tubulin
1a	(2-Cl, 5-Cl benzyl)	(4-Br phenyl)	CH_3	5.640165	5.71388	-8.31	-9.11
2a	(2-Cl, 5-Cl benzyl)	(4-OCH₃ phenyl)	CH_3	6.309804	6.30659	-7.99	-9.11
3a	(2-Cl, 5-Cl benzyl)	(2-OCH₃, 4-OCH₃ phenyl)	CH_3	5.428291	5.56097	-7.59	-8.88
4a	(2-Cl, 5-Cl benzyl)	(4-I phenyl)	CH_3	6.180456	6.01277	-8.62	-9.75
5a	(2-Cl, 5-Cl benzyl)	(naphthyl)	CH_3	5.655608	5.56843	-9.08	-10.34
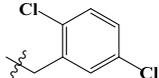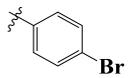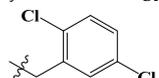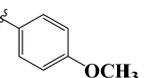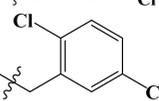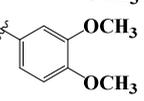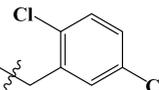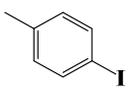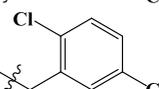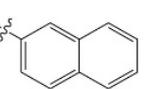
Table 1. Cntd
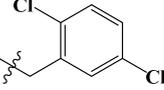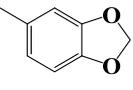
No.	R1	R2					
6a	(2,5-dichlorobenzyl)	(benzodioxole)	CH$_3$	5.545155	5.59254	-8.69	-9.12
7a	n-C$_6$H$_{13}$	(4-Br-phenyl)	CH$_3$	4.398918	4.43685	-7.23	-9.04
8a	n-C$_6$H$_{13}$	(4-OCH$_3$-phenyl)	CH$_3$	4.699405	4.81959	-6.81	-8.6
9a	n-C$_6$H$_{13}$	(2,4-diOCH$_3$-phenyl)	CH$_3$	3.998223	4.08720	7.15	-8.67
10a	n-C$_6$H$_{13}$	(4-I-phenyl)	CH$_3$	4.897566	4.60381	-7.33	-9.71
11a	n-C$_6$H$_{13}$	(naphthyl)	CH$_3$	4.191721	4.15625	-7.66	-9.84
12a	n-C$_6$H$_{13}$	(benzodioxole)	CH$_3$	4.541211	4.51461	-7.2	-9.18
13a	(ethyl-CF$_3$-phenyl)	(4-Br-phenyl)	CH$_3$	5.022734	4.84393	-7.65	-9.07
14a	(ethyl-CF$_3$-phenyl)	(4-OCH$_3$-phenyl)	CH$_3$	5.692504	5.56468	-7.3	-8.47
15a	(ethyl-CF$_3$-phenyl)	(2,4-diOCH$_3$-phenyl)	CH$_3$	4.771087	4.92903	-7.37	-7.89
16a	(ethyl-CF$_3$-phenyl)	(4-I-phenyl)	CH$_3$	5.455932	5.56469	-7.91	-9.86
17a	(ethyl-CF$_3$-phenyl)	(naphthyl)	CH$_3$	4.735182	4.95924	-8.56	-10.23
18a	(ethyl-CF$_3$-phenyl)	(benzodioxole)	CH$_3$	4.832092	4.70962	-8.09	-8.62
19a	(diOCH$_3$-ethylphenyl)	(4-Br-phenyl)	CH$_3$	8.454693	8.39460	-7.87	-8.54
20a	(diOCH$_3$-ethylphenyl)	(4-OCH$_3$-phenyl)	CH$_3$	8.59176	8.75116	-7.19	-6.5
21a	(diOCH$_3$-ethylphenyl)	(2,4-diOCH$_3$-phenyl)	CH$_3$	8.583359	8.41657	-7.23	-8.47
22a	(diOCH$_3$-ethylphenyl)	(4-I-phenyl)	CH$_3$	8.917215	8.84899	-8.02	-8.7
23a	(diOCH$_3$-ethylphenyl)	(naphthyl)	CH$_3$	7.77963	8.24032	-8.16	-9.59
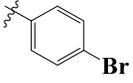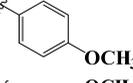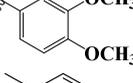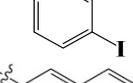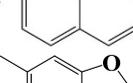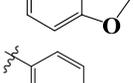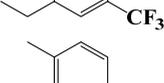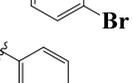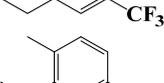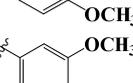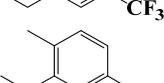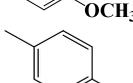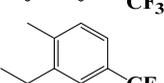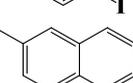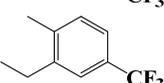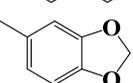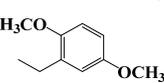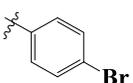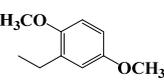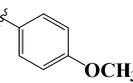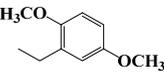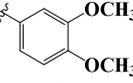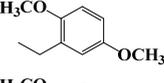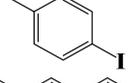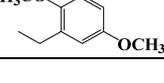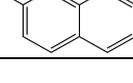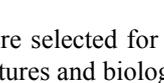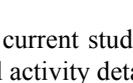
inhibitors were selected for the current study [18]. The structural features and biological activity details of these compounds are listed in Table 1. The anti-proliferative activity against SKBR-3 breast cancer cell line, as IC$_{50}$ values, were used for the QSAR modeling studies.

Table 1. Cntd
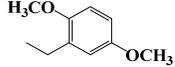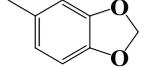
24a		CH₃	8.123782	8.10676	-7.52	-7.98
25a		H	8.018634	7.77940	-7.78	-9.47
26a		H	7.872895	7.97442	-7.06	-9.03
27a		H	7.958607	7.67759	-7.38	-7.56
28a		H	7.927015	8.26384	-7.54	-8.68
29a		H	7.877129	7.69029	-7.98	-9.3
30a		H	7.452841	7.72158	-7.64	-8.52
31a		CH₃	6.657577	6.58020	-8.56	-8.89
32a		CH₃	6.823909	6.92608	-8.02	-8.79
33a		CH₃	6.721246	6.65901	-7.71	-8.49
34a		CH₃	6.886057	6.74699	-8.57	-9.38
35a		CH₃	6.677781	6.22912	-9.09	-9.99
36a		CH₃	6.69897	6.79937	-8.44	-8.67
37a		CH₃	5.498941	5.93092	-8.03	-9.98
38a		CH₃	6.327902	6.29100	-7.05	-8.16
39a		CH₃	6.958607	6.89926	-6.37	-8.27
40a		CH₃	6.259637	6.17080	-6.39	-7.31
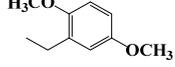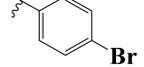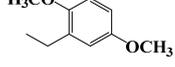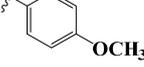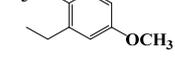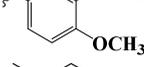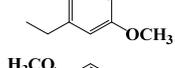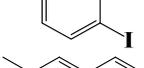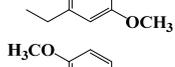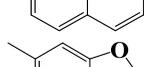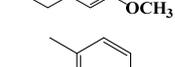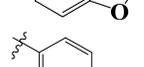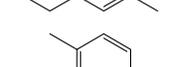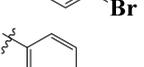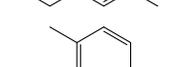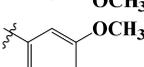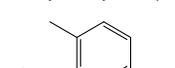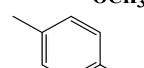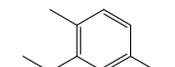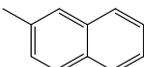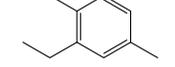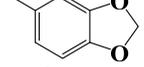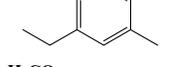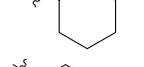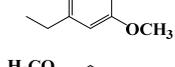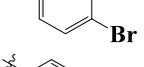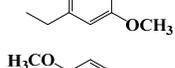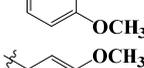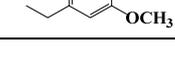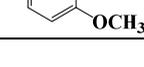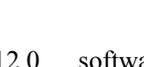
Molecular descriptors

Using ChemBioDraw 12.0 software, two-dimensional structures of the ligands were produced. Each ligand was optimized with different minimization methods including molecular mechanics (MM⁺) and quantum-based semi-empirical method (AM1) using an

in-house TCL script [19-21] using Hyperchem. To calculate a large number of molecular descriptors, Hyperchem, Gaussian 98 [22] and Dragon packages [23] were applied. Also, to calculate chemical parameters such as molecular surface area (SA), molecular volume (V), hydration energy (HE),

Table 1. Cntd
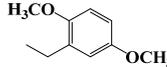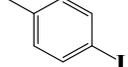
41a			CH3	6.29243	6.31578	-7.2	-6.52
42a			CH3	6.69897	6.34115	-7.46	-9.46
43a			CH3	6.148742	5.95048	-7.12	-6.86
44a			CH3	4.438422	4.53988	-7.8	-8.49
45a			CH3	5.274088	5.30246	-7.27	-8.23
46a			CH3	4.405829	4.74026	-7.51	-8.6
47a			CH3	4.869988	4.83518	-7.69	-8.16
48a			CH3	4.383	4.74955	-8.36	-9.55
49a			CH3	4.661942	4.41994	-7.69	-8.56
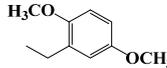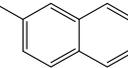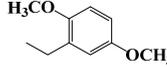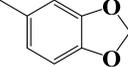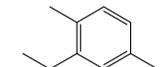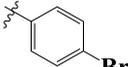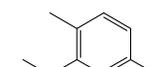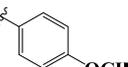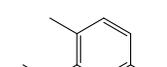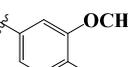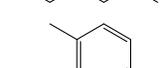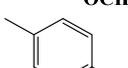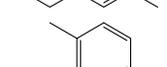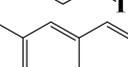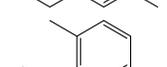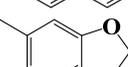
[1] docking binding energy

hydrophobicity (LogP), and molecular polarizability (MP), Hyerchem software (Version 8, Hypercube Inc., Gainesville, FL, USA) was used. Similarly, Gaussian 98 software was applied to calculate the most positive and the negative net atomic charges, highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies, the average absolute atomic charge and molecular dipole moment in x, y and z direction. Using the equations developed by Thanilaivelan et al. quantum chemical indices including hardness ($\eta = 0.5$ (HOMO+LUMO)); softness ($S = 1/\eta$), electrophilicity ($\omega = \chi^2/2\eta$) and electronegativity ($\chi = -0.5$) (HOMO-LUMO) were calculated [24]. Different geometrical, topological, empirical, charge and constitutional descriptors for each molecule and also 2D autocorrelations, aromaticity indices, atom-centered fragments and functional groups were calculated by Dragon.

Model development

In a data matrix with the number of molecules and descriptors as the number of rows and columns respectively, the calculated descriptors were illustrated. For producing QSAR equations, four different regression methods such as factor analysis as the data processing step for variable selection (FA-MLR), genetic algorithm-partial least squares (GA-PLS), principal component regression analysis (PCRA) and simple multiple linear regression with stepwise variable selection (MLR), were used. These known procedures are well explained in our previous QSAR studies [19, 25, 26].

To develop QSAR models, stepwise selection and elimination of variables was done by SPSS software (version 21; SPSS Inc., IBM, Chicago, IL, USA). To check the predictability, validity, and robustness of the models, leave-one-out cross-validation procedure using MATLAB 2015 software (version 8.5; Math work Inc., Natick, MA, USA) was obtained.

FA-MLR was also conducted on the dataset. To reduce the number of variables and to detect structure in the relations between them, factor analysis (FA) was performed. Also, to identify the important predictor variables and to avoid co-linearity among them, this data-processing step was used [27]. Along with FA-MLR, PCRA, was also applied to database. Among X variables, co-linearities as a distributing factor are not included in PCRA and also the number of variables was not more than the number of observations [28]. Factor scores obtained from FA, are played the role of predictor variables. All descriptors in PCRA are important, and detecting the relevant descriptors is the end of factor analysis [25].

The PLS regression method was applied to the

NIPALS-based algorithm existed in the chemometrics toolbox of MATLAB software. Also, to obtain the desirable number of factors, according to Haaland and Thomas F-ratio criterion [29, 30], leave-one-out cross-validation procedure was applied. For PLS and GA modeling, MATLAB PLS toolbox was applied. Variable important in objection (VIP) process was done using XLSTAT 2017 software [31].

All calculations were run on a core i7 personal computer (CPU at 16 MB) with Windows 7 operating system.

Variable importance in the projection (VIP)

To investigate the relative importance of the variable in the final model in GA-PLS method, variable important in objection (VIP) was applied [32]. The importance of variables in PLS method is represented by VIP values. According to Erikson et al. it is possible for X-variables (predictor variables) to be categorized regarding their relevance in explaining y (predicted variable). Then, VIP>1.0 and 0.8<VIP<1.0 are highly and moderately influential and VIP<0.8 is less influential [33].

Model validation

To validate the regression equation, statistical parameters including correlation coefficient (R^2), root mean square error of cross-validation (RMScv), leave-one-out cross-validation correlation coefficient (Q^2), and the variance ratio (F) with certain degrees of freedom were applied. 20% of the molecules were selected as test set (prediction set) molecules to test the developed model performance. The predictive value of a QSAR model that has not been taken into account during the process of developing the model should be tested on an external set of data.

Applicability domain

Precise prediction ability of QSAR model for new compounds has been made it widely used studies [19, 25]. It should be noted that no matter how significant and validated a QSAR may be, it cannot be expected to predict the modeled property for the entire space of chemicals reliably. Hence, the domain of application of QSAR before it is put into use for screening chemicals must be defined and predictions should be considered reliable for only those chemicals that fall in this domain. The applicability domain is appraised by the leverage values for each compound of our dataset. A Williams plot (the plot of standardized residuals versus leverage values (h)) can then be used for an immediate and simple graphical detection of both the response outliers (Y outliers) and structurally influential chemicals (X

outliers) in our model. The applicability domain for the graph is defined in a squared area within ±x (standard deviations) and leverage threshold h*.

The certain features of the numerical value of leverage include 1) being greater than zero and 2) the lower the value, the higher is the confidence in the prediction. Value of 1 equals to very poor prediction, and value of zero equals to perfect prediction that will not be reached. The other factor to analysis the results is warning leverage (h*). The threshold h* are centered on $3(k + 1)/n$, with k = the number of model parameter and n= number of training set (calibration set) compounds while x=2 or 3. Prediction of high leverage value (h > h*) compounds may not be reliable. If leverage is higher than warning leverage h*, it means the predicted response is the consequence of substantial extrapolation of the model. So, it is unreliable. In other perspective, being lower the compound leverage value than the threshold one means possibility of agreement between the values observed and predicted for compounds is as high as for the calibration set of compounds [34, 35].

Docking procedure

Molecular docking was carried out using an *in house* batch script (DOCKFACE) [25, 36] of AutoDock 4.2; each ligand was subjected to MM^+ and AM1 minimization methods using Hyperchem 8 package. With Gasteiger-Marsili procedure implanted in the AutoDock Tools program [37], the partial charges of atoms were calculated. Having non-polar hydrogens of compounds merged and the rotatable bonds assigned, the output structures were changed to PDBQT using MGLtools 1.5.6 [38].

The three-dimensional crystal structure of Hsp27 (PDB ID: 4MJH) and tubulin (PDB ID: 4YJ2) were retrieved from protein data bank (http://www.rcsb.org/pdb/home/home.do). All water molecules were removed, missing hydrogens were added, and after determining the Kollman united atom charges, non-polar hydrogens were merged into their corresponding carbons using AutoDock Tools [37]. Finally, desolvation parameters were assigned to each protein atom. Among the three different search algorithms performed by AutoDock 4.2, the commonly used Lamarckian Genetic Algorithm (LGA) was applied [39-41]. Subsequently, the enzymes were converted to PDBQT using MGLTOOLS 1.5.6.

A maximum number of 2,500,000 energy evaluations, 150 population size, 27000 maximum generations, 100 runs, a gene mutation rate of 0.02 and a crossover rate of 0.8 were used for Lamarckian GA. The grid maps of the receptors were calculated using AutoGrid tools of AutoDock 4.2. The size of grid

includes both the active site and considerable proteins of the encircling surface. A grid box of 82×66×94 and 60×60×60 points in x, y, and z directions was built and centered on the center of the ligand in the complex with a spacing of 0.375 Å for 4MJH and 4YJ2, respectively. A number of points for 4MJH in x, y and z were 10.26, 27.197 and 40.918, and for 4YJ2 was -14.704, 6.403 and 25.505, consequently. Using AutoDock Tools the grid and docking parameter files, gpf and dpf was produced. With a root mean square deviation (RMSD) tolerance of 2Å, cluster analysis was done on the docked results. All the docking protocols were done on validated structures with RMSD values below 2 Å.

According to docking results, ligand-receptor interactions were detected using AutoDock tools program (ADT, Version 1.5.6), VMD software [42] and PyMOL molecular graphics program [43].

Results and Discussion

Here, we developed a detailed QSAR study using a combination of chemical, electronic, quantum, geometrical, topological, 2D autocorrelation, and substituent constant, to explore structural parameters affecting anti-proliferative activity of methanesulfonamide derivatives as potent dual Hsp27 and tubulin inhibitors against SKBR-3 breast cancer cell line. Among the different chemometrics tools available for modeling the relationship between the biological activity and molecular descriptors, four methods (i.e., stepwise MLR, FA-MLR, PCRA, and GA-PLS) were applied.

MLR modeling

In the first step, separate stepwise selection-based MLR analyses were performed using different types of descriptors, and then, an MLR equation was obtained utilizing the pool of all calculated descriptors. The

Table 2. The results of different QSAR models with different type of dependent variables.

Model	Eq	MLR Equation	N	R^2_C	F	Q^2	RMScv	SE	R^2_P
MLR	1	PCI$_{50}$=-2.928GATS1e($\pm$0.736)-5.828GATS6v($\pm$0.396)-80.895MATS3m($\pm$6.459)+68.014PW5($\pm$11.944)-15.936MATS8p($\pm$1.343)-6.429MATS2e($\pm$0.702)-24.570G(N..N)($\pm$9.796)+26.809MATS8m($\pm$7.562)-1.328ASP($\pm$0.625)+206.913($\pm$57.987)	39	0.93	223.19	0.82	0.26	0.19	0.91
FA-MLR	2	PCI$_{50}$=89.213PW5($\pm$8.929)-24.410MATS8v($\pm$0.996)+53.444X0Av($\pm$2.472)-208.490X4Av($\pm$15.104)+13.790MATS6p($\pm$1.684)+2.971MATS3v($\pm$0.704)+0.004VOL($\pm$0.001) +3.585ElectroNEG($\pm$0.933)-1.411ASP($\pm$0.562)-29.366($\pm$1.986)	39	0.96	199.06	0.80	0.28	0.19	0.94
PCR	3	PCI$_{50}$=0.805FAC4($\pm$0.063)-0.748FAC8($\pm$0.069)+0.555FAC3($\pm$0.060)-0.293FAC2($\pm$0.058)-0.167FAC9($\pm$0.067)+0414FAC14($\pm$0.080)-0.316FAC7($\pm$0.073)+0.215FAC5($\pm$0.059)-0.239FAC11($\pm$0.105)+6.214($\pm$0.063)	39	0.91	60.532	0.78	0.35	0.36	0.93
GA-PLS	4	PCI$_{50}$=210.547PW5($\pm$14.859)-90.006MATS3m($\pm$4.650)+17.837MATS6v($\pm$1.229)-95.430PW2($\pm$5.648)-4.180GATS8e($\pm$0.326)-1.454TI2($\pm$0.187)-2.107ASP($\pm$0.496)+0.561nN(CO)$_2$($\pm$0.160)+3.056HARDNESS($\pm$0.978)+146.819($\pm$6.318)	39	0.98	195.22	0.89	0.31	0.21	0.98

[1] Number of molecules of training set used to derive the QSAR models
R^2: Regression Coefficient for Calibration set
Q^2: Regression Coefficient for Leave One Out Cross Validation
RMS$_{cv}$: Root Mean Square Error of cross validation
R^2p: Regression Coefficient for prediction set

results of QSAR models are summarized in Table 2. The statistical parameters calculated for each target such as R^2, correlation coefficient (R^2p) of the test set, standard error (SE), F at specified degrees of freedom, Q^2 and RMScv were used for validating the goodness of fit of the resulted QSAR equations are represented in Table 2. Equation 1 (in Table 2) was selected as the best equation in the MLR model. The selected variables in Table 2 demonstrate that topological (PW5), 2D autocorrelations (GATS1e, GATS6v, MATS3m, MATS8p, MATS2e, MATS8m) and geometrical (G(N..N), ASP) descriptors affect the anti-proliferative activity of the studied compounds against SKBR-3 breast cancer cell line.

FA-MLR and PCRA

Table 3 shows the nine-factor loadings of the variables (after VARIMAX rotation) for the compounds based on their anti-proliferative activity on SKBR-3 (factor 2, 3, 4, 5, 7, 8, 9, 11 and 14). As it is observed in the table, about 68% of variances in the original data matrix could be explained by selected nine-factors.

Table 3 revealed that descriptors such as PW2, MATS8m, MATS8p, GATS6v, GATS1e and ASP are the highest loading values for factor 2. The highest loading values for factor 3 are associated with TI2, MATS2e and PW5 whereas GATS1e, PW2 and $nN(CO)_2$ are the highly loaded descriptors of factor 4. As it was shown in Table 3, the highest loading values for factor 7 is associated with MATS3m, for factor 8 is associated with MATS8v, for factor 9 is associated with

MATS6p whereas ASP, G(N..N) and hardness are the highly loaded descriptors of factor 11. Table 3 revealed that factors 2 and 3 are moderately loaded with the anti-proliferative activity of the studied compounds against SKBR-3 cell line. Interestingly, the former possessed the highest loadings from 2D autocorrelations (MATS8m, MATS8p, GATS6v, GATS1e), topological (PW2) and geometrical (ASP) whereas the latter is containing the information from topological (TI2, PW5) and 2D autocorrelations (MATS2e) descriptors. The subsequent FA-MLR equation using highly loaded descriptors is shown in Table 2, Eq.2.

PCRA

Equation 3 is obtained from factor scores as predictor parameters in multiple regression equations via forwarding selection method (PCRA). Unlike selected descriptors, factor scores include information from different descriptors. Thus the risk for data missing is reduced. By principle component method, nine factors scores (Table 3) were used as independent parameters for developing QSAR equations. The variables used in Eq. 3 shows statistical quantities similar to those obtained by the FA-MLR method. But, it indicates partly higher calibration and lower cross-validation statistics concerning Eq.2.

In Table 2, Factor score 2 signifies the importance of PW2, MATS8m, MATS8p, GATS6v, GATS1e and ASP descriptors. Factor score 3 indicates the importance of TI2, PW5, MATS2e descriptors. Factor score 4 indicates the importance of GATS1e, PW2, $nN(CO)_2$

Table 3. Factor loadings of some significant descriptors after VARIMAX rotation.

Descriptor	F2	F3	F4	F5	F7	F8	F9	F11	F14	Communalities
Vol.	-0.301	-0.061	0.159	0.402	0.036	0.196	0.055	0.017	0.007	0.980
TI2	-0.439	0.536	0.153	-0.063	0.198	0.198	0.020	-0.016	-0.057	0.994
X0Av	0.443	-0.110	0.269	0.271	-0.136	0.041	0.093	0.053	0.016	0.996
X4Av	0.309	-0.317	-0.106	0.322	-0.048	-0.011	-0.038	-0.181	0.111	0.987
PW2	0.569	-0.067	-0.448	-0.087	-0.272	-0.067	0.003	-0.002	-0.024	0.997
PW5	-0.211	0.813	0.381	0.212	0.085	-0.016	-0.051	-0.047	-0.042	0.990
MATS3m	0.235	0.053	-0.437	-0.426	0.551	0.328	-0.071	-0.251	0.049	0.989
MATS8m	0.761	-0.250	-0.121	-0.229	0.327	0.148	0.073	0.127	-0.052	0.989
MATS6v	0.231	0.301	0.224	-0.197	0.030	-0.415	-0.492	-0.064	0.247	0.982
MATS8v	0.414	0.013	-0.172	-0.048	-0.020	0.581	0.185	0.015	-0.533	0.993
MATS2e	0.373	-0.445	-0.751	0.150	0.088	-0.029	-0.035	-0.041	-0.076	0.994
MATS6p	0.103	0.394	-0.035	-0.370	-0.095	-0.212	-0.508	-0.121	0.225	0.966
MATS8p	0.594	-0.084	-0.138	0.068	-0.105	0.478	0.145	0.019	-0.242	0.986
GATS6v	-0.735	-0.151	-0.180	-0.146	-0.176	0.427	0.278	0.000	-0.128	0.989
GATS1e	-0.560	0.185	0.685	-0.288	0.049	-0.208	0.132	0.098	-0.001	0.989
GATS8e	-0.422	0.086	0.188	0.234	-0.082	0.173	-0.076	-0.018	-0.100	0.959
ASP	0.546	0.042	-0.058	0.425	0.050	-0.209	0.014	0.363	0.022	0.834
G(N..N)	0.044	0.252	-0.252	-0.067	-0.270	0.007	0.172	0.268	0.213	0.548
nN(CO)2	0.171	-0.292	-0.433	0.085	-0.470	0.314	0.264	0.138	0.187	0.992
Hardness	0.194	0.397	0.029	-0.021	-0.086	-0.350	-0.228	-0.205	-0.182	0.984
ElectroNeg	0.293	0.310	-0.065	-0.068	-0.121	-0.278	-0.183	-0.191	-0.188	0.979

descriptors. Factor score 5 indicates the importance of VOL, MATS3m and ASP descriptors. Factor score 7 indicates the importance of MATS3m. Factor score 8 indicates the importance of MATS8v. Factor score 9 indicates the importance of MATS6p. Factor score 11 indicates the importance of ASP, G(N..N) and hardness descriptors. Factor score 14 signifies the importance of MATS8v descriptor.

GA-PLS

In PLS analysis, the descriptors data matrix is decomposed to orthogonal matrices with an inner relationship between the dependent and independent variables. Therefore, unlike MLR analysis, the multi-colinearity problem in the descriptors is omitted by PLS analysis. Since a minimal number of latent variables are used for modeling in PLS; this modeling method coincides with noisy data better than MLR. So, many different GA-PLS runs were done using the different initial set of populations. The statistical parameters calculated for this model are shown in Table 2.

Table 2 represents that a combination of 2D autocorrelations (MATS3m, MATS6v, GATS8e), topological (TI2, PW2, PW5), geometrical (ASP), quantum (hardness) and functional (nN(co)$_2$) descriptors have been selected by GA-PLS to account for the anti-

proliferative activity of the studied compounds against SKBR-3 cell line. In this table, Eq.4 with high statistical quality parameters was obtained from the pool of calculated descriptors (i.e., $R^2 = 0.98$ and $Q^2 = 0.89$) and, the predictive R^2 value for the test set was found to be 0.98. Table 2 shows that none of the suggested QSAR models were obtained by chance and the best set of calculated descriptors was selected by genetic algorithm because of its greatest statistical parameters. Therefore the best predictive results were observed from this model.

The most convenient GA-PLS model that resulted in the best fitness contained 62 indices. The PLS estimate of coefficients for the descriptors are given in Figure 1. As it observed, a combination of chemical, topological, geometrical, quantum, 2D-autocorrelations and functional descriptors have been selected by GA-PLS to account the anti-proliferative activity of methanesulfonamide derivatives against SKBR-3 cell line. VIP was calculated for each descriptor to determine the importance of 62 selected GA-PLS descriptors. Figure 2 shows the VIP analysis of PLS equation. VIP shows that 25 descriptors with VIP>1.0 such as PW5, nRORph, MATS3m, MATS6v, X2sol, logP, GATS7v, MATS2v, X5sol, nNCO2, MATS8m, MATS7v, GATS8v, X3Av, GNar‹ pol, LP1, GATS6v,
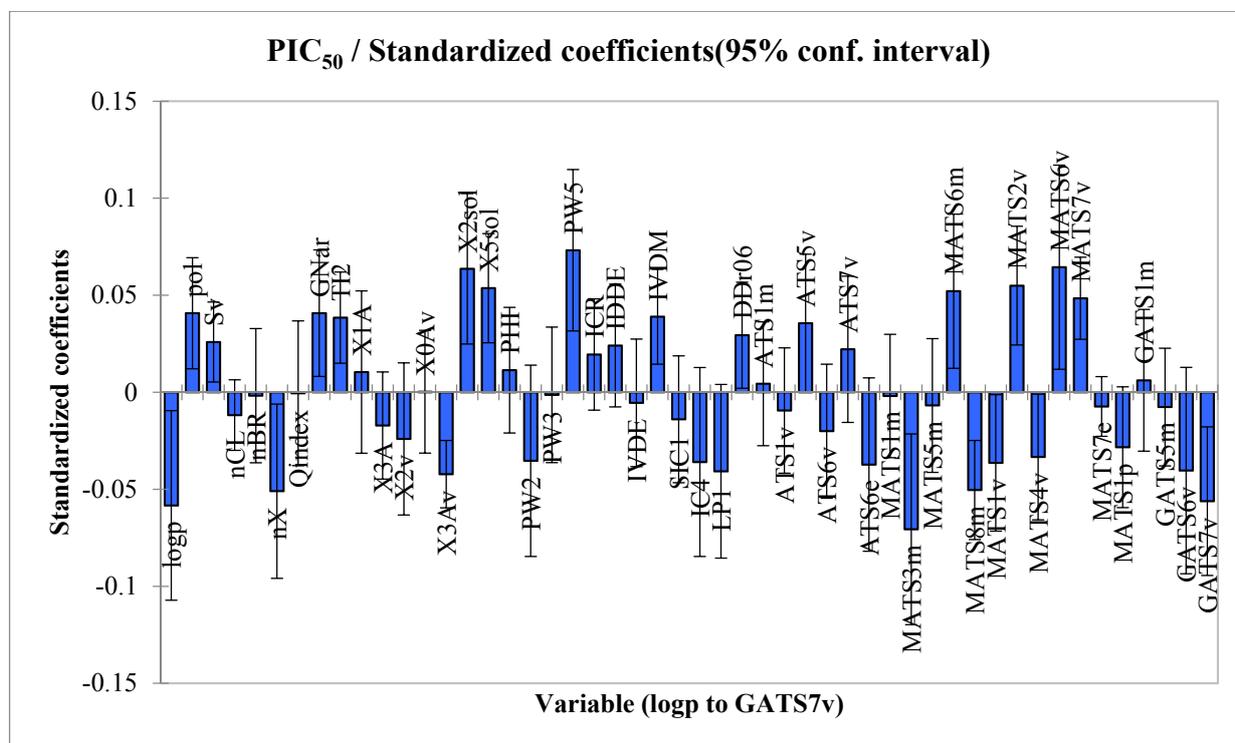

**Figure 1.** Regression coefficients for the variables used in GA-PLS model.
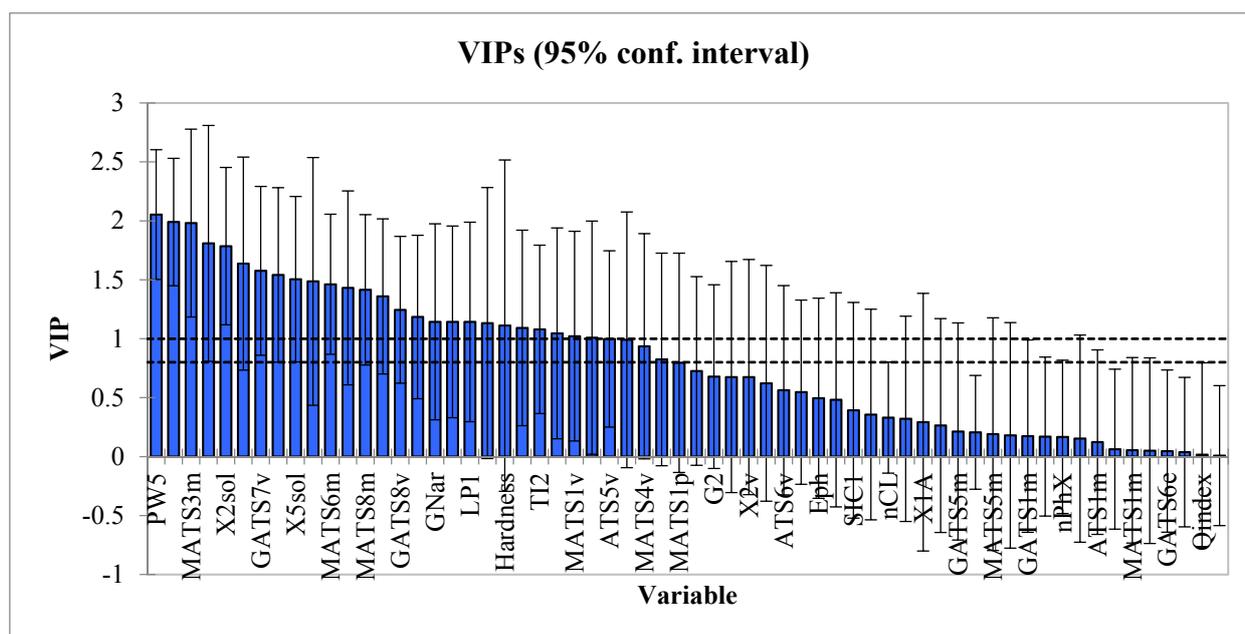
Figure 2. Variable importance in the projection (VIP) for the variables used in GA-PLS model.

hardness , IVDM, Ti2, ATS6e and MATS1v are the most important indices in the QSAR equation derived by PLS analysis. Also, descriptors such as IC4, ATS5v, PW2, MATS4v and DDr06 with $0.8 < VIP < 1.0$ are moderately influential and descriptor with $VIP < 0.8$ which was shown in Figure 2 are less influential.

The brief description of the descriptors which were inserted in the QSAR models is listed in Table 4.

In silico screening

Instead of time-consuming and expensive *in vivo* experiments, it is possible to apply *in silico* screening in initial steps of drug development because it can accelerate the speed of discovery, anticipate and explore new pharmaceutical agents. To explore and detect active compounds among molecular databases, a strong technique, virtual screening, may be used. This technique is used through deletion, insertion, and

Table 4. Definitions of molecular descriptors present in the models

| Descriptors	Brief description
MATS2e	Moran autocorrelation - lag 2 / weighted by atomic Sanderson electronegativities
MATS3m	Moran autocorrelation- lag 3 / weighted by atomic masses
MATS8m	Moran autocorrelation - lag 8 / weighted by atomic masses
MATS6p	Moran autocorrelation - lag 6 / weighted by atomic polarizabilities
MATS8p	Moran autocorrelation - lag 8 / weighted by atomic polarizabilities
MATS3V	Moran autocorrelation - lag 3 / weighted by atomic van der Waals volumes
MATS6V	Moran autocorrelation - lag 6 / weighted by atomic van der Waals volumes
MATS8V	Moran autocorrelation - lag 8 / weighted by atomic van der Waals volumes
GATS1e	Geary autocorrelation - lag 1 / weighted by atomic Sanderson electronegativities
GATS8e	Geary autocorrelation - lag 8 / weighted by atomic Sanderson electronegativities
GATS6v	Geary autocorrelation - lag 6 / weighted by atomic van der Waals volumes
PW2	Path/walk 2- Randic shape index
PW5	Path/walk 5- Randic shape index
G(N..N)	Sum of geometrical distances between N..N
ASP	Asphericity
TI2	Second Mohar index TI2
Nn(CO)₂	Number of imides
X0Av	Average valence connectivity index chi-0
X4Av	Average valence connectivity index chi-4
Hardness	Quantum chemical indices of hardness
VOL	Molecular volume
ElectroNeg	Electronegativity
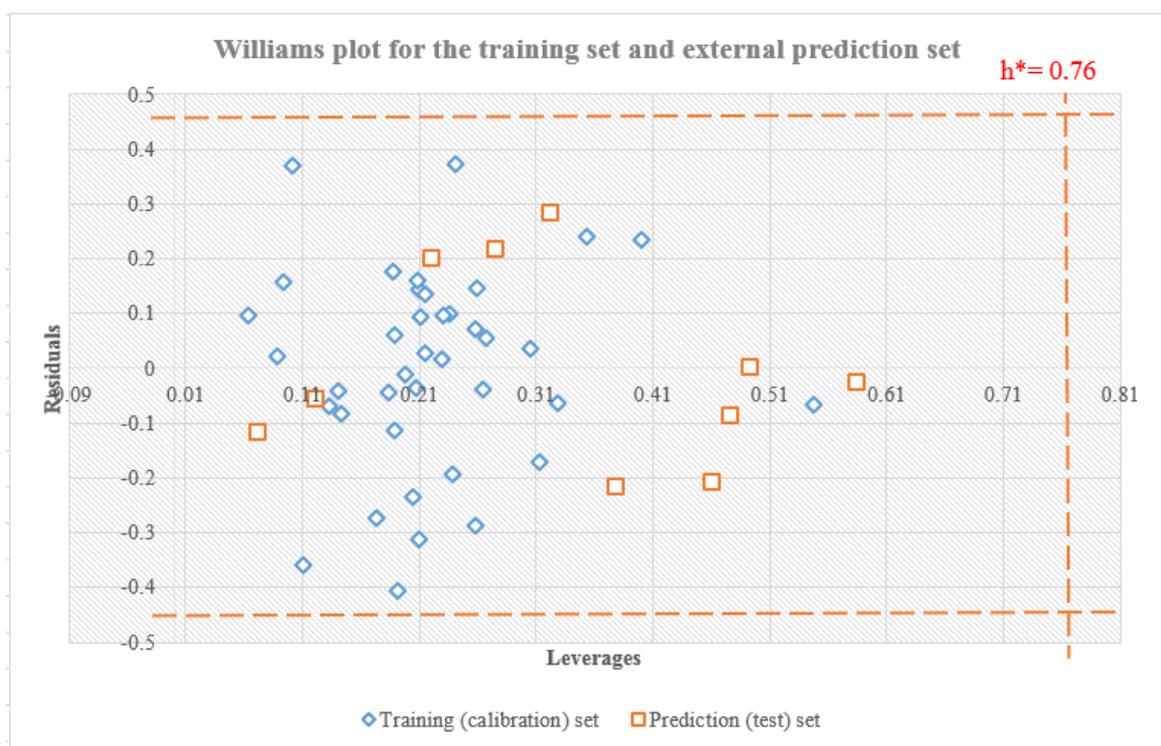
Figure 3. Williams plot for the calibration set and external prediction set.

substitution of different substitutes on the parent molecules and it is able to investigate the influences of the structural modifications on the biological activity [19, 25].

For the use of the model in screening new compounds, the domain application of QSAR model was determined. The applicability domain (AD) of QSAR model was applied to verify the prediction reliability, to recognize the troublesome compounds and

to predict the compounds with an acceptable activity that falls within this domain. We employed the important descriptors selected from GA-PLS model (as the best-studied model because of its greatest statistical parameters) for designing new active compounds.

According to the analysis done on AD model in the Williams plot of the MLR model (Figure 3) using the whole dataset, it was demonstrated that neither one of compounds is an obvious outliner. As it was depicted in

Table 5. Designed compounds based on *in silico* screening and their predicted activities and docking binding energies on Hsp27 and Tubulin.

Name	R1	R2	R3	pIC$_{50}$ pred	leverage	Hsp27 ΔE (kcal/mol)	Tubulin ΔE (kcal/mol)
1b			-	9.047	0.286	-6.98	-8.74
2b			-	9.051	0.356	-8.24	-8.20

231

Figure 3, none of the compounds have leverage (h) values greater than the threshold leverages (h*). The warning leverage (h*), was found to be 0.76. To the best of our knowledge, the compounds that had a standardized residual more than three times of the standard deviation units were considered to be outliers. For both the calibration set and test set of targets, the presented model matches the high-quality parameters with good fitting power and the capability of assessing external data. Moreover, almost all of the compounds were within the applicability domain of the proposed model and were evaluated accurately. While chemicals with a leverage value higher than h* were considered to

be influential or high leverage chemicals [29, 30].

In the studied QSAR model, the *in silico* screening should be used to design new lead compounds with improved the potential anti-proliferative activity against SKBR-3 cell line. Then, the *in silico* screen was applied by substituting diverse groups in different places using bioisosterism rules. Table 5 shows the results of *in silico* screening. 25 novel compounds were designed, and their predicted activities for anti-proliferative activity based on GA-PLS equation, as well as their docking binding energies on HSP27 and tubulin were obtained (Table 5). Leverage values show that all of the designed compounds were within the applicability domain.

Table 5. Cntd
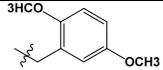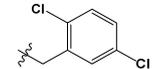
No.	R_1	R_2	R_3				
3b	3HCO-/OCH3 benzyl	pyridine	-	9.062	0.395	-6.97	-7.93
4b	Cl/Cl benzyl	CH_3	CH_3	9.021	0.787	-7.96	-6.93
5b	n-C_6H_{13}	CH_3	CH_3	8.001	0.749	-7.09	-7.23
6b	CF3 benzyl	CH_3	CH_3	8.195	0.672	-7.73	-6.17
7b	3HCO-/OCH3 benzyl	CH_3	CH_3	11.979	0.735	-7.49	-6.06
8b	3HCO-/OCH3 benzyl	CH_3	H	11.254	0.702	-7.37	-6.17
9b	3HCO-/OCH3 benzyl	CF_3	CH_3	7.433	0.537	-7.20	-5.91
10b	dimethylbenzyl	CF_3	CH_3	5.932	0.655	-7.97	-6.02
11b	3HCO-/OCH3 benzyl	CF_3	CH_3	9.974	0.351	-6.85	-6.07
12b	3HCO-/OCH3 benzyl	CF_3	H	12.478	0.383	-6.54	-5.28
13b	Cl/Cl benzyl	CH_3	CH_3	9.075	0.337	-7.57	-8.32
14b	3HCO-/OCH3 benzyl	CH_3	CH_3	12.012	0.373	-7.10	-7.40
15b	3HCO-/OCH3 benzyl	CH_3	H	10.924	0.384	-6.89	-7.16
16b	dimethylbenzyl	CH_3	CH_3	10.015	0.361	-7.45	-7.83
17b	Cl/Cl benzyl	I	CH_3	9.473	0.359	-7.87	-6.75
18b	n-C_6H_{13}	I	CH_3	8.894	0.418	-6.77	-6.40
19b	CF3 benzyl	I	CH_3	8.414	0.336	-7.56	-6.30
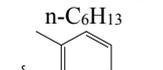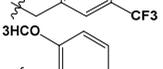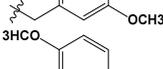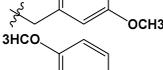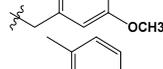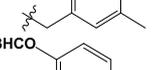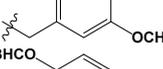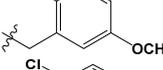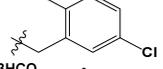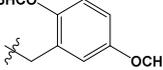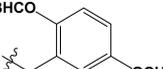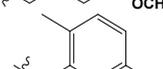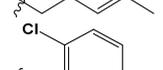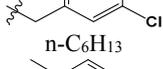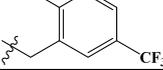
Table 5. Cntd
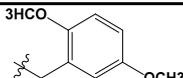
20b	$3HCO$ ... OCH_3	I	CH_3	13.184	0.314	-7.22	-5.54
21b	$3HCO$ ... OCH_3	I	H	12.245	0.450	-7.19	-5.96
22b		I	CH_3	11.156	0.472	-7.63	-6.63
23b	$3HCO$ ... OCH_3	OCH_3	CH_3	11.395	0.405	-6.83	-7.59
24b	$3HCO$ ... OCH_3	OCH_3	H	11.155	0.368	-6.87	-7.38
25b		OCH_3	H	9.490	0.422	-7.21	-5.67
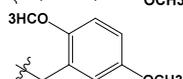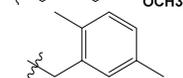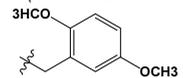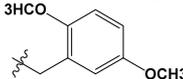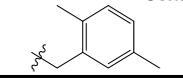
Among these molecules, 9 compounds such as 7b, 8b, 12b and 20-24b showed the highest pIC_{50} and hence are the best anti-proliferative compounds. These compounds have a good potentially for becoming potential anticancer agent.

In Figure 4, the data for predicted activities are plotted against experimental ones to be taken into account the cross-validated prediction results. As it was mentioned above, the least scattering of data was obtained from GA-PLS model. High regression ratio (R^2 = 0.9796) in this plot shows good agreement between the experimental activity and cross-validated predicted values of activity.

Docking Studies

Here, molecular docking simulations were performed on 49 compounds of the dataset as well as 25 designed compounds, to elucidate their interactions and to gain some insight into their molecular binding mode with Hsp27 and tubulin. The results obtained from molecular docking containing the estimated free binding energy values (ΔG_{bind}) for the best position of the docked compounds expressed in Kcalmol^{-1}, which are summarized in Table 1 for data set compounds and Table 5 for designed compounds, along with the corresponding favorable interactions with the key amino acid residues at the active site of Hsp27 and tubulin are depicted in Figures 5-9.


$$y = 0.9796x + 0.1259$$
$$R^2 = 0.9796$$
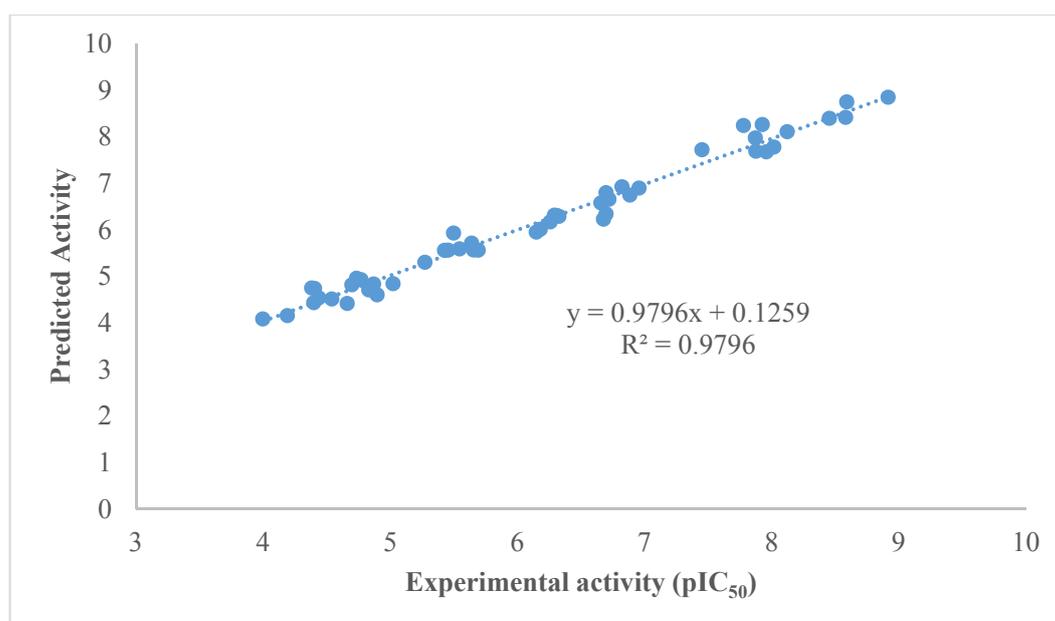
Figure 4. Plots of cross-validated predicted values of activity by GA-PLS against the experimental values

Table 1 shows that, the ΔG_{bind} values of the best-docked poses of these compounds are within the range of -6.37 to -9.09 Kcal.mol^{-1} for Hsp27 and -6.52 to -10.34 Kcal.mol^{-1} for tubulin binding. The best docking binding energies in binding to Hsp27 belongs to the compound 35a, whereas the compound 5a shows the great docking binding energies in binding to tubulin in compared to the others. As it was shown in Table 5, the best docked poses of designed compounds are within the range of -6.54 to -8.24 Kcal.mol^{-1} for Hsp27 and -5.54 to -8.74 Kcal.mol^{-1} for tubulin. Here, The best docking binding energies in binding to Hsp27 is belong

to the compound 2b, whereas the compound 1b shows the great docking binding energies in binding to tubulin.

Types of the interaction of these compounds to both of these targets were also explored. As it was shown in Figure 5A, the methyl group attached to sulfonamide moiety of compound 7a is involved in hydrophobic interaction with residue Phe104 in the active site of Hsp27. The oxygen groups of sulfonamide moiety interact via acceptor hydrogen bonds with residues Val101, and His103. The NH group of the amide bond of this compound is involved in hydrogen bond interaction with residue Cys137. The methyl attached to
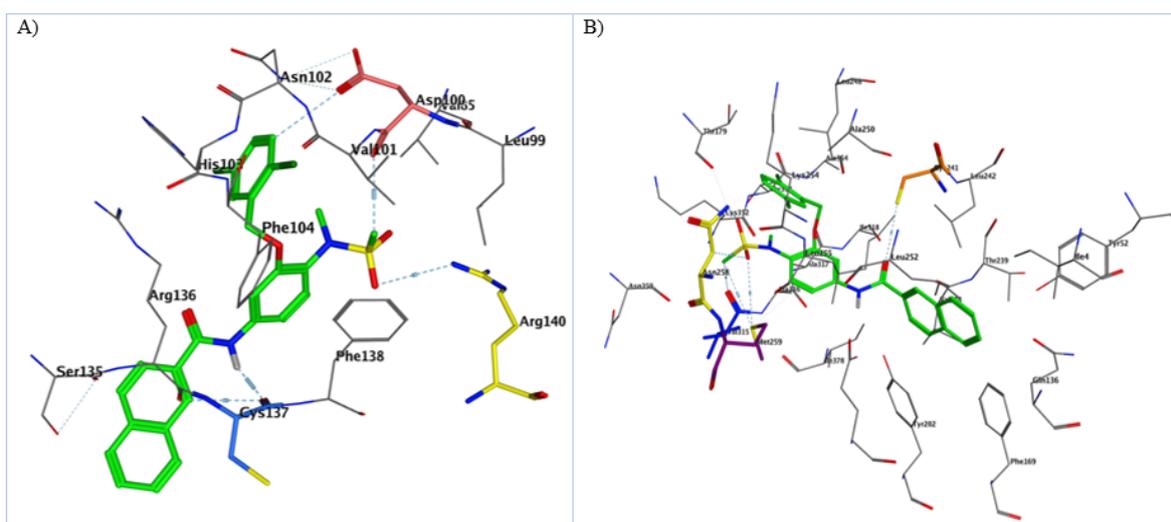
Figure 5. A) The structure of 5a surrounded by the key residues in the active site of HSP27. B) Molecular docking simulation studies of the interaction between 5a and tubulin.
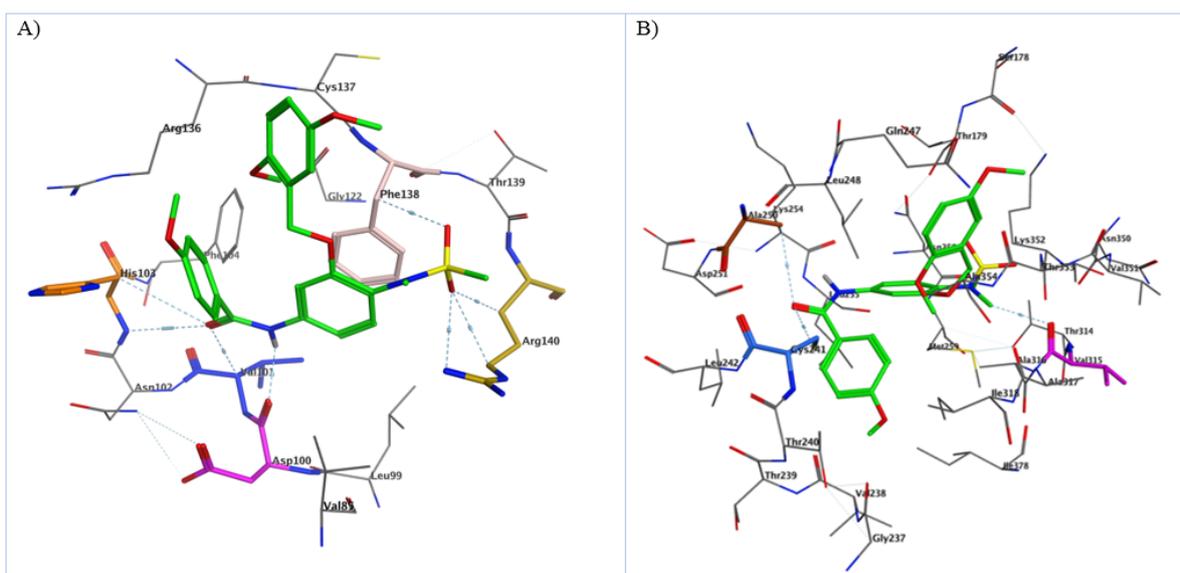

Figure 6. A) The structure of 20a surrounded by the key residues in the active site of HSP27. B) Molecular docking simulation studies of the interaction between 20a and tubulin.

the nitrogen of sulfonamide is involved in an interaction with residue His103. The chloro group on the phenyl ring is involved in a weak hydrogen bond with Thr139. The most favorable energetically conformation of the docked pose of 7a in interaction with tubulin is shown in Figure 5B. The carbonyl group of this compound interacts via acceptor hydrogen bonding with Cys241. The methyl group attached to sulfonamide moiety of this compound is involved in interaction with residue Val315. The oxygen group of sulfonamide moiety interacts via acceptor hydrogen bonds with residues Met259 and Asn258. There also exist a weak hydrogen bonds between chloro group on the phenyl ring with Thr353.

Types of interactions of compound 20a with both targets are depicted in Figure 6. The oxygen groups of sulfonamide moiety interact through acceptor hydrogen bonding with residues Phe138 and Arg140. The carbonyl group of the amide bond of this compound is involved in hydrogen bond interaction with residues Val101, and His103. The NH group of the amide bond interact via hydrogen bonding with residue Asp100. The π-π (arene-arene) interaction between the phenyl bearing sulfonamide moiety and residue Phe138 is also observed. Figure 6B shows the interaction of compound 20a with tubulin. The methyl group attached to sulfonamide moiety of this compound interacts with residue Met259. The oxygen group of sulfonamide moiety is involved in acceptor hydrogen bonding with residue Asn258. The carbonyl group of the amide bond of this compound is involved in hydrogen bond interaction with residues Cys241. The methyl attached

to the nitrogen of sulfonamide interacts with residue Val315.

As indicated in Figure 7A, the oxygen groups of sulfonamide moiety of compound 36a interact with residues Arg136 in the active site of Hsp27. The NH group of the amide bond interact via hydrogen bonding with residue Cys137. There existed an arene-hydrogen bond interaction between the benzodioxole group with residue Thr139. The π-π interaction between the phenyl of benzyloxy moiety and residue Phe138 is also observed. The most favorable energetically conformation of the docked pose of 36a in interaction with tubulin is shown in Figure 7B. It interacts via the oxygen group of sulfonamide moiety with residues Asn258 and Leu255. The methyl attached to the nitrogen of sulfonamide is involved in an interaction with residue Met259.

The carbonyl group of the amide bond of this compound is involved in hydrogen bond interaction with residues Thr353. The oxygen of dioxolane group is forming a hydrogen bond with residue Arg221.

Types of interaction of the designed compounds based on *in silico* screening using GA-PLS model with both targets were also determined.

As depicted in Figure 8A, the oxygen groups of sulfonamide moiety of compound 7b interact with residues Phe138 and Thr139 in the active site of Hsp27. The nitrogen of phenanthridine ring in this compound is involved in acceptor hydrogen bonding with Arg140. The hydroxyl group attached to the phenanthridine ring of this compound interacts via hydrogen bonding with Asp100. The π-π interaction between the pyridine ring
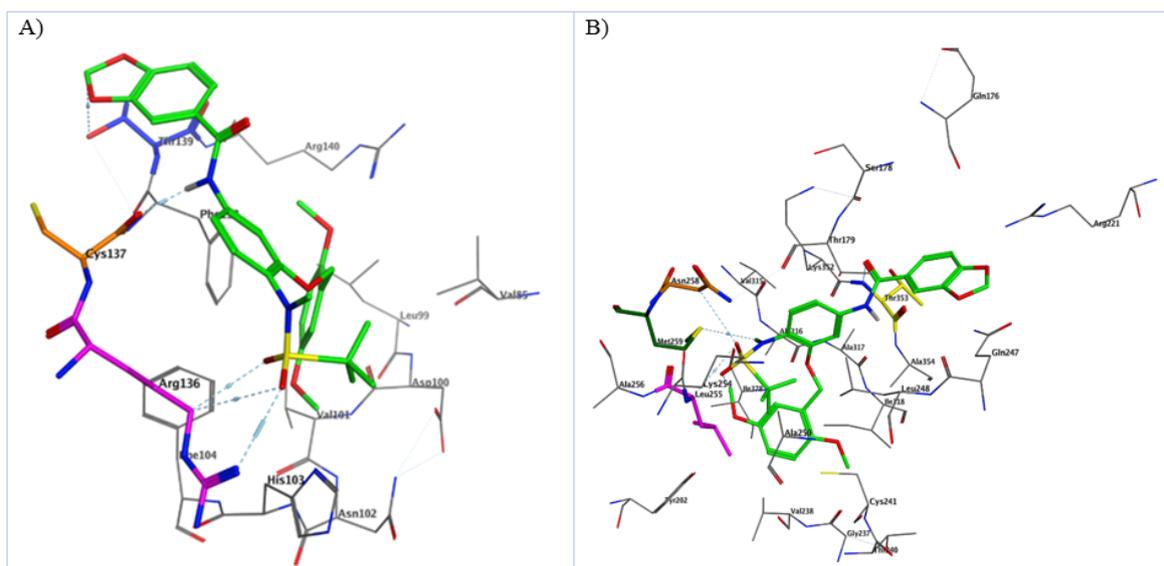

Figure 7. A) The structure of 36a surrounded by the key residues in the active site of HSP27. B) Molecular docking simulation studies of the interaction between 36a and tubulin.

Figure 8. A) The structure of 7b surrounded by the key residues in the active site of HSP27. B) Molecular docking simulation studies of the interaction between 7b and tubulin.


Figure 9. A) The structure of 14b surrounded by the key residues in the active site of HSP27. B) Molecular docking simulation studies of the interaction between 14b and tubulin.

of phenanthridine moiety and residue Phe138 is also observed. The interaction of 7b with tubulin is shown in Figure 8B. The oxygen group of sulfonamide interacts with residues Asn258. The methyl group attached to sulfonamide moiety of this compound interacts with residue Thr179. The methoxy group on its benzyloxy group is involved in hydrogen bonding with Val238. The hydroxyl group attached to the phenanthridine ring interacts via hydrogen bonding with Asn350.

As it was shown in Figure 9A, the oxygen groups of sulfonamide moiety of compound 14b interact with residues Phe138 and Thr139. The hydroxyl group on its

quinoline ring interacts via hydrogen bonding with Asp100. There existed an arene-hydrogen bond interaction between the benzyloxy group with residue Arg140. The π-π interaction between the pyridine of quinoline moiety with residue Phe138 is also observed. The most favorable energetically conformation of the docked pose of 14b in interaction with tubulin is shown in figure 9B. There existed a hydrogen bond interaction between the oxygen group of its sulfonamide moiety with residue Cys241. The methyl attached to the nitrogen of sulfonamide is involved in an interaction with residue Cys241.

The hydroxyl group on the quinoline ring interacts via hydrogen bonding with Ser178. The methoxy group on the benzyloxy group is involved in hydrogen bonding with Asn258.

The results obtained from this docking study indicate that the important amino acids inside the active site cavity that are in charge of essential interactions with HSP27 are Arg140, Thr139, Phe138, Cys137, Arg136, Phe104, His103, Val101, and Asp100. And the most important amino acids inside the active site cavity being responsible for essential interactions with tubulin are Asn258, Val238, Cys241, Asn350, Leu255, Met259, Val315, Thr353, Arg221, Thr179 and Ser178.

One of the great importance steps in a molecular docking studies should be exploring the validity of the process. The application of relative operating characteristic curve (ROC) as a helpful metric tool to weigh the validity of docking procedures was first reported by Triballeau et al. in computational medicinal chemistry [44]. Nowadays, it was widely used as a validating procedure [45]. First of all, about 71 Hsp27 inhibitors and 65 tubulin inhibitors were retrieved from ChEMBL database as SMILES format [46-48]. The structures based on their experimental activities categorized into two subsets of ligands (active) and decoys (inactive). 19 ligands and 52 decoys for Hsp27 and 21 ligands and 44 decoys for tubulin were generated. Subsequently, through a shell script using openbabel 2.3.2, the primary 3D generation of the structures as the mol2 format was provided [49]. For all structures, the ionization states at PH=7 were also calculated. Using batch scripting in windows operating
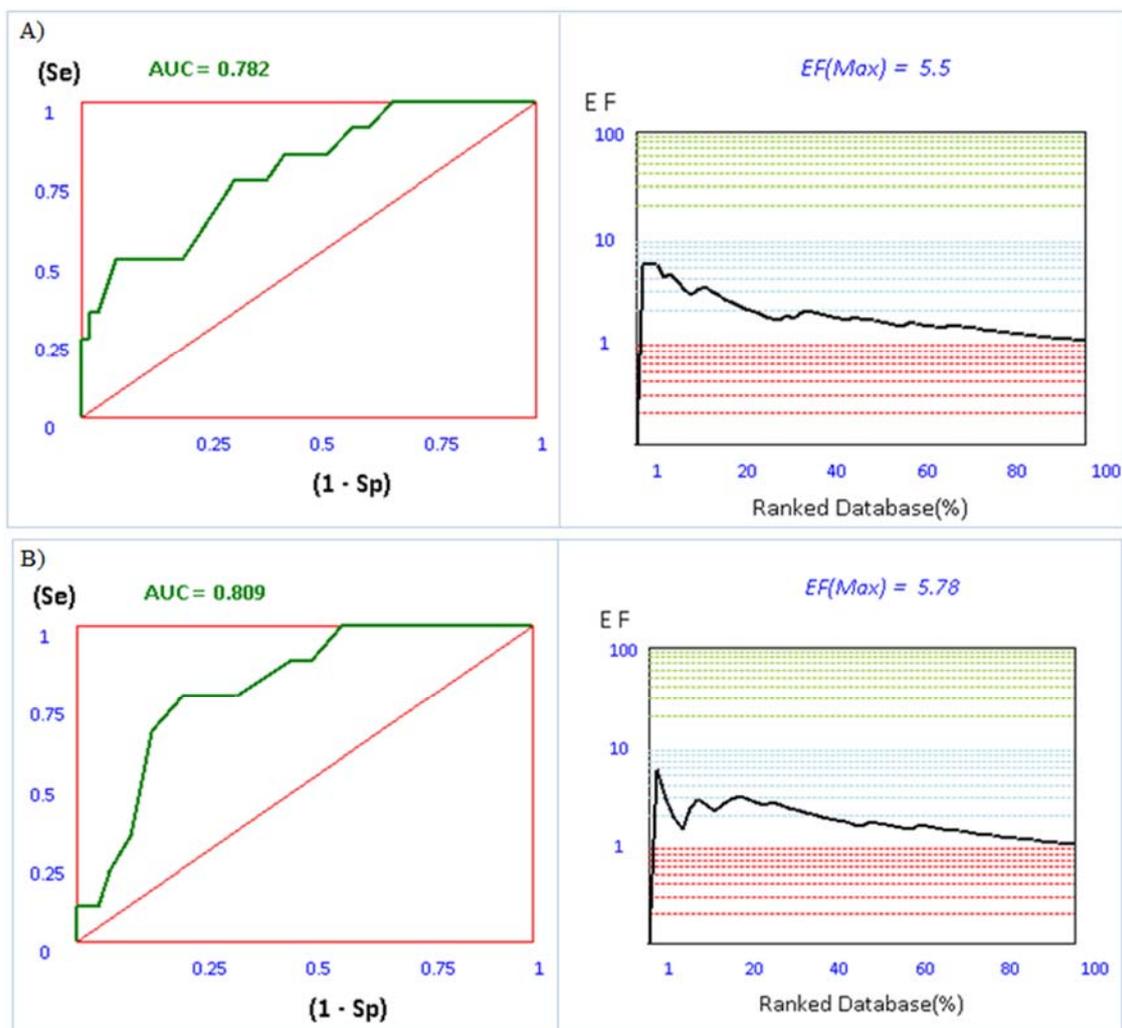

Figure 10. A) ROC and EF diagrams for docking with HSP27. B) ROC and EF diagrams for docking with tubulin. ROC value is the area under the curve (AUC) for the plot of the true positive rate (TPR or sensitivity) against the false positive rate (FPR or 1- specificity) at various threshold settings.

system, the shell script was obtained. A valid docking should be able to discriminate between inactive decoys and active ligands. ROC value is the area under the curve (AUC) for the plot of the true positive rate (TPR or sensitivity) against the false positive rate (FPR or 1-specificity) at various threshold settings. The ROC curve is thus the sensitivity as a function of 1-specificity. The AUC for ROC is calculated by trapezoidal integration method as implemented in our *in house* ROC application [45]. The more ROC$_{AUC}$ value means that the docking protocol is more able to discriminate between ligands and decoys. As it was shown in Figure 10, the AUC of 0.782 for Hsp27 and 0.809 for tubulin shows that our applied docking protocol was a validated protocol.

Another tool which was used to evaluate the efficiency and validity of docking protocol was Enrichment Factor (EF$_{max}$). Its calculations were based on the Li et al. work [50]. EF$_{max}$ factor in comparison to ROC curves is highly dependent to the number of actives in a data set [44]. It means that early enrichment can be easily obtained if the number of active ligands is increasing in a dataset. Hence, the ROC should be considered most importantly in the validation of docking methods.

Conclusion

Here, with the aim of comparative QSAR analysis such as MLR, GA-PLS, FA-MLR and PCRA, quantitative relationships between molecular structure and anti-proliferative activity against SKBR-3 breast cancer cell line for a series of methanesulfonamide derivatives as dual HSP27 and tubulin inhibitors were acquired. Using different criteria such as cross-validation, validation through Y-randomization and root mean square error of cross-validation (RMSECV), the reliability, accuracy, and predictability of the proposed models were assessed. The role of different descriptors such as charge, geometrical, topological, functional, 2D autocorrelations, quantum and chemicals on anti-proliferative activity was obtained. A comparison between the different statistical methods employed indicated that GA-PLS represented superior results. According to the developed QSAR model, *in silico* screening was applied and new compounds such as 7b, 8b, 12b and 20-24b with potential anti-proliferative activity were designed and suggested for synthesis. Molecular docking simulations were also performed on these compounds to elucidate their interactions and to gain some insight into their molecular binding site and binding mode with HSP27 and tubulin as their targets.

References
1. Zhu Z., Xu X., Yu Y., Graham M., Prince ME., Carey TE., and Sun D. Silencing Heat Shock Protein 27 Decreases Metastatic Behavior of Human Head and Neck Squamous Cell Cancer Cells in Vitro. *Mol. Pharm.* **7**: 1283-1290 (2010).
2. Oya-Ito T., Naito Y., Takagi T., Handa O., Matsui H., Yamada M., Shima K., and Yoshikawa T. Heat-shock protein 27 (Hsp27) as a target of methylglyoxal in gastrointestinal cancer. *Biochim. biophys. acta, Mol. basis dis.* **1812**:769-781 (2011).
3. Hsu HS., Lin JH., Huang WC., Hsu TW., Su K., Chiou SH., Tsai YT., and Hung SC. Chemoresistance of lung cancer stemlike cells depends on activation of Hsp27. *Cancer.* **117**: 1516-1528 (2011).
4. Singh MK., Sharma B., and Tiwari PK. The small heat shock protein Hsp27: Present understanding and future prospects. *J. Therm. Biol.* **69**:149-154 (2017).
5. Kimura H., Mikami D., Kamiyama K., Sugimoto H., Kasuno K., Takahashi N., Yoshida H., and Iwano M. Telmisartan, a possible PPAR-δ agonist, reduces TNF-α-stimulated VEGF-C production by inhibiting the p38MAPK/HSP27 pathway in human proximal renal tubular cells. *Biochem. Biophys. Res. Commun.***454**:320-327 (2014).
6. Temajo NO., Howard N. The virus-induced HSPs regulate the apoptosis of operatus APCs that results in autoimmunity, not in homeostasis. *Autoimmun Rev.***13**: 1013-1019 (2014).
7. Jego G., Hazoumé A., Seigneuric R., and Garrido C. Targeting heat shock proteins in cancer. *Cancer lett.* **332**:275-285 (2013).
8. Gonzalez-Mejia M., Voss O., Murnan E., and Doseff A. Apigenin-induced apoptosis of leukemia cells is mediated by a bimodal and differentially regulated residue-specific phosphorylation of heat-shock protein–27. *Cell. Death. Dis.* **1**: e64 (2010).
9. Baylot V., Andrieu C., Katsogiannou M., Taieb D., Garcia S., Giusiano S., Acunzo J., Iovanna J., Gleave M., and Garrido C. OGX-427 inhibits tumor progression and enhances gemcitabine chemotherapy in pancreatic cancer. *Cell. Death. Dis.* **2**: e221 (2011).
10. Goloudina AR., Demidov ON., and Garrido C. Inhibition of HSP70: A challenging anti-cancer strategy. *Cancer Lett.* **325**:117-124 (2012).
11. Lianos GD., Alexiou GA., Mangano A., Mangano A., Rausei S., Boni L., Dionigi G., Roukos DH. The role of heat shock proteins in cancer. *Cancer letters.* **360**:114-118 (2015).

12. Langer R., Ott K., Specht K., Becker K., Lordick F., Burian M., Herrmann K., Schrattenholz A., Cahill MA., and Schwaiger M. Protein expression profiling in esophageal adenocarcinoma patients indicates association of heat-shock protein 27 expression and chemotherapy response. *Clin. Cancer Res.* **14**: 8279-8287 (2008).

13. Wu J., Liu T., Rios Z., Mei Q., Lin X., Cao S. Heat Shock Proteins and Cancer. *Trends Pharmacol. Sci.* **38**:226-256 (2017).

14. Heinrich J-C., Tuukkanen A., Schroeder M., Fahrig T., and Fahrig R. RP101 (brivudine) binds to heat shock protein HSP27 (HSPB1) and enhances survival in animals and pancreatic cancer patients. *J. Cancer Res. Clin. Oncol.* **137**: 1349-1361 (2011).

15. Zhong B., Lama R., Kulman DG., Li B., and Su B. Lead optimization of dual tubulin and Hsp27 inhibitors. *Eur. J. Med. Chem.* **80**:243-253 (2014).

16. Zhong B., Lama R., Smith KM., Xu Y., and Su B. Design and synthesis of a biotinylated probe of COX-2 inhibitor nimesulide analog JCC76. *Bioorg. Med. Chem. Lett.* **21**: 5324-5327 (2011).

17. Yi X., Zhong B., Smith KM., Geldenhuys WJ., Feng Y., Pink JJ., Dowlati A., Xu Y., Zhou A., and Su B. Identification of a class of novel tubulin inhibitors. *J. Med. Chem.* **55**: 3425-3435 (2012).

18. Zhong B., Chennamaneni S., Lama R., Yi X., Geldenhuys WJ., Pink JJ., Dowlati A., Xu Y., Zhou A., and Su B. Synthesis and anticancer mechanism investigation of dual Hsp27 and tubulin inhibitors. *J. Med. Chem.* **56**: 5306-5320 (2013).

19. Fereidoonnezhad M., Faghih Z., Jokar E., Mojaddami A., Rezaei Z., and Khoshneviszadeh M. QSAR, Molecular Docking and protein ligand interaction fingerprint studies of N-phenyl dichloroacetamide derivatives as anticancer agents. *Trends in Pharmaceutical Sciences.* **2**: 159-176 (2016).

20. Fereidoonnezhad M., Faghih Z., Mojaddami A., Sakhteman A., and Rezaei Z. A Comparative Docking Studies of Dichloroacetate Analogues on Four Isozymes of Pyruvate Dehydrogenase Kinase in Humans. *Indian J. Pharm Edu. Res.* **50**: S32-S8 (2016).

21. Fereidoonnezhad M., Faghih Z., Mojaddami A., Tabaei S., and Rezaei Z. Novel Approach Synthesis, Molecular Docking and Cytotoxic Activity Evaluation of N-phenyl-2, 2-dichloroacetamide Derivatives as Anticancer Agents. *J. Sci., Islamic Repub. Iran.* **27**: 39-49 (2016).

22. Frisch MJ., Trucks GW., Schlegel HB., Scuseria GE., Robb MA., Cheeseman JR., Scalmani G., Barone V., Mennucci B., Petersson GA., Nakatsuji H., Caricato M., Li X., Hratchian HP., Izmaylov AF., Bloino J., Zheng G., Sonnenberg JL., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Vreven T., Montgomery Jr. JA., Peralta JE., Ogliaro F., Bearpark MJ., Heyd J., Brothers EN., Kudin KN., Staroverov VN., Kobayashi R., Normand J., Raghavachari K., Rendell AP., Burant JC., Iyengar SS., Tomasi J., Cossi M., Rega N., Millam NJ., Klene M., Knox JE., Cross JB., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann RE., Yazyev O., Austin AJ., Cammi R., Pomelli C., Ochterski JW., Martin RL., Morokuma K., Zakrzewski VG., Voth GA., Salvador P.,

Dannenberg JJ., Dapprich S., Daniels AD., Farkas Ö., Foresman JB., Ortiz JV., Cioslowski J., and Fox DJ., *Gaussian 09*. 2009, Gaussian, Inc.: Wallingford, CT, USA.

23. A. Mauri VC., M. Pavan., R. Todeschini: D,. RAGON Software: An Easy Approach to Molecular Descriptor Calculations. *MATCH Commun. Math. Comput. Chem.* **56**: 237-248 (2006).

24. Thanikaivelan P., Subramanian V., Raghava Rao J., and Unni Nair B. Application of quantum chemical descriptor in quantitative structure activity and structure property relationship. *Chem. Phys. Lett.* **323**: 59-70 (2000).

25. Zare S., Fereidoonnezhad M., Afshar D., and Ramezani Z. A comparative QSAR analysis and molecular docking studies of phenyl piperidine derivatives as potent dual NK 1 R antagonists/serotonin transporter (SERT) inhibitors. *Comp. Biol. Chem.* **67**: 22-37 (2017).

26. Fereidoonnezhad M., Faghih Z., Mojaddami A., Rezaei Z., and Sakhteman A. A comparative QSAR analysis, molecular docking and PLIF studies of some N-arylphenyl-2,2-dichloroacetamide analogues as anticancer agents. *Iran J. Pharm. Res.* **16**: 981-998 (2017).

27. Adhikari N., Maiti MK., Jha T. Predictive comparative QSAR modelling of (phenylpiperazinyl-alkyl) oxindoles as selective 5-HT 1A antagonists by stepwise regression, PCRA, FA-MLR and PLS techniques. *Eur. J. Med. Chem.* **45**:1119-1127 (2010).

28. Bhattacharya P., and Roy K. QSAR of adenosine A 3 receptor antagonist 1, 2, 4-triazolo [4, 3-a] quinoxalin-1-one derivatives using chemometric tools. *Bioorg. Med. Chem. Lett.* **15**: 3737-3743 (2005).

29. Asadollahi T., Dadfarnia S., Shabani AMH., Ghasemi JB., and Sarkhosh M. QSAR Models for CXCR2 Receptor Antagonists Based on the Genetic Algorithm for Data Preprocessing Prior to Application of the PLS Linear Regression Method and Design of the New Compounds Using In Silico Virtual Screening. *Molecules.* **16**: 1928-1955 (2011).

30. Khoshneviszadeh M., Edraki N, Miri R., Foroumadi A., and Hemmateenejad B. QSAR Study of 4-Aryl-4H-Chromenes as a New Series of Apoptosis Inducers Using Different Chemometric Tools. *Chem. Biol. Drug. Des.* **79**: 442-458 (2012).

31. Addinsoft S. XLstat 2012: Leading data analysis and statistical solution for microsoft excel. *Addinsoft SRL.* (2012)

32. Olah M., Bologa C., and Oprea TI. An automated PLS search for biologically relevant QSAR descriptors. *J. Comput. Aided. Mol. Des.* **18**: 437-449 (2004).

33. Galindo-Prieto B., Eriksson L., and Trygg J. Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS). *J. Chemometrics.* **28**:623-632 (2014).

34. Roy K., Kar S., and Das RN., Chapter 7 - Validation of QSAR Models, in Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment., Academic Press: Boston. p. 231-89(2015).

35. Weaver S., and Gleeson MP. The importance of the domain of applicability in QSAR modeling. *J. Mol. Graph. Model.* **26**: 1315-1326 (2008).

36. Mojaddami A., Sakhteman A., Fereidoonnezhad M.,

Faghih Z., Najdian A., Khabnadideh S., Sadeghpour H., and Rezaei Z. Binding mode of triazole derivatives as aromatase inhibitors based on docking, protein ligand interaction fingerprinting, and molecular dynamics simulation studies. *Res. Pharm. Sci.* **12**: 21-30 (2017).

37. Helgren TR., Hagen TJ. Demonstration of AutoDock as an Educational Tool for Drug Discovery. *J. Chem. Educ.* **94**: 345-349 (2017).

38. Fereidoonnezhad M., Kaboudin B., Mirzaee T., Babadi Aghakhanpour R., Golbon Haghighi M., Faghih Z., Faghih Z., Ahmadipour Z., Notash B., and Shahsavari HR. Cyclometalated Platinum (II) Complexes Bearing Bidentate O, O′-Di (alkyl) dithiophosphate Ligands: Photoluminescence and Cytotoxic Properties. *Organometallics.* **36**:1707-1717 (2017).

39. Morris GM., Goodsell DS., Halliday RS., Huey R., Hart WE., Belew RK., and Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**: 1639-1662 (1998).

40. Hamedi A., Khoshnoud MJ., Tanideh N., Abbasi F., Fereidoonnezhad M., and Mehrabani D. Reproductive toxicity of Cassia absus seeds in female rats: possible progesteronic properties of chaksine and b-sitosterol. *Pharm. Chem. J.* **49**: 268-274 (2015).

41. Faghih Z., Fereidoonnezhad M., Tabaei SMH., Rezaei Z., and Zolghadr AR. The binding of small carbazole derivative (P7C3) to protofibrils of the Alzheimer's disease and β-secretase: Molecular dynamics simulation studies. *Chem. Phys.* **459**: 31-39 (2015).

42. Humphrey W., Dalke A., and Schulten K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**: 33-8, 27-8 (1996).

43. DeLano WL. PyMOL. *DeLano Scientific, San Carlos, CA.*

700: (2002)

44. Triballeau N., Acher F., Brabet I., Pin JP., and Bertrand HO. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J. Med. Chem.* **48**: 2534-2547 (2005).

45. Rezaei Z., Fereidoonnezhad M., Faghih Z., Sadeghpur H., Mojaddami A., and Sakhteman A. Comparison of docking procedures and its efficiency for Betasecretase, Aromatase and Pyruvate dehydrogenase kinase inhibitors. *Trends. Pharm. Sci.* **3**: 31-42 (2017).

46. Gaulton A., Bellis LJ., Bento AP., Chambers J., Davies M., Hersey A., Light Y., McGlinchey S., Michalovich D., Al-Lazikani B., and Overington JP. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**: D1100-1107 (2012).

47. Wassermann AM., and Bajorath J., BindingDB and ChEMBL: online compound databases for drug discovery. *Expert Opin Drug Discov.* **6**: 683-687 (2011).

48. Willighagen EL., Waagmeester A., Spjuth O., Ansell P., Williams AJ., Tkachenko V., Hastings J., Chen B., and Wild DJ. The ChEMBL database as linked open data. *J. Cheminform.* **5**: 23 (2013).

49. O'Boyle NM., Banck M., James CA., Morley C., Vandermeersch T., and Hutchison GR. Open Babel: An open chemical toolbox. *J Cheminform.* **3**: 33 (2011).

50. Li H., Zhang H., Zheng M., Luo J., Kang L., Liu X., Wang X., and Jiang H. An effective docking strategy for virtual screening based on multi-objective optimization algorithm. *BMC bioinformatics.* **10**: 58 (2009).