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

A. Mostoufi1,2, H. Eucefifar1,2, F. Beygi1, 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

Received: 11 August 2017 / Revised: 12 September 2017 / Accepted: 17 October 2017

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 survival 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 anti-apoptotic properties which come from its molecular chaperone functions at

* Corresponding author: Tel: +986133112278; Fax: +986133118381; Email: Fereidoonnezhad-m@ajums.ac.ir
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 autocorrelation, 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>
<thead>
<tr>
<th>Name</th>
<th>R₁</th>
<th>R₂</th>
<th>R₃</th>
<th>Exp.pIC₅₀</th>
<th>Pred.pIC₅₀</th>
<th>∆E¹ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hsp27</td>
<td>Tubulin</td>
<td></td>
</tr>
<tr>
<td>1a</td>
<td>Cl</td>
<td>Cl</td>
<td>Br</td>
<td>CH₃</td>
<td>5.640165</td>
<td>-3.51</td>
</tr>
<tr>
<td>2a</td>
<td>Cl</td>
<td>Cl</td>
<td>OCH₃</td>
<td>CH₃</td>
<td>6.309804</td>
<td>-7.99</td>
</tr>
<tr>
<td>3a</td>
<td>Cl</td>
<td>Cl</td>
<td>OCH₃</td>
<td>CH₃</td>
<td>5.428291</td>
<td>-7.59</td>
</tr>
<tr>
<td>4a</td>
<td>Cl</td>
<td>Cl</td>
<td>I</td>
<td>CH₃</td>
<td>6.180456</td>
<td>-8.62</td>
</tr>
<tr>
<td>5a</td>
<td>Cl</td>
<td>Cl</td>
<td></td>
<td>CH₃</td>
<td>5.655608</td>
<td>-9.08</td>
</tr>
</tbody>
</table>

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.
In Silico Screening Studies on Methanesulfonamide Derivatives as Dual Hsp27 and …

...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 IC₅₀ values, were used for the QSAR modeling studies.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6a</td>
<td>CH₃</td>
<td>5.545155</td>
<td>5.59254</td>
<td>-8.69</td>
<td>-9.12</td>
<td></td>
</tr>
<tr>
<td>7a</td>
<td>n-C₆H₁₃</td>
<td>CH₃</td>
<td>4.398918</td>
<td>4.43685</td>
<td>-7.23</td>
<td>-9.04</td>
</tr>
<tr>
<td>8a</td>
<td>n-C₆H₁₃</td>
<td>CH₃</td>
<td>4.699405</td>
<td>4.81959</td>
<td>-6.81</td>
<td>-8.6</td>
</tr>
<tr>
<td>9a</td>
<td>n-C₆H₁₃</td>
<td>CH₃</td>
<td>3.998223</td>
<td>4.08720</td>
<td>7.15</td>
<td>-8.67</td>
</tr>
<tr>
<td>10a</td>
<td>n-C₆H₁₃</td>
<td>CH₃</td>
<td>4.897566</td>
<td>4.60381</td>
<td>-7.33</td>
<td>-9.71</td>
</tr>
<tr>
<td>11a</td>
<td>n-C₆H₁₃</td>
<td>CH₃</td>
<td>4.191721</td>
<td>4.15625</td>
<td>-7.66</td>
<td>-9.84</td>
</tr>
<tr>
<td>12a</td>
<td>n-C₆H₁₃</td>
<td>CH₃</td>
<td>4.541211</td>
<td>4.51461</td>
<td>-7.2</td>
<td>-9.18</td>
</tr>
<tr>
<td>13a</td>
<td>CH₃</td>
<td>5.022734</td>
<td>4.84393</td>
<td>-7.65</td>
<td>-9.07</td>
<td></td>
</tr>
<tr>
<td>14a</td>
<td>CH₃</td>
<td>5.692504</td>
<td>5.56468</td>
<td>-7.3</td>
<td>-8.47</td>
<td></td>
</tr>
<tr>
<td>15a</td>
<td>CH₃</td>
<td>4.771087</td>
<td>4.92903</td>
<td>-7.37</td>
<td>-7.89</td>
<td></td>
</tr>
<tr>
<td>16a</td>
<td>CH₃</td>
<td>5.455932</td>
<td>5.56469</td>
<td>-7.91</td>
<td>-9.86</td>
<td></td>
</tr>
<tr>
<td>17a</td>
<td>CH₃</td>
<td>4.735182</td>
<td>4.95924</td>
<td>-8.56</td>
<td>-10.23</td>
<td></td>
</tr>
<tr>
<td>18a</td>
<td>CH₃</td>
<td>4.832092</td>
<td>4.70962</td>
<td>-8.09</td>
<td>-8.62</td>
<td></td>
</tr>
<tr>
<td>19a</td>
<td>CH₃</td>
<td>8.454693</td>
<td>8.39460</td>
<td>-7.87</td>
<td>-8.54</td>
<td></td>
</tr>
<tr>
<td>20a</td>
<td>CH₃</td>
<td>8.59176</td>
<td>8.75116</td>
<td>-7.19</td>
<td>-6.5</td>
<td></td>
</tr>
<tr>
<td>21a</td>
<td>CH₃</td>
<td>8.583359</td>
<td>8.41657</td>
<td>-7.23</td>
<td>-8.47</td>
<td></td>
</tr>
<tr>
<td>22a</td>
<td>CH₃</td>
<td>8.917215</td>
<td>8.84899</td>
<td>-8.02</td>
<td>-8.7</td>
<td></td>
</tr>
<tr>
<td>23a</td>
<td>CH₃</td>
<td>7.77963</td>
<td>8.24032</td>
<td>-8.16</td>
<td>-9.59</td>
<td></td>
</tr>
</tbody>
</table>
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),
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 ($\eta = 0.5 \times \text{HOMO+LUMO}$); softness ($S=\frac{1}{\eta}$), electrophilicity ($\omega = \chi^2 / 2\eta$) and electronegativity ($\chi = -0.5 \times \text{HOMO-LUMO}$) 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

---

**Table 1. Cntd**

| 41a | HCO | CH3 | 6.29243 | 6.31578 | -7.2 | -6.52 |
| 42a | HCO | CH3 | 6.69897 | 6.34115 | -7.46 | -9.46 |
| 43a | HCO | CH3 | 6.148742 | 5.95048 | -7.12 | -6.86 |
| 44a | Br | 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
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 $0.8 < \text{VIP} < 1.0$ 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 $\pm x$ (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 \(82 \times 66 \times 94\) and \(60 \times 60 \times 60\) 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.

<table>
<thead>
<tr>
<th>Model</th>
<th>Eq</th>
<th>(N)</th>
<th>(R^2_C)</th>
<th>(F)</th>
<th>(Q^2)</th>
<th>RMScv</th>
<th>SE</th>
<th>(R^2_P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLR 1</td>
<td>(PCI_{50}=-2.928GATS6v(±0.736)-5.828GATS6v(±0.396)-80.895MATS3m(±6.459)+68.014PW5(±11.944)-15.936MATS8p(±1.343)+6.429MATS2s(±0.707)-24.570G(N..N)(±9.796)+26.809MATS8m(±7.562)-1.328ASP(±0.625)+206.913(±57.987))</td>
<td>39</td>
<td>0.93</td>
<td>223.19</td>
<td>0.82</td>
<td>0.26</td>
<td>0.19</td>
<td>0.91</td>
</tr>
<tr>
<td>FA-MLR 2</td>
<td>(PCI_{50}=89.213PW5(±8.929)-24.410MATS8v(±0.996)+53.444X0Av(±2.472)-208.490X4Av(±15.104)+13.790MATS6p(±1.684)+2.971MATS3v(±0.704)+0.004VOL(±0.001)+3.585ElectroNEG(±0.933)-1.411ASP(±0.562)+29.366(±1.986))</td>
<td>39</td>
<td>0.96</td>
<td>199.06</td>
<td>0.80</td>
<td>0.28</td>
<td>0.19</td>
<td>0.94</td>
</tr>
<tr>
<td>PCR 3</td>
<td>(PCI_{50}=0.805FAC4(±0.063)-0.748FAC8(±0.069)+0.555FAC3(±0.060)-0.293FAC2(±0.058)-0.167FAC9(±0.067)+0.414FAC14(±0.080)-0.316FAC7(±0.073)+0.215FAC5(±0.059)-0.293FAC11(±0.105)+6.214(±0.063))</td>
<td>39</td>
<td>0.91</td>
<td>60.532</td>
<td>0.78</td>
<td>0.35</td>
<td>0.36</td>
<td>0.93</td>
</tr>
<tr>
<td>GA-PLS 4</td>
<td>(PCI_{50}=210.547PW5(±14.859)-90.006MATS3m(±6.650)+17.837MATS6v(±1.229)-95.430PW2(±5.648)-4.180GATS8e(±0.326)-1.454TI2(±0.187)-2.107ASP(±0.496)+0.561nN(CO)2(±0.160)+3.056HARDNESS(±0.978)+146.819(±6.318))</td>
<td>39</td>
<td>0.98</td>
<td>195.22</td>
<td>0.89</td>
<td>0.31</td>
<td>0.21</td>
<td>0.98</td>
</tr>
</tbody>
</table>

1 Number of molecules of training set used to derive the QSAR models

- \(R^2_C\): Regression Coefficient for Calibration set
- \(Q^2\): Regression Coefficient for Leave One Out Cross Validation
- RMScv: Root Mean Square Error of cross validation
- \(R^2_P\): Regression Coefficient for prediction set

---

227
results of QSAR models are summarized in Table 2. The statistical parameters calculated for each target such as $R^2$, correlation coefficient ($R^2_p$) 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, GATS1e and ASP descriptors. Factor score 3 indicates the importance of TI2, PW5 and nN(CO)2 descriptors. Factor score 4 indicates the importance of GATS1e, PW2, nN(CO)2

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

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
<th>F8</th>
<th>F9</th>
<th>F11</th>
<th>F14</th>
<th>Communalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vol.</td>
<td>-0.301</td>
<td>-0.061</td>
<td>0.159</td>
<td>0.402</td>
<td>0.036</td>
<td>0.196</td>
<td>0.055</td>
<td>0.017</td>
<td>0.007</td>
<td>0.980</td>
<td></td>
</tr>
<tr>
<td>TI2</td>
<td>-0.439</td>
<td>0.536</td>
<td>0.153</td>
<td>-0.063</td>
<td>0.198</td>
<td>0.198</td>
<td>0.020</td>
<td>-0.016</td>
<td>-0.057</td>
<td>0.994</td>
<td></td>
</tr>
<tr>
<td>X0Av</td>
<td>0.443</td>
<td>-0.110</td>
<td>0.269</td>
<td>0.271</td>
<td>-0.136</td>
<td>0.041</td>
<td>0.093</td>
<td>0.053</td>
<td>0.016</td>
<td>0.996</td>
<td></td>
</tr>
<tr>
<td>X4Av</td>
<td>0.309</td>
<td>-0.317</td>
<td>-0.106</td>
<td>0.322</td>
<td>-0.048</td>
<td>-0.011</td>
<td>-0.038</td>
<td>-0.181</td>
<td>0.111</td>
<td>0.987</td>
<td></td>
</tr>
<tr>
<td>PW2</td>
<td>-0.569</td>
<td>0.067</td>
<td>-0.448</td>
<td>-0.087</td>
<td>-0.272</td>
<td>-0.067</td>
<td>0.003</td>
<td>-0.002</td>
<td>-0.024</td>
<td>0.997</td>
<td></td>
</tr>
<tr>
<td>PW5</td>
<td>-0.211</td>
<td>0.813</td>
<td>0.381</td>
<td>0.212</td>
<td>0.085</td>
<td>-0.016</td>
<td>-0.051</td>
<td>-0.047</td>
<td>-0.042</td>
<td>0.990</td>
<td></td>
</tr>
<tr>
<td>MATS3m</td>
<td>-0.235</td>
<td>0.053</td>
<td>-0.437</td>
<td>-0.426</td>
<td>0.551</td>
<td>0.328</td>
<td>-0.071</td>
<td>-0.251</td>
<td>0.049</td>
<td>0.989</td>
<td></td>
</tr>
<tr>
<td>MATS8m</td>
<td>-0.761</td>
<td>-0.250</td>
<td>-0.121</td>
<td>-0.229</td>
<td>0.327</td>
<td>0.148</td>
<td>0.073</td>
<td>0.127</td>
<td>-0.052</td>
<td>0.989</td>
<td></td>
</tr>
<tr>
<td>MATS6v</td>
<td>0.231</td>
<td>0.301</td>
<td>0.222</td>
<td>-0.197</td>
<td>0.030</td>
<td>-0.415</td>
<td>-0.492</td>
<td>-0.064</td>
<td>0.247</td>
<td>0.982</td>
<td></td>
</tr>
<tr>
<td>MATS8v</td>
<td>0.414</td>
<td>0.013</td>
<td>-0.172</td>
<td>-0.048</td>
<td>-0.020</td>
<td>0.581</td>
<td>0.185</td>
<td>0.015</td>
<td>-0.533</td>
<td>0.993</td>
<td></td>
</tr>
<tr>
<td>MATS2e</td>
<td>0.373</td>
<td>-0.445</td>
<td>-0.751</td>
<td>0.150</td>
<td>0.088</td>
<td>-0.029</td>
<td>-0.035</td>
<td>-0.041</td>
<td>-0.076</td>
<td>0.994</td>
<td></td>
</tr>
<tr>
<td>MATS6p</td>
<td>0.103</td>
<td>0.394</td>
<td>-0.035</td>
<td>-0.370</td>
<td>-0.095</td>
<td>-0.212</td>
<td>-0.508</td>
<td>-0.121</td>
<td>0.225</td>
<td>0.966</td>
<td></td>
</tr>
<tr>
<td>MATS8p</td>
<td>0.594</td>
<td>-0.084</td>
<td>-0.138</td>
<td>0.068</td>
<td>-0.105</td>
<td>0.478</td>
<td>0.145</td>
<td>0.019</td>
<td>-0.242</td>
<td>0.986</td>
<td></td>
</tr>
<tr>
<td>GATS6v</td>
<td>-0.735</td>
<td>-0.151</td>
<td>-0.180</td>
<td>-0.146</td>
<td>-0.176</td>
<td>0.427</td>
<td>0.278</td>
<td>0.000</td>
<td>-0.128</td>
<td>0.989</td>
<td></td>
</tr>
<tr>
<td>GATS1e</td>
<td>-0.560</td>
<td>0.185</td>
<td>0.685</td>
<td>-0.288</td>
<td>0.049</td>
<td>-0.208</td>
<td>0.132</td>
<td>0.098</td>
<td>-0.001</td>
<td>0.989</td>
<td></td>
</tr>
<tr>
<td>GATS8e</td>
<td>-0.422</td>
<td>0.086</td>
<td>0.188</td>
<td>0.234</td>
<td>-0.082</td>
<td>0.173</td>
<td>-0.076</td>
<td>-0.018</td>
<td>-0.100</td>
<td>0.995</td>
<td></td>
</tr>
<tr>
<td>ASP</td>
<td>0.546</td>
<td>0.042</td>
<td>-0.058</td>
<td>0.425</td>
<td>0.050</td>
<td>-0.209</td>
<td>0.014</td>
<td>0.363</td>
<td>0.022</td>
<td>0.834</td>
<td></td>
</tr>
<tr>
<td>G(N..N)</td>
<td>0.044</td>
<td>0.252</td>
<td>-0.252</td>
<td>-0.067</td>
<td>-0.270</td>
<td>0.007</td>
<td>0.172</td>
<td>0.268</td>
<td>0.213</td>
<td>0.548</td>
<td></td>
</tr>
<tr>
<td>nN(CO)2</td>
<td>0.171</td>
<td>-0.292</td>
<td>-0.433</td>
<td>0.085</td>
<td>-0.470</td>
<td>0.314</td>
<td>0.264</td>
<td>0.138</td>
<td>0.187</td>
<td>0.992</td>
<td></td>
</tr>
<tr>
<td>Hardness</td>
<td>0.194</td>
<td>0.397</td>
<td>0.029</td>
<td>-0.021</td>
<td>-0.086</td>
<td>-0.350</td>
<td>-0.228</td>
<td>-0.205</td>
<td>-0.182</td>
<td>0.984</td>
<td></td>
</tr>
<tr>
<td>ElectroNeg</td>
<td>0.293</td>
<td>0.310</td>
<td>-0.065</td>
<td>-0.068</td>
<td>-0.121</td>
<td>-0.278</td>
<td>-0.183</td>
<td>-0.191</td>
<td>-0.188</td>
<td>0.979</td>
<td></td>
</tr>
</tbody>
</table>
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 MATS6p. Factor score 9 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 multicollinearity 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)) 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.](image-url)
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 DD06 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

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

---

**Table 4.** Definitions of molecular descriptors present in the models

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATS2e</td>
<td>Moran autocorrelation - lag 2 / weighted by atomic Sanderson electronegativities</td>
</tr>
<tr>
<td>MATS3m</td>
<td>Moran autocorrelation- lag 3 / weighted by atomic masses</td>
</tr>
<tr>
<td>MATS8m</td>
<td>Moran autocorrelation - lag 8 / weighted by atomic masses</td>
</tr>
<tr>
<td>MATS6p</td>
<td>Moran autocorrelation - lag 6 / weighted by atomic polarizabilities</td>
</tr>
<tr>
<td>MATS5p</td>
<td>Moran autocorrelation - lag 5 / weighted by atomic polarizabilities</td>
</tr>
<tr>
<td>MATS3V</td>
<td>Moran autocorrelation - lag 3 / weighted by atomic van der Waals volumes</td>
</tr>
<tr>
<td>MATS6V</td>
<td>Moran autocorrelation - lag 6 / weighted by atomic van der Waals volumes</td>
</tr>
<tr>
<td>MATS8V</td>
<td>Moran autocorrelation - lag 8 / weighted by atomic van der Waals volumes</td>
</tr>
<tr>
<td>GATS1e</td>
<td>Geary autocorrelation - lag 1 / weighted by atomic Sanderson electronegativities</td>
</tr>
<tr>
<td>GATS8e</td>
<td>Geary autocorrelation - lag 8 / weighted by atomic Sanderson electronegativities</td>
</tr>
<tr>
<td>GATS6v</td>
<td>Geary autocorrelation - lag 6 / weighted by atomic van der Waals volumes</td>
</tr>
<tr>
<td>PW2</td>
<td>Path/walk 2- Randic shape index</td>
</tr>
<tr>
<td>PW5</td>
<td>Path/walk 5- Randic shape index</td>
</tr>
<tr>
<td>G(N..N)</td>
<td>Sum of geometrical distances between N..N</td>
</tr>
<tr>
<td>ASP</td>
<td>Asphericity</td>
</tr>
<tr>
<td>T12</td>
<td>Second Mohar index T12</td>
</tr>
<tr>
<td>Nn(CO)2</td>
<td>Number of imides</td>
</tr>
<tr>
<td>X0Av</td>
<td>Average valence connectivity index chi-0</td>
</tr>
<tr>
<td>X4Av</td>
<td>Average valence connectivity index chi-4</td>
</tr>
<tr>
<td>Hardness</td>
<td>Quantum chemical indices of hardness</td>
</tr>
<tr>
<td>VOL</td>
<td>Molecular volume</td>
</tr>
<tr>
<td>ElectroNeg</td>
<td>Electronegativity</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Name</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>pIC50 pred</th>
<th>leverage</th>
<th>Hsp27 ΔE (kcal/mol)</th>
<th>Tubulin ΔE (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1b</td>
<td></td>
<td></td>
<td></td>
<td>9.047</td>
<td>0.286</td>
<td>-6.98</td>
<td>-8.74</td>
</tr>
<tr>
<td>2b</td>
<td></td>
<td></td>
<td></td>
<td>9.051</td>
<td>0.356</td>
<td>-8.24</td>
<td>-8.20</td>
</tr>
</tbody>
</table>

Figure 3. Williams plot for the calibration set and external prediction set.
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>
<thead>
<tr>
<th>Compound</th>
<th>Cntd</th>
<th>3b</th>
<th>4b</th>
<th>5b</th>
<th>6b</th>
<th>7b</th>
<th>8b</th>
<th>9b</th>
<th>10b</th>
<th>11b</th>
<th>12b</th>
<th>13b</th>
<th>14b</th>
<th>15b</th>
<th>16b</th>
<th>17b</th>
<th>18b</th>
<th>19b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.395</td>
<td>0.787</td>
<td>0.749</td>
<td>0.672</td>
<td>0.735</td>
<td>0.702</td>
<td>0.537</td>
<td>0.655</td>
<td>0.351</td>
<td>0.383</td>
<td>0.337</td>
<td>0.373</td>
<td>0.384</td>
<td>0.361</td>
<td>0.359</td>
<td>0.418</td>
<td>0.336</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.97</td>
<td>-2.96</td>
<td>-2.09</td>
<td>-2.73</td>
<td>-2.49</td>
<td>-2.37</td>
<td>-2.20</td>
<td>-2.97</td>
<td>-6.85</td>
<td>-5.54</td>
<td>-7.57</td>
<td>-7.10</td>
<td>-6.89</td>
<td>-7.45</td>
<td>-7.87</td>
<td>-6.77</td>
<td>-7.56</td>
</tr>
</tbody>
</table>

Table 5. Cntd
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 ($\Delta G_{\text{bind}}$) 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.

<table>
<thead>
<tr>
<th>Compound</th>
<th>H</th>
<th>CH$_3$</th>
<th>CH$_3$</th>
<th>$\Delta G_{\text{bind}}$ (Kcalmol$^{-1}$)</th>
<th>pIC$_{50}$</th>
<th>pIC$_{50}$</th>
<th>pIC$_{50}$</th>
<th>pIC$_{50}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20b</td>
<td>I</td>
<td>CH$_3$</td>
<td>13.184</td>
<td>0.314</td>
<td>-7.22</td>
<td>-5.54</td>
<td>-5.54</td>
<td>-5.54</td>
</tr>
<tr>
<td>21b</td>
<td>I</td>
<td>H</td>
<td>12.245</td>
<td>0.450</td>
<td>-7.19</td>
<td>-5.96</td>
<td>-5.96</td>
<td>-5.96</td>
</tr>
<tr>
<td>22b</td>
<td>I</td>
<td>CH$_3$</td>
<td>11.156</td>
<td>0.472</td>
<td>-7.63</td>
<td>-6.63</td>
<td>-6.63</td>
<td>-6.63</td>
</tr>
<tr>
<td>23b</td>
<td>OCH$_3$</td>
<td>CH$_3$</td>
<td>11.395</td>
<td>0.405</td>
<td>-6.83</td>
<td>-7.59</td>
<td>-7.59</td>
<td>-7.59</td>
</tr>
<tr>
<td>24b</td>
<td>OCH$_3$</td>
<td>H</td>
<td>11.155</td>
<td>0.368</td>
<td>-6.87</td>
<td>-7.38</td>
<td>-7.38</td>
<td>-7.38</td>
</tr>
<tr>
<td>25b</td>
<td>OCH$_3$</td>
<td>H</td>
<td>9.490</td>
<td>0.422</td>
<td>-7.21</td>
<td>-5.67</td>
<td>-5.67</td>
<td>-5.67</td>
</tr>
</tbody>
</table>

**Figure 4.** Plots of cross-validated predicted values of activity by GA-PLS against the experimental values
Table 1 shows that, the \( \Delta G_{\text{bind}} \) 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 residues 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 benzoxo 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.
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.

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.
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](image)

**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 (EFmax). Its calculations were based on the Li et al. work [50]. EFmax factor in comparison to various threshold settings. The ROC curve 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.

Acknowledgment
The authors would like to thank research deputy Ahvaz Jundishapur University of medical sciences who support this work. Collaboration of medicinal chemistry, faculty of pharmacy, Ahvaz Jundishapur University of Medical Sciences, in providing the required facilities for this work is greatly acknowledged. This paper is a part of thesis by Heydar Euceffiar.

References
In Silico Screening Studies on Methanesulfonamide Derivatives as Dual Hsp27 and ...


35. Mojaddami A., Sakhteman A., Fereidoonnezhad M.,...


