

# 3D-QSAR Studies of Isatin Derivatives with Anti-Cancer *In Vitro*: Advanced CoMFA, CoMSIA and Docking Methods

Elidrissi B<sup>1\*</sup>, Ousaa A<sup>1</sup>, Aouidate A<sup>1</sup>, Zaki H<sup>1</sup>, Ajana MA<sup>1</sup>, Lakhlifi T<sup>1</sup> and Bouachrine M<sup>2</sup>

<sup>1</sup>Molecular Chemistry and Natural Substances Laboratory, Faculty of Science, University Moulay Ismail, Meknes, Morocco

<sup>2</sup>ESTM, University Moulay Ismail, Meknes, Morocco

## Abstract

Three-dimensional quantitative structure–activity relationship (3D-QSAR) and docking methods were performed to study 47 tubulin inhibitors, isatin derivatives with anticancer activity against human monocyte-like histiocytic lymphoma human U937 cells. The established 3D-QSAR model from Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Index Analysis (CoMSIA) approaches show a significant statistical quality and a satisfying predictive ability, with high correlation coefficient value ( $R^2=0.936$ ,  $R^2=0.970$ ) and cross-validation coefficient value ( $Q^2=0.821$ ,  $Q^2=0.884$ ) of CoMFA and CoMSIA, respectively. The predictive ability of the CoMFA ( $R^2_{\text{test}}=0.607$ ) and CoMSIA ( $R^2_{\text{test}}=0.650$ ) model was confirmed by a test set. The CoMFA, CoMSIA contour maps and docking analyses indicate that the substituent  $R_1$  should be oxygen which better than azotes to forms a more potent inhibitor against tubulin enzyme and  $R_2$ ,  $R_3$ ,  $R_4$  and  $R_5$  substituents should be higher electronegative but  $R_6$  should be a bulky aromatic group like a methyl naphthol monosubstituted or disubstituted by  $\text{OCH}_3$ . The interaction information between target and ligand was presented such the useful theoretical references for analyzing to understand the action mechanism, designing new more potent inhibitors and optimizing their activities prior to synthesis.

**Keywords:** Tubulin enzyme; Anti-cancer activity; CoMFA; CoMSIA; Docking study

## Introduction

Tubulin is the structural protein of microtubules and a major component of the cytoskeleton in eukaryotic cells; it is involved in many important cellular processes including mitosis. This protein is composed of  $\alpha$ - and  $\beta$ -tubulin, microtubules have emerged as a strategic target in anticancer drugs [1,2]. On the monomeric unpolymerized  $\alpha/\beta$ -tubulin exists the colchicine-binding site which represents another potential tubulin target for the development of apoptosis inducing chemotherapeutic agents [3]. There is only few examples reached so far clinical and commercial success, despite existence of the broad range of antimetabolic agents [2]. The failure of the plurality of these molecules could be attributed to poor therapeutic, related to pharmacokinetics, to the problems of solubility, low affinity products by the therapeutic agents to tubulin molecules *in vivo* trials and other unrecognized factors. The construction of therapeutic molecules in accordance with the concept of multivalency is one way to improve some pharmacological parameters [4]. Combretastatin A-4 phosphate (CA-4P) [5], water-soluble prodrug of combretastatin A-4 (CA-4), inhibits tubulin polymerization by binding at the colchicine site and is currently used in clinical trials. The CA-4 is characterized by its potent cytotoxicity, for that it is one of the few tubulin targeting agents reported to have selective vascular-disrupting activity [6,7]. Recently, Vine and Matesic have designed and reported a series of 47 indole-2,3-dione derivatives and assessed their anticancer activities against U937 (human monocyte-like histiocytic lymphoma) cells [8-10]. The results found by these authors demonstrated that the recent compounds prevent the polymerization of the  $\alpha/\beta$ -tubulin by binding to the colchicine site and all have pronounced cytotoxicity, demonstrating the great potential of developing isatin derivatives as a new class of anticancer drug.

The correlation between variations of biological activity and properties or molecular structures supposed an important approach for understanding action mechanisms drugs and designing new drugs, Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Index Analysis (CoMSIA) have been utilized to find the steric, electrostatic, hydrophobic, hydrogen-bond donor and

hydrogen-bond acceptor contour maps can helpful to understanding the detail of interaction between the ligand and the active site of receptor [11,12]. The interaction mechanism studied by docking analysis as also an effective method, that is means it can offer vivid interaction picture between a ligand and an acceptor [13,14]. In this report, CoMFA and CoMSIA and docking studies of 47 tubulin inhibitors, isatin derivatives with anticancer activity against U937 cells *in vitro*, were carried out. To construct an optimal 3D-QSAR model for these compounds, we have used CoMFA and CoMSIA methods and explored the inhibition mechanism via docking analysis.

## Material and Methods

### The data of the studied compounds and their biological activity

The experimental values *in vitro*  $\text{IC}_{50}$  of 47 isatin or indole-2,3-dione derivatives for inhibition of  $\beta$ -tubulin colchicine, were taken from the literature [10-12]. They are defined as the value of the necessary molar concentration of compound to cause 50% growth inhibition against the human monocyte-like histiocytic lymphoma U937 cells. The compounds and their  $\text{pIC}_{50}$  ( $-\log\text{IC}_{50}$ ) values are shown in Figure 1 and Table 1.

### Molecular modeling

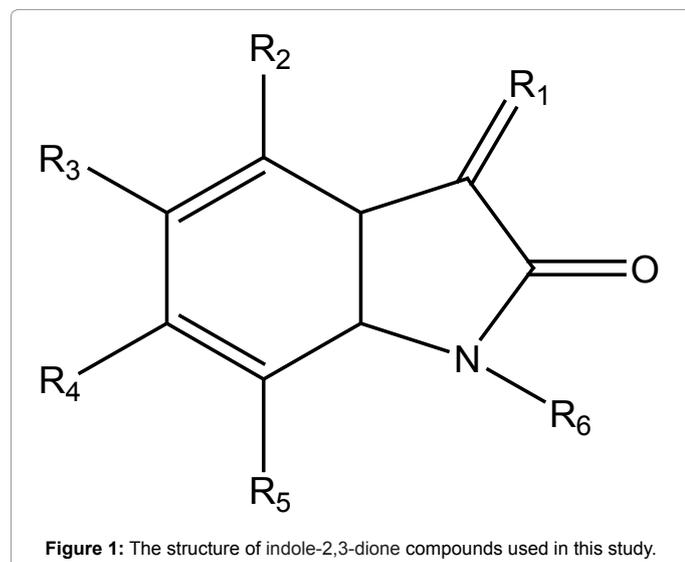
CoMFA and CoMSIA studies were performed on Windows 7 computer operating system using SYBYL-X2.0 molecular modeling software. 3D structures were built using the SKETCH option in

\*Corresponding author: Elidrissi B, Molecular Chemistry and Natural Substances Laboratory, Faculty of Science, University Moulay Ismail, Meknes, Morocco, Tel: +212607662438; E-mail: elidrissi.info@gmail.com

Received May 25, 2017; Accepted June 01, 2017; Published June 12, 2017

**Citation:** Elidrissi B, Ousaa A, Aouidate A, Zaki H, Ajana MA, et al. (2017) 3D-QSAR Studies of Isatin Derivatives with Anti-Cancer *In Vitro*: Advanced CoMFA, CoMSIA and Docking Methods. Chem Sci J 8: 158. doi: 10.4172/2150-3494.1000158

**Copyright:** © 2017 Elidrissi B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



SYBYL. Each structure of 47 isatin compounds is fully geometry-optimized using the standard tripos molecular mechanics force field and energy gradient convergence criterion (0.01 kcal/mol), the partial atomic charges required for calculation of the electrostatic potential are assigned using the Gasteiger\_Huckel formation [15,16].

### Alignment

The selected template molecule is typically one of the following [17]: the most active and the lead and/or commercial compound which containing the greatest number of functional groups. Generally, the low energy conformation of the most active compound is set as a reference. In this study the compound number 41 is the most potent inhibitor, taken as the template molecule, and each molecule has to be superimposed onto it Figure 2.

### CoMFA and CoMSIA studies

**CoMFA:** In order to establish the 3D-QSAR towards their anti-cancer activity, we have performed and evaluated the Comparative Molecular Fields Analysis (CoMFA) in SYBYL-X 2.0. The steric (van der Waals) and electrostatic (Coulombic with  $1=r$  dielectric) fields for the CoMFA were based on both Lennard-Jones and Coulombic potentials [18]. They are calculated at every grid point using a sp<sup>3</sup>-hybridized carbon probe with van der Waals radius of 1.52 Å, charge of +1.0 and grid spacing of 2.0 Å. A 30 kcal/mol energy cutoff is applied, which means steric and electrostatic energies greater than 30 kcal/mol are truncated to that value [19].

**CoMSIA:** We have carried out 3D-QSAR studies employing the Comparative Molecular Similarity Indices Analysis (CoMSIA) technique in SYBYL-X 2.0 [20], and using the same training and test sets, the same grid box as used in CoMFA calculation. In this recent paper, five physicochemical properties have been calculated and evaluated: steric, electrostatic, hydrophobic, and hydrogen-bond donor or acceptor properties to develop a CoMSIA model. A sp<sup>3</sup> carbon with a charge, hydrophobic interaction, hydrogen-bond donor and acceptor properties of +1.0 was used as a probe at each grid point to measure the five above-mentioned fields. These fields are selected to cover the major contributions to ligand binding.

**Partial Least Squares (PLS) analysis and validation of 3D-QSAR model:** Partial least squares (PLS) analysis was used to construct a linear

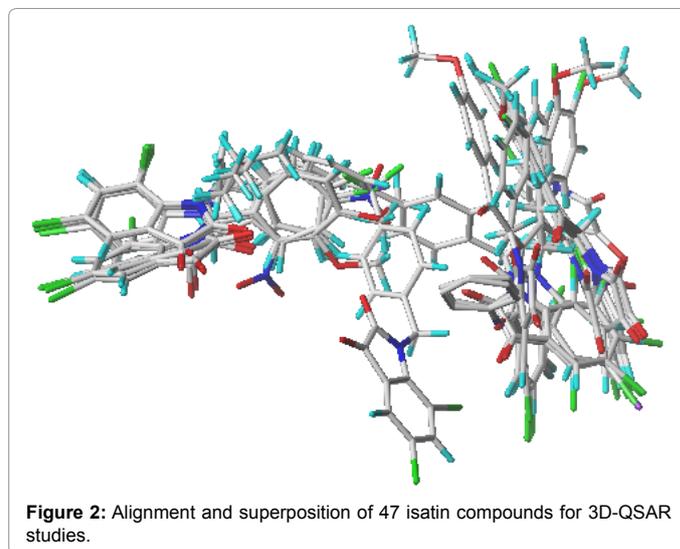
correlation between the CoMFA, CoMSIA fields and the anticancer activity values. To select the best model, the cross-validation analysis was performed using the leave-one-out (LOO) method in which one compound was removed from the data set and its activity was predicted using the model built from rest of the data set [21], the same way is repeated until all compounds have been eliminated once. It results in the cross-validation correlation coefficient ( $Q^2$ ), cross-validation standard error of estimate ( $S_{cv}$ ) and the optimum number of components ( $N$ ). The non-cross-validation ( $R^2$ ) was performed with a column filter value of 2.0 kcal/mol to speed up the analysis and reduce the noise.

**Validation of model:** Finally, in order to evaluate the predictive abilities of the CoMFA and CoMSIA models derived by the training set, an external test set composed of eight compounds was used to predict its biological proprieties [22]. By using the same methods described above, these molecules were aligned and their activities were predicted using the generated CoMFA and CoMSIA models from the training set.

**Y-randomization test:** The models were also evaluated against chance correlation by Y-randomization [23]. Property values were randomized within the training set by many iterations. From each new randomized data set, a new model QSAR was computed again, with performances expected to have lower  $Q^2$  and  $R^2$  values than those the original models. Finally, the average values of the  $Q^2$  and  $R^2$  were calculated to check that the original model was strongly more performant than the randomized ones.

**Molecular docking:** To locate the appropriate binding orientations and conformations of these indole-2,3-dione derivatives interacting with tubulin, docking study was performed with the Autodock vina and Autodock tools 1.5.4 programs. The ligand and protein preparation steps for the docking protocol were carried out in Autodock tools 1.5.4 from MGL Tools package [24], a grid box was set to cover the folic acid binding site in the tubulin protein PDB code: 1AS0, the bioactive conformations were simulated using Autodock vina [25]. The results were analyzed using Discovery studio 2017 [26] and PyMol [27].

**Protein and ligand preparation:** The crystal structure of tubulin enzyme was retrieved from the RCSB Protein Data Bank (PDB entry code: 1SA0). The ligand already present in 1SA0 (CN2: 2-mercapto-N-[1,2,3,10-tetramethoxy-9-oxo-5,6,7,9-tetrahydro-benzo[A]heptalen-7-yl] acetamide) was extracted out, and the most active compounds were docked in the active site of the studied enzyme as ligand. All water



molecules were deleted and the polar hydrogen atoms were added. The selected ligands were modeled for docking in the same way for the 3D-QSAR studies [15,16] (Tripos standard force field, Gasteiger-Hückel atomic partial charges, convergence criterion of 0.01 Kcal/mol).

## Results and Discussion

The 3D-QSAR model was established from CoMFA and CoMSIA analysis and its calculated statistical parameters using SYBYL-X 2.0 are listed in Tables 2-4 ( $Q^2$ : Coefficient of determination for cross validation; N: Optimum number of components obtained from cross-validated PLS analysis and the same used in final non cross validation analysis;  $R^2$ : Coefficient of determination for non-cross validation;  $S_{cv}$ : Standard error of the estimate; F-t: F -test value;  $R^2_{test}$ : Coefficient of determination for external validation). For a reliable predictive model, the Coefficient of determination for cross validation  $Q^2$  should be greater than 0.5.

### CoMFA model

The statistic indicators of the CoMFA model are:  $Q^2$  (0.821),  $R^2$  (0.936),  $S_{cv}$  (0.241), F (78.606) with the optimum number of component (6) which means that six components are sufficient to contain most of the information from the original descriptors. Those statistical parameters indicate that the model has believable predictive ability. Finally, the prediction ability of the proposed model was confirmed using the external test of 8 compounds, the  $R^2_{test}$  value obtained is 0.607. Those statistic results indicated the good stability and the powerful predictive ability of CoMFA model. The steric field descriptor of the CoMFA model explains 68.3% of the total variance and 31.7% has been explained by electrostatic descriptor. This calculated result indicates that both the steric and electrostatic fields contribute to the biological activity, but that the contribution of the steric field is predominant. Thus alteration of the volume and polarity of the compounds is the principal way to improving their anticancer activities.

### CoMSIA model

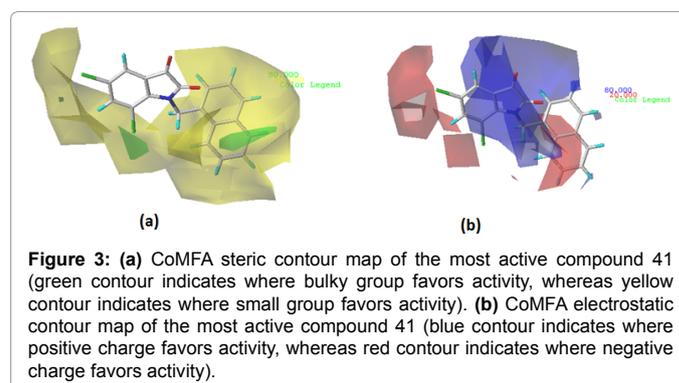
In the developed CoMSIA model, the contributions of steric, electrostatic, hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor fields were found to be 22.1%, 24.3%, 10.1%, 1.2%, and 42.3%, respectively. The purpose of using five different fields is to explain the effects of substituents of a series of 39 of isatin derivatives on the anticancer activity against U937 in the training set. In CoMSIA model, the statistic indicators are:  $Q^2$  (0.884),  $R^2$  (0.970),  $S_{cv}$  (0.166), F (172.498) with the optimum number of component (6). Those statistical parameters indicate that the model has believable predictive ability. Finally, the prediction ability of the proposed model was confirmed using the external test of 8 compounds, the  $R^2_{test}$  value obtained is 0.650.

### Contour maps analysis

The results of CoMFA and CoMSIA can be displayed as vivid 3D contour maps, providing an opportunity to explain the observed variance in the anticancer activity expressed by pIC50. The 3D contour maps were generated from the model 3D-QSAR to get information about the favorable and unfavorable regions for biological activity of the studied compounds.

### CoMFA contour maps

The CoMFA model is used also for generating 3D contour maps to imply 3D-QSAR model on the target, these contours displayed in Figure 3. The steric interactions are represented by yellow and green contours, while the electrostatic interactions are represented by red and



**Figure 3:** (a) CoMFA steric contour map of the most active compound 41 (green contour indicates where bulky group favors activity, whereas yellow contour indicates where small group favors activity). (b) CoMFA electrostatic contour map of the most active compound 41 (blue contour indicates where positive charge favors activity, whereas red contour indicates where negative charge favors activity).

blue contours. The fractions of the steric and electrostatic fields were 73.2%, and 26.8%, respectively.

In CoMFA steric contour maps Figure 3a, the sterically favorable and sterically unfavorable regions are represented by green and yellow contours, respectively. A small green contour is found at  $R_6$  position of isatin ring, which is covered by other big yellow contour maps, that means that the substituent  $R_6$  is imposed to be an adaptive bulk substituent as a naphthylmethyl group, that is a possible reason why compounds 41 ( $R_6=1$ -naphthylmethyl, pIC50=6.72) and 42 ( $R_6=2$ -naphthylmethyl, pIC50=6.13) have the higher activities than other compounds and not to be more steric than that like methyl, because if it is, it will fall within the yellow contour, which is an unfavorable region for steric bulk, consequently the activity will decrease, that is also a possible reason why compound 35 ( $R_6=CH_3$ , pIC50=3.62) has a lower activity than other compounds. The big yellow contour near the  $R_3$ ,  $R_4$  and  $R_5$  groups indicates that these substituents should be small as observed in compounds 27 ( $R_3=OCH_3$ , pIC50=3.38) and 31 ( $R_5=NO_2$ , pIC50=3.59). These contour maps give us some general insight into the nature of the receptor-ligand binding region.

In the CoMFA electrostatic contour maps Figure 3b, electronegative charge favorable and electropositive charge favorable regions are represented by red and blue contours, respectively. On the first hand, two red contours are found, one is near the benzene isatin ring specially near the substituent  $R_3$  and  $R_5$ , it indicates that the compounds with high electronegative groups on these positions would exhibit good anticancer activity. For example, compound 13 ( $R_3=Br$ ,  $R_4=H$ ,  $R_5=Br$ , pIC50=6.11) and 14 ( $R_3=H$ ,  $R_4=Br$ ,  $R_5=H$ , pIC50=5.28), those compounds have almost the same structure except for the position  $R_3$ ,  $R_4$  and  $R_5$  but the substituent  $R_4$  has not important influence to the activity. And the second red contour around the substituent  $R_6$  not near of its first atom. For example, compound 28 ( $R_6=H$ , pIC50=4.98) and 37 ( $R_6=H_2CCH_2C_6H_4Br$ , pIC50=6.11) have the same structure except the substituent  $R_6$ . On the second hand, a large blue contour is found around the first atom of substituent  $R_6$  and near  $R_1$  group, indicating that negatively charged substituent in the area is unfavorable but these parts should be positively charged in favor of the activity as in compound 19 ( $R_1=O$ , pIC50=3.25) and 33 ( $R_1=N-C_6H_5$ , pIC50=4.12) which have the same structure except  $R_1$ .

### CoMSIA contour maps

The CoMSIA steric, electrostatic, hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor contour maps are shown in Figure 4, the molecule in all the maps is the most active inhibitor (compound 41). The fractions of the steric, electrostatic, hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor fields were 22.1%, 24.3%, 42.3%,

N°	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	R <sub>6</sub>	pIC <sub>50</sub>
1	O	H	Br	H	Br	H <sub>2</sub> CCH=CH <sub>2</sub>	5.18
2	O	H	Br	H	Br	H <sub>2</sub> CCH <sub>2</sub> OCH <sub>3</sub>	5.46
3	O	H	Br	H	Br	H <sub>2</sub> CCH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	5.62
4	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>5</sub>	5.94
5	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> CH <sub>3</sub> <sup>c</sup>	6.31
6	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> OCH <sub>3</sub> <sup>c</sup>	5.74
7	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> OCH <sub>3</sub> <sup>d</sup>	5.75
8	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> NO <sub>2</sub> <sup>c</sup>	6.05
9	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> NO <sub>2</sub> <sup>e</sup>	5.64
10	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> Cl <sup>c</sup>	6.01
11	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> Br <sup>c</sup>	6.20
12	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> I <sup>c</sup>	5.64
13	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> CF <sub>3</sub> <sup>c</sup>	6.10
14	O	H	H	Br	H	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> CF <sub>3</sub> <sup>c</sup>	5.28
15	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> COOCH <sub>3</sub> <sup>c</sup>	5.92
16	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> C(CH <sub>3</sub> ) <sub>3</sub> <sup>c</sup>	5.95
17	O	H	Br	H	Br	H <sub>2</sub> CCH=CHC <sub>6</sub> H <sub>5</sub>	5.63
18	O	H	Br	H	Br	H <sub>2</sub> CC <sub>6</sub> H <sub>4</sub> C <sub>6</sub> H <sub>5</sub> <sup>c</sup>	6.12
19	O	H	H	H	H	H	3.25
20	O	Br	H	H	H	H	3.67
21	O	H	Br	H	H	H	4.19
22	O	H	H	Br	H	H	4.13
23	O	H	H	H	Br	H	4.08
24	O	H	F	H	H	H	4.01
25	O	H	I	H	H	H	4.27
26	O	H	NO <sub>2</sub>	H	H	H	3.88
27	O	H	OCH <sub>3</sub>	H	H	H	3.38
28	O	H	Br	H	Br	H	4.98
29	O	H	Br	Br	H	H	4.94
30	O	H	I	H	I	H	5.11
31	O	H	Br	H	NO <sub>2</sub>	H	3.59
32	O	H	Br	Br	Br	H	5.17
33	N-C <sub>6</sub> H <sub>5</sub>	H	H	H	H	H	4.12
34	N-C <sub>6</sub> H <sub>5</sub>	H	Br	H	Br	H	4.86
35	O	H	H	H	H	CH <sub>3</sub>	3.62
36	O	H	Br	H	Br	H <sub>2</sub> CCH <sub>2</sub> C <sub>6</sub> H <sub>5</sub>	6.11
37	O	H	Br	H	Br	H <sub>2</sub> CCH <sub>2</sub> C <sub>6</sub> H <sub>4</sub> Br <sup>d</sup>	6.11
38	O	H	Br	H	Br	H <sub>2</sub> CCH <sub>2</sub> C <sub>6</sub> H <sub>4</sub> Br <sup>e</sup>	6.06
39	O	H	Br	H	Br	H <sub>2</sub> CCH <sub>2</sub> C <sub>6</sub> H <sub>4</sub> OCH <sub>3</sub> <sup>d</sup>	5.97
40	O	H	Br	H	Br	H <sub>2</sub> CCH <sub>2</sub> C <sub>6</sub> H <sub>4</sub> OCH <sub>3</sub> <sup>c</sup>	5.63
41	O	H	Br	H	Br	CH <sub>2</sub> C <sub>10</sub> H <sub>7</sub> <sup>f</sup>	6.72
42	O	H	Br	H	Br	CH <sub>2</sub> C <sub>10</sub> H <sub>7</sub> <sup>g</sup>	6.13
43	O	H	Br	H	Br	CH <sub>2</sub> COC <sub>6</sub> H <sub>5</sub>	5.00
44	O	H	Br	H	Br	CH <sub>2</sub> COC <sub>6</sub> H <sub>4</sub> Br <sup>d</sup>	5.20
45	O	H	Br	H	Br	CH <sub>2</sub> COC <sub>6</sub> H <sub>4</sub> Br <sup>e</sup>	5.04
46	O	H	Br	H	Br	CH <sub>2</sub> COC <sub>6</sub> H <sub>4</sub> OCH <sub>3</sub> <sup>d</sup>	5.33
47	O	H	Br	H	Br	CH <sub>2</sub> COC <sub>6</sub> H <sub>4</sub> CH <sub>3</sub> <sup>c</sup>	5.27

**Table 1:** Chemical structures and experimental activity anti-cancer of studied molecules. <sup>c</sup>Test set: <sup>c</sup>Substitutions at para position, <sup>d</sup>Substitutions at meta position, <sup>e</sup>2-naphthylmethyl, <sup>f</sup>Substitutions at ortho position, <sup>g</sup>1-naphthylmethyl.

Model	Q <sup>2</sup>	R <sup>2</sup>	S <sub>cv</sub>	F-t	N	R <sup>2</sup> <sub>test</sub>	Fractions				
							Ster	Elec	Hyd	Don	Acc
CoMFA	0.821	0.936	0.241	78.606	6	0,607	0.683	0.317	-	-	-
CoMSIA	0.884	0,970	0.166	172.498	6	0,650	0.221	0.243	0.101	0.012	0.423

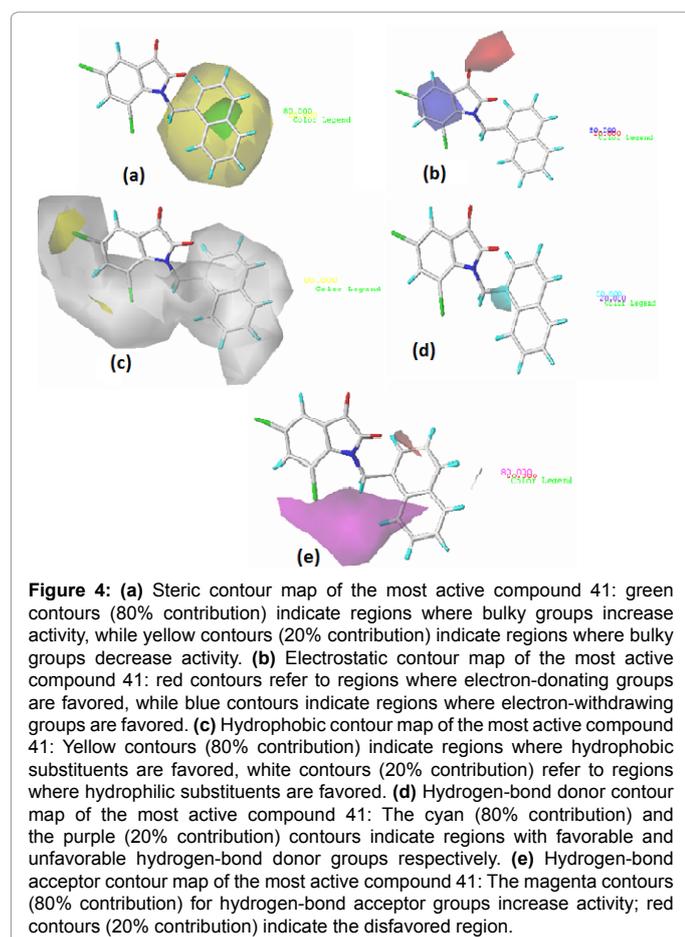
**Table 2:** PLS Statistic indicators of CoMFA and CoMSIA models.

1.2%, and 10.1%, respectively. The CoMSIA steric contour maps Figure 4a were placed almost similar to that of the CoMFA model. In Figure 4b, the electrostatic contour maps are summarized, the areas where negatively charged group enhance anti-cancer activity are contoured by red maps and

where positively charged groups increase activity are surrounded by blue maps. One red contour is found at R<sub>1</sub> position of pyrrole ring of isatin. One blue area is above the aromatic ring of isatin, which means positive charge in this region corresponds to the more active inhibitor.

N°	pIC <sub>50</sub> (obs)	pIC <sub>50</sub> (pred)				N°	pIC <sub>50</sub> (obs)	pIC <sub>50</sub> (pred)			
		CoMFA	Residu	CoMSIA	Residu			CoMFA	Residu	CoMSIA	Residu
1	5.18	5.023	0.157	5.172	0.008	25	4.27	3.721	0.549	3.175	1.095
2	5.46	5.298	0.162	5.521	-0.061	26	3.88	3.734	0.146	3.346	0.534
3	5.62	5.645	-0.025	5.605	0.015	27	3.38	3.73	-0.35	3.13	0.25
4	5.94	5.886	0.054	5.849	0.091	28	4.98	4.536	0.444	4.776	0.204
5	6.31	6.164	0.146	5.99	0.32	29	4.94	4.672	0.268	4.806	0.134
6	5.74	5.769	-0.029	5.952	-0.212	30	5.11	4.998	0.112	5.387	-0.277
7	5.75	5.926	-0.176	5.89	-0.14	31	3.59	4.531	-0.941	3.573	0.017
8	6.05	6.033	0.017	6.092	-0.042	32	5.17	5.285	-0.115	5.241	-0.071
9	5.64	5.805	-0.165	5.652	-0.012	33	4.12	4.032	0.088	3.99	0.13
10	6.01	5.93	0.08	5.95	0.06	34	4.86	4.944	-0.084	4.994	-0.134
11	6.2	6.079	0.121	6.058	0.142	35	3.62	3.518	0.102	3.687	-0.067
12	5.64	5.863	-0.223	5.958	-0.318	36	6.11	6.04	0.07	5.954	0.156
13	6.1	6.023	0.077	6.035	0.065	37	6.11	6.133	-0.023	6.022	0.088
14	5.28	4.732	0.548	4.316	0.964	38	6.06	6.014	0.046	5.973	0.087
15	5.92	4.42	1.5	3.932	1.988	39	5.97	5.943	0.027	5.906	0.064
16	5.95	4.176	1.774	3.744	2.206	40	5.63	5.803	-0.173	5.775	-0.145
17	5.63	5.003	0.627	4.422	1.208	41	6.72	6.609	0.111	6.458	0.262
18	6.12	6.033	0.087	6.176	-0.056	42	6.13	6.141	-0.011	6.341	-0.211
19	3.25	4.14	-0.89	3.512	-0.262	43	5	5.088	-0.088	5.196	-0.196
20	3.67	3.404	0.266	3.538	0.132	44	5.2	5.233	-0.033	5.205	-0.005
21	4.19	3.838	0.352	4.338	-0.148	45	5.04	5.171	-0.131	5.111	-0.071
22	4.13	4.356	-0.226	3.983	0.147	46	5.33	5.262	0.068	5.244	0.086
23	4.08	4.231	-0.151	3.95	0.13	47	5.27	5.24	0.03	5.181	0.089
24	4.01	3.739	0.271	3.209	0.801						

**Table 3:** Actual and predicted pIC<sub>50</sub> along with residual of training and test sets using CoMFA and CoMSIA models.



The hydrophobic substances are soluble in non-polar solvents like benzene but only sparingly or not soluble in water. The CoMSIA hydrophobic contour maps are shown in Figure 4c. In general, yellow contours indicate that hydrophobic substituents are 'good' for increasing the activity for anticancer, while hydrophilic substituents are beneficial to the activity at white contours. There are two yellow contours above the benzene ring of compound 41 (near R<sub>3</sub> and R<sub>5</sub> positions), which mean favorable for hydrophobic substituents, for this reason, compounds 27 (R<sub>5</sub>=OCH<sub>3</sub>, pIC<sub>50</sub>=3.38) and 31 (R<sub>5</sub>=NO<sub>2</sub>, pIC<sub>50</sub>=3.59) are less potent than compound 21 (R<sub>3</sub>=Br, pIC<sub>50</sub>=4.19) and 28 (R<sub>5</sub>=Br, pIC<sub>50</sub>=4.98), in that case an alkyl halogenated or a long hydrocarbon group are supposed a hydrophobic group. The same region has been indicated in the CoMFA steric maps Figure 3a in green to favor a more bulky group; this indicates that bulky substituents with hydrophobic character are preferred in this region. Presence of a big white contour near R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>6</sub> substituents of isatin ring shows the importance of hydrophilic groups on the anti-cancer activity in this region. In CoMSIA hydrogen bond donor and acceptor contour maps (Figure 4d and 4e), the cyan color indicates favorable donor group regions, purple (not appeared) color indicates unfavorable donor group regions, magenta color indicates favorable acceptor group regions and red color indicates unfavorable acceptor group regions. On the first hand, presence of a small cyan contour near R<sub>6</sub> substituent indicated that hydrogen bond donor atoms were favored at this position. On the other hand, an important magenta contour near R<sub>6</sub> substitution indicating this area is favorable for hydrogen-bond acceptors.

### Validation of 3D-QSAR models

In order to check the reliability and the stability of the 3D-QSAR model elaborated by the CoMFA and CoMSIA methods, the authors have used the internal and external validations. The leave-one-out cross-validation Q<sup>2</sup> of the CoMFA and CoMSIA analysis, showing the

Iteration	CoMFA		CoMSIA	
	Q <sup>2</sup>	R <sup>2</sup>	Q <sup>2</sup>	R <sup>2</sup>
1	0.421	0.854	0.435	0.876
2	0.347	0.807	0.379	0.89
3	0.291	0.701	0.299	0.721
4	0.111	0.651	0.123	0.654
5	0.369	0.764	0.317	0.792

**Table 4:** Y-Randomization validation results of the CoMFA and CoMSIA models (Q<sup>2</sup> and R<sup>2</sup> values after several Y-randomization tests).

good robustness of the model. True predictive power of a QSAR model is to test their ability to predict accurately the anti-cancer activity of isatin compounds from an external test set: 14-15-16-17-24-25-2-27 (compounds which were not used for the model development). The comparison of the values of pIC<sub>50</sub> (test) to pIC<sub>50</sub> (obs.) shows that a good prediction has been obtained for the eight compounds. The R<sup>2</sup><sub>test</sub> values from the CoMFA and CoMSIA models using test set were found to be 0.607 and 0.650, respectively. The main performance parameters of the two models are shown in Table 5.

### Y-randomization

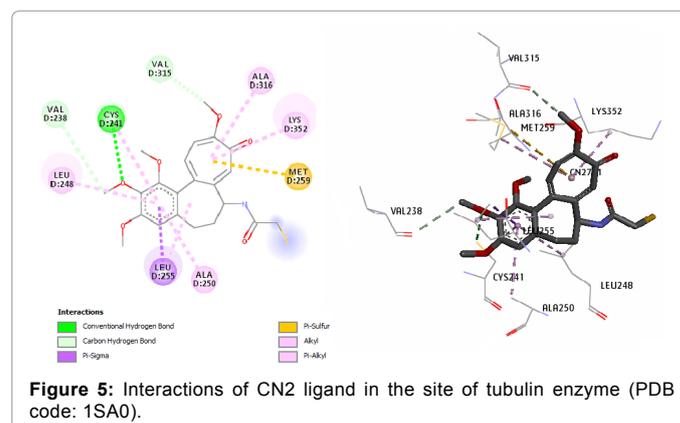
In this test, random CoMFA and CoMSIA models are generated by randomly shuffling the dependent variable while keeping the independent variables as it is. The new QSAR models are expected to have significantly low R<sup>2</sup> and Q<sup>2</sup> values for several trials, which confirm that the developed QSAR models are robust and the results of the CoMFA and CoMSIA methods are not due to a chance correlation of the training set.

### Docking analysis

The anticancer mechanism of this kind of compounds can be preliminarily regarded as the inhibitors against tubulin colchicine (PDB code: 1SA0) [28]. The docking study could offer more insight into understanding the protein-inhibitor interactions and the structural features of active site of protein. First of all, we adopted the known X-ray structure of tubulin in complex with the molecular ligand CN2 (2-mercapto-N-[1,2,3,10-tetramethoxy-9-oxo-5,6,7,9-tetrahydro-benzo[A]heptalen-7-yl] acetamide) to validate the docking reliability. The RMSD (Root Mean Square Distance) of the docked ligand was within the reliable range of 2 Å [28]. The redocked CN2 and crystal CN2 are almost at the same position in the active site of tubulin. Therefore, the docking protocol suggesting that CN2 could interact with the crystal structure of 1SA0 similarly to the crystallized CN2. Docking studies showing 3 hydrogen bond interactions of CN2 with 1SA0 enzyme at CYS 241 which has hydrophilic characters, also at VAL 238 and VAL 315 whose have hydrophobic characters (Figure 5). The methoxy groups of ligand CN2 formed two H-bonds with oxygen atom which exists in carbonyl groups of VAL 238 and VAL 315. The oxygen atom which exists in methoxy group of CN2 formed one H-bond with thiol group (-SH) of SYS 241. The most potent inhibitor compound 41 and the least potent inhibitor compound 19 were used as template to elucidate the interaction mechanism using. Then they were docked at the binding site of 1SA0 enzyme, using the Lamarckian Genetic Algorithm (LGA) available in Autodock Vina. The obtained results were then analyzed for detailed interactions in Discovery Studio Visualizer 4.5. The most possible interacting model between compound 41 and 1SA0 receptor, and the main residues involved were generally depicted in Figure 6. The more potent inhibitor (compound 41) is suitably situated at the tubulin binding site and there are many non-covalent interactions between it and the binding region of the enzyme. There are stabilizing hydrophobic interactions, originating from Pi-alkyl with LEU 248, LYS 352 and LEU

255 amino acid. The benzene ring of isatin corresponding to the most active compound 41 showed two Pi-alkyl interactions with LEU 248 and LYS 352 amino acids at distances 5.34 Å and 4.94 Å, respectively. The Bromine atom in position R<sub>5</sub> was buried under the hydrophobic pocket Figure 4c and formed a Pi-alkyl interaction with methyl group of LYS 352 amino acid at distance 3.87 Å. One part of naphthalene in position R<sub>6</sub> showed a Pi-alkyl interaction with LEU 255 amino acid at distance 4.94 Å. Three Pi-alkyl interactions were formed with methyl group of LEU 248, LYS 352 and LEU 255 amino acid, the aminoacids residues of LEU 248 and LEU 255 are hydrophobic in nature, which consistent with the yellow hydrophobic contour maps in the CoMSIA model Figure 4c, while the amino acid of LYS 352 is positively charged. The explanation for a small difference in activity could be presence and affected by the position of Bromine group in benzene ring of isatin (the binding mode of 3-Br is not the same as 5-Br and such small difference in activity). Here, the Figure 7 represents the interaction model of the compound 19 (compound with lowest activity has no substituent groups) with tubulin, its aromatic ring showed a Pi-donor hydrogen bond with residue SYS 241 which has hydrophilic characters at distance of 3.48 Å, a Pi-sigma interaction with residue LEU 255 aminoacid at distance of 3.49 Å, and a Pi-alkyl interaction with residue ALA 250 aminoacid at distance of 4.83 Å. In the other side, the pyrrole ring of isatin showed a Pi-donor hydrogen bond with residue SYS 241 at distance of 3.84 Å and three Pi-alkyl interactions with VAL 318, ALA 316 and LEU 255 aminoacids whose have hydrophobic characters at distances of 5 Å, 4.28 Å and 5.23 Å respectively.

The docking studies and the overall contour maps provide significant relationships between structural features and anti-cancer activity, it is found that the binding site of the isatin inhibitor is located in a hydrophobic pocket in the side of the tubulin enzyme. As aromatic and pyrrole rings, whose were capable of playing a principal role of the



**Figure 5:** Interactions of CN2 ligand in the site of tubulin enzyme (PDB code: 1SA0).

Designed compounds			Predicted pIC <sub>50</sub>		Log P	H-A	H-D	P.S	R.B	MW
R1	R2	R3	CoMFA	CoMSIA						
Br	OCH <sub>3</sub>	OCH <sub>3</sub>	6.58	5.02	0.87	7	0	193.85	5	613.95
OCH <sub>3</sub>	H	OCH <sub>3</sub>	6.52	5.22	0.20	7	0	199.41	5	534.05
Br	H	OCH <sub>3</sub>	6.17	5.03	5.58	6	1	132.22	4	600.05
Br	Br	H	6.11	5.69	1.99	5	0	159.55	3	557.93

**Table 5:** Chemical structure of newly designed molecules and their predicted pIC<sub>50</sub> based on CoMFA and CoMSIA 3D-QSAR models.



9. Vine KL, Locke JM, Ranson M, Benkendorff K (2007) An investigation into the cytotoxicity and mode of action of some novel N-alkyl-substituted isatins. *J Med Chem* 50: 5109-5117.
10. Matesic L, Locke JM, Bremner JB, Pyne SG, Skropeta D, et al. (2008) N-phenethyl and N-naphthylmethyl isatins and analogues as in vitro cytotoxic agents. *Bioorg Med Chem* 16: 3118-3124.
11. Zhou YJ, Zhu LP, Tang Y, Ye DY (2007) *Eur J Med Chem* 42: 977-984.
12. Liao SY, Qian L, Chen JC, Sheng Y, Zheng KC (2008) *J Theor Comput Chem* 7: 287-301.
13. Ravindra GK, Achaiah G, Sastry GN (2008) Molecular modeling studies of phenoxyrimidinyl imidazoles as p38 kinase inhibitors using QSAR and docking. *Eur J Med Chem* 43: 830-838.
14. Yi P, Qiu M (2008) 3D-QSAR and docking studies of aminopyridine carboxamide inhibitors of c-Jun N-terminal kinase-1. *Eur J Med Chem* 43: 604-613.
15. Clark M, Cramer RD, Van ON (1989) Validation of the general purpose tripos 5.2 force field. *J Comput Chem* 10: 982-1012.
16. Purcell WP, Singer JA (1967) A brief review and table of semiempirical parameters used in the Hueckel molecular orbital method. *J Chem Eng Data* 12: 235-246.
17. Agarwal A, Taylor EW (1993) 3-D QSAR for intrinsic activity of 5-HT<sub>1A</sub> receptor ligands by the method of comparative molecular field analysis. *J Comput Chem* 14: 237-245.
18. Cramer RD III, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110: 5959-5967.
19. Stahle L, Wold S (1988) Multivariate data analysis and experimental design in biomedical research. *Prog Med Chem* 25: 291-338.
20. Klebe G, Abraham U, Mietzner T (1994) Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J Med Chem* 37: 4130-4146.
21. Tetko IV, Tanchuk VY, Villa AE (2001) Prediction of n-octanol/water partition coefficients from PHYSPROP database using artificial neural networks and E-state indices. *J Chem Inf Comput Sci* 41: 1407-1421.
22. Golbraikh A, Tropsha A (2002) Beware of q<sup>2</sup>!. *J Mol Graph Model* 20: 269-276.
23. Rücker C, Rücker G, Meringer M (2007)  $\gamma$ -Randomization and Its Variants in QSPR/QSAR. *J Chem Inf Model* 47: 2345-2357.
24. Morris WE, Goodsell GM, Halliday DS, Huey RS, Hart R, et al. (1988) Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J Comput Chem* 19: 1639-1662.
25. Trott O, Olson AJ (2009) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31: 455-461.
26. <http://accelrys.com/products/collaborative-science/biovia-discovery-studio/>
27. Wren LD (2002) The PyMOL Molecular Graphics System DeLano Scientific. *PyMOL* 6: 1236-1238.
28. Liao SY, Qian L, Miao TF, Lu HL, Zheng KC (2009) *European Journal of Medicinal Chemistry* 44: 2822-2827.