Structure Requirements for 4-Aryl-4*H*-Chromenes as Apoptosis Inducers Using 3D QSAR Methods and Docking Studies

Zhen Liu[1], Yan Li[1,*], Hong Ren[2,3], Shuwei Zhang[1], Yonghua Wang[4], Guohui Li[3] and Ling Yang[5]

[1]School of Chemical Engineering, Dalian University of Technology, Dalian 116012, Liaoning, P.R. China
[2]Department of Ophthalmology, Qi Lu Hospital, Medical School of Shandong University, Jinan 250012, Shandong, P.R. China
[3]Laboratory of Molecular Modeling and Design, State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, Liaoning, P.R. China
[4]Center of Bioinformatics, Northwest A&F University, Yangling 712100, Shanxi, P.R. China
[5]Laboratory of Pharmaceutical Resource Discovery, Dalian Institute of Chemical Physics, Graduate School of the Chinese Academy of Sciences, Dalian 116023, Liaoning, P.R. China

*Corresponding author: Fax: +86 411 84896063; Tel: +86 411 84896062; E-mail: yling@dicp.ac.cn

Presently, a computational study based on the combinational use of 3D-quantitative structure-activity relationship analyses (QSAR) methods including both the comparative molecular field analysis (CoMFA) and comparative molecular similarity analysis (CoMSIA) approaches and molecule docking was conducted on a series of 124 types of 4-aryl-4*H*-chromenes of 24 diverse structural scaffolds as promising novel apoptosis inducers, with purpose to explore the requisite structural features influencing their activity of caspase-3 activation in human breast tumor cells. The obtained 3D-QSAR models exhibited proper reliability and predictivity, where the optimal comparative molecular field analysis and comparative molecular similarity analysis ones gave leave-one-out cross-validation coefficient Q^2 of 0.508 and 0.477, conventional cross-validation coefficient R^2_{ncv} of 0.888 and 0.816 for the training set and predictive correlation coefficients R^2_{pre} of 0.604 and 0.150 for the independent test set, respectively. Analyses of the derived contour maps reveal that steric substituents at positions 4, 9-13, 15 and 16 favour the apoptosis inducing activity and electron-withdrawing groups at 6 and 10 positions or electron-donating groups at position-7 enhance the activity. Further docking study validates that 4-aryl-4*H*-chromenes bind at the colchicine site of tubulin and several hydrogen bonds serve to stabilize the ligand-tubulin complex. These models and the derived information, would be of value for further exploration of the apoptosis inducing mechanism and the screening of novel potent chromene-based apoptosis inducers.

Key Words: Chromenes, Apoptosis inducer, CoMFA, CoMSIA, Caspase-3, Tubulin.

INTRODUCTION

As a normal physiological cell suicide process, apoptosis presents in the embryonic and adult development of many species and tissues, functioning as a balance retainer between the cell proliferation and death. Since the defects of programmed cell death contribute greatly to multiple diseases including cancers[1,2], the understanding and manipulation of the cell death machinery is always a hotspot of medical research[3,4]. Up to now, two major pathways of apoptosis have been identified[5], where the extrinsic pathway involves the interaction of a death receptor such as the TNF (tumor necrosis factor) receptor-1 or Fas receptor with its ligands, while the intrinsic (mitochondrial) pathway relates to the apoptosis initiated by chemotherapeutic agents, genotoxic stress and other death stimuli. In both pathways, the upstream caspases, the aspartate-specific cysteine proteases, are believed to cause the apoptotic phenotype by self-activation and consequently the activation of the downstream effectors, *i.e.*, the executioner caspases[5]. Actually, in animal cells caspases present in an inactive form, but when cleaving or degrading the substrates at aspartic acid (Asp) residues, they are self-activated[5,6]. Of the caspase family, caspases 3, 6 and 7 are apoptosis effectors, in which caspase-3 is especially the most prevalent one which is ultimately responsible for the majority of cleavage[5,6].

As another promising anticancer target, tubulin also attracts much attention of research in recent years, since series of microtubule-interfering agents like vinca alkaloids and taxanes have been demonstrated binding to tubulin and leading to the dynamic instability of the microtubule, cell cycle arrest at M phase[7,8], which effects are usually believed to be the reason triggering the molecular signalling for the mitochondrial

pathway of apoptosis[7-10]. In addition, although some tubulin-targeting chemotherapeutic agents (such as taxanes and vinca alkaloids) have gained certain success in clinical use, the emergency of drug resistance, dose-limiting neurologic and bone marrow toxicity has also become the bottleneck of these drugs[11]. Therefore, the development of more potent apoptosis inducers is still an urgent task which may present a new strategy for anticancer therapy.

From 2004 to 2008, Kemnizer and co-workers[11-14] synthesized a great number of 4-aryl-4*H*-chromenes using a cell- and caspase-based high-throughput assay and evaluated their capase-3 activation activities in cancer cells and reported the SAR results of the compounds. The 4-aryl-4*H*-chromenes possess vascular-disrupting activity[11,14] and have ability of inhibiting the tubulin from polymerization. Binding at or close to the binding site of colchicines[11], they are not only active in the multidrug resistant MESSA/DX5 tumor cells, but also exhibit high activity both as single agent and in combination with other antitumor agents in several animal tumor models[11,14]. Hence, it is very promising that these chromene-based compounds to be developed to novel anticancer agents as apoptosis inducers.

Due to above reasons, in silico studies, which have been successfully demonstrated as an alternative to the traditional time-consuming and costly trial and error approach of medicines development[15-19] were carried out on the chromene-based molecules. Up to now, only several computational studies including the quantitative structure-activity relationship (QSAR) analysis on this kind of molecules were published[20-23]. Afantitis *et al.*[20] and Fatemi and Gharaghami[21] established two QSAR models by using multiple linear regression (MLR) and support vector machine (SVM) methods on a same dataset consisting of 43 varities of 4-aryl-4*H*-chromenes, respectively. Based on a series of novel 3D descriptors, *i.e.*, the triplets of pharmacophoric point (TOPP), Sciabola *et al.*[22] published a partial least square (PLS) model by investigating 80 varities of 4-aryl-4*H*-chromenes. Liao *et al.*[23] also conducted a 3D-QSAR study using two widely used QSAR methods, *i.e.*, the comparative molecular field analysis (CoMFA)[24] and comparative molecular similarity analysis (CoMSIA)[25] on 36 varities of 4-aryl-4*H*-chromenes. In spite of these efforts, the interaction between these chromenes and the apoptosis target is, yet, still unclear.

In order to further explore the structure-activity relationships of the 4-aryl-4*H*-chromenes as apoptosis inducers, in the present work, CoMFA and CoMSIA 3D-QSAR methods were exploited on 124 varities of 4-aryl-4*H*-chromenes with 24 different structural scaffolds which is, to our best knowledge, the largest dataset of this kind of molecules up-to-date, to derive models. In addition to these ligand-based 3D-QSAR analyses, another receptor-based study by molecular docking was also performed on the molecules to deeply investigate their possible apoptosis inducing mechanism and explore the ligand-receptor interactions at the active site. This work would be of value for deeply exploration of the apoptosis inducing mechanism as well as the designing and forecasting of new chromene-based apoptosis inducers.

EXPERIMENTAL

Building of the dataset: In order to build a dataset as large as possible, but also keep the consistency of the structure and bioactivity to derive the computational models, a total of 124 types of 4-aryl-4*H*-chromenes with apoptosis inducing activities in human breast cancer tumor cell line T47D were collected from the work of Kemnitzer *et al.*[11-14] in the present study (Tables S1-S14, supplemental data). These chemicals belong to 24 kind of diverse structures, the representative structures of which are listed in Table-1. In a proportion of approximately 4:1, this dataset was divided into a training (96) and a test (28 molecules) sets respectively, the selection of which ensures that the test compounds represent as large structural diversity as possible and a range of biological activities similar to that of the training set. The caspase-3 activation activities (EC_{50}, half maximal effective concentration, μM) were converted into the corresponding pEC_{50} (-log EC_{50}) values, which span over a range from -0.86 to 2.7 and used as the dependent variable in the modeling process.

Molecular modeling: During the modeling process, the 3D structures of all compounds in the dataset were subjected to a full geometry optimization using SYBYL 6.9[26] with the Gasteiger-Marsili charges[27] assigned. Energy minimizations on these molecules were performed using the Tripos force field[28] with a distance-dependent dielectric and repeated minimization was performed using Powell conjugate gradient algorithm till the convergence criterion of 0.05 kcal/mol Å was achieved. For the molecular alignment of the compounds, a crucial step for the development of CoMFA and CoMSIA models[29], the most potent molecule of the dataset (compound 113) was chosen as a template to fit the remaining training and test sets of compounds by using database alignment option presently. Fig. 1 depicts the common substructure for the alignment in bold and Fig. 2 shows the resultant ligand-based alignment model.
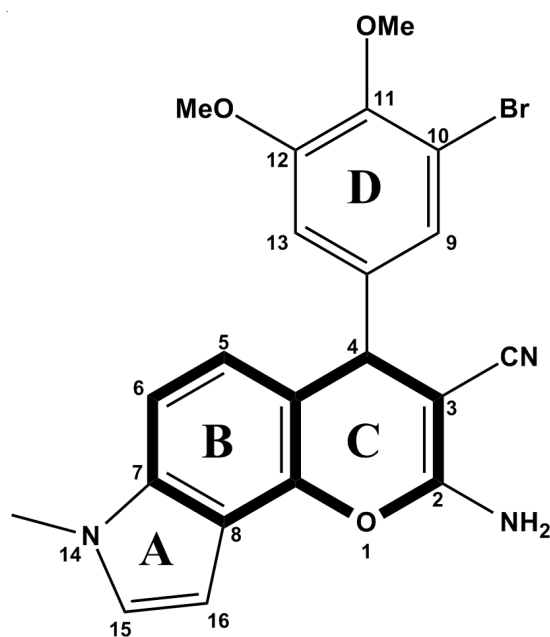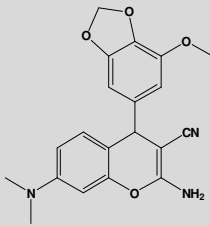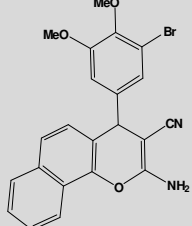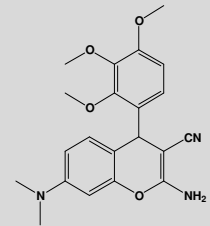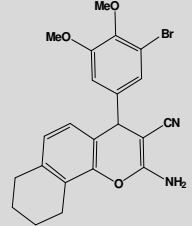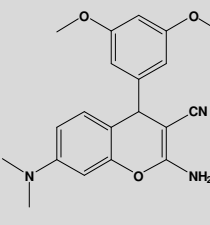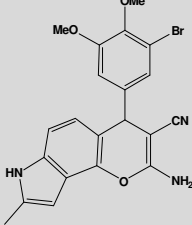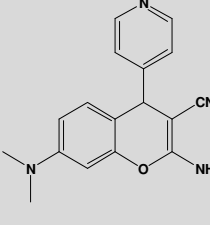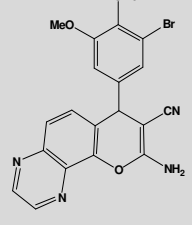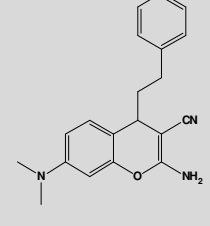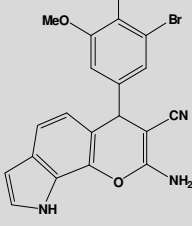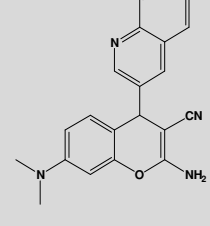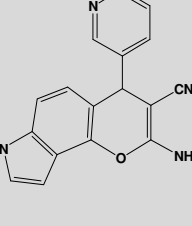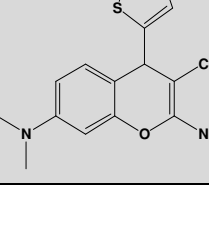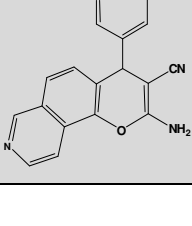

Fig. 1.　Compound 113 used as a template for alignment. The common substructure is shown in bold

	TABLE-1				
	REPRESENTATIVE STRUCTURES AND CASPASE-3 ACTIVATION ACTIVITIES OF THE DATASET
No	Structure	EC$_{50}$ (µM)	No	Structure	EC$_{50}$ (µM)
1		0.073	81		0.073
8		3.16	85		0.190
11		0.023	89		0.030
30		0.381	86		0.460
35		0.43	90		0.016
36		1.82	98		0.110
37		0.20	111		0.120
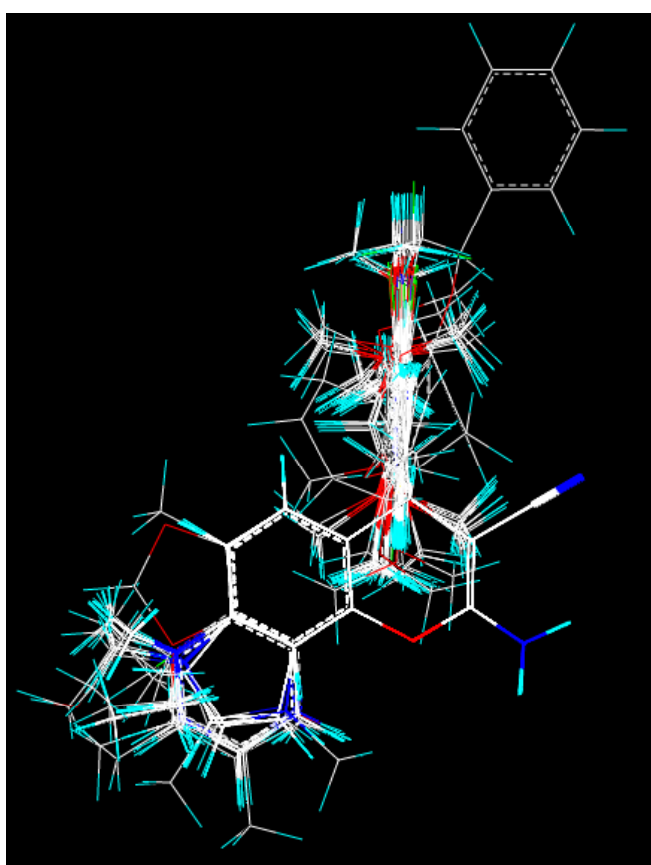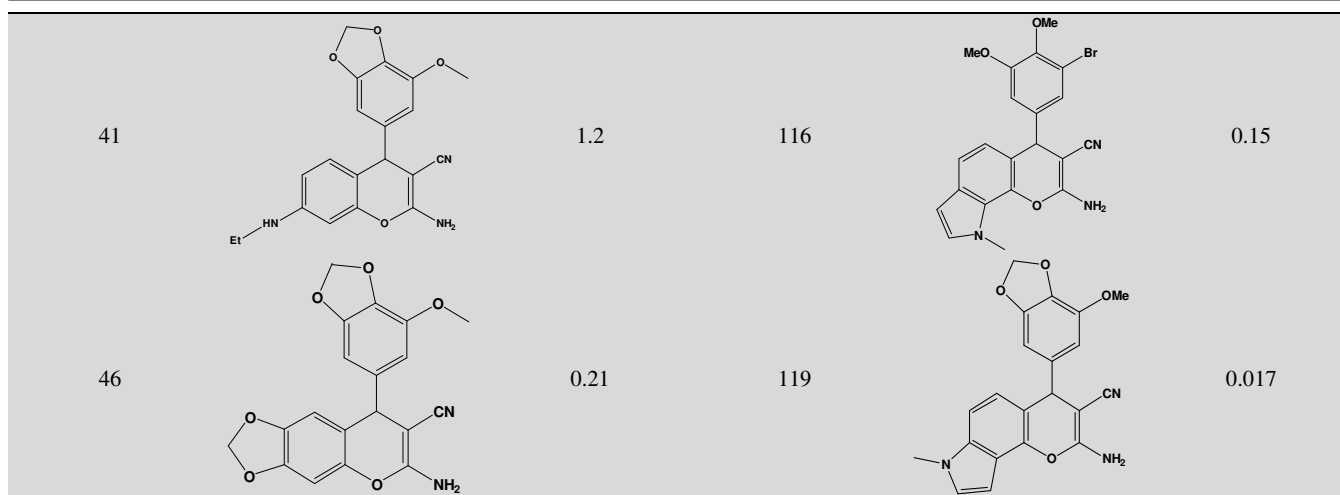
Fig. 2. Alignment of all compounds in the dataset

CoMFA and CoMSIA calculations: All the molecules were placed in a three-dimensional regular lattice with grid spacing of 2 Å in x, y and z directions. The steric (Van der Waals potentials) and electrostatic (Coulombic potential) fields were calculated separately for each molecule using standard Tripos force field. A sp^3-hybridized carbon atom probe with a charge of +1.00 and a van der Waals radius of 1.52 Å was employed as a probe atom to generate steric (Lennard-Jones 6-12 potential) and electrostatic (Coulombic potential) field energies. The energy cut-off values are 30 kcal/mol for both fields. The probe atom was placed at each lattice point and their steric and electrostatic interactions with each atom in the molecule were computed using the CoMFA standard scaling.

The same lattice boxes as those used in the CoMFA calculations were employed to derive the CoMSIA similarity index descriptors. In CoMSIA, the steric, electrostatic, hydrophobic and hydrogen-bond donor and acceptor descriptors were calculated using a probe atom of radius 1.0 Å, charge +1.0 and hydrophobicity +1.0. A Gaussian function is used to evaluate the mutual distance between the probe atom and each molecule atom. CoMSIA similarity indices (A_F) for molecule j with atom i at grid point q are calculated by eqn. 1:

$$A_{F,k}^q(j) = -\sum \omega_{probe,k} \omega_{ik} e^{-\alpha r_{iq}^2} \qquad (1)$$

where k represents the steric, electrostatic, hydrophobic, or hydrogen-bond donor or acceptor descriptor, $\omega_{probe,k}$ is the probe atom with radius 1.0 Å, charge +1, hydrophobicity +1, hydrogen bond donating +1, hydrogen bond accepting +1, ω_{ik} is the actual value of the physicochemical property k of atom i and r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule, respectively. The attenuation factor 'a' was set to 0.7.

In the end, MLogP, a descriptor as the logarithm of the 1-octanol/water partition coefficient frequently used in the medical and biological studies, was also applied in this work, which values for the dataset were calculated from Dragon software[30].

Partial least square (PLS) analyses and validations: To develop 3D-QSAR models, partial least squares regression was used to analyze the training set by correlating their pEC_{50} values (the dependent variable) with variations in the CoMFA and CoMSIA interaction fields (the independent variables). To evaluate the reliability of the models generated from the PLS analysis, the leave one out (LOO) cross-validation method in which one compound was omitted from the dataset and its activity was predicted by using the model derived from the rest of the dataset was used, which approach provides a cross-validated correlation coefficient Q^2 for evaluation of the quality of the resultant model. Then, the number of components identified in the LOO cross-validation process was used in the non-cross-validated PLS analysis, where the non-cross-validation coefficient (R_{ncv}^2) and standard error of estimates (SEE) were calculated. Finally, the CoMFA and CoMSIA results were graphically represented by way of field contour maps, where the coefficients were generated using the field type "Stdev*Coeff".

To further test the real predictive ability of the optimal models generated by the CoMFA and CoMSIA analyses using the training set, the pEC_{50} values of the test set compounds are further calculated by the obtained models to evaluate the prediction error by comparison with the experimental activities. A predictive R^2 value was thus obtained according to the following formula (2):

$$R^2_{pre} = 1 - \frac{\sum (Y_{pre} - Y_{exp})^2}{\sum (Y_{exp} - \overline{Y})^2} \qquad (2)$$

where Y_{pre} and Y_{exp} represent the predicted and observed activity values, of the test set, respectively and $\overline{Y}$ represents the mean activity value of the training set.

Molecular docking: According to Kemnitzer *et al.*[11,14], the 4-aryl-4*H*-chromenes are tubulin inhibitors which bind at or close to the colchicines site of β-tubulin and these compounds induce apoptosis through the inhibition of tubulin polymerization. However, up to now no molecule docking research has been reported on this kind of molecules. Therefore, in our work, a docking study on these 4-aryl-4*H*-chromenes was carried out by use of Surflex-dock module (V 2.51) of another advanced version of SYBYL package (X 1.1)[31] to validate the apoptosis inducing mechanism of these molecules. An X-ray structure of tubulin in complex with colchicine and a stathmin-like domain obtained by Ravelli *et al.*[32] with a protein data bank entry of 1SA0 and a resolution 3.58 Å from the RCSB Protein Data Bank (http://www.pdb.org/pdb/home/home.do) was employed to conduct the docking process, which structure has been successfully applied in several studies to explore the antitumor mechanisms of different compounds[33,34]. The docking was executed following such procedures: Firstly, the crystal structure of tubulin was imported into Surflex-dock with its original small molecules except the ligands were removed. Secondly, the protomol was generated using a ligand-based approach. During this procedure, two parameters critical for forming an appropriate binding pocket, *i.e.*, the protomol_bloat which determines how far from a potential ligand the site should extend and the protomol_threshold which determines how deep into the protein the atomic probes used to define the protomol can penetrate, are adjusted repeatedly. In the end, the protomol_bloat value was set to 1 and the protomol_threshold value with 0.47 when a reasonable binding pocket was obtained. Finally, all the 4-aryl-4*H*-chromenes were docked into the binding pocket and each of them got 10 different conformations.

RESULTS AND DISCUSSION

CoMFA and CoMSIA statistical results: To obtain an effective 3D-QSAR model, a number of statistical parameters, including the cross-validated correlation coefficient (Q^2), non-cross-validated correlation coefficient (R^2_{ncv}), standard error estimate (SEE) and F-statistic values should be analyzed. Generally speaking, high Q^2, R^2_{ncv} and F values along with the low standard error estimate values are considered proof of the high internal predictive ability[35,36]. However, sometimes it is found that, by virtue of a chance correlation or structural redundancy, some models derived by modeling of the training

set with randomized activity possess high Q^2 values, but show unfavorable predictive power for unknown molecules[36]. Hence, only by evaluating the value of the LOO cross-validated Q^2 is insufficient to assess the predictive power of the QSAR models[36]. In light of such risks, we further validate the models obtained from analysis of the training set by predicting the apoptosis inducing activities of the external test set compounds that are different from those used in the training set. Table-2 summarizes the statistical results obtained from the optimal CoMFA and CoMSIA models based on the same training set.

TABLE-2
SUMMARY OF CoMFA AND CoMSIA RESULTS*

PLS statistics	CoMFA		CoMSIA	
	1	2	3	4
Q^2	0.508	0.494	0.477	0.464
R^2_{ncv}	0.888	0.894	0.816	0.794
SEE	0.264	0.259	0.340	0.360
R^2_{pre}	0.604	0.623	0.150	0.137
SEP	0.409	0.403	0.607	0.599
F	76.505	42.422	38.241	33.106
N	9	10	10	10
	Contribution %			
S	0.466	0.467	0.202	0.191
E	0.534	0.527	0.798	0.804
M	–	0.006	–	0.006

*Q^2: Cross-validated correlation coefficient after the leave-one-out procedure. R^2_{ncv}: non-cross-validated correlation coefficient. SEE: standard error of estimate. R^2_{pre}: predicted correlation coefficient for the test set of compounds. SEP: standard error of prediction. F: ratio of R^2_{ncv} explained to unexplained = $R^2_{ncv}/(1-R^2_{ncv})$. N: optimal number of principal components. S: steric field. E, electrostatic field. M: M log P.

To avoid the possibility of omitting the best QSAR model, all combinations of the CoMFA and CoMSIA fields *i.e.*, the steric, electrostatic, hydrogen bond (H-bond) donor, H-bond acceptor and hydrophobic field descriptors and MLogP are used to seek for the optimal models. In this study, a compound with a high residual between the observed and predicted biological activity as an apoptosis inducer was taken as an outlier. For this reason, compounds 27, 32, 40 and 115 are considered as outliers. Although microgubule-targeting agents (MATs) such as taxanes and vinca-alkaloids are generally believed to bind to the tubulin/microtubule system and consequently induce cell cycle arrest and apoptosis, they may also target the mitochondria[10]. Green *et al.*[37] showed that numbers of key events in apoptosis focus on mitochondria, so both microtubule and mitochondria should be investigated as targets for microbtubule-targeting agents[10], so different mechanisms of the 4-aryl-4*H*-chromenes as apoptosis inducers may be the reason for the emergency of outliers.

During the modeling process, to our disappointment, most of the models exhibited poor predictive ability with Q^2 less than 0.4 and R^2_{ncv} lower than 0.75, respectively. The reason is might due to the too many diverse types of scaffolds for this series of molecules. However, after omitting the above outliers, several CoMFA and CoMSIA models with proper robustness and reliability were also obtained (Table-2).

An interesting phenomenon was observed that only those models containing the combinational use of the steric and electrostatic fields exhibited statistically satisfactory results (with

a Q^2 above 0.45), which indicates the importance of these two kind of interactions for the apoptosis inducing activity of the molecules.

For example, the optimum model established in the present study is the CoMFA-SE model, where S and E represent the steric and electrostatic field descriptors employed to correlate with the biological activities when building the models, respectively. For this model, a cross-validated correlation coefficient $Q^2 = 0.508$, a non-cross-validated $R^2_{ncv} = 0.888$, a standard error of estimate SEE = 0.264 and an F = 76.505 was obtained using 9 optimum components. When validated by the independent test set, an $R^2_{pre} = 0.608$ and SEP = 0.404 was achieved, proving its good predictivity. For CoMSIA analysis, same phenomenon was observed that only models making use of the steric and electrostatic fields obtained satisfactory outcomes, where the CoMSIA-SE model gives a Q^2 value of 0.477, an R^2_{ncv} value of 0.816, an SEE value of 0.340 and an F value of 38.241 with 10 optimum components for the training set and an R^2_{pre} of 0.150 and SEP of 0.607 when validated by the external test set.

Besides, by comparison of the relative contributions of the two fields to each models in Table-2, it is easily to draw a conclusion that electrostatic field is a more vital factor influencing the apoptosis inducing activities by contributing a few larger than the steric field for building the models, since for both the two optimal CoMFA models (CoMFA-SE and CoMFA-SEM), 0.47: 0.53 was the calculation result as the ratio of the respective contribution of steric to electrostatic field and for both the two optimal CoMSIA models (CoMSIA-SE and CoMSIA-SEM), this value is about 0.21:0.79. However, despite the discovery of the less contribution of steric field than the electrostatic one, it is found that, actually, both the two fields are crucial and requisite for the apoptosis inducing activity of the dataset, since without any of the fields none CoMFA or CoMSIA models can be established with approving statistical results in present studies.

Many studies including our previous investigations[18,38] have demonstrated that the hydrophobicity, a term interpreted to be the association of non-polar groups or molecules in an aqueous environment arising from the tendency of water to exclude non-polar molecules[39], is one of the most important properties related to many biomolecular, especially the ligand-receptor interactions. Naturally, in present study, a question comes out that whether at this time, despite the contributions coming from the steric and electrostatic interactions, the apoptosis inducing activities of the 4-aryl-4*H*-chromene molecules are still tightly correlated with their hydrophobic properties accounting for their aqueous solubility, cell permeability and affinity to the target. Therefore, presently an additional descriptor of MLogP, the logarithm of the 1-octanol/water partition coefficient, a widely computational study used hydrophobicity parameter was also employed in the attempt of building QSAR models for the 4-aryl-4*H*-chromene dataset, with purpose to explore the effect of hydrophobic/lipophilic properties of the molecules on the apoptosis inducing activity.

As a result, it is found that the apoptosis inducing activities used in the study are also correlated to this hydrophobicity variable. Both the derived CoMFA-SEM and CoMSIA-SEM

models (where M represent the MLogP parameter employed) show that their results are statistical acceptable. However, the addition of this special descriptor does not obviously improve the internal predictivity, but results in, contradictorily a slightly decrease of the cross-validated Q^2 values for the CoMFA and CoMSIA models (from 0.508-0.494 and 0.477-0.464, respectively) and a slight improvement of the external predictive power when validated by the test set for the CoMFA models (from 0.604-0.623). By consideration of the comprehensive impact of MLogP in the models, we have to say that for 4-aryl-4*H*-chromene derivatives, their capase-3 activation activities are related to a small extent, to their hydrophobic properties.

Fig. 3 depicts the respective plots of observed *versus* calculated pEC_{50} values for the training (filled squares) and the test (filled dots) set molecules of the four optimal CoMFA and CoMSIA models. All of these plots show a uniformly distribution of the points around the regression line and good correlations between the calculated activities and the experimental ones, proving the good reliability and predictability of the models.

3D-QSAR contour maps: As one of the attractive characteristics of CoMFA and CoMSIA methods, the visualization of the 3D contour plots are beneficial to identify important regions where changes in the steric and electrostatic fields may
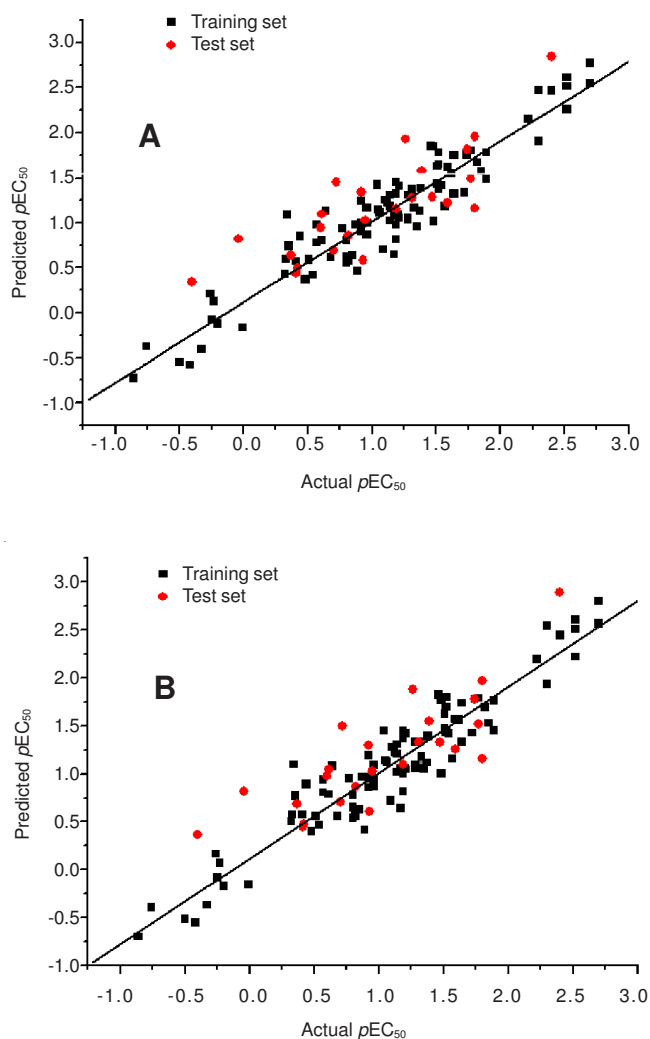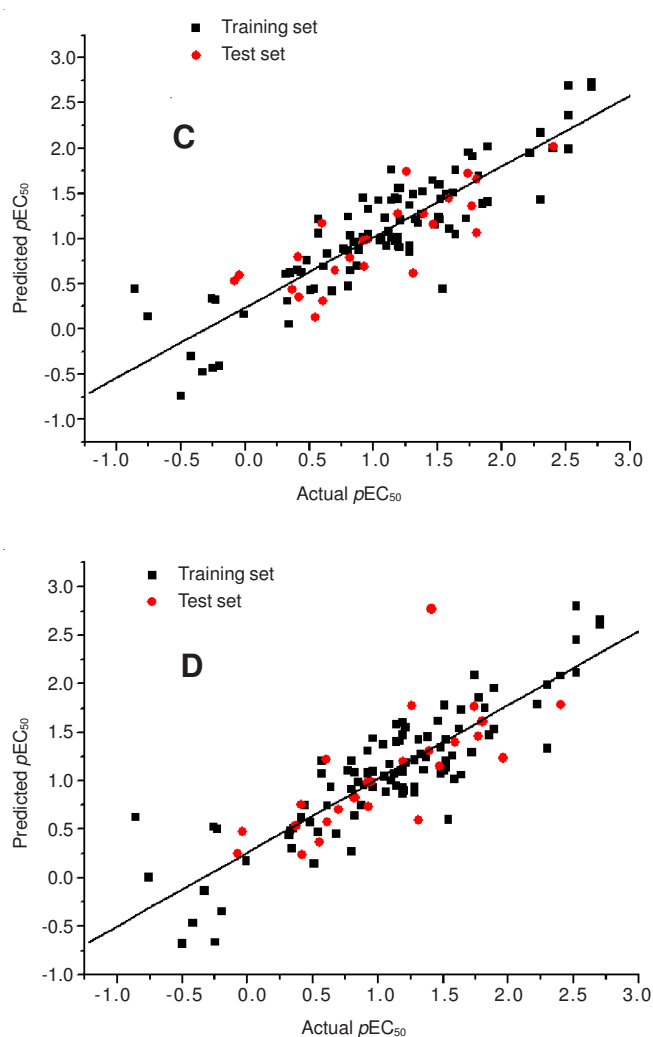
Fig. 3. Correlation plots of the observed *versus* the calculated pEC_{50} values using the training (filled squares) and the test sets (filled dots) based on the A: CoMFA-SE model, B: CoMSIA-SEM model, C: CoMSIA-SE model and D: CoMFA-SEM model

affect the biological activity and in addition may also help to identify the possible interaction sites. Since the resultant CoMSIA models are inferior to the CoMFA models both in internal and external predictabilities, so only the contour maps derived from the two CoMFA models are displayed in this study. To aid in visualization, one of the most active compounds (compound 113) in the series is shown superimposed with the contour maps (Figs. 4 and 5).

Fig. 4 depicts the steric field contours of CoMFA-SE model (Fig. 4A) and CoMFA-SEM model (Fig. 4B).These two contours are almost the same, so only the contours from the CoMFA-SE model is illustrated here. The green and yellow contours describe regions of space around the molecules where an increase in steric bulk enhances (80 % contribution) and diminishes (20 % contribution) the apoptosis inducing activity, respectively.

In the contour plots, yellow isopleths can be found adjacent to positions-4, 9-13, 15 and 16 of the template compound 113. These results are consistent with the reported experimental data. The addition of a methyl group of compound 39 at position-4 may encounter unfavorable steric interaction with the
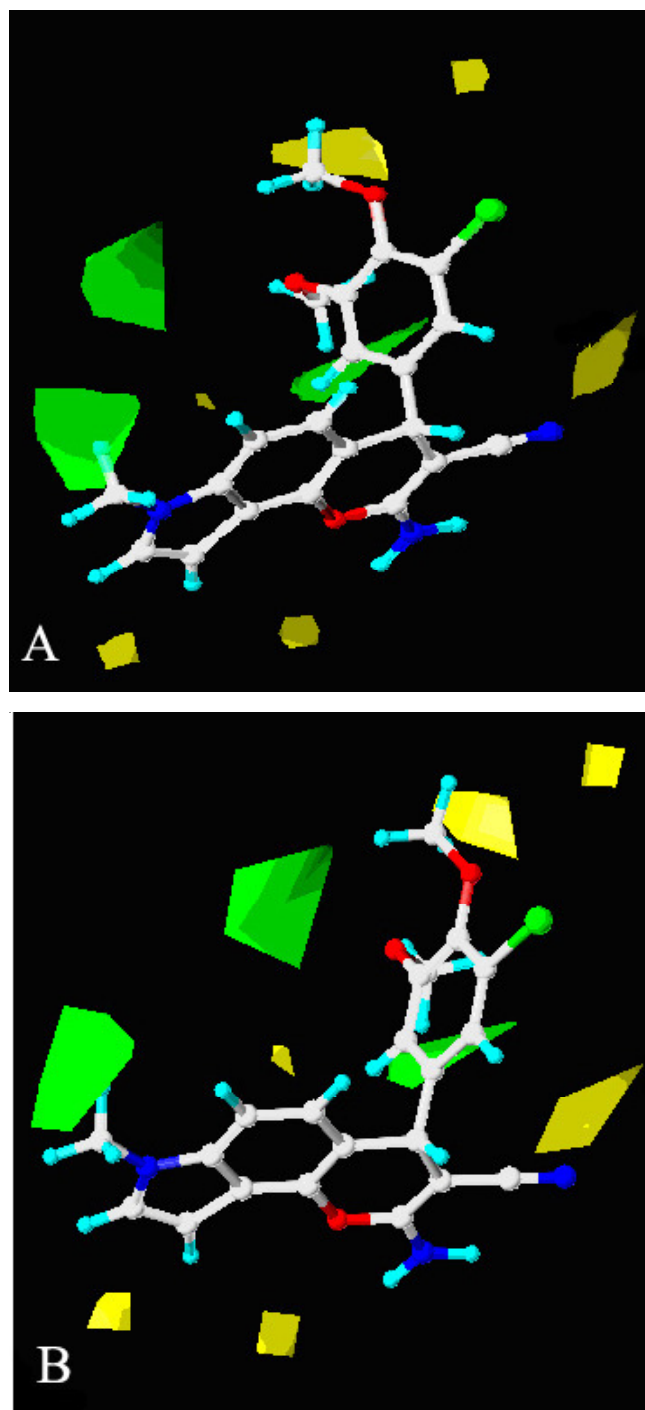
Fig. 4. Steric field contour maps of the 3D-QSAR models shown with the template compound **113**. (A): CoMFA-SE model; (B): CoMFA-SEM model. The compound is presented in ball and stick style. Green and yellow isopleths enclose regions where steric interaction is favored and disfavored, respectively

receptor, which is the cause of the 178-fold decrease in activity compared to compound 3. In Fig. 4A, a yellow contour appears near the 9 position and this fact does explain the better activity of compound 25 as compared to compound 27. Compound 2 retains an OMe group which is more bulky than the Br, Cl and I groups of 3, 4 and 5, respectively, at position 10, so compound 2 is less potent than the other three compounds. Substitution of the H at position 10 of compound 25 with a more bulky OCF_3 group results in *ca.* 4 fold decrease in activity of
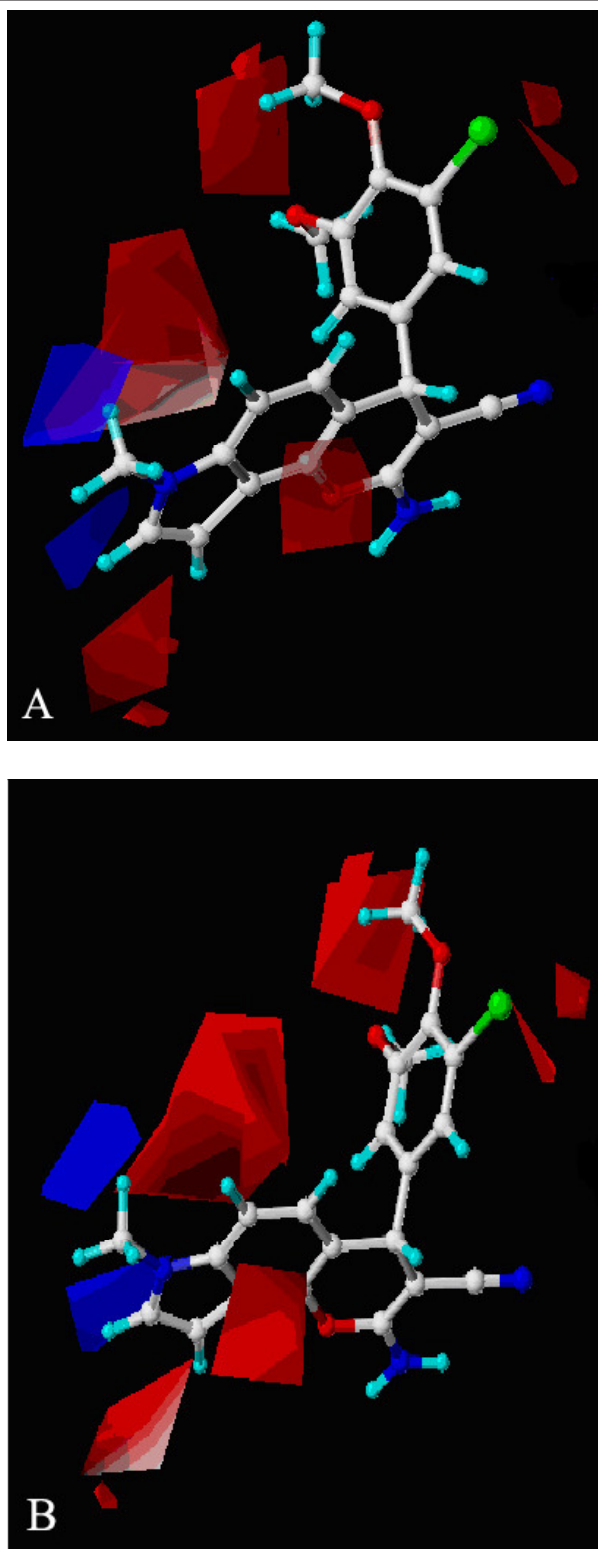
Fig. 5.   Electrostatic contour maps of the 3D-QSAR models shown with the template compound **113**. (A): CoMFA-SE model; (B): CoMFA-SEM model. Blue and red contours correspond to regions where an increase in positive or negative charges will increase activity, respectively

compound 23. Comparison of the activity of compound 36 (1.82 μM) with that of 31 (0.033 μM) and 32 (0.011 μM) reveals that the steric disfavored fused ring structure at the 10 and 11 positions compared to the Br or CH$_3$ groups at the 10 position of 31 and 32, respectively, of 36 accounts for the low activity.

The relatively low activity of compound 2-11 is due to the OMe group of compound 2 encounters disfavored steric interaction at the 11 position. Similarly, this can explain the lower activities of compounds 8, 26 and 60 in comparison to compounds 15, 25 and 61, respectively. The ring structures at the 11 and 12 positions of compounds 1, 6 and 119 extend to the steric disfavored region, so this is the reason why they are less potent to compounds 2, 3 and 120, respectively. The contour maps also present a yellow contour near position 13 indicating steric bulky groups may decrease activity in this position. So compound 8 with an OMe group at this position is less potent than 14 can be explained as more bulky groups are not tolerated at this position. An addition of a more bulky methyl group to compound 88 at the 15 position accounts for the 6-fold decrease in activity of compound 89. The same reason can explain the decreased activity of compound 116 compared to compound 90.

In CoMFA electrostatic field contours, the red (negatively charged favored) and white (positively charged favored) contours represent 80 and 20 % level contributions, respectively (Fig. 5A-B for CoMFA-SE and CoMFA-SEM model, respectively). In both contours, negatively charged favored contours are observed at position 10, indicating electron-withdrawing groups which can enhance activity at this position. This can be illustrated by compounds 3, 4 and 5 *versus* 2, for the Br, Cl, I groups of the formers are more negatively charged than the OMe group of the latter. Besides, the less potency of compound 66 compared to compounds 63, 64 and 65 and 95 compared to 94, 96 and 97 also demonstrates that electron-withdrawing groups are favored for activity at the 10 position. In Fig. 5A, positively charged blue contour can also be seen near the 7 position so compound 50 is more active than compounds 53 and 54. This can be explained as an electron-donating methoxy group is more tolerant than the electro-withdrawing Br and Cl groups in this position. Besides, Fig. 8A also depicts a red contour adjoining the 6 position, so an electron-donating methyl group of compound 45 decrease the activity relative to compound 41.

Docking studies: To authenticate the credibility of the docking method to predict the bioactive conformation, the CN2 (2-mercapto-N-[1,2,3,10-tetramethoxy-9-oxo-5,6,7,9-tetrahydro-benzo[A]heptalen-7-yl]acetamide) molecule was extracted from the X-ray structure of tubulin (pdb id: 1SA0) and subsequently docked into the active site of crystal structure of tubulin. The resultant conformation corresponding to the highest total score was selected as the most probable bioactive conformation. Surflex-dock predicted conformation of CN2 was perfectly superimposed on the X-ray crystallographic structure of the same ligand as shown in Fig. 6, suggesting high reliability of Surflex-dock in terms of reproducing the experimental binding mode of CN2.

After the validation of the accuracy of the docking procedure, all the 124 types of 4-aryl-4*H*-chromenes were docked to tubulin based on the same protocol. After docking each compound has ten different conformations. There was no precise correlation between the highest total docking scores and the pEC$_{50}$ values. However, after manually selection, a correlation analysis between the total scores and the pEC$_{50}$ values presents an R$^2_{ncv}$ of 0.235, revealing a proper correlation between the
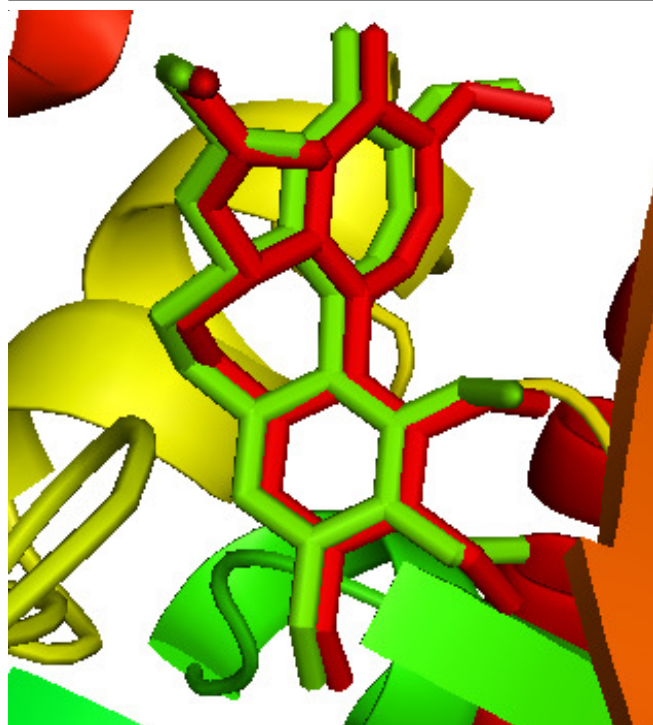
Fig. 6. Superposition of the crystal ligand (the green molecule) and the Surflex-dock predicted best docked conformation (the red molecule) of CN2

Fig. 7. Docked conformation of compound **113** within the colchicine-binding site of tubulin. Hydrogen-bond is shown as dotted white lines. The active amino acid residues Ala 317 and Val 318 and the ligand are presented as sticks

Fig. 8. Binding site of compound **113** at tubulin. (A) Compound **113** in the binding pocket of tubulin; (B) Key residues are labeled around the compound 113

docking conformation and the apoptosis inducing activities of the dataset as well as a reasonable docking model. The selection was conducted in such a way that docking conformations with comparatively higher scores were chosen for those compounds with higher apoptosis inducing activities and *vice versa*.

All the 4-aryl-4*H*-chromenes bind at the colchicines site of β-tubulin in a similar mode. One of the most potent compounds, compound 113 is taken as an example to explain the binding mode of the 4-aryl-4*H*-chromenes. It can be observed that the ligand core is anchored in the binding site *via* some hydrogen bonds (Fig. 7). The cyano group at position 3 forms a hydrogen bond to the backbone NH of Thr353 as a hydrogen bond acceptor (d = 3.24 Å, θ = 60.58º). The -NH$_2$ group of compound 113 at the 2 position forms several hydrogen bonds with backbone carbonyl oxygen atom of Ala316 (-NH···O=C, d = 3.05 Å, θ$_1$ = 136.53º), backbone carbonyl oxygen atom and NH of Ala317 (-NH···N, d$_1$ = 2.35 Å, θ$_1$ = 103.18º, -NH···O=C, d$_2$ = 2.52 Å, θ$_2$ = 75.27º, -NH···N, d$_3$ = 1.97 Å, θ$_3$ = 106.50º, -NH···O=C, d$_4$ = 2.59 Å, θ$_4$ =89.48º, N···HN, d$_5$ = 2.70 Å, θ$_5$ = 83.54º), the backbone NH of Val318 (-NH···N, d = 3.07 Å, θ = 115.98°) and the backbone NH of Ala354 (-NH···N , d = 3.36 Å, θ = 128.09°) of the receptor. These results are in line with the CoMSIA results of Liao *et al.*[23].

Fig. 8 depicts the binding conformation of 113 in the binding pocket of the tubulin. Rings A, B and C fit snugly into a narrow pocket composed by Val238, Cys241, Leu242, Leu248, Ala250, Leu252, Leu255, Ala316, Val318, Ala354 and Ile378. The substituents at positions 15 and 16 may be involved in steric clash with the side chains of residues Val238, Leu242 and Leu255, which can be evidenced by the presence of steric disfavored yellow contours near these positions by the CoMFA model.
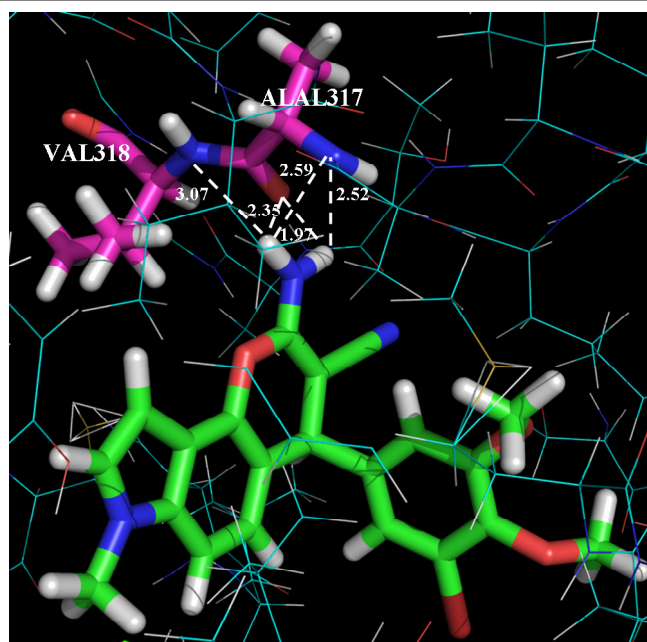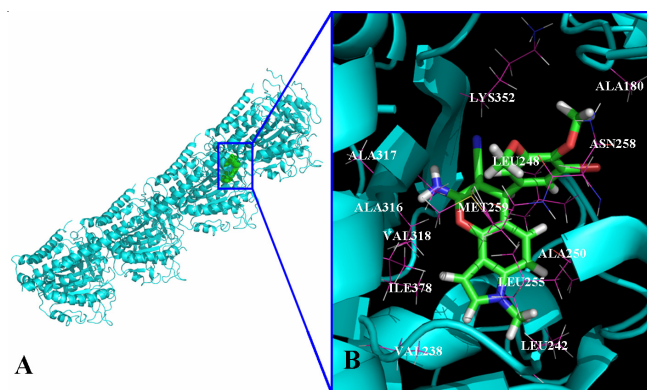
The substituent at position 4 and those on ring D may also be involved in steric interactions with the side chains of residues constituting the active site. These observations are also corroborated by the above CoMFA results. The long side chain of residue Leu248 resides adjacent to positions 4 and 9. The substituent appended to position-4 is blocked by the side chain of this residue, so the additional methyl group of compound 39 at position-4 may encounter unfavorable steric interaction with the receptor, which is the cause of the 178-fold decrease in activity compared to compound 3. For the same reason, a more space-filling group at position-9 will be deleterious to the apoptosis inducing activity. A steric group at position-10 may be involved in unfavorable steric conflict with Lys254. The side chain of residue Asn258 appears between positions-11 and 12, suggesting large substitutes, especially the fused ring structure (compound 1, 6 and 119) at these positions will conflict with this residue. Besides, the side chains of Ala180, Met259 and Ala316 may also engage in steric

conflicts with the substituent at positions 11, 12 and 13, respectively.

The two negatively charged favored red contours adjacent to positions-6 and 10 correspond to the positively charged side chains of residues Ala250, Lys254 and Leu255, so only electron-withdrawing groups at these positions can effectively interact with above residues. The positively charged favored blue contours corresponds to the electron-rich C=O group of Val238, so the substituent at position 7 should be electron-donating groups.

As a consequence, the structural insights obtained from molecular docking and those from 3D-QSAR modeling can validate each other very well. What's more, the consistency of the available experimental activity data indicates the molecular docking is reliable and that the developed 3D-QSAR models are reasonable.

Comparison with previous work: Previous to our work, several QSAR studies on 4-aryl-4*H*-chromenes has been reported[20-23] in which different in silico methods has been exploited to correlate the structure diversifications to the differences in the apoptosis inducing activities. From the work of Afantitis *et al.*[20] and Fatemi *et al.*[21], it can be concluded that both steric parameters and electronic interactions can affect the apoptosis inducing activities of 4-aryl-4*H*-chromenes, which is in form with our CoMFA results.

Besides, the previous studies[20-23] also highlight the hydrophobicity and hydrogen-bonding abilities of the 4-aryl-4*H*-chromenes also contribute to the apoptosis inducing activities. In our study, the introduction of MLogP, a hydrophobic parameter, which contributes least (0.006) to the apoptosis inducing activities, results decrease in Q^2 values but increase in R^2_{pre} values. So our results demonstrate that hydrophobicity really contribute to the apoptosis inducing activities for the chromenes.

The CoMFA and CoMSIA studies of Liao *et al.*[23] provide some key factors responsible for the apoptosis inducing activities. According to their work, the steric interaction is the decisive factor in determining the apoptosis inducing activities. Steric interaction is disfavored at the 6 position. A sterically moderate, highly electropositively and hydrophobic alkyl terminal at the position 7 is favored. A hydrophobic substituent at the 8 position is favored. A less bulky and weak electron-withdrawing group is favored at the 10 position. Simultaneous introduction of hydrogen bond acceptors at the first atoms at the 11 and 12 positions is favorable to the activity but a link between these positions may be adverse. The steric field analyses at the 11 and 12 positions are consistent with our results but the electrostatic field analysis at the 10 position is adverse. This phenomenon is probably due to the differences in the dataset or molecular diversities between the two studies. However, taken all these results together, a more comprehensive understanding of the mechanism for the 4-aryl-4*H*-chromenes as apoptosis inducers could be obtained.

Conclusion

The 4-aryl-4*H*-chromenes were found to be a promising series of novel apoptosis inducers. In this study, a combination of three-dimensional quantitative structure-activity relationship analyses and molecule docking is conducted on 124 types of 4-aryl-4*H*-chromenes with their activity of caspase-3 activation in human breast cancer tumor cell line using CoMFA and CoMSIA.

The derived best CoMFA model gives a Q^2 of 0.508, R^2_{ncv} of 0.888, SEE of 0.264, F value of 76.505, R^2_{pre} of 0.604, SEP of 0.409. The best CoMSIA model gives a Q^2 of 0.477, R^2_{ncv} of 0.816, SEE of 0.340, F value of 38.241, R^2_{pre} of 0.150, SEP of 0.607. These results demonstrate the robustness and predictive power of the established models.

Analyses of the 3D contour maps reveal that at positions 4, 9-13, 15 and 16, steric interactions are disfavored for apoptosis inducing activities, besides modifications of electron-withdrawing groups at 6 and 10 positions or electron-donating groups at the 7 position will enhance apoptosis inducing activities. Docking study shows that 4-aryl-4*H*-chromenes bind at the colchicine site of tubulin and the ligand-tubulin complex is stabilized by several hydrogen bonds between the chromenes and the receptor residues. In addition, a good consistency was also found between our docking models and the 3D-QSAR models. All these results should provide information for a better understanding of the apoptosis inducing mechanism of these chromenes and thus be helpful in design of the new chromenes analogues in the future.

REFERENCES

1. J.J. Bright and A. Khar, *Biosci. Rep.*, **14**, 67 (1994).
2. S. Kasibhatla and B. Tseng, *Mol. Cancer Ther.*, **2**, 573 (2003).
3. J.C. Reed and K.J. Tomaselli, *Curr. Opin. Biotechnol.*, **11**, 586 (2000).
4. W. Hu and J.J. Kavanagh, *Lancet Oncol.*, **4**, 721 (2003).
5. K.C. Zimmermann and D.R. Green, *J. Allergy Clin. Immunol.*, **108**, S99 (2001).
6. R.M. Friedlander, *N. Engl. J. Med.*, **348**, 1365 (2003).
7. A.L. Risinger, F.J. Giles and S.L. Mooberry, *Cancer Treat. Rev.*, **35**, 255 (2009).
8. K.N. Bhalla, Microtubule-Targeted Anticancer Agents and Apoptosis, Nature Publishing Group, Basingstoke, Royaume-Uni (2003).
9. F. Mollinedo and C. Gajate, *Apoptosis*, **8**, 413 (2003).
10. M.A. Esteve, M. Carre and D. Braguer, *Curr. Cancer Drug Targets*, **7**, 713 (2007).
11. W. Kemnitzer, S. Kasibhatla, S. Jiang, H. Zhang, J. Zhao, S. Jia, L. Xu, C. Crogan-Grundy, R. Denis, N. Barriault, L. Vaillancourt, S. Charron, J. Dodd, G. Attardo, D. Labrecque, S. Lamothe, H. Gourdeau, B. Tseng, J. Drewe and S.X. Cai, *Bioorg. Med. Chem. Lett.*, **15**, 4745 (2005).
12. W. Kemnitzer, J. Drewe, S. Jiang, H. Zhang, C. Crogan-Grundy, D. Labreque, M. Bubenick, G. Attardo, R. Denis, S. Lamothe, H. Gourdeau, B. Tseng, S. Kasibhatla and S.X. Cai, *J. Med. Chem.*, **51**, 417 (2008).
13. W. Kemnitzer, J. Drewe, S. Jiang, H. Zhang, Y. Wang, J. Zhao, S. Jia, J. Herich, D. Labreque, R. Storer, K. Meerovitch, D. Bouffard, R. Rej, R. Denis, C. Blais, S. Lamothe, G. Attardo, H. Gourdeau, B. Tseng, S. Kasibhatla and S.X. Cai, *J. Med. Chem.*, **47**, 6299 (2004).

14. W. Kemnitzer, J. Drewe, S. Jiang, H. Zhang, J. Zhao, C. Crogan-Grundy, L. Xu, S. Lamothe, H. Gourdeau, R. Denis, B. Tseng, S. Kasibhatla and S.X. Cai, *J. Med. Chem.*, **50**, 2858 (2007).

15. I.M. Kapetanovic, *Chem. Biol. Interact.*, **171**, 165 (2008).

16. Q. Li, X. Kong, Z. Ma, L. Zhang, F. Wang, H. Zhang and Y. Wang, *J. Mol. Model.*, (2011). DOI: 10.1007/s00894-011-1293-z

17. H.-X. Zhang, Y. Li, X. Wang and Y. Wang, *J. Mol. Model.*, 2011. DOI: 10.1007/s00894-011-1042-3.

18. H. Zhang, Y. Li, X. Wang, Z. Xiao and Y. Wang, *Curr. Med. Chem.*, **18**, 4019 (2011) In press.

19. J. Liu, F. Wang, Z. Ma, X. Wang and Y. Wang, *Int. J. Mol. Sci.*, **12**, 947 (2011).

20. A. Afantitis, G. Melagraki, H. Sarimveis, P.A. Koutentis, J. Markopoulos and O. Igglessi-Markopoulou, *Bioorg. Med. Chem.*, **14**, 6686 (2006).

21. M.H. Fatemi and S. Gharaghani, *Bioorg. Med. Chem.*, **15**, 7746 (2007).

22. S. Sciabola, E. Carosati, L. Cucurull-Sanchez, M. Baroni and R. Mannhold, *Bioorg. Med. Chem.*, **15**, 6450 (2007).

23. S.Y. Liao, L. Qian, T.F. Miao, Y. Shen and K.C. Zheng, *J. Theor. Comput. Chem.*, **8**, 143 (2009).

24. R. Cramer, D. Patterson and J. Bunce, *J. Am. Chem. Soc.*, **110**, 5959 (1988).

25. G. Klebe, U. Abraham and T. Mietzner, *J. Med. Chem.*, **37**, 4130 (1994).

26. SYBYL6.9, Molecular Modeling Software, Tripos Associates, Inc.: 1669, South Hanley Road, Suite 303, St. Louis, Missouri, MO 63144-2913, USA.

27. J. Gasteiger and M. Marsili, *Tetrahedron*, **36**, 3219 (1980).

28. M. Clark, R.D. Cramer III and N. Van Opdenbosch, *J. Comput. Chem.*, **10**, 982 (1989).

29. M.D. AbdulHameed, A. Hamza, J. Liu and C.G. Zhan, *J. Chem. Inf. Model.*, **48**, 1760 (2008).

30. http://www.disat.unimib.it/chm

31. A.N. Jain, *J. Med. Chem.*, **46**, 499 (2003).

32. R.B.G. Ravelli, B. Gigant, P.A. Curmi, I. Jourdain, S. Lachkar, A. Sobel and M. Knossow, *Nature*, **428**, 198 (2004).

33. S.Y. Liao, L. Qian, T.F. Miao, H.L. Lu and K.C. Zheng, *Eur. J. Med. Chem.*, **44**, 2822 (2009).

34. S. Fortin, L. Wei, E. Moreau, P. Labrie, E. Petitclerc, L.P. Kotra and RC.G., *Bioorg. Med. Chem.*, **17**, 3690 (2009).

35. Q. Wang, R.H. Mach and D.E. Reichert, *J. Chem. Inf. Model.*, **49**, 1963 (2009).

36. A. Golbraikh and A. Tropsha, *J. Mol. Graph. Model.*, **20**, 269 (2002).

37. D.R. Green and J.C. Reed, *Science*, **281**, 1309 (1998).

38. Y. Wang, K.-L. Han, S.-L. Yang and L. Yang, *J. Mol. Struc. (Theochem.)*, **710**, 215 (2004).

39. M. Wiese and I.K. Pajeva, *Curr. Med. Chem.*, **8**, 685 (2001).