

ISSN: 0975-7538 Research Article

Pharmacophore Modeling and 3D QSAR analysis of flavonoids and congeners active against A549 cell line

Sudhan Debnath*1, Utpal Ch. De², Debanjan Sen³, Biswanath Dinda²

¹Department of Chemistry, M.B.B. College, Agartala – 799 004, India ^{2,3}Department of Chemistry, Tripura University, Suryamaninagar – 799 022, India, ³BIPS, Kalyani, W.B, India

ABSTRACT

In our present study we have developed a ligand based 3D QSAR of 27 flavonoids and congeners active against human lung cancer cell line (A549) using the Phase module of Schrödinger suit. A five feature pharmacophore (AAARR) hypothesis yielded a statistically significant 3D-QSAR model having $r^2 = 0.9058$, $q^2 = 0.6545$, Pearson R = 0.8223, *RMSD* = 0.373 and $r_{pred}^2 = 0.65512$. The results of this pharmacophore based 3D-QSAR model gives a structural insight of selected flavonoids and congeners as A549 inhibitors which will provide guidance for the potential anti lung cancer drug design.

Keywords: Pharmacophore; 3D QSAR; flavonoid; cytotoxicity

INTRODUCTION

At present, cancer is one of the leading causes of death in the developed as well as in developing countries [http://www.medicalnewstoday.com/info/lungcancer]. Lung cancer is the most frequent cause of cancer-related death and accounts for more than a million deaths yearly worldwide with non-small cell lung cancer (NSCLC) accounting for 75-85% of lung cancer [Greenlee et al., 2001; Greenlee et al., 2000]. The pathogenesis of lung cancer involves the accumulation of multiple molecular abnormalities. These alterations include the genetic mutations, chromosomal changes with consequent inactivation of tumor suppressor genes and over activity of signal transduction cascades. Effective treatment of cancer is a result of the cytotoxicity of a drug against proliferating cells. Most of the clinically available anticancer drugs interfere with DNA function to exert their cytotoxicity. Apoptosis (programmed cell death), not only occurs naturally during development and differentiation, but also acts to eliminate damaged cells after injury [Evans, 1993]. Molecular studies of lung cancer have provided new avenues for early diagnosis and therapeutic strategies; however, certain patients are still plagued by rapid disease recurrence and progression and there has been no significant improvement in their overall survival. Therefore, it remains a disease with poor

* Corresponding Author Email: bcsdebnath@gmail.com Contact: +91-9436518210 Received on: 15-11-2011 Revised on: 17-02-2012 Accepted on: 11-03-2012 prognosis and the primary cause of cancer-related death in both men and women despite recent advances made in drug development. The development or presence of resistance to chemotherapeutic agents is a major obstacle to the effective treatment of lung cancer. Identification of the molecular determinants of sensitivity and resistance to chemotherapy is expected to improve the therapeutic efficacy. A literature survey [Harborne, 1994; Harbored, 2000] reveals that flavonoid derivatives are active against many tumor cell lines. Flavonoids are a group of polyphenolic compounds that occur naturally in foods of plant origin. These compounds possess a common phenylbenzopyrone structure (C6-C3-C6), and they are categorized according to the saturation level and opening of the central pyran ring, mainly into flavones, flavanols, isoflavones, flavonols, flavanones, and flavanonols [Harborne, 1994; Harbored, 2000]. These polyphenolic compounds display remarkable spectrum of biological activities including those that might be able to influence processes that are deregulated during cancer development. These include, for example, antiallergic, antiinflammatory, antioxidant, antimutagenic and specific cytotoxic activity towards different cancer cell lines which have generated large interest in developing flavonoid based cytostatics for anticancer therapy [Middleton, 2000; Galati et al., 2000; Yang et al., 2001; Sharma et al., 2006; Galati et al., 2004]. In the light of their biological significance, we have undertaken molecular modeling studies on a collection of flavonoids and congeners which were acctive against A549 cell line and built a legand based 3D QSAR model. It is our aim to develop a model which depicts the crucial structural features responsible for anti-lung cancer activity. The models developed for the receptor have

.OH









H₃C

но









16

HO.







12





been analyzed to identify potential areas of selectivity in the hyperspace of 3D pharmacophores that may lead to the discovery of anti-lung cancer drug or such compounds which could serve as templates for the design of new molecules as potential anti lung cancer agents.

MATERIAL AND METHODS

Data set

In the present study for designing of potential anti-lung cancer drug we have collected 27 flavonoids and congeners (Fig. 1) from different literature [Seelinger et al., 2008; Xinya et al., 2011; Ding et al., 2004; Caltagirone et al., 1997; Bible et al., 1996; Li et al., 2008; Bonesi et al., 2008] having precise IC₅₀ values against human lung cancer cell line A549. The IC₅₀ values of the compounds were cited in the literature in μ M/nM and the values were converted into their corresponding negative logarithm of IC_{50} (pIC₅₀) ranged from 6.602 to 4.005. The activity above 5.495 was arbitrarily taken as active during model generation. From the entire dataset of 27 compounds a subset of eight compounds was randomly excluded as an independent validation set and the remaining 19 compounds were taken as training set (Table -1).

Pharmacophore Modeling

Pharmacophore is an important and unifying concept in rational drug design which embodies the notion that

molecules are active at a particular enzyme or receptor when they possess a number of chemical features that favorably interact with the target possessing geometry complementary to it [Marriott *et al.*, 1999]. A pharmacophore hypothesis collects common features distributed in three-dimensional space representing groups in a molecule that participate in favourable interactions between drug and active site of the receptor. Pharmacophore modeling and 3D database searching are now recognized as integral components of lead discovery and lead optimization. In the present study pharmacophore modeling was carried out using PHASE [Schrödinger, LLC, NY 2010] running on windows 7 OS associated with Intel processor.

Ligand Preparation

The 2D structures of all the compounds were drawn in ChemDraw Ultra 8.0 and geometrically refined using 'Ligprep' module which produces low-energy 3D structure with correct chiralities for each successfully proposed input structure. While performing this step, chiralities were determined from 3D structure and original states of ionization were retained. Conformers were generated using Macro Model search method as implemented in phase module of maestro version 9.2 using OPLS-2005 force field discarding current conformers [Schrödinger, LLC, NY 2010]. For each molecule, a set of conformers with a maximum energy difference of 10 kcal/mol relative to the global energy minimum

Compound No.	Training/Test	Experimental Activity $IC_{50}(\mu m)$	Predicted Activity IC ₅₀ (μm)	Fitness Score
1	training	4.241	4.2	1.72
2	test	4.75	5.28	2.05
3	training	6.276	6.12	2.47
4	training	4.857	4.78	2.09
5	test	5.495	5.21	2.75
6	test	6	5.4	3
7	training	6.108	5.71	2.73
8	training	6.602	6.84	2.38
9	training	4.398	4.88	2.47
10	test	5.066	4.74	2.16
11	training	4.42	4.65	2.15
12	training	4.337	4.4	1.29
13	training	5.004	4.91	1.44
14	training	4.381	4.11	1.33
15	training	4.005	4.03	1.33
16	test	4.207	4.1	1.27
17	test	4.403	4.68	1.57
18	training	4.012	3.95	1.45
19	test	4.371	4.73	1.89
20	training	4.411	4.41	2.22
21	training	5.252	5	2.41
22	training	5.131	4.97	2.4
23	training	4.879	4.73	2.4
24	training	4.223	4.72	2.38
25	training	4.348	4.28	2.52
26	test	4.07	4.34	2.53
27	training	4.273	4.45	2.41

Table 1: Experimental and predicted activity of the molecules based on hypothesis AARRR.26

Table 2: Parameters of pharmacophore hypothesis

Hypothesis	Survival score	(Survival –inactive) score	Hypothesis	Survival	(Survival –inactive)
טו		•	טו	score	score
AARRR.26	4.047	2.071	AAARR.114	3.984	1.995
AAARR.72	4.046	1.96	AAAAR.262	3.983	2.042
AAARR.120	4.046	1.96	AAAAR.228	3.983	2.042
AAARR.26	4.019	2.034	AAAAR.251	3.983	2.042
AAARR.25	4.019	2.034	AAAAR.240	3.983	2.042
AARRR.4	4.018	2.073	AAAAR.227	3.983	2.042
AARRR.3	4.018	2.073	AADRR.8	3.766	2.432
AAAAR.188	4.000	1.930	AAADR.12	3.464	1.772
AAAAR.45	4.000	1.930	AADRR.7	3.075	1.878
AAARR.38	3.998	2.014	AAAAR.49	2.956	1.450
AARRR.25	3.987	2.043	AAAAR.195	2.956	1.450
AAARR.117	3.987	1.990	AAAAR.191	2.956	1.450
AAARR.71	3.987	1.990	AAARR.41	2.867	1.549
AAARR.116	3.986	2.151			

conformer was retained. The conformational searches were done for aqueous solution using the distancedependent dielectric model [Chung et al., 2010].

Hypothesis Generation

In the present study, an initial analysis of the compounds revealed that three chemical features i.e., hydrogen-bond acceptor (A), hydrogen-bond donor (D) and aromatic ring (R) could effectively map all critical chemical features of all molecules in the data set. The minimum and maximum sites for all the features were kept 3 and 5, respectively. These features were selected and used to build up a series of hypothesis with the 'find common pharmacophore' option in PHASE which provides a standard set of six pharmacophore features *viz.* hydrogen bond acceptor (A), hydrogen

Hypothesis ID	Survival score	(Survival –inactive) score	Hypothesis ID	Survival score	(Survival – inactive) score
AARRR.26	4.047	2.071	AADRR.8	3.766	2.432
AAARR.72	4.046	1.96	AAADR.12	3.464	1.772
AAAAR.188	4	1.93			

Table 3: Parameters of five best pharmacophore hypotheses

Table 4: Angles between different pharmacophoric sites of model AARRR.26

Site1	Site2	Site3	Angle	Site1	Site2	Site3	Angle
A3	A1	R13	9.7	A3	R13	R14	45.9
A3	A1	R14	23.3	A3	R13	R15	69.5
A3	A1	R15	47.0	R14	R13	R15	23.5
13	A1	R14	13.6	A1	R14	A3	108.5
13	A1	R15	37.3	A1	R14	R13	15.4
R14	A1	R15	23.7	A1	R14	R15	127.7
A1	A3	R13	7.2	A3	R14	R13	93.1
A1	A3	R14	48.2	A3	R14	R15	123.7
A1	A3	R15	83.2	R13	R14	R15	143.2
R13	A3	R14	41.0	A1	R15	A3	49.8
R13	A3	R15	76.0	A1	R15	R13	15.3
R14	A3	R15	35.0	A1	R15	R14	28.6
A1	R13	A3	163.1	A3	R15	R13	34.5
A1	R13	R14	151.0	A3	R15	R14	21.2
A1	R13	R15	127.5	R13	R15	R14	13.3

Table 5: Distances between different pharmacophoric sites of AARRR.26

Site1	Site2	Distance	Site1	Site2	Distance
A1	A3	6.408	A3	R14	2.671
A1	R13	2.764	A3	R15	6.133
A1	R14	5.037	R13	R14	2.438
A1	R15	8.330	R13	R15	6.355
A3	R13	3.712	R14	R15	4.233



Figure 2: Pharmacophore AARRR.26

bond donor (D), hydrophobic group (H), negatively ionizable (N), positively ionizable (P), and aromatic ring (R). 'Find Common Pharmacophore' option in PHASE generated 27 different variants (Table 2). Hypothesis scores were calculated for both active and inactive compounds by 'Score actives' and 'Score inactive' option using an overall maximum root mean square deviation (RMSD) value of 1.2 Å. The quality of hypotheses was measured based on the survival score and five different variants were selected (Table 3). Angles between different pharmacophoric sites of the hypothesis AARRR are shown in Table 4 where as the distances between different pharmacophoric sites of the same hypothesis

ID	AARRR.26	AAARR.72	AADRR.8	AAAAR.188	AAADR.12
PLS Factors	3	3	3	3	3
SD	0.2595	0.257	0.1715	0.1972	0.252
R ²	0.9058	0.9062	0.9516	0.9456	0.9133
F	48.1	51.5	98.2	92.6	56.2
Р	6.25e-008	1.92e-08	4.37e-10	2.52e-10	1.02e-08
Stability	-0.153	-0.0805	-0.0342	0.0484	-0.0934
RMSE	0.373	0.3891	0.4951	0.4231	0.3393
q^2	0.6545	0.5967	0.5856	0.5037	0.6386
Pearson-R	0.8223	0.8193	0.7976	0.7134	0.8103
R _{pred} ²	0.655158	0.606699	0.600546	0.504567	0.637307

 Table 6: Validation results of best five pharmacophore hypothesis



Figure 3: Experimental vs Predicted Activity



Figure 4: Contour maps of Hydrogen bond donor substitution, (a) flavones skeleton (b) compound 8 (c) compound 26



Figure 5: Effect of hydrophobic/non-polar substitution (a) compound 8 (b) compound 26



Figure 6: Effect of electron withdrawing (a) compound 8 (b) compound 26

AARRR are shown in Table 5. The best pharmacophoric hypothesis (AARRR.26) is displayed in fig. 2.

3D QSAR

Among the 27 generated hypotheses only 5 different hypotheses were selected based on survival score and on the basis of the pharmacophore based alignment (fig. 3), 3D QSAR model was generated. In 'Build QSAR' [19] option, random training set was kept as 70% (Table-1) where remaining 30% as test set and a model was generated by keeping 1Å grid spacing, PLS factor 3 and eliminate variables with (t-value) < 2.00. The validation results of this 3D QSAR model was quite good and its external validation further proved its statistical significance (Table 6). The result at PLS factor 3 was $r^2 = 0.9058$, $q^2 = 0.655158$ for statistically best model and this model can add an edge to the development of potent anti-lung cancer agents.

RESULTS AND DISCUSSION

To establish the reliability and the true predictive power of QSAR models, it is necessary to demonstrate that the model should be able to predict accurately the activities of test set compounds. Generally, the model would be accepted provided $R^2 > 0.6$, Pearson-R value ~1 for training set and q^2 >0.5 and R_{pred}^2 >0.6 for test set. In this study we have developed a model which provides useful pharmacophoric features necessary for the selected compounds to function as cytotoxic agent against lung cancer A549 cell line. The whole dataset was divided into training set consisting of 19 compounds, and test set of 8 compounds (Table -1). The pharmacophoric features selected for creating sites were hydrogen bond acceptor (A), hydrogen bond donor (D) and aromatic ring (R). The three and four featured pharmacophore hypotheses were rejected due to low value of survival score and as they were unable to define the complete binding space of the selected molecules. Two featured five point pharmacophore hypotheses were selected and subjected to stringent scoring function analysis. Among 27 hypotheses (survival score range 4.047 to 2.867), AARRR.26, AAARR.72,

AADRR.8, AAAAR.188 and AAADR.12 were selected on the basis of highest survival score for the development of pharmacophore based 3D QSAR model. These pharmacophoric features were used to align all the molecules for 3D QSAR analysis and different models were developed. All the generated models showed good correlation between experimental and predicted activity $(q^2 = 0.5037 - 0.6545)$ and Pearson R (0.7134 -0.8223). To evaluate the external predictive power of the models, the predicted activities (pIC₅₀ values) of the eight compounds set aside as external validation set were compared with their observed activities. The model with the highest predictive power was obtained for the training set of nineteen compounds represented in Table –1 and correlation between experimental vs phase predicted pIC₅₀ values of all the compounds for best model are shown in fig. 3. All the models were subjected to pass the same validation criteria to calculate the predicted activity. The results clearly demonstrated the high prediction accuracy, R_{pred}^2 values ranges from 0.504567 to 0.655158 (Table 6).

Interpretations of QSAR results

The 3D colour maps around the ligands obtained by PHASE QSAR indicated the important features and effects of various substitution patterns upon flavonoid scaffold for their cytotoxicity against human lung cancer cell line A549. The effectiveness of flavones scaffold (RRR) is indicated by the moderate experimental pIC₅₀ values of compounds 14, 15, 16, 18, 19 in which the RRR features are partially satisfied.

H-Bond Donor

The location of Hydrogen bond donor feature has been obtained from the QSAR model when applied to the most active ligand **8**; low active ligand **26** and is shown in Fig. 4. The blue colour indicates the favourable and orange colour indicates an unfavourable region for hydrogen bond donor group to show A549 cytotoxicity. As glimpse in fig. 4, the blue region stretched out from 7, 8 position of flavonoid scaffold suggested that these positions are very much prone to hydrogen bond donor substituent for enhancing anti lung cancer activity whereas presence of such groups at 5-position may have detrimental effect on activity. In the compounds 3, 6, 7 and 8 (highest active compounds, pIC50 >6), these regions are occupied by hydrogen bond donor substituents and showed higher activity which clearly demonstrated the robustness of the model. On the other hand in compounds 1, 24 and 26 no such hydrogen bond donor groups are present either at 7, 8 and position 4' being occupied by hydrophobic group (like-OR) played a negative role on activity. Therefore during the design of new molecule of this class hydrogen bond donor group incorporation at position 7, 8 & 4' and avoidance of such groups at position 5 may be considered for enhancing the activity. A small blue and orange region stretched out simultaneously from 4'position indicates that free -OH group at this position is unlikely to have better activity of the compounds against A549 cell line. In compound 3, 6 and 7 this position is occupied by free -OH group and hence showed comparatively less activity than compound 6 in which this position is free.

Effect of Hydrophobic/Non-polar substitution

The green colour regions indicate the favourable and purple colour region indicates the unfavourable hydrophobic/non-polar interaction sites (Fig. 5). It is clear from the hydrophobic contour diagram (fig. 5) that presence of hydrophobic or non-polar group in the vicinity of 4' position and absence of non-polar group in the region near to 6, 7 and 8 is likely to enhance the anti-lung cancer activity of these classes of compounds. These observations further suggested that a free –OH group at 4' will disfavour the activity whereas a hydrophilic environment (like sugar molecules) near to 6, 7 and 8-positions is expected to be favourable for desired activity. The experimental pIC₅₀ values of compounds 3, 8 and 21 (7-position more hydrophilic), as well as 22, 23, 24, 25 and 26 (7-position is comparatively less hydrophilic) corroborated these observations. Hydrophilic feature at 3-position also favours the activity which clearly explain the pIC₅₀ values of compounds 3, 7 and 20. It may be noted that green cubes above and below the two benzenoid ring (R13 and R15 features in fig. 2) suggested that these rings undergoes a strong hydrophobic interaction with receptor. This clearly demonstrated the larger activity of the compounds 3, 5, 6, 7, 8, 10, 21 and 22 than the compounds 14, 15, 16, 17, 18 and 19 (fig. 1)

Effect of Electron Withdrawing Group

The cyan colour (Fig. 6) around the ligand indicates the addition of electron withdrawing group will favour the anti-lung cancer activity and presence of electron withdrawing groups at red colour region disfavours the anti-lung cancer activity. The cyan colour around the region stretched out from position 3, 8 and 2' indicates that electron withdrawing effect at this position is likely to favour the in-vitro cell growth inhibition of A549

cell line for these classes of compounds. The experimental plC_{50} values of 4, 5, 6 and 8 clearly explained this feature. 3, 8 and 2' positions of compound 8, being occupied by good electron withdrawing group showed highest plC_{50} value. On the other hand at position 6 and 4', electron withdrawing group showed detrimental effect on the activity. So these positions should be handled with care during design of new inhibitors of these classes for A549.

CONCLUSION

It emphasizes that before attempting to predict target properties of untested compounds one should exhaustively validate QSAR models developed for the training sets both internally and externally such that only models that pass rigorous validation criteria should be used for the prediction. In this study, we have developed and thoroughly validated QSAR models for a series of flavonoids and congeners that have been studied as potential anticancer agents against A549 cell line. We have demonstrated that the validated 3D QSAR modeling workflow was successful in generating models with high internal and external accuracy. These models may further be exploited for the design and discovery of new potent anticancer agents by means of virtual screening of available chemical databases.

ACKNOWLEDGEMENT

The authors are thankful to Schrödinger, LLC, NY for providing software support, Department of chemistry, M.B.B. College, Department of Chemistry, Tripura University for providing computer and library facility.

REFERENCES

- Bible K. C., Kaufmann S. H: Flavopiridol: A cytotoxic flavone that Induces cell death in noncycling A549 human lung carcinoma cells, Cancer Res 56: (1996) 4856-4861.
- Bonesi M., Tundis R., Deguin B., Loizzo M.R., Menichini F., Tillequin F., Menichini Fr: In vitro biological evaluation of novel 7-O-dialkylaminoalkyl cytotoxic pectolinarigenin derivatives against a panel of human cancer cell lines, Bioorganic & Medicinal Chemistry Letters 18: (2008) 5431–5434.
- Caltagirone S., Ranelletti F.O., Rinelli A., Maggiano N., Colasante A., Musiani P., Aiello F.B., Piantelli M: Interaction with type II estrogen binding sites and antiproliferative activity of tamoxifen and quercetin in human non-small-Cell lung cancer, American Journal of Respiratory Cell and Molecular Biology 17: (1997) 51–59.
- Chung K.S., Choi J.H., Back N.I., Choi M.S., Kang E.K., Chung H.G., Jeong T.S., Lee K.T: Eupafolin, a flavonoid isolated from Artemisia princeps, induced apoptosis in human cervical adenocarcinoma HeLa cells, *Molecular Nutrition and Food Res*earch 54: (2010) 1318-1328.

- Ding P., Chen D., Bastow K. F., Nyarko A. K., Wang X, Lee HK: Cytotoxic isoprenylated flavonoids from the roots of *Sophora flavescens*, Helvetica Chimica Acta 87: (2004) 2574-2580.
- Evans V.-G: Multiple pathways to apoptosis, Cell Biology International 17: (1993) 461-476.
- Galati G, Teng S, Moridani M.Y., Chan TS, O'Brien P.J.: Cancer chemoprevention and apoptosis mechanisms induced by dietary polyphenolics, Drug Metabolism and Drug Interactions 17: (2000) 311–349.
- Galati G., O'Brien P. J: Potential toxicity of flavonoids and other dietary phenolics: significance for their chemopreventive and anticancer properties, Free Radical Biology & Medicine 37: (2004) 287–303.
- Greenlee R. -T., Hill-Harmon M. B., Murray T., Thun M: Cancer Statistics. CA: A Cancer Journal for Clinicians 51: (2001) 15-36.
- Greenlee R. -T., Murray T., Bolden S, Wingo P. -A: Cancer Statistics. CA: A Cancer Journal for Clinicians 51: (2000) 7-33.
- Harborne J. B. ed, The flavonoids: Ad- vances in research since 1986 1st Edition. London: Chapman & Hall (1994) 619–652
- Harborne J. B., Williams C. A: Advances in flavonoid research since 1992. Phytochemistry 55: (2000) 481– 504.
- http://www.medicalnewstoday.com/info/lung-cancer
- Li F., Awale S., Tezuka Y., Kadota S: Cytotoxic constituents from Brazilian red propolis and their structureactivity relationship, Bioorganic & Medicinal Chemistry 16: (2008) 5434–5440.
- Maestro, Version 9.2, Schrödinger, LLC, NY 2010.
- Marriott D. P., Dougall I. G., Meghani P., Liu Y. J., Flower D. R: Lead generation using pharmacophore mapping and three dimensional database searching: application to muscarinic M₃ receptor antagonists, Journal of Medicinal Chemistry 42: (1999) 3210-3216.
- Middleton E, Kandaswami Jr. C., Theoharides T. C: The effects of plant flavonoids on mammalian cells: Implications for inflammation, heart disease and cancer, Pharmacological Reviews 52: (2000) 673–751.
- Seelinger G, Merfort I, Wölfle U, Schempp C. M: Anticarcinogenic effects of the flavonoid luteolin, molecules 13: (2008) 2628-2651; DOI: 10.3390/molecules13102628.
- Sharma D. K: Pharmacological properties of flavonoids including flavonolignans-Integration of petrocrops with drug developments from plants, Journal of Scientific & Industrial Research 65: (2006) 477-484.

- Xi X., Xie H., Hao J., Jiang Y., Wei X.: Flavonoid glycosides from the seeds of *Litchi chinensis*, Journal of Agricultural & Food Chemistry 59: (2011) 1205-1209.
- Yang C.S., Landau J.M., Huang M.T., Newmark H.L: Inhibition of carcinogenesis by dietary polyphenolic compounds, Annual Review of Nutrition 21: (2001) 381–406.