



Received on 27 September, 2016; received in revised form, 29 November, 2016; accepted, 08 January, 2017; published 01 April, 2017

A COMPARATIVE GROUP-QSAR AND MOLECULAR DOCKING STUDIES OF 4 - THIAZOLIDINONE CONTAINING INDOLIN-2-ONE MOIETY AS VEGFR INHIBITORS

Pragya Nayak* and Monica Kachroo

Department of Pharmaceutical Chemistry, Al-Ameen College of Pharmacy, Bangalore, Karnataka, India.

Keywords:

VEGFR, QQSAR,
Thiazolidinone, Anti-tumour

Correspondence to Author:

Pragya Nayak


Research scholar
Department of pharmaceutical
chemistry, Al-ameen college of
pharmacy, Bangalore, Karnataka,
India.

E-mail: prgy.nyk@gmail.com

ABSTRACT: Vascular endothelial growth factor receptors (VEGFR) are kinase based receptors reported as a promising target in anti tumour therapy. VEGFR inhibitors are being investigated which can have important contribution in anti-angiogenic therapy for treatment of cancer. In the present study, an attempt has been made to develop a site-specific QSAR model in order to explore the definite sites of substitution of a series of 4-thiazolidinone derivatives having reported antitumor activity against h460 cell lines. Each molecule of the series was divided into seven fragments for varying substituent at the positions R1, R2, R3, R4, R5, R6 and R7 of the parent nucleus. QQSAR was performed using MLR, PCR, PLS and KNN methods of variable selection. Amongst these methods, PCR has come out with promising result as compared to other methods. A comparative docking study was performed to explore the particular sites of interactions within the binding cavity of VEGFR protein. (PDB id: 1Y6A). The important substitutions contributing towards the biological activity by interpreting the developed QQSAR equation using virtual studies were found out which are as follows. position R1 should be substituted with groups with low electro negativity and higher atomic mass, oxygen count at R5 should be increased which would act as hydrogen bond acceptors and total polar surface area at R7 has to be decreased by making substitutions of non polar groups to promote hydrophobic interactions.

INTRODUCTION: Cancer is the uncontrolled growth of cells, which can invade and spread to distant sites of the body. Cancer can have severe health consequences, and is a leading cause of death. 8.2 million People die each year from cancer, an estimated 13% of all deaths worldwide. There are more than 100 types of cancer, including breast cancer, skin cancer, lung cancer, colon cancer, prostate cancer, and lymphoma¹. Vascular endothelial growth factor (VEGF) is a signal protein produced by cells which stimulates angiogenesis.

It is a part of the system that restores the oxygen supply to tissues when blood circulation is inadequate^{2, 3}. It is up regulated in cases of tumours. It acts as an essential growth factor and has important contributions towards tumour angiogenesis. VEGF receptors (VEGFR) are also expressed on tumour cells. Anti-VEGF strategies to treat cancer were designed to target the pro-angiogenic function of VEGF and thereby inhibit neovascularisation it has been suggested that direct stimulation of tumour cells by VEGF may protect the cells from apoptosis and increases their resistance to conventional chemotherapy and radiotherapy^{4, 5}. Anti-VEGF therapies are therefore likely to target both the pre-angiogenic activity of VEGF and the anti-apoptotic functions of VEGF. Combination therapies using anti-VEGF therapies with chemotherapy and radiotherapy are effective against many types of tumours.

<p>QUICK RESPONSE CODE</p> 	<p>DOI: 10.13040/IJPSR.0975-8232.8(4).1796-05</p> <hr/> <p>Article can be accessed online on: www.ijpsr.com</p> <hr/> <p>DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.8(4).1796-05</p>
---	--

VEGFR blockage renders tumour cells more susceptible to conventional treatment^{6,7}.

Quantitative structure activity relationships are the most important application of chemo-metrics giving useful information for the design of new compounds active on a specific target⁸. A good QSAR model enhances our understanding of the specifics of drug action and provides a theoretical basis for lead optimization. Conventional QSAR approaches provide information about the necessary features for desired activity but not about the site where the substitution has to be made⁹. Group based QSAR (G-QSAR) is a new approach to investigate site-specific structure activity relationship between activity and descriptors calculated for various molecular groups of interest. Group based lead design has shown promise in current drug discovery and lead optimization efforts¹⁰⁻¹⁴. Docking is an important tool in molecular modeling that contributes in exploring the receptor binding cavity and understanding the important ligand-receptor interaction sites¹⁵.

Thiazolidinone an important member of heterocyclic compounds is a saturated form of thiazole having a carbonyl group at fourth carbon atom^{16, 17}. It has been considered as a magic moiety by many researchers because of its important contribution in the field of drug

discovery and its property of possessing several various types of biological activities including antitumor activity. It has been a nucleus of interest because of its usefulness as intermediate for the synthesis of many new heterocyclic compounds. It provides different sites for substitution so that a large combinatorial library can be designed around it with a large chemical space¹⁸⁻²⁰.

METHODS:

The Data Set: In the present study, a congeneric series of 38 thiazolidinone derivatives contained indolin-2-one moiety was selected with reported antitumor activity in terms of IC₅₀ (μm) against non small-cell lung cancer cell lines (H460)¹ for the development of G-QSAR models (Table 1)²¹. The inhibitory concentration [IC₅₀ (μm)] values were converted to the negative logarithmic scale pIC₅₀ (M) [pIC₅₀= log (10⁶/IC₅₀ (μM))] for the present work.

The whole dataset was divided into a training set and a test set considering even distribution of structure and activity. It was confirmed that the test molecules are a subset of training molecules by uni-column statistics (the minimum and maximum value for test set should be higher and lower, respectively than that of training set).

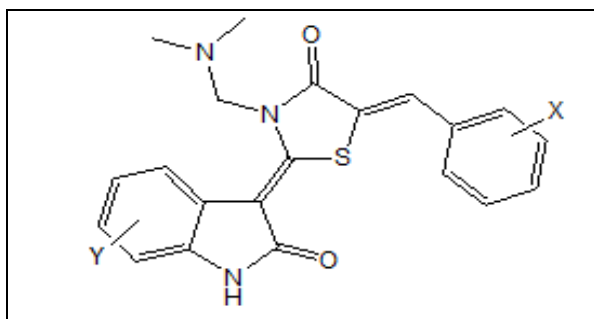


FIG. 1: 5A-5S

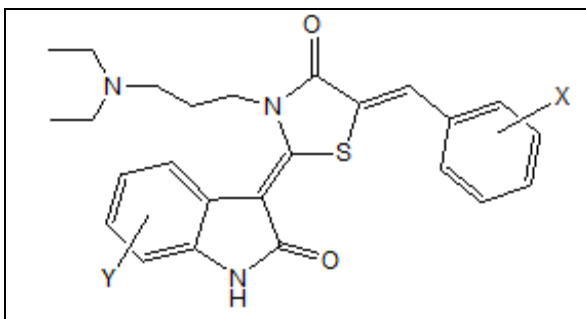


FIG. 2: 10A-10S

TABLE 1: DEVELOPMENT OF G-QSAR

Sl. No.	Compound	X	Y
1.	5a	4-hydroxy	H
2.	5b	4-hydroxy	5-CH ₃
3.	5c	4-hydroxy	5-F
4.	5d	4-hydroxy	6-F
5.	5e	4-hydroxy	5-Cl
6.	5f	4-hydroxy	5-Br
7.	5g	4-methoxy	6-F
8.	5h	3,4-difluoro	5-F
9.	5i	3,4-difluoro	6-F
10.	5j	3,4-difluoro	5-CH ₃

11.	5k	2,4-difluoro	6-F
12.	5l	2,4-dimethoxy	5-Cl
13.	5m	2,4-dimethoxy	6-F
14.	5n	2,5-dimethoxy	6-F
15.	5o	2,5-dimethoxy	H
16.	5p	2,5-dimethoxy	H
17.	5q	3,4-dioxymethylene	5-CH ₃
18.	5r	3,4-dioxymethylene	6-F
19.	5s	3,4,5-trimethoxy	5-F
20.	10a	3,4-dioxymethylene	H
21.	10b	3,4-dioxymethylene	5-CH ₃
22.	10c	3,4-dioxymethylene	5-F
23.	10d	3,4-dioxymethylene	6-F
24.	10e	3,4-dioxymethylene	5-Cl
25.	10f	3,4-difluoro	5-F
26.	10g	3,4-difluoro	6-F
27.	10h	2,4-difluoro	H
28.	10i	2,4-difluoro	6-F
29.	10j	2,4-difluoro	5-Cl
30.	10k	3,4,5-trimethoxy	H
31.	10l	3,4,5-trimethoxy	5-CH ₃
32.	10m	3,4,5-trimethoxy	5-F
33.	10n	3,4,5-trimethoxy	6-F
34.	10o	3,4,5-trimethoxy	5-Cl
35.	10p	4-hydroxy	5-CH ₃
36.	10q	2-hydroxy	H
37.	10r	2-hydroxy	5-CH ₃
38.	10s	2-hydroxy	5-Br

G-QSAR: GQSAR modelling was performed using the Molecular Design Suite (VLife MDS software package, version 4.1, from VLife Sciences Technologies Pvt. Ltd., India) on a Windows 7 operating system.

Molecular Fragmentation and Descriptor Calculation: The GQSAR tool begins with the fragmentation of each molecule in the dataset,

based on a set of predefined rules, before calculating their corresponding fragment descriptors. Each molecule of the series was divided into seven fragments for varying substituent at the positions R1, R2, R3, R4, R5, R6 and R7 of the parent nucleus. The common scaffold in all the molecules was selected as the template and all other substitutions at different positions were defined as different fragments. (Fig. 3)

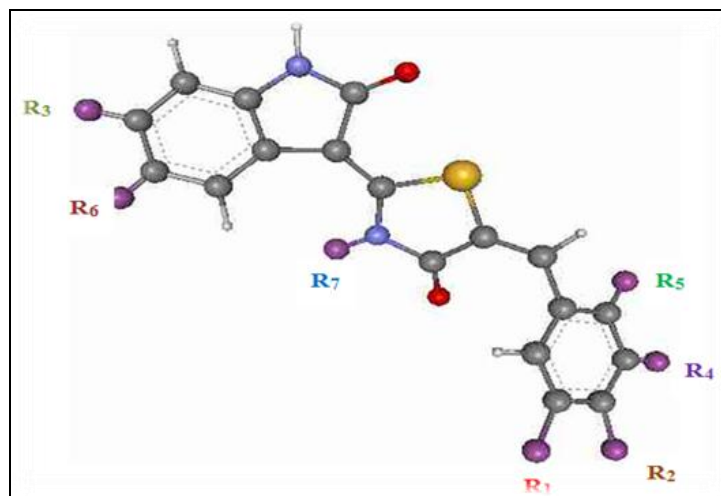


FIG. 3: TEMPLATE USED FOR GQSAR

Fragment-based molecular descriptors calculation: Individual physicochemical descriptors like molecular weight, hydrogen bond donors and acceptors, retention index (chi), path count, estate numbers, atomic valence connectivity index (chiv), polar surface area, oxygen count, fluorine count, hydrophilic and hydrophobic surface area, etc were calculated for all fragments. All descriptors with constant values among the dataset were deleted, resulting in 564 different descriptors (independent variables) which were used in the QSAR analysis.

Model building:

Variable Selection Method: In order to select a subset of descriptors (variables) from the descriptor pool, a variable selection method known as stepwise forward backward selection was used^{22, 23}. The following techniques were used to develop the QSAR models.

Multiple Regression Analysis (MLA): Multiple regressions is the standard method for multivariate data analysis. Linear regression models a linear relationship between two variables or vectors, x & y. Thus in two dimensions, the relationship can be describes by a straight line given by equation 1.

$$y = ax + b \text{ Equation 1}$$

Where 'a' is the slope of the line and b is the intercept of the line at y-axis. The goal of linear regression is to adapt the values of the slope and of the intercept so that the line gives the best prediction of y from x. This is achieved by minimising the sum of the squares of the vertical distances of the points from the line. While simple linear regression uses only one independent variable for modelling, MLA uses more variables at a time.

This method of regression estimates the values of the regression coefficients by applying least squares curve fitting method. For getting reliable results, dataset having typically 5 times as many data points (molecules) as independent variables (descriptors) is required. The regression equation takes the form

$$Y = a_0 + a_1 x_1 + a_2 x_2 \dots \dots + a_n x_n \dots \dots \text{Equation 2}$$

Where Y is the dependent variable, the 'a's are regression coefficients for corresponding x (independent variable), a_0 is a regression constant or intercept²⁴.

Partial Least Square (PLS) Analysis: Partial least squares regression, also called projection to latent structure can be applied to establish a predictive model even if the features are highly correlated. This makes PLS an attractive method for QSAR. The goal of PLS is to establish relationship between the two matrices x & y. The procedure is as follows: first, the principle component for x & y is calculated separately. The scores of the matrix x are then used for a regression model to predict the scores of y, which can be used to predict y. It is an extension of the multiple linear regressions²⁵.

Principal Component Analysis: It is a frequently used method which is applied to extract the systematic variance in a data matrix. It helps to obtain an overview over dominant patterns and major trends in the data. The aim of PCA is to create a set of latent variables which is smaller than the set of original variables. In mathematical terms, PCA transforms a number of correlated variables into a smaller number of uncorrelated variables, the so-called principal components. An advantage of PCA is its ability to cope with almost any kind of data matrix²⁶.

k- Nearest Neighbour (k-NN) Analysis: The k-NN method was also used to develop a QSAR model using continuous variable i.e. using activity as pIC50 values. In this case, by using a developed k-NN QSAR model the activity of a molecule can be predicted using weighted average activity (Eq. (1)) of the k most similar molecules in the training set.

$$\hat{y}_i = \sum w_i y_i \text{ Equation 3}$$

Where \hat{y}_i and y_i are the actual and predicted activity of the molecule respectively, and w_i are weights calculated using equation 4.

$$w_i = \frac{\exp(-d_j)}{\sum_{j=1}^k \exp(-d_j)} \text{ Equation 4}$$

The similarities were evaluated as the inverse of Euclidean distances (d_j) between molecules using only the subset of descriptors corresponding to the

model. Where, k is number of nearest neighbours in the model^{27, 28}.

Model validation:

Internal validation of training set: To evaluate the robustness of the generated QSAR models, internal validation was performed on the training set using the leave-one-out (LOO) method²⁹. The compounds from the training set were removed individually, and the activity of each was predicted using the model fitted to the remaining molecules. The process is repeated until all compounds in the training set are exhausted, and the cross-validated coefficient of determination (q^2) was found from the equation 5:

$$q^2 = \frac{\sum_{i=1}^{n_{\text{training}}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{\text{training}}} (y_i - \bar{y})^2} \dots \dots \dots \text{Equation 5}$$

Where y_i , \hat{y}_i , and \bar{y} denote the actual, predicted, and average activity of training set molecules.

External validation of test set: The predictive power of the developed models was further validated using the squared correlation coefficient (pred-r^2) of the test set³⁰. The model was generated from training set data, and the pIC_{50} values of test set compounds were predicted from the model and pred-r^2 was determined from the following equation:

$$\text{pred-r}^2 = \frac{\sum_{i=1}^{n_{\text{test}}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{\text{test}}} (y_i - \bar{y})^2} \dots \dots \dots \text{Equation 5}$$

Where y_i and \hat{y}_i are the actual and predicted activities of test set compounds and \bar{y} represents the average activity of training set molecules.

Model evaluation criteria: The model was considered to have a significant productivity when the squared correlation coefficient (r^2) between descriptors and activity (pIC_{50}) was more than 0.7. Similarly, the models were considered to possess significant internal and external productivity when the cross-validated correlation coefficient of the leave-one-out method (q^2) > 0.5 and the correlation coefficient of the training set (pred_r^2) > 0.5³¹.

Molecular docking: In addition to the QSAR analysis, molecular docking studies were also performed and the ligands were docked into the binding cavity of VGEFR to explore the important

binding sites within the cavity which contributes in the better binding of ligand with receptor as well as the interactions which contributes negatively towards the ligand-receptor binding. The docking results were compared with the sites predicted by QSAR analysis.

Validation of tool: The docking software (Autovina using PyRx) is validated prior to docking analysis by re-docking the co-crystallised ligand into the VFEGR (PDB Id: 1Y6A) cavity and calculating root mean square deviation between the atoms of co-crystallised ligand and the docked pose. If RMSD value comes within 2, the software could be considered suitable for the selected receptor.

Protein and ligand preparation: The crystal structure of VEGFR (PDB Id: 1Y6A) protein was downloaded from the official website of protein data bank (rcsb.org) and prepared using auto dock tools. The missing atoms and hydrogen were added, polar hydrogen atoms and charges of the Gasteiger-type were assigned and it was saved in 'pdbqt' format. (Fig. 2)

The ligand structure were drawn in 2D format using vLife 2D drawer tool and converted to 3D. The structures were energy minimized using Merck Molecular Force Field (MMFF) using convergence criterion (RMS gradient) of 0.01 kcal/mol and maximum number of cycles of 100. The molecules were saved in 'mol' format.

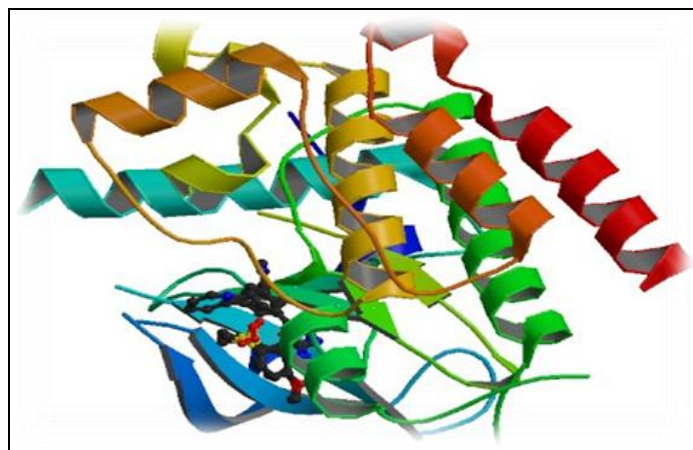


FIG. 4: TEMPLATE USED FOR QSAR

Docking: A rectangular grid of dimensions 10Å was used to define the binding Cavity of the receptor, the docking application was started and the binding scores were observed. The two

dimensional interaction of the protein and the docked conformation of the ligand were obtained and the important interacting sites were found out.

RESULTS AND DISCUSSION:

Group-based quantitative structure relationship (GQSAR): Thirty eight compounds with antitumor activity were used for fragment-based descriptor calculation. After removing invariable descriptors, a pool of 564 fragment based molecular descriptors

remained and their contribution to activity variation was evaluated. Sphere exclusion algorithms with a dissimilarity value of +0.5 resulted in a reasonable rational division of the data into a training set (n=29) and test set (n=8). The calculated uni column statistics of both sets (**Table 2**) show that activity was evenly distributed within both tests and the selected sets fulfilled the main characteristics of valid data selection³².

TABLE 2: STATISTICAL PARAMETERS OF ACTIVITY DISTRIBUTION WITHIN THE SELECTED TRAINING AND TEST SETS

	N	Average	Max	Min	SD	Sum
Training set	29	2.6920	5.160	0.4628	1.4337	112.6320
Test set	08	3.4397	3.9810	0.7280	1.0701	31.9571

Statistical evaluation and validation of the developed QSAR models:

Models 1 (SWFB/PCR):

$$\text{Activity} = -0.85025(\text{R5-OxygensCount}) + 1.56069(\text{R1-MomInertiaX}) - 1.83255(\text{R1-SAMostHydrophilic}) - 0.731537(\text{R7-BalabanCdash}) - 0.221826$$

The model was generated using stepwise forward-backward algorithms followed by principal component regression. The statistical parameters of model 1 are shown in **Table 3**. The regression equation of the developed model 1 explains ~71%

($r^2=0.71$) of the total variance in the training set and has an internal and external predictive ability of approximately 61% ($q^2=0.61$) and approximately 61% ($\text{pred}_r^2=0.61$), respectively.

Models 2 (SWFB/PLS):

$$\text{Activity} = -1.18451(\text{R5-OxygensCount}) + 1.82268(\text{R1-MomInertiaX}) - 0.651913(\text{R1-FluorinesCount}) + 0.339449(\text{R7-chi4}) - 0.243135$$

The model was generated using stepwise forward-backward algorithms followed by partial least square regression. The statistical parameters of model 1 are shown in **Table 3**. The regression equation of the developed model 1 explains ~71%

($r^2=0.71$) of the total variance in the training set and has an internal and external predictive ability of approximately 56% ($q^2=0.56$) and approximately 58% ($\text{pred}_r^2=0.58$), respectively.

Models 3(SWFB/MLR):

$$\text{Activity} = -1.16829(\text{R5-OxygensCount}) + 1.78244(\text{R1-MomInertiaX}) - 0.571869(\text{R1-FluorinesCount}) + 0.0316311(\text{R7-SAHydrophilicArea}) - 0.243135$$

The model was generated using stepwise forward-backward algorithms followed by multiple linear regressions. The statistical parameters of model 1 are shown in **Table 3**. The regression equation of the developed model 1 explains ~71% ($r^2=0.71$) of

the total variance in the training set and has an internal and external predictive ability of approximately 55% ($q^2=0.55$) and approximately 65% ($\text{pred}_r^2=0.65$), respectively.

TABLE 3: STATISTICAL PARAMETERS FOR THE DEVELOPED QQSAR MODELS

Statistical parameters	Model 1	Model 2	Model 3
Variable selection method	PCR	PLS	MLR
N (training/test)	29/8	29/8	29/8
Degree of freedom	25	26	24
F-test	20.6547	32.4479	15.1362
r ²	0.7125	0.7140	0.7161
q ²	0.6118	0.5608	0.5508
pred-r ²	0.6131	0.5824	0.6562
r ² _{se}	0.3922	0.3836	0.3977
q ² _{se}	0.4557	0.4753	0.5003
pred_r ² _{se}	0.2333	0.2423	0.2199
Z Score R ²	7.78813	6.09869	7.20495
Z Score Q ²	1.37110	1.12865	1.09676
Best Rand R ²	0.36885	0.45304	0.34233
Best Rand Q ²	0.16044	-0.00442	0.06470
Alpha Rand R ²	0.00000	0.00000	0.00000
Alpha Rand Q ²	0.10000	99.00000	99.00000
Z Score Pred R ²	1.03895	1.30630	1.39394
best Rand Pred R ²	0.72531	0.72196	0.64183
alpha Rand Pred R ²	0.00000	0.10000	0.10000
K(number of components)	3	2	

Comparing the calculated parameters of the all the models, it can be presumed that variable selection by multiple linear regression (model 3) was capable of developing a more robust and predictive QQSAR model in terms of r², q², and pred-r². However all the three models have statistically significant robustness and predictive power and therefore can be used to explain the structural requirements. Model 3 is interpreted further and is used to explain the important structural requirement for desired activity.

Interpretation of model 3: Fig. 3 depicts the distribution of training and test set and indicates

how the training and test sets are distributed over the regression line. The contribution plot (Fig. 4) tells about the contribution of each selected descriptor towards the prediction of activity. It can be seen from the contribution plot that at position R1, moment of inertia which is related to the mass of component, contributes positively while at the same position fluorine counts contributes negatively. At R5 position, oxygen count is playing an important role and related inversely with the activity. At R7 position the hydrophilic surface area of the substituted group is related negatively with activity.

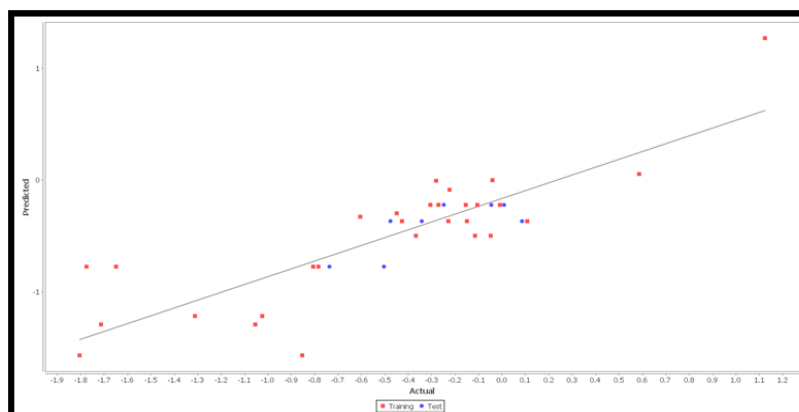


FIG. 5: FITNESS PLOT OF MODEL 3 THAT WAS GENERATED BY STEPWISE FORWARD-BACKWARD ALGORITHMS COUPLED WITH MULTIPLE LINEAR REGRESSIONS

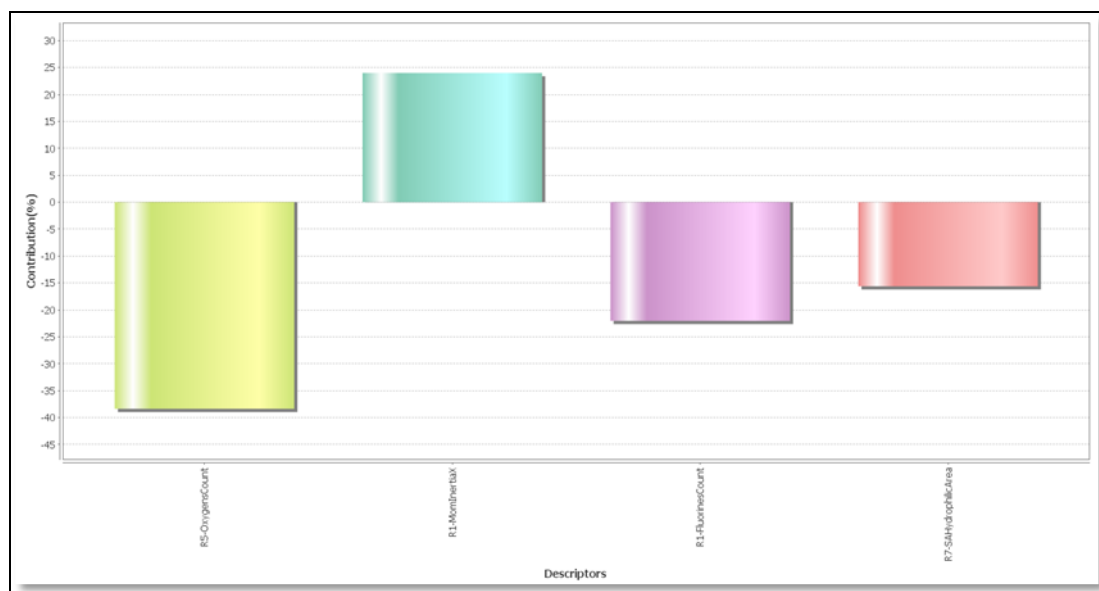


FIG. 6: CONTRIBUTION PLOT OF MODEL 3

Docking: The docking analysis showed the ligand exhibited a binding score in a range of -8.3 to -3.4. It was observed that the ligands 5a and 5k which lacks R5 oxygen as well as R1- fluorine substitutions exhibited highest binding scores while two of the lowest binding compounds 5j and 10f possess these substitutions. Secondly, the position

r7 is surrounded with hydrophobic amino acids namely Leu 1038, Phe 1045, Val 846, Cys 917, etc and hence no hydrophilic polar interaction is seen between R7 and the binding site of the receptor and also hydrophobic interaction has been observed to have contribution towards binding score.(Fig. 7)

TABLE 4: DOCKING ANALYSIS

Compound	Docking score
5a	-8.3
5k	-8.3
10l	-7.4
10i	-7.1
5f	-7.0
10a	-7.0
10s	-6.4
5e	-6.6
10q	-6.4
5c	-6.2
10n	-6.2
10c	-6.2
5o	-6.1
10d	-6.1
10r	-6.0
5g	-5.9
10o	-5.9
5p	-5.8
5d	-5.8
10g	-5.3
5r	-5.3
5b	-5.2
10e	-5.1
5m	-5.0
5i	-4.7
10k	-4.7
10m	-4.7
5n	-4.6

10b	-4.3
5l	-4.3
5s	-4.2
5h	-4.2
5j	-4.1
10h	-3.9
10f	-3.7
10j	-3.7
10p	-3.4

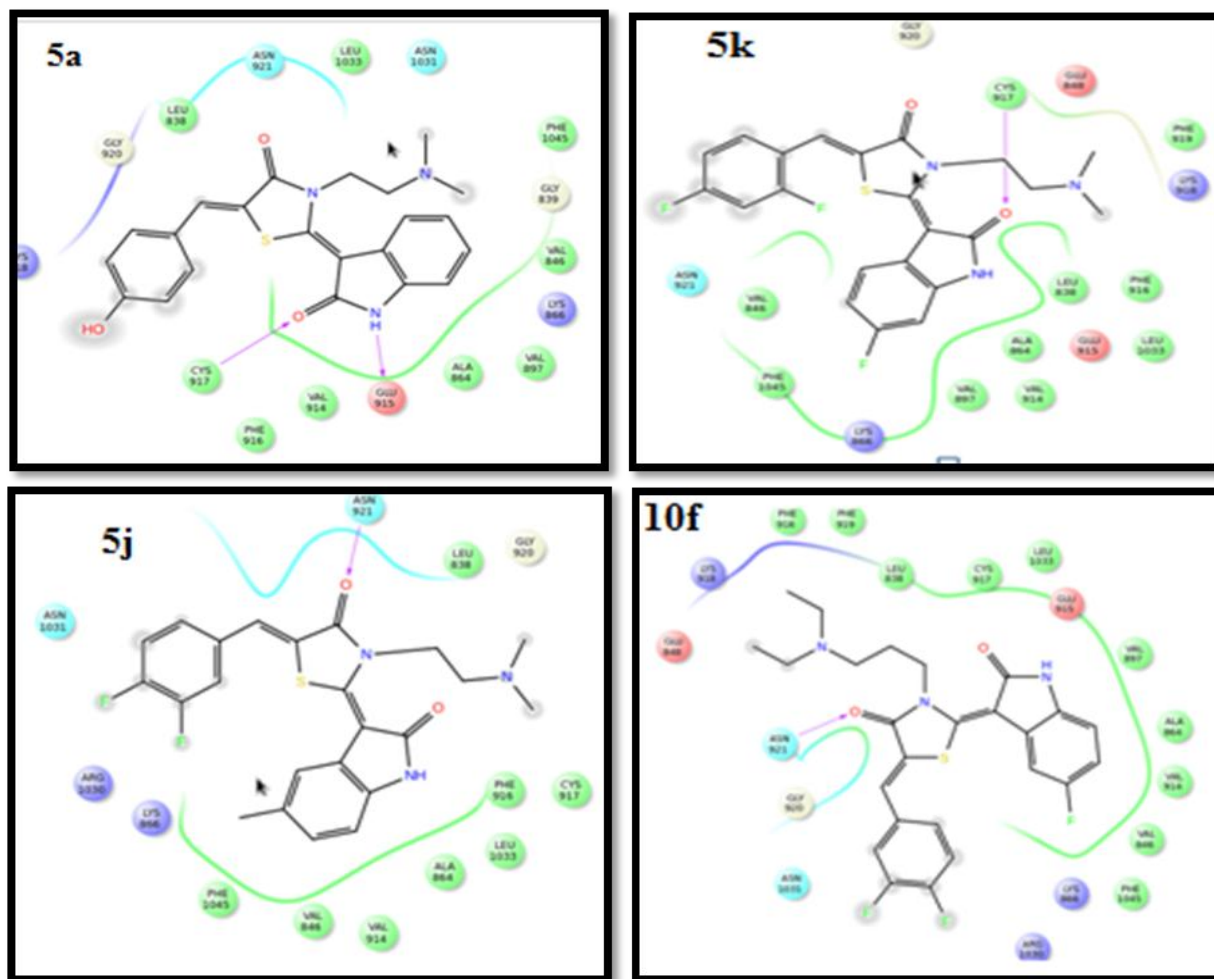


FIG. 7: 2 DIMENSIONAL INTERACTION OF LIGANDS WITH PROTEIN RECEPTOR 1Y6A (5A AND 5K: GOOD BINDING SCORE; 5J 10F: POOR BINDING)

CONCLUSION: From the QSAR equation, Oxygen count and fluorine count at **R5** and **R1** positions were found to contribute negatively towards activity. Docking study reveals that compounds showing good docking score **5A** and **5K** lacks these two substitutions while compound with low docking score possess fluorine atom at **R1** position.

G-QSAR equations tell that hydrophilic surface area at **R7** should be lowered in order to get good activity. It is also clear from the 2D drug receptor

interaction of best scoring compound **5A** that **R7** is surrounded with hydrophobic amino acid indicated by green color and no hydrophilic interaction is observed at the particular site. And hydrophobic interactions were also found to contribute towards the overall ligand-receptor binding.

From the above study it can be concluded that new and better thiazolidinone derivatives can be designed with enhanced target specific antitumor activity by modifying the structure according to the features found from the above studies.

ACKNOWLEDGEMENT: Authors would like to acknowledge Department of science and technology, Govt. of India for funding this research project. (Author is DST-INSPIRE fellow). We would also like to thank Rajiv Gandhi University of health sciences, Bangalore, for promoting and encouraging research facilities. We are thankful to Mr. Lokesh Pathak (Al-Ameen college of Pharmacy) for his support.

REFERENCES:

1. <http://www.who.int/cancer/en/> Accessed 24/08/2016.
2. Senger DR, Galli SJ, Dvorak AM, Perruzzi CA, Harvey VS, Dvorak HF: Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid, *Science* 1983; 219 (4587): 983-5.
3. Palmer BF, Clegg DJ: Oxygen sensing and metabolic homeostasis, *Molecular and Cellular Endocrinology* 2015; 397: 51-57.
4. Harmey JH, bouchier-hayes O: VEGF-a survival factor for tumour cells, implications for anti-angiogenic therapy, *Bioassays* 2002; 24:280-283.
5. Gorski DH, Beckett NA, Jaskoviak NT, Blockage of the vascular endothelial growth factor stress response increases the antitumor effects of ionizing radiation, *Cancer research* 1999; 59:3374-3378.
6. Angele MD, Denial J, bouchier-hayes O, Judith HH: Vascular endothelial growth factor (VEGF) and its role in non-endothelial cells: autocrine signalling by VEGF: Madame Curie's bioscience database.
7. Shibuya M: Vascular endothelial growth factor (VEGF) and its receptor (VEGFR) signalling in angiogenesis, *Genes Cancer* 2011; 2(12): 1097-1105.
8. H. Kubinyi (Ed.): *QSAR: Hansch Analysis and Related Approaches*, VCH, 1993.
9. Leonard JT, Roy K: On selection of training and test sets for the development of predictive QSAR models, *QSAR Comb Sci* 2006; 25:235-251.
10. Ajmani S, Jadhav K, Kulkarni SA: Group-Based QSAR (G-QSAR): Mitigating Interpretation Challenges in QSAR, *QSAR Comb Sci* 2009; 28(1): 36-51.
11. Golbraikh AI, Shen M, Xiao Z, Xiao YD, Lee KH, Tropsha A: Rational selection of training and test sets for the development of validated QSAR models, *J Comput Aided Mol Des* 2003; 17(2): 241-253.
12. Rabal O, Urbano-Cuadrado M & Oyarzabal J: Computational medicinal chemistry in fragment-based drug discovery: what, how and when, *Future Med. Chem* 2011; 3(1):95-134.
13. Abdullahi AD, Abdulkader AM, Samat NHA, Mohamed F, Muhammad BY, Mohammed HA: Novel Insight into the Structural Requirements of P70S6K Inhibition Using Group-based Quantitative Structure Activity Relationship (GQSAR), *Journal of Applied Pharmaceutical Science* 2014; 4(06):16-24.
14. V Life MDS 3.0, Molecular Design Suite Developed by V Life Sciences Technologies Pvt. Ltd. 2007
15. Glide, version 5.5, Schrödinger, LLC, 2009.
16. Mulay abhinit: Exploring potential of 4-thiazolidinone: a brief review: *IJPPS* 2009; 1(1): 47-64.
17. Metzger J V: *Comprehensive Heterocyclic Chemistry*. Pergamon: Oxford 6 1984; 236-330.
18. Singh T P et-al: Pharmacological Evaluation of Thiazolidinone Derivatives: A Prospective Review, *Der Pharma Chemica* 2011; 31: 194-206.
19. Singh S P, Parmar S S and Raman K: Chemistry and biological activity of thiazolidinones, *Chem. Rev.* 1981; 175.
20. Jain AK, Vaidya A, Ravichandran V, Kashaw SK, Agrawal RK: Recent developments and biological activities of thiazolidinone derivatives: A review, *Bioorg Med Chem* 2012;20:3378-95.
21. Wang S, Zhao Y, Zhang G, Yingxiang Lv, Zhan N, Gong P: Design, synthesis and biological evaluation of novel 4-thiazolidinones containing indolin-2-one moiety as potential antitumor agent, *European journal of medicinal chemistry* 2011; 46: 3509-3518.
22. Foster, Dean P. and Edward I. George: The Risk Inflation Criterion for Multiple Regression, *Annals of Statistics* 1994; 22(4) 1947-75
23. Wilkinson L. and Dallal GE: Tests os Significance in Forward Selection Regression with an F-to-Enter Stopping Rule, *Technometrics* 1981: 23:377-380.
24. Gasteiger J and Engel T. *Cheminformatics: A Textbook*. Wiley-VCH, Verlag GmbH & Co. KGaA. 2003.
25. Wold S: *QSAR-Chemometric Methods in Molecular Design*, vol. 2, Wiley VCH, 1995.
26. Miranda AA, Le Borgne YA, and Bontempi G: New Routes from Minimal Approximation Error to Principal Components, *Neural Processing Letters* 2008; 27(3): 29.
27. Kutner MH, Nachtsheim CJ, and Neter J: *Applied Linear Regression Models*, McGraw-Hill, fourth edition 2004.
28. Virupakshl B, Prashant K: Analysis of naphthoquinone derivatives as topoisomerase-I inhibitors using fragment based QSAR, *Int J Curr Res Aca Rev.* 2015; 3(5): 288-307.
29. Tropsha A, Gramatica P, Gombar VK: The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, *QSAR & Comb Sci* 2003; 22:69-77.
30. Abdullahi AD, Abdulkader AM, Samat NHA, Mohamed F, Muhammad BY, Mohammed HA: Novel insight into the structural requirements of p70s6k inhibition using group-based quantitative structure activity relationship (GQSAR), *Journal of Applied Pharmaceutical Science* 2014; 4 (06):016-024.
31. Abdullahi AD, Abdulkader AM, Abdulsamat NB, Ingale K: Application of group-based qsar and molecular docking in the design of insulin-like growth factor antagonists, *Trop J Pharm Res.* 2015; 14 (6): 941-951.
32. Scior T, Medina-Franco J, Do QT, Martínez-Mayorga K, Yunes Rojas J, Bernard P: How to recognize and workaroud pitfalls in QSAR studies: a critical review, *Cur Med Chem*, 2009; 16:4297-4313.

How to cite this article:

Nayak P and Kachroo M: A comparative group-QSAR and molecular docking studies of 4 -thiazolidinone containing indolin-2-one moiety as vgef inhibitors. *Int J Pharm Sci Res* 2017; 8(4): 1796-05. doi: 10.13040/IJPSR.0975-8232.8(4).1796-05.