WORLD JOURNAL OF PHARMACEUTICAL RESEARCH
SJIF Impact Factor 5.045
Volume 3, Issue 6, 890-901. Research Article ISSN 2277 – 7105

3D QSAR STUDIES ON THIAZOLIDINONE DERIVATIVES AS POTENTIAL EPIDERMAL GROWTH FACTOR RECEPTOR INHIBITORS

Sanmati K Jain*, Ravi Tripathi and Piyush Ghode

Drug Discovery and Research Laboratory, Deptt. of Pharmaceutical Chemistry, SLT Institute of Pharmaceutical Sciences, Guru Ghasidas Vishwavidyalaya, Bilaspur (C.G.), India

Article Received on
16 June 2014,

Revised on 11 July 2014,
Accepted on 06 August 2014

***Correspondence for Author**
Dr. Sanmati K. Jain
Drug Discovery and Research Laboratory, Deptt. of Pharmaceutical Chemistry, SLT Institute of Pharmaceutical Sciences, Guru Ghasidas Vishwavidyalaya, Bilaspur (C.G.), India.

ABSTRACT

Three dimensional quantitative structure activity relationship (3D QSAR) analysis was performed on thiazolidinone analogues for their Epidermal Growth Factor inhibitory activity (EGFR inhibitors) by k Nearest Neighbor Molecular Field Analysis (kNN-MFA) method using Molecular Design Suite (VLife MDS) software. kNN-MFA coupled with stepwise variable selection (forward-backward) method was applied to derive QSAR models and these models were validated for statistical significance by internal and external validation. Among the models generated, two models with optimum values of validation parameters had q2 and predr2 values of 0.8032 and 0.7843 for model 1 and 0.7933 and 0.7700 for model 2 respectively. The negative value of steric field point at S_374 in both the models indicates the need of less bulky group at this point for favorable biological activity. The positive value of electrostatic field point at E_555 suggests the need of positive ionizable group. The present work may help in providing guidance for further lead optimization and designing of potent anticancer agents.

KEYWORDS: 4-Thiazolidinone, EGFR, QSAR, anticancer.

INTRODUCTION

The conventional chemotherapeutic agents are responsible for severe toxicities thanks to the cytotoxic events resulting from the disruption of various aspects of DNA synthesis and repair or as a consequence of the disturbance of mitosis. The increased understanding of transmembrane signaling by growth factor receptor kinases (RTKs) has paved a way towards

the development of novel targeted therapy specific to cancer cells. The inhibition of proteins expressed by the ErbB gene family is proving to be the promising target for cancer therapy in modern days since the aberrant expression of these receptors may result in many types of solid tumors[1]. The epidermal growth factor receptor (EGFR) and human epidermal growth factor receptor (HER-2) are two such receptor tyrosine kinases, which belong to ErbB family and their over-expression has been proven to play key roles in cell proliferation and differentiation[2].The over-expression, mutations or autocrine expression of ligand may result in hyperactivation of EGFR. Inhibition of EGFR activity may be effected by two mechanisms (i) by blocking binding of native EGF ligand to the receptor e.g. by monoclonal antibodies such as Cetuximab and Panitumumab and (ii) tailoring small molecules that compete with ATP in the intracellular tyrosine kinase domain (TKD) and block the receptor activity regardless of endogenous ligand binding, e.g. Geftinib, Erlotinib and Lapatinib[3,4]. However, these drugs may be very effective but at the same time they may not be considered to be the "boon" against cancer because of the reported mutations in EGFR rendering it resistant to Geftinib and Erlotinib[5].

The Quantitative Structure Activity Relationship (QSAR) technique belongs to the paradigm of rational drug design. Also known as indirect drug design or ligand based drug design; this technique establishes a statistical relationship between the properties of amolecule (physicochemical, structural and conformational)and the biological response; Thus providing an understanding of the driving forces for the drug's action and helping to predict the biological activities of newly designed analogues and hence, contributing to the drug discovery processes[6, 7]. Based on the types of descriptors (physicochemical, structural and conformational properties) involved, QSAR may be one dimensional (1D), 2D, 3D, 4D, 5D and 6D. Out of these, the most popular are 2D and 3D QSAR methodologies. Whereas the 2D QSAR involves the descriptors correlating activity with structural patterns like connectivity indices, 2D- pharmacophores etc., without taking into account the 3D-representation of these properties; the 3D QSAR involves descriptors correlating activity with non-covalent interaction fields surrounding the molecules[8].

In view of the biological activities of heterocyclic compounds, the thiazolidinones and their derivatives have a variety of biological activities[9,10] such as antibacterial[11], antitubercular[12], antiinflammatory[13],antihistaminic[14], antiviral[15] etc. Thiazolidinones have also been reported to possess anticancer activity[16, 17].Recently, thiazolidinones were

reported to possess potential EGFR and HER-2 kinase inhibitory activities[18]. Thus, the literature offers the indication of possible candidature of different substituted thiazolidinones derivatives against EGFR and the present work is an attempt to investigate the 3D QSAR on a series of thiazolidinones in order to find the structural features that may prove to be useful in development of superior EGFR inhibitors.

MATERIALS AND METHODS

3D QSAR studies using k nearest neighbor molecular field analysis (kNN-MFA) approach were performed using the Molecular Design Suite (VLife MDS software package, version 3.5; from VLife Sciences, Pune, India)[19], user interface implemented on HCl computer with a Intel (R) dual core processor and Windows XP operating system.

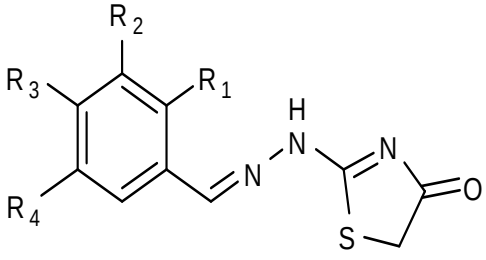
Data set for analysis

The data set used for this study contains 19 thiazolidinone derivatives as reported by Cheng *et al*[18]**(Table 1)** in which the EGFR inhibitory activities of molecules were calculated as concentration required to inhibit 50% (IC_{50}) of EGFR using solid phase ELISA assay. The IC_{50} values were converted to the negative logarithmic scale [pIC_{50} (moles)].

Molecular modeling for 3D QSAR

The structures of all the 19 molecules were built using the 2D draw application of VLife Molecular Design Suite 3.5 software and were converted to 3D structures for further analysis. All the compounds were batch optimized for the minimization of energies and geometry optimization using Universal Force Field (UFF) followed by considering distance-dependent dielectric constant of 1.0, convergence criterion or root-mean-square (RMS) gradient at 0.01kcal/mole and the iteration limit to 10,000.For electrostatic and steric energies, cutoffs were set to default values of 30.0 and 10.0 Kcal/mole respectively.

Table 1: Series of thiazolidinones with their pIC50 values



S. No.	Comp. No.	R_1	R_2	R_3	R_4	EGFR kinase	
						IC50	pIC50
1	3	H	H	F	H	2.18	5.66
2	4	H	H	Cl	H	1.94	5.71
3	5	H	H	Br	H	1.72	5.76
4	6	H	H	CH3	H	3.26	5.49
5	7	H	H	OCH3	H	5.78	5.24
6	8	H	H	NO2	H	6.82	5.17
7	9	H	H	OH	H	4.71	5.33
8	10	OH	H	H	F	0.97	6.01
9	11	OH	H	H	Cl	0.56	6.25
10	12	OH	H	H	Br	0.09	7.05
11	13	OH	Cl	H	Cl	1.46	5.84
12	14	OH	Br	H	Br	1.24	5.91
13	15	H	F	H	H	13.53	4.87
14	16	H	Cl	H	H	10.78	4.97
15	17	H	Br	H	H	7.46	5.13
16	18	H	CH3	H	H	23.49	4.63
17	19	H	OCH3	H	H	31.25	4.51
18	20	H	NO2	H	H	26.61	4.57
19	21	H	OH	H	H	19.32	4.71

Alignment of molecules

Molecular alignment is a crucial step in 3D-QSAR study in order to obtain significant results. In this study, the molecules of the data set were aligned by template-based method, where a template structure is defined and used as a basis for alignment of a set of molecules, and a reference molecule is chosen on which the other molecules of the data set get aligned considering the chosen template. After defining the template, the most active compound in the series i.e. compound 12 was chosen as the reference molecule and all the other molecules were aligned over it using the template based alignment (**Fig 1**).

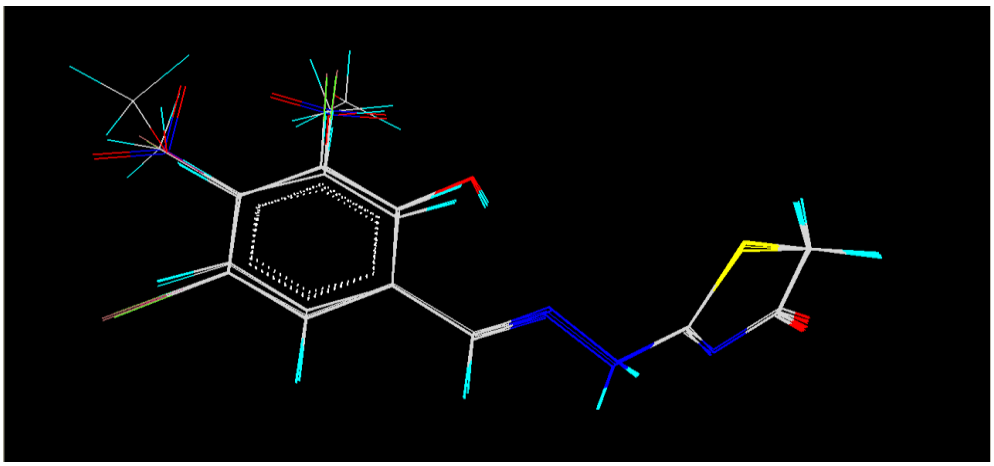


Fig. 1 Template based alignment of all the molecules of data set

Descriptor Selection

After the alignment of molecules, the molecular field was computed on a grid of points in space around the molecule. This field provides a description of how each molecule will tend to bind in the active site. Descriptors representing the steric and electrostatic interaction energies were computed at the lattice points of the grid using a methyl probe of charge +1.

Division of data in training and test set

The validation of QSAR models is an essential criterion so as to prove their statistical significance. The QSAR models can be subjected to internal validation and external validation. The internal validation can be performed using leave one out cross validation procedure. The result of this procedure is obtained in the form of the cross validated R^2 (q^2) which can serve as the basis for acceptance or rejection of the model(s) (the model is accepted if $q^2>0.5$ and vice versa). But since in this procedure, no external data set is involved, it cannot be agreed upon as the absolute criterion for either accepting or rejecting a model. Therefore, for the purpose of external validation, the data set should be divided into training and test sets. The training set is used to develop the model and the test set is used for the validation of the model. For the purpose of this study, the data set was divided into training and test sets by random selection method. In this method, the user can define the percentage of compounds he wants in the training set. The software randomly places the compounds in the training set and rest of the compounds are automatically set out in the test set. The training set comprised of 80% of the compounds and rest 20% of the compounds belonged to the test set.

Feature selection and model development

The 3D QSAR models were generated using k Nearest Neighbor Molecular Field Analysis (kNN-MFA) module with stepwise variable selection (SW) as variable selection method. The parameter settings used for SW-kNN MFA are cross correlation limit as 0.5, maximum number of variables in final equation as n/5 (n is the number of compounds in training set), term selection criteria as q2, Ftest in as 4 and Ftest out as 3.99, variance cut-off as 0 and Scaling as Auto Scaling, number of maximum neighbors as 5, number of minimum neighbors as 2 and distance based weighted average as prediction method.

RESULTS AND DISCUSSION

A series of 19 thiazolidinone derivatives having EGFR inhibitory activity was selected for 3D QSAR studies. This was achieved by dividing the compounds in training and test sets (80% of compounds in training set and rest 20% in test set).Various models were constructed from the different combinations of training sets and test sets with random selection method which resulted in different models among which some were statistically significant (q2 > 0.5, predr2 > 0.5) (**Table 2**). The models having optimum values of q2 and predr2 are reported here. For the development of these models, the training and test sets were selected in such a way that they follow the Unicolumn statistics, i.e. the maximum value of pIC50 test set is less than maximum value of pIC50 of training set and the minimum value of pIC50 of test set is higher than the value of pIC50 of the training set. Unicolumn statistics for the most optimum models in terms of q2 and predr2 are presented in **Table 3**. These models are described in **Table 4**.

Table 2: Results of 3D-QSAR analysis using kNN (k nearest neighbor) method by using Random data selection (80% training set)

Trial	Test Set	SW-kNN MFA			
		q2	q2_se	Predr2	pred_r2se
1	17, 20, 5, 9	0.7307	0.3611	-0.7988	0.7840
2	16, 17, 21, 9	0.6944	0.3858	-0.4993	0.7460
3	16, 17, 21, 9	0.6944	0.3858	-0.4993	0.7460
4	15, 17, 19, 9	0.7674	0.3244	-1.8379	1.2614
5	17, 18, 5, 9	0.7491	0.3512	-0.9815	0.7764
6	15, 16, 17, 9	0.7061	0.3847	0.3480	0.4288
7	16, 20, 7, 9	0.6820	0.3886	0.8418	0.2612
8	16, 18, 21, 9	0.6488	0.3935	0.4865	0.5717
9	18, 4, 8, 9	0.8128	0.3045	0.7122	0.2875
10	14, 17, 5, 9	0.8437	0.2849	0.4045	0.3125
11	21, 6, 8, 9	0.7380	0.3651	0.2267	0.4239
12	10, 15, 7, 9	0.7180	0.3769	0.4004	0.3722
13	15, 20, 6, 9	0.7104	0.3695	-4.5438	1.5514
14	4, 6, 8, 9	0.7621	0.3589	-2.2910	0.4199
15	13, 14, 21, 9	0.7857	0.3229	0.0335	0.5486
16	10, 11, 6, 9	0.6638	0.3960	0.6361	0.4102
17	15, 21, 8, 9	0.8032	0.3075	0.7843	0.2960
18	15, 4, 6, 9	0.7933	0.3292	0.7700	0.1760
19	17, 3, 5, 9	0.8287	0.3023	-0.4745	0.3694
20	17, 20, 4, 9	0.6951	0.3850	-1.6981	0.9511

Table 3: Unicolumn statistics for the described models

Model No.	Data Set	Average	Max	Min	Std Dev	Sum
1 (Trial 17)	Training	5.52	7.05	4.51	0.693334	82.73
	Test	5.02	5.33	4.71	0.281188	20.08
2 (Trial 18)	Training	5.43	7.05	4.51	0.724228	81.41
	Test	5.02	5.33	4.87	0.355903	20.08

Table 4: Optimum models generated using kNN MFA method

Parameters	Model-1	Model-2
Training Set Size (n)	15	15
Test set size	4	4
K nearest neighbor	2	2
Degree of freedom	12	12
q2	0.8032	0.7933
q2_se	0.3075	0.3292
Pred_r2	0.7843	0.7700
Pred_r2se	0.2960	0.1760
Descriptor range	S_374 -0.4341 -0.4185 S_209 -0.4723 -0.4669	S_374 -0.4237 -0.4185 E_555 0.0949 0.1431

Interpretation of 3D QSAR models

The careful analysis of comparison of biological activities (pIC50) and predicted activities for training set and test set molecules in the most significant models indicate very less significant differences (lower values of residuals as shown in **Table 5.**

Table 5: Observed, predicted and residual values of the best models

S. No.	Compound No.	Actual Values	Model 1		Model 2	
			Predicted Value	Residual Values	Predicted Value	Residual Values
1	3	5.66	5.83	-0.17	5.50	0.16
2	4	5.71	5.75	-0.04	5.50*	0.21*
3	5	5.76	5.58	0.19	5.45	0.31
4	6	5.49	5.45	0.04	5.45*	0.04*
5	7	5.24	5.31	-0.07	5.71	-0.47
6	8	5.17	5.58*	-0.41*	5.50	-0.33
7	9	5.33	5.58*	-0.25*	5.46*	-0.13*
8	10	6.01	5.81	0.20	6.09	-0.08
9	11	6.25	6.48	-0.23	6.48	-0.23
10	12	7.05	6.08	0.97	6.13	0.92
11	13	5.84	5.69	0.15	6.08	-0.24
12	14	5.91	5.71	0.20	6.04	-0.13
13	15	4.87	5.05*	-0.18*	5.05*	-0.18*
14	16	4.97	5.31	-0.34	4.92	0.05
15	17	5.13	5.23	-0.10	5.07	0.06

16	18	4.63	4.77	-0.14	4.84	-0.21
17	19	4.51	4.60	-0.09	4.67	-0.16
18	20	4.57	4.57	0.00	4.57	0.00
19	21	4.71	4.79*	-0.08*	4.80	-0.09

Test set compounds

Model 1 (Trial 17)

k Nearest Neighbor= 2, $N_{training}$ = 15 , N_{test} = 4, Degree of freedom = 12, q2 = 0.8032, q2_se = 0.3075, Predr2 = 0.7843, pred_r2se = 0.2960

In this model, the lattice points at S_374 and S_209 (green points) are having a range of -0.434085 to -0.418465 and -0.472346 to -0.466916 respectively. These indicate that the steric contribution i.e. the steric bulk of the substituents is important for structure activity relationships. The negative contribution of these steric field points indicates that the substitutions at these points should be done with less bulky groups as the more bulky groups may interfere with attainment of proper orientation in the receptor binding site (**Fig. 2**). The data fitness plot and a comparison between actual and predicted activities of all the compounds is shown in **Fig. 4**.

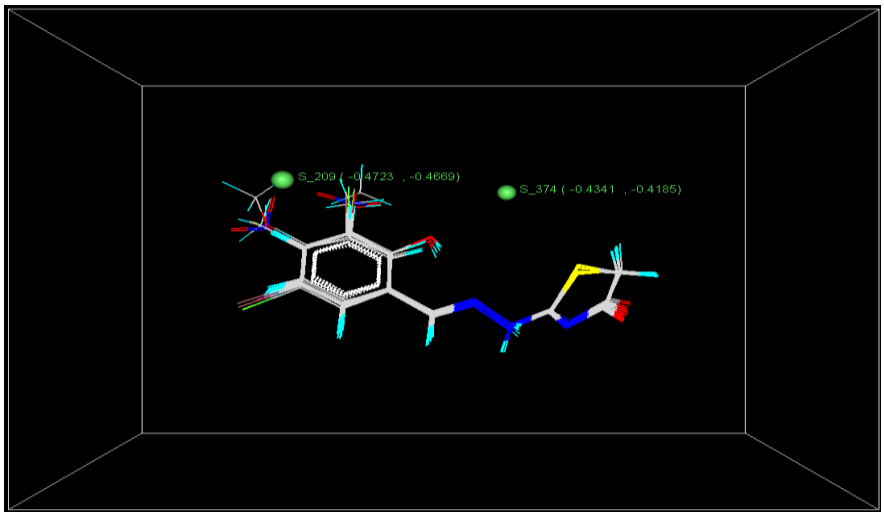


Fig. 2: Contribution 3D plot for Model 1 (SW-kNN MFA)

Model 2 (Trial 18)

k Nearest Neighbor= 2, $N_{training}$ = 15 , N_{test} = 4, Degree of freedom = 12, q2 = 0.7933, q2_se = 0.3292, Predr2 = 0.7700, pred_r2se = 0.1760

The lattice points in this model are S_374 and E_555 showing the effects of steric and electrostatic fields at lattice points 374 and 555 respectively. The range of S_374 is -0.4237 to -0.4185 and E_555 is 0.0949 to 0.1431. Again, the negative value of steric field point

range indicates the need of less bulky substituents for favorable biological activity. The positive contribution of electrostatic field point suggests the requirement of a positive ionizable group at this position (**Fig. 3**). The data fitness plot and a comparison between actual and predicted activities of all the compounds is shown in **Fig. 5.**

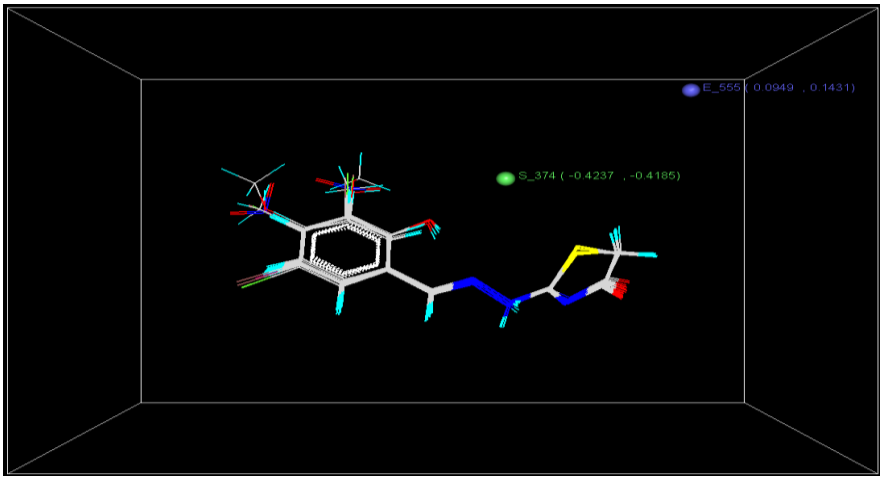
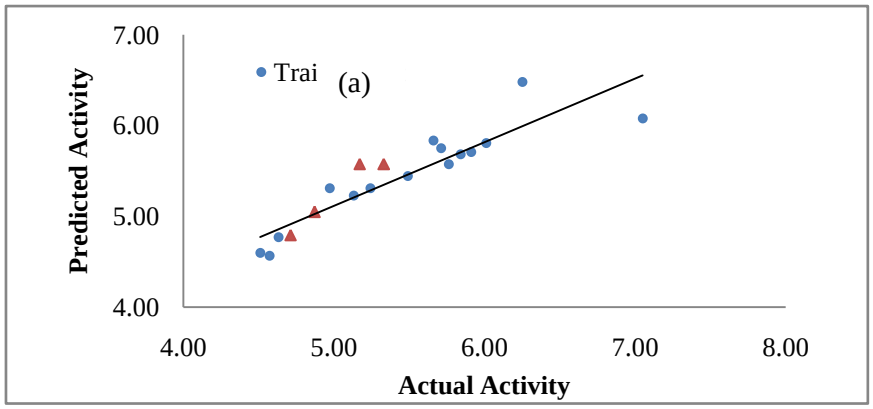
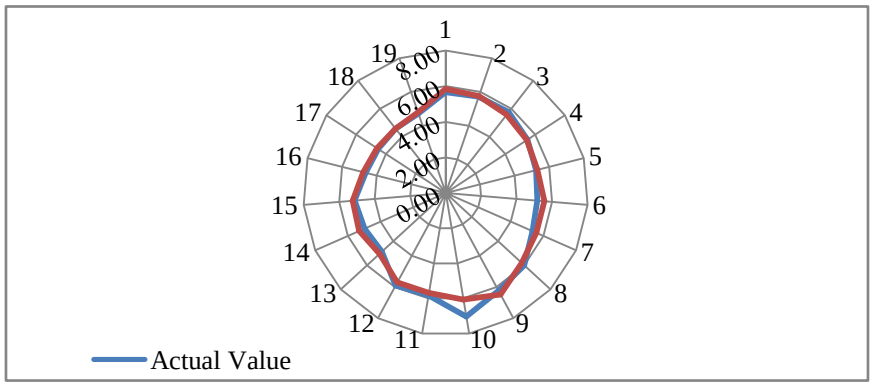


Fig. 3: Contribution 3D plot for Model 2 (SW-kNN MFA)



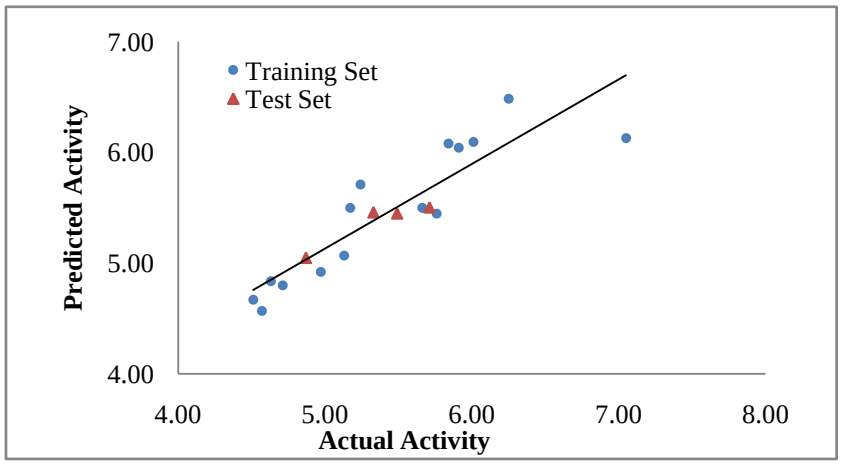
a



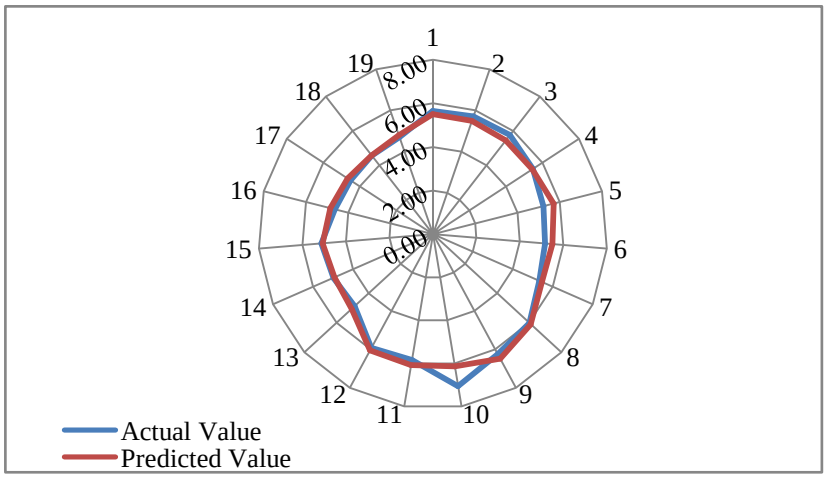
b

Fig. 4: (a) Graph of observed vs. predicted activities for model 1 (SW-kNN-MFA)

(b) Comparison of actual and predicted activities for all the compounds for the same model (blue line: actual activity, red line: predicted activity)



a



b

Fig. 5: (a) Graph of observed vs. predicted activities for model 2 (SW-kNN-MFA)

(b) Comparison of actual and predicted activities for all the compounds for the same model (blue line: actual activity, red line: predicted activity)

SUMMARY AND CONCLUSION

In the present study, an attempt has been made to establish the quantitative structure activity relationship between a set of 19 thiazolidin-4-one derivatives and their EGFR inhibitory activity. 3D QSAR studies were performed using kNN-MFA method. For model development, steric and electrostatic descriptors were considered. The compounds were divided into training and test sets by random selection method (80% in training set and 20% in test set). Twenty trials were run. The variable selection method applied was stepwise forward backward. Out of all the developed models, two models were selected for description based on the values of internal and external validation parameters (q2 and predr2 respectively). These models (model 1 and model 2) belong to trial 17 and 18 and have q2 values of 0.8032 and 0.7933 and predr2 values of 0.7843 and 0.7700 respectively. The

negative range of steric fields at lattice point S_374 in both the models strongly suggests the need of less bulky group at this position while the positive electrostatic contribution of electrostatic field point in model 2 suggests the requirement of a positive ionizable group at this position. Therefore the substituents having the desired properties at these positions may result in compounds with higher biological activity and the structural information attained from this study can be used for designing of the newer EGFR inhibitors with better anticancer activity.

REFERENCES

1. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol, 2001; 2(2):127–37.
2. Olayioye MA, Neve RM, Lane HA, Hynes NE. The ErbB signaling network: receptor heterodimerization in development and cancer. EMBO J, 2000; 19(13): 3159-67.
3. Fry DW. Mechanism of action of erbB tyrosine kinase inhibitors. Exp Cell Res, 2003; 284(1): 131-9.
4. Lin J, Shen W, Xue J, Sun J, Zhang X, Zhang C. Novel oxazolo[4,5-g]quinazolin-2(1H)-ones: Dual inhibitors of EGFR and Src protein tyrosine kinases. Eur J Med Chem, 2012; 55(9): 39-48.
5. Cruz-López O, Conejo-García A, Núñez MC, Kimatrai M, García-Rubiño ME, Morales F, Gómez-Pérez V, Campos JM. Novel substituted quinazolines for potent EGFR tyrosine kinase inhibitors. Curr Med Chem, 2011; 18(7): 943-63.
6. Ferreira MMC. Multivariate QSAR. J BrazChemSoc, 2002; 13(6):742–53.
7. Kurup A, Garg R, Carini DJ, Hansch C. Comparative QSAR: angiotensin II antagonists. Chem Rev, 2001; 101(9):2727–50.
8. Verma J, Khedkar VM, Coutinho EC. 3D-QSAR in Drug Design, Curr Top Med Chem, 2010; 10(1): 95-115.
9. 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(11): 3378–95.
10. Verma A, Saraf SK. 4-Thiazolidinone: A biologically active scaffold. Eur J Med Chem, 2008; 43(5): 897-905.
11. Liesen, AP, Aquino TM, Carvalho CS, Lima VT, Araujo JM, Lima JG, Faria AR, Melo EJT, Alves AJ, Alves EW, Alves AQ, Goes AS. Synthesis and evaluation of anti-

Toxoplasma gondii and antimicrobial activities of thiosemicarbazides, 4-thiazolidinones and 1,3,4-thiadiazoles. Eur J Med Chem, 2010; 45(9): 3685-91.

12. Kucukguzel SG, Oruc EE, Rollas S, Sahin F, Ozbek A. Synthesis, characterisation and biological activity of novel 4-thiazolidinones, 1,3,4-oxadiazoles and some related compounds. Eur J Med Chem 2002; 37(3): 197-206.
13. Kumar A, Rajput CS, Bhati SK. Synthesis of 3-[4′-(*p*-chlorophenyl)-thiazol-2′-yl]-2-[(substituted azetidinone/thiazolidinone)-aminomethyl]-6-bromoquinazolin-4-ones as anti-inflammatory agent. Bioorg Med Chem, 2007; 15(8): 3089-96.
14. Diurno MV, Mazzoni O, Piscopo E, Caliganao A, Giordano F, Bolognese A. Synthesis and antihistaminic activity of some thiazolidin-4-ones. J Med Chem, 1992 35(15): 2910-12.
15. Terzioglu N, Karalı N, Gursoy A, Pannecouque C, Leysen P, Paeshuyse J, Neyts J, De Clercq E. Synthesis and primary antiviral activity evaluation of 3-hydrazono-5-nitro-2-indolinone derivatives. ARKIVOC, 2006; I: 109-118.
16. Chandrappa S, Kavitha CV, Shahabuddin MS, Vinaya K, Anandakumar CS, Ranganatha SR, Raghavan SC, Rangappa KS. Synthesis of 2-(5-((5-(4-chlorophenyl)furan-2-yl)methylene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid derivatives and evaluation of their cytotoxicity and induction of apoptosis in human leukemia cells. Bioorg Med Chem, 2009; 17(6): 2576-84.
17. Zhou H, Wu S, Zhai S, Liu A, Sun Y, Li R, Zhang Y, Ekins S, Swaan PW, Fang B, Zhang B, Yan B. Design, Synthesis, Cytoselective Toxicity, Structure–Activity Relationships, and Pharmacophore of Thiazolidinone Derivatives Targeting Drug-Resistant Lung Cancer Cells. J Med Chem, 2008; 51(5): 1242-51.
18. Lv PC, Zhou CF, Chen J, Liu PG, Wang KR, Mao WJ, Li HQ, Yang Y, Xiong J, Zhu HL. Design, synthesis and biological evaluation of thiazolidinone derivatives as potential EGFR and HER-2 kinase inhibitors. Bioorg Med Chem, 2010; 18(1): 314–9.
19. VLife MDS: Molecular Design Suite, VLife Sciences Technologies Pvt. Ltd., Pune, India, 2010 (www.vlifesciences.com).