QSAR and Molecular Docking Studies on a Series of Cinnamic Acid Analogues as Epidermal Growth Factor Receptor (EGFR) Inhibitors

4 authors, including:

Shaik Bashirulla
National Institute of Technical Teachers' Training and Research, Bhopal
45 PUBLICATIONS   **90** CITATIONS

SEE PROFILE

Omar Deeb
Al-Quds University
43 PUBLICATIONS   **494** CITATIONS

SEE PROFILE

Satya Prakash Gupta
National Institute of Technical Teachers' Training and Research, Bhopal
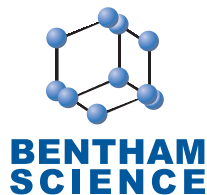**225** PUBLICATIONS   **1,845** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   QSAR and Molecular Modeling View project

RESEARCH ARTICLE

QSAR and Molecular Docking Studies on a Series of Cinnamic Acid Analogues as Epidermal Growth Factor Receptor (EGFR) Inhibitors

Basheerulla Shaik[1], Omar Deeb[2], Vijay K. Agrawal[3] and Satya P. Gupta[1,*]

[1]*Department of Applied Sciences, NITTTR, Shamla Hills, Bhopal-462002, Madhya Pradesh, India;* [2]*Faculty of Pharmacy, Al-Quds University, P.O. Box 20002, Jerusalem, Palestine;* [3]*Department of Chemistry, A. P. S. University, Rewa-486003, Madhya Pradesh, India*

Abstract: Quantitative structure-activity relationship (QSAR) and docking studies have been performed on a large series of cinnamic acid analogues studied by various authors as Epidermal Growth Factor Receptor (EGFR) inhibitors. A multiple linear regression (MLR) analysis has shown that electronic properties of these compounds are the governing factors of their activity and docking study has shown that compounds can form hydrogen bonds with the receptor and have effective steric interactions involving dispersion forces. Using the MLR model, some new compounds were proposed that have higher potency than the existing ones.

INTRODUCTION

Presently, it is felt that there is a serious need for new targets to develop anticancer drugs. It has been found that a protein tyrosine kinase (PTK), epidermal growth factor receptor (EGFR), has been associated with many human cancers, such as breast and liver cancers. This led many to consider that EGFR can be an attractive target for the development of anticancer drugs. Conservative cytotoxic drugs for cancer chemotherapy have been usually found to be associated with severe toxic side effects, while drugs intended to inhibit molecular targets are found to show high selectivity and low toxicity. Therefore, searching small molecules to inhibit EGFR activities shows great promise in developing new anticancer agents.

A naturally occurring aromatic fatty acid of low toxicity, cinnamic acid, is quite well exposed to human and its amide derivatives with cyano and fluoro substituents have been found to be of particular value due to their inhibitory effect in mitochondrial pyruvate transport. They were reported to have many different biological activities such as anticancer and antioxidant effects [1-3]. The simple substituents on their N-phenyl ring affect the EGFR inhibitory activities. Nitroimidazole derivatives have attracted significant attention as they showed affinity to penetrate and accumulate in regions of tumors. They can experience bioreduction to yield electrophilic substances which can damage protein and nucleic acids. The toxicology and metabolism of nitroimidazoles, particularly metronidazole, have been characterized. As a novel class of bioreductively activated nitroimidazole compounds, a series of cinnamic acid metronidazole ester derivatives were synthesized and evaluated for their biological activity to find that some of them had anticancer activity [4]. A combination of 4-anilinoquinazoline and cinnamic acid was found to have a novel mode of binding to the EGFR tyrosine kinase [5]. EGFR inhibitors are used to treat non-small-cell lung cancer, pancreatic cancer, breast cancer, colon cancer and some other cancers that are caused by epidermal growth factor receptor up-regulation.

In designing new drugs, biological activity estimation forms the basis for compound selection and optimization. While several experimental methods are available for screening the biological activity of compounds, they are somehow too expensive and time consuming. Quantitative structure-activity relationship (QSAR) analysis provides an effective and powerful tool for achieving the same goal with much lower cost. The aim of this work is to apply different statistical methods to explore the crucial structural properties of the compounds that are associated with their EGFR inhibitory activities. QSAR models were constructed for a combined data set of 54 compounds having EGFR inhibitory activity, which was retrieved from the literature [3-5]. The strength and the predictive performance of the proposed models were verified using both internal (cross-validation and Y-scrambling) and external statistical validations. In order to explore the scope of further modification in the structures of the

*Address correspondence to this author at the Department of Applied Sciences, NITTTR, Shamla Hills, Bhopal-462002, Madhya Pradesh, India; Tel: + 91-4449534745; Fax: +91-755-2661996; E-mail: spgbits@gmail.com

1

2

3
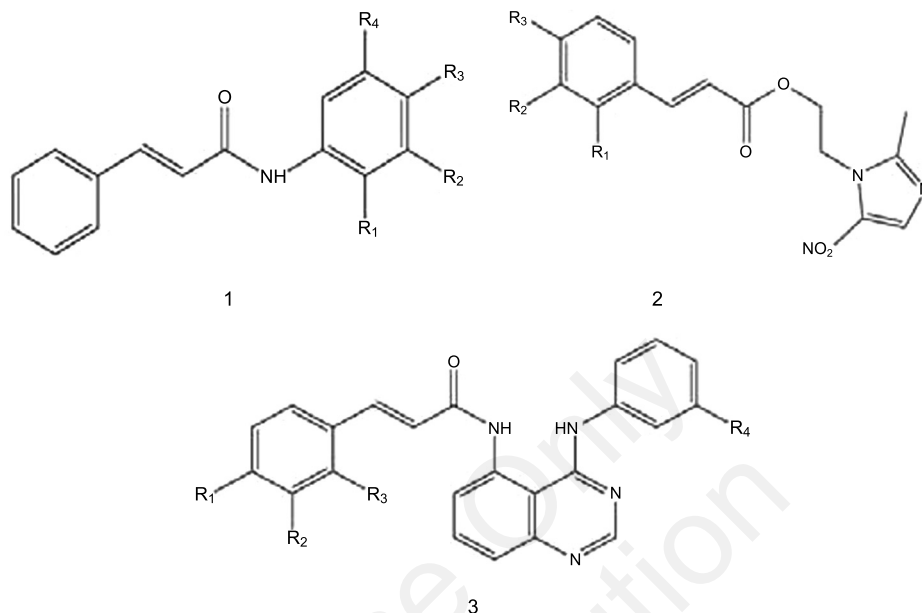
compounds leading to further increase in the potency, we have performed docking studies on some predicted compounds.

MATERIALS AND METHODS

We have taken three similar series of cinnamic acid analogues (**1**-**3**) that were synthesized and evaluated for their EGFR inhibitory activity by various workers [3-5]. A combination of these series is listed in Table **1** along with the EGFR inhibitory activities of the compounds. The chemical structures were drawn using ChemDoodle software [6]. Table **1** also lists the topological parameters of the compounds that were found to govern their potency. These topological parameters used have been calculated using Dragon software [7]. Among the thousands of parameters that were calculated, Table **1** lists only those parameters that were found to be important. These were frequency of C-O at topological distance 10, *(F10[C-O])*; spectral mean absolute deviation from Burden matrix weighted by ionization potential, *(SpMADB(i))*; Geary autocorrelation of lag 8 weighted by mass, *(GATS8m)*; and Moran autocorrelation of lag 4 weighted by Sanderson electronegativity, *(MATS4e)*.

RESULTS & DISCUSSION

Of the total dataset comprising of 54 compounds (Table **1**), 40 compounds were selected for training set and 14 compounds were used for the test set to evaluate the predictability of the developed models. The test set compounds are indicated in the table by a bold superscript '**b**'. The QSAR models were obtained using QSARINs software [8-10]. The generation of training and test sets is done using split option present in QASRINs software by choosing the molecules on random basis. For obtaining QSAR models, Genetic Algorithm-Multi Linear Regression (GA-MLR) was employed using default settings in QSARINS. For statistical validation, variety of statistical parameters were calculated.

The best GA-MLR equation based on four descriptors along with statistical parameters were as follows.

pIC_{50} = 0.2019(±0.0339) *F10[C-O]* + 3.7300(±1.5732)*SpMADB(i)* − 0.7758(±0.3664)*GATS8m* + 0.8473(±0.5829)*MATS4e* − 0.1609 (1)
N = 39, R^2_{tr} = 0.842, $R^2_{adj.}$ = 0.824, S= 0.182, F= 45.401, Kxx = 0.253, ΔK = 0.089,
$RMSE_{tr}$ = 0.170, MAE_{tr} = 0.142, RSS_{tr} = 1,133, CCC_{tr} = 0.914

Q^2_{LOO} = 0.776, $RMSE_{cv}$ = 0.203, MAE_{cv} = 0.168, $PRESS_{CV}$ = 1.612, CCC_{CV} = 0.878,
Q^2_{LMO} = 0.766, R^2Y_{scr} =0.105, Q^2Y_{scr} = -0.195, R^2_{Pred} = 0.411, $RMSE_{ext}$ = 0.348,
MAE_{ext} = 0.284, $PRESS_{ext}$ = 1.571, CCC_{ext} = 0.594, $R^2_m\Delta$ = 0.327

pIC_{50} = 0.1873(±0.0358) *F10[C-O]* + 3.1318(±1.6759)*SpMADB(i)* − 0.6918(±0.3994)*GATS8m* + 0.6745 (2)
N = 39, R^2_{tr} = 0.802, $R^2_{adj.}$ = 0.785, S= 0.202, F= 47.204, Kxx = 0.201, ΔK = 0.137,
$RMSE_{tr}$ = 0.191, MAE_{tr} = 0.158, RSS_{tr} = 1.423, CCC_{tr} = 0.890

Q^2_{LOO} = 0.721, $RMSE_{cv}$ = 0.227, MAE_{cv} = 0.183, $PRESS_{CV}$ = 2.003, CCC_{CV} = 0.845,
Q^2_{LMO} = 0.705, R^2Y_{scr} =0.081, Q^2Y_{scr} = -0.154, R^2_{Pred} = 0.400, $RMSE_{ext}$ = 0.351,
MAE_{ext} = 0.296, $PRESS_{ext}$ = 1.602, CCC_{ext} = 0.550, $R^2_m\Delta$ = 0.4110

pIC_{50} = 0.1642(±0.0380) *F10[C-O]* + 3.4820(±1.9064)*SpMADB(i)* − 0.5804 (3)
N = 39, R^2_{tr} = 0.732, $R^2_{adj.}$ = 0.717, S= 0.213, F= 49.116, Kxx = 0.104, ΔK = 0.363,
$RMSE_{tr}$ = 0.222, MAE_{tr} = 0.173, RSS_{tr} = 1.926, CCC_{tr} = 0.845

$Q^2_{LOO} = 0.671$, $RMSE_{cv} = 0.246$, $MAE_{cv} = 0.191$, $PRESS_{CV} = 2.367$, $CCC_{CV} = 0.806$,
$Q^2_{LMO} = 0.655$, $R^2Y_{scr} = 0.053$, $Q^2Y_{scr} = -0.118$, $R^2_{Pred} = 0.337$, $RMSE_{ext} = 0.368$,
$MAE_{ext} = 0.327$, $PRESS_{ext} = 1.764$, $CCC_{ext} = 0.507$, $R^2_m\Delta = 0.328$

$$pIC_{50} = 0.1714(\pm0.0438)\ F10[C\text{-}O] + 4.8101 \qquad (4)$$

$N = 39$, $R^2_{tr} = 0.630$, $R^2_{adj.} = 0.620$, $S = 0.268$, $F = 62.887$, $Kxx = 0.000$, $\Delta K = 0.794$,
$RMSE_{tr} = 0.261$, $MAE_{tr} = 0.208$, $RSS_{tr} = 2.661$, $CCC_{tr} = 0.7727$

$Q^2_{LOO} = 0.590$, $RMSE_{cv} = 0.274$, $MAE_{cv} = 0.220$, $PRESS_{CV} = 2.9479$, $CCC_{CV} = 0.742$,
$Q^2_{LMO} = 0.584$, $R^2Y_{scr} = 0.0272$, $Q^2Y_{scr} = -0.0847$, $R^2_{Pred} = 0.357$, $RMSE_{ext} = 0.364$,
$MAE_{ext} = 0.317$, $PRESS_{ext} = 1.7214$, $CCC_{ext} = 0.4934$, $R^2_m\Delta = 0.436$

Among the statistical parameters, N is the number of data points (compounds), R^2_{tr} is the correlation coefficient of training set compounds, Q^2_{LOO} is the square of the cross-validated correlation coefficient obtained from leave-one-out (LOO), S is the standard deviation, and F is Fischer ratio between the variances of calculated and observed activities. The remaining symbols have their usual meaning given by various researchers [11-17].

Of the above four models, the very first model (eq. 1) is found to be statistically most significant. The validation of models using different statistical parameters indicates that the models have not occurred by chance. In models 1-4, the descriptor *F10[C-O]*, which is a 2D atom pair descriptor, has positive coefficient means that the frequency of C-O at topological distance 10 is conducive to the inhibitory activity. Similarly, the positive coefficients of *SpMADB(i)* and *MATS4e* suggest that ionization potential and electronegativity of the molecules have positive effect on the potency of the compounds. The negative coefficient of *GATS8m*, however, indicates that massive molecule will have adverse effect. Thus, this model suggested that the potency of the compounds are governed by their electronic properties. All the parameters used in the models appear to be significant, because if they are removed one by one the significance of the model decreases (eqs. 1-4).

The model expressed by eq. 1 was found to have good predictive ability when the activities of both the training and test set compounds were compared with their corresponding observed activities. A graph drawn between calculated and observed activities also verifies this fact (Fig. **1**). Using this model we predicted some new compounds of high EGFR inhibition potency as shown in Table **2**. Each predicted compound has higher potency than any compound in the existing series. Their potency is higher even than Erlotinib (compound 23, Table **1**, also given in Table **2** for direct comparison). All these predicted compounds follow the Lipinski rule of 5, according to which an orally active drug molecule should not have its hydrogen-bond donors (NH + OH) > 5,

hydrogen-bond acceptors (N + O) > 10, molecular weight > 500, and logP value > 5 (Table **3**).

DOCKING STUDY

A docking study was performed on the predicted compounds using Lead IT FlexX software to see the binding of these compounds with the EGFR. The ability of a molecule to interact with an enzyme decides its potency. For the study of docking, the crystal structure of the related enzyme is required which can now be retrieved from RCSB protein data bank. We selected the Erlotinib-enzyme complex with PDB entry code 1M17 (http://www.pdb.org). Erlotinib is the FDA approved EGFR inhibitor. Its structure is given in Table **2**, but for the sake of convenience it is shown as **4** given below. The docking results are shown in Table **4**. This table shows that all the
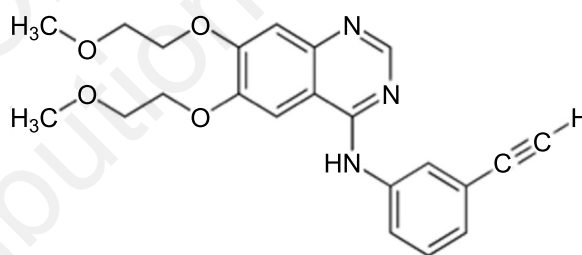
4, Erlotinib

predicted compounds, except 12, have hydrogen bonds equal to or greater than Erlotinib (**4**) and all of them have comparable docking scores. It can be seen that among all the predicted compounds, compound 7 has the highest activity ($pIC_{50} = 7.34$), and much higher than Erlotinib ($pIC_{50} = 6.47$). We made a comparative study of the interactions of these two compounds with the enzyme. The interactions of Erlitinib are shown in Figure **2** and that of compound 7 in Figure **3**. Figure **2** shows that in addition to forming 3 hydrogen bonds as described in Table **4**, Erlotinib also interacts with the enzyme through its various rings and side chains. Its aryl ring with acetylenic moiety is engulfed in a deep pocket of the enzyme formed by Pro770, Leu 694, and Leu768 residues and undergoes steric interaction. Similarly, heterocyclic ring of quinazoline interacts with a shallow trough of the enzyme constituted of only Met769, and its benzene ring faces a convex surface of the enzyme. An ester chain at the 5-position of the quinazoline ring tries to penetrate the shallow gulf formed by Gly 772 and Cys773 residues. In Figure **3**, that shows the binding of predicted compound **7**, it is observed that there are deep penetrations of crucial moieties of the compound in three deep pockets of the enzyme. A pocket formed by Asp831 and Phe699 engulfs a large portion of the pendent substituent at position 1 of the central ring, a pocket formed by Leu768, Leu694, Met769, and Leu820 engulfs the fluoro-substituted alkyl moiety of side chain, and a pocket formed by Thr766, Lys721, Val702, and Ala719 interacts with the alkyl moiety of the other side chain. The central ring also faces a convex surface of the residue Thr830. All these interactions are steric interactions, involving dispersion forces. Additionally, the molecule also forms three hydrogen bonds with the receptor as described in Table **4**.

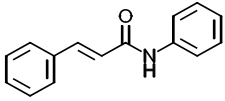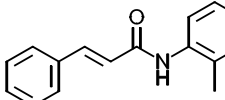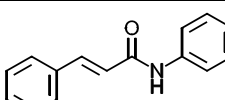**Table 1.** **List of Compounds Acting as EGFR Inhibitors and Their Potency and Related Structural Variables.**
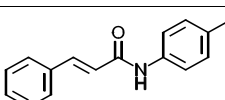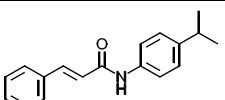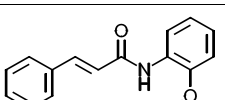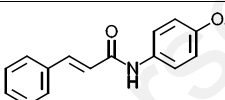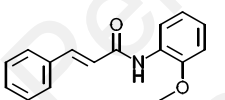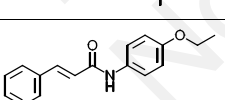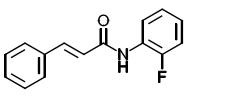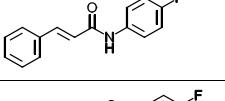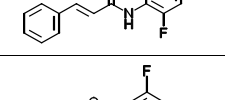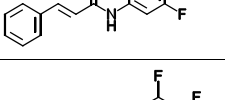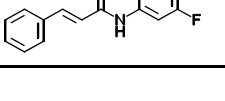
No	Structure	F10[C-O]	SpMADB(i)	GATS8m	MATS4e	pIC$_{50}$[a]		
						Obsd.	Cald. eq. 1	LOO[d]
1[b]		0	1.61	1.04	-0.02	4.64	5.03	-
2		0	1.58	0.92	-0.10	4.91	4.95	4.95
3		0	1.56	1.01	0.03	4.71	4.91	4.92
4		0	1.58	1.28	-0.01	4.90	4.74	4.71
5		0	1.54	1.00	0.00	4.84	4.82	4.82
6		1	1.58	0.90	-0.12	5.29	5.13	5.12
7[b]		2	1.58	0.98	-0.09	5.04	5.30	-
8		1	1.55	1.04	-0.12	5.03	4.93	4.93
9		2	1.55	1.08	-0.09	4.99	5.13	5.13
10		0	1.58	0.89	-0.14	5.13	4.93	4.91
11		0	1.58	1.06	-0.09	5.00	4.84	4.83
12		0	1.56	0.95	0.07	4.93	4.99	4.99
13		0	1.55	1.11	0.33	4.88	5.03	5.10
14		0	1.54	1.17	0.19	4.89	4.82	4.81

Table 1. contd…
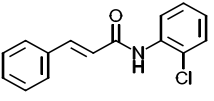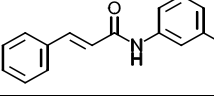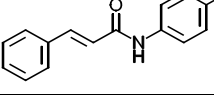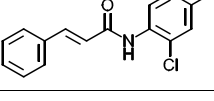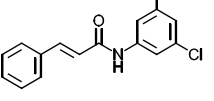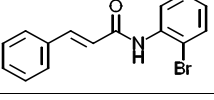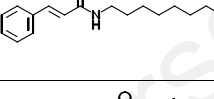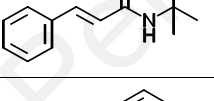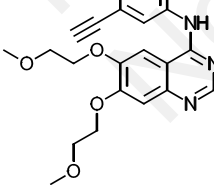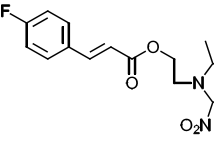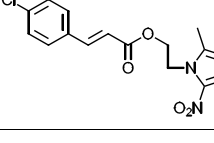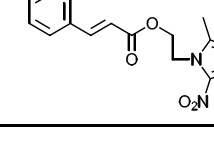
No	Structure	F10[C-O]	SpMADB(i)	GATS8m	MATS4e	pIC$_{50}$[a]		
						Obsd.	Cald. eq. 1	LOO[d]
15		0	1.58	0.83	-0.13	5.01	4.98	4.98
16		0	1.57	1.04	0.07	4.84	4.93	4.94
17		0	1.58	1.04	-0.09	4.86	4.85	4.85
18		0	1.55	0.92	-0.01	4.79	4.88	4.89
19		0	1.52	1.07	0.29	4.86	4.93	4.95
20		0	1.58	0.85	-0.11	4.80	4.98	5.00
21		1	1.42	1.01	-0.02	4.41	4.55	4.62
22		0	1.42	1.37	0.17	4.53	4.23	4.04
23		8	1.53	1.14	0.04	6.47	6.30	6.20
24[b]		2	1.52	1.02	-0.12	5.39	5.02	-
25[b]		2	1.51	0.85	-0.11	5.44	5.12	-
26		2	1.51	0.73	-0.10	5.49	5.24	5.17

Table 1. contd…
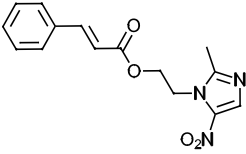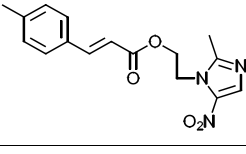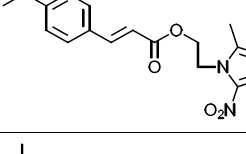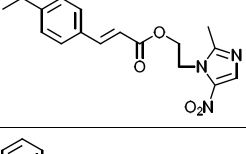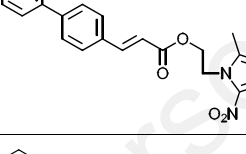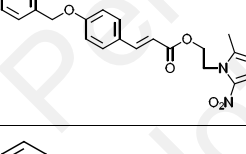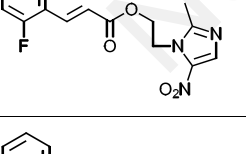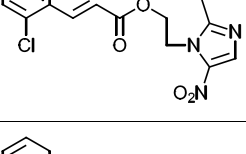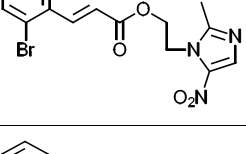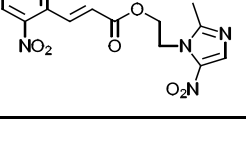
No	Structure	F10[C-O]	SpMADB(i)	GATS8m	MATS4e	pIC$_{50}$[a]		LOO[d]
						Obsd.	Cald. eq. 1	
27		2	1.55	1.19	-0.07	5.11	5.03	5.03
28[b]		2	1.51	1.04	-0.03	5.09	5.06	-
29		3	1.52	1.01	-0.09	4.96	5.27	5.29
30		2	1.50	1.00	0.00	5.17	5.06	5.05
31[b]		6	1.58	1.05	-0.03	6.21	6.09	-
32		5	1.55	0.98	-0.08	5.90	5.81	5.80
33		2	1.52	1.30	-0.17	4.57	4.77	4.80
34		2	1.51	1.33	-0.15	4.68	4.72	4.73
35		2	1.51	1.29	-0.13	4.61	4.78	4.80
36[b]		2	1.54	1.20	-0.17	4.60	4.93	-

Table 1. contd…
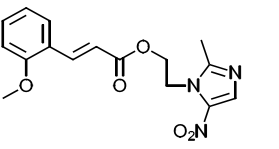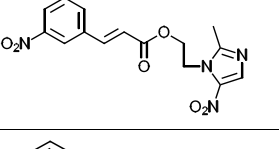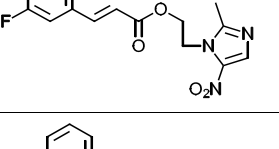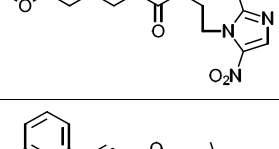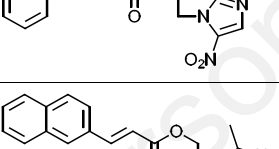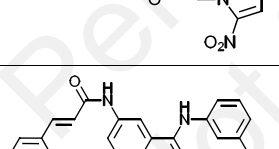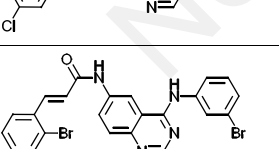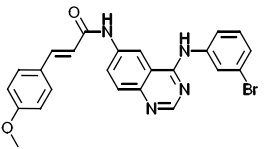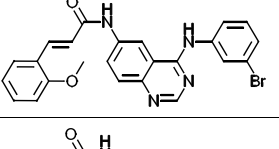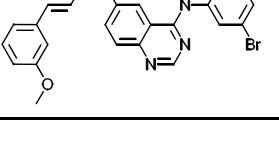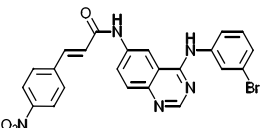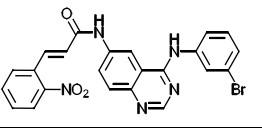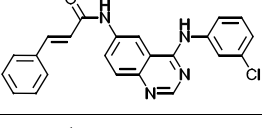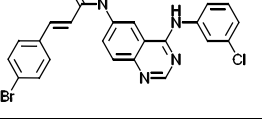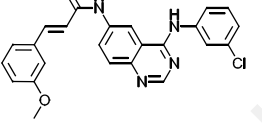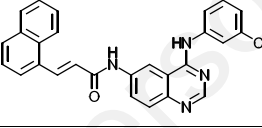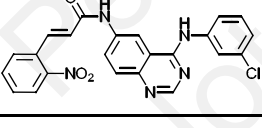
No	Structure	F10[C-O]	SpMADB(i)	GATS8m	MATS4e	pIC$_{50}$[a]		LOO[d]
						Obsd.	Cald. eq. 1	
37[c]		4	1.52	1.43	-0.13	-	-	-
38[c]		4	1.54	1.06	-0.16	-	-	-
39[b]		2	1.52	1.12	-0.10	4.91	4.96	-
40[b]		2	1.52	1.07	-0.08	4.86	5.03	-
41		2	1.58	1.31	-0.04	5.41	5.07	5.04
42		2	1.58	1.09	-0.04	5.24	5.25	5.25
43[b]		2	1.57	1.28	-0.06	5.92	5.05	-
44		2	1.57	1.02	-0.07	5.37	5.24	5.24
45[b]		4	1.58	1.32	-0.06	5.14	5.45	-
46		4	1.58	1.36	-0.09	5.11	5.39	5.43
47[b]		4	1.58	1.34	-0.03	5.22	5.46	-

Table 1. contd…

No	Structure	F10[C-O]	SpMADB(i)	GATS8m	MATS4e	pIC$_{50}$[a]		
						Obsd.	Cald. eq. 1	LOO[d]
48		4	1.59	1.37	-0.12	5.29	5.42	5.45
49		6	1.59	1.37	-0.11	5.99	5.83	5.80
50		2	1.60	1.23	0.01	5.44	5.25	5.23
51		2	1.57	0.70	-0.03	5.46	5.53	5.54
52		4	1.58	1.18	-0.02	5.18	5.59	5.62
53[b]		2	1.62	1.24	0.01	5.55	5.31	-
54		6	1.59	1.22	-0.10	6.03	5.96	5.94

[a]All observed data were taken from refs. [3-5], IC$_{50}$ values were in μM concentration. [b] Test set compounds. [c]Outliers. [d]Leave-one-out (Jackknife procedure)

Table 2.   Some Proposed Compounds Belonging to the Series of Table 1 and their Predicted Activity.
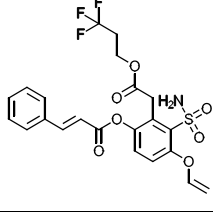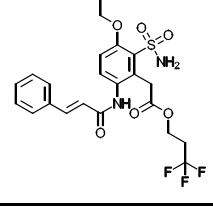
No.	Structure	F10[C-O]	SpMAD_B(i)	GATS8m	MATS4e	Pred Act. eq.1
1		11	1.449	0.954	0.08	6.79
2		11	1.449	0.952	0.125	6.83
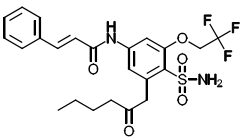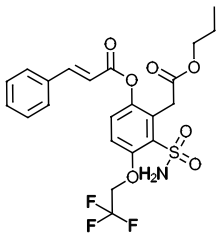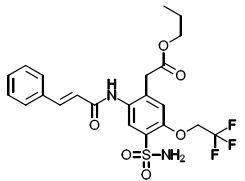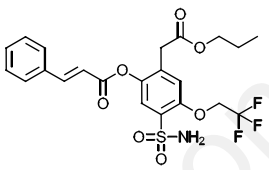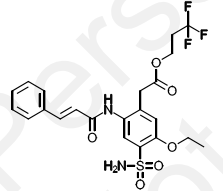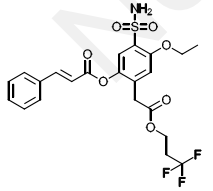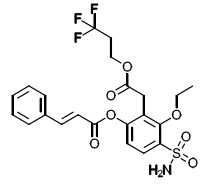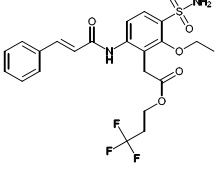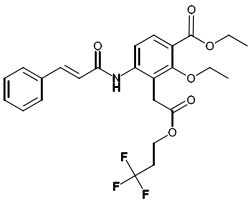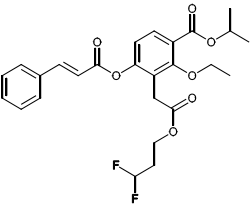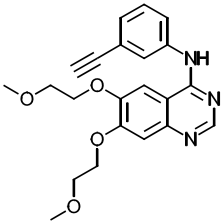
No.	Structure	F10[C-O]	SpMAD_B(i)	GATS8m	MATS4e	Pred Act. eq.1
3		11	1.462	0.899	-0.113	6.72
4		11	1.462	0.867	-0.144	6.72
5		13	1.462	0.823	-0.085	7.21
6		13	1.462	0.792	-0.12	7.20
7		13	1.449	0.872	0.173	7.34
8		13	1.449	0.874	0.122	7.29
9		11	1.453	0.893	0.167	6.93
10		11	1.454	0.891	0.187	6.95

Table 2. contd…

No.	Structure	F10[C-O]	SpMAD_B(i)	GATS8m	MATS4e	Pred Act. eq.1
11		12	1.456	0.932	0.2	7.14
12		13	1.451	0.929	0.13	7.27
13	Erlotinib	8	1.53	1.14	0.04	6.47[a]

[a]As reported in ref [3].

Table 3. **Data Related to Lipinski Rules and Activity Values of Predicted Compounds.**

Compd	HD	HA	MW	(AlogP)[a]	pIC$_{50}$
1	2	8	499.457	3.112	6.79
2	3	8	500.488	2.729	6.83
3	3	8	500.488	3.541	6.72
4	2	8	501.473	4.188	6.72
5	3	8	500.488	3.541	7.21
6	2	8	501.473	4.188	7.20
7	3	8	500.488	2.729	7.34
8	2	8	501.473	3.376	7.29
9	2	8	501.473	3.376	6.93
10	3	8	500.488	2.729	6.95
11	1	7	493.472	4.228	7.14
12	0	7	490.493	5.302	7.27

[a]Atom based calculated logP.

Table 4.   Docking Results of Predicated Molecules with Reference to FDA Approved Molecule.

Compd. No.	No. of Hydrogen Bonds	H-bonds	H-bonds Length (Å)	Score
1	2	O(28)-Asp831 N(30)-Lys721	-4.70 -8.30	-8.0881
2	3	O(28)-Asp831 N(30)-Lys721 O(25)-Lys721	-2.68 -8.30 -3.46	-13.4961
3	3	N(30)-Lys721 O(20)-Lys721 O(29)-Asp831	-8.30 -2.76 -4.70	-12.9483
4	3	N(30)-Lys721 O(20)-Lys721 O(29)-Asp831	-8.30 -4.21 -4.70	-9.4459
5	3	O(21)-Met769 H(42)-Met769 O(32)-Cys773	-4.18 -4.70 -4.70	-10.8692
6	3	O(33)-Asp831 N(34)-Lys721 O(17)-Lys721	-4.70 -8.30 -4.70	-11.5870
7	3	O(30)-Asp831 H(42)-Asp831 N(31)-Lys721	-4.12 -2.81 -8.30	-14.2482
8	3	O(17)-Lys721 N(31)-Lys721 O(30)-Asp831	-4.70 -8.30 -4.70	-12.9263
9	3	O(17)-Cys773 O(33)-Met769 H(56)-Thr766	-3.07 -4.70 -4.70	-8.3986
10	3	O(21)-Cys773 H(42)-Met769 H(57)-Thr830	-3.61 -4.70 -4.03	-8.2246
11	2	H(43)-Asp776 O(17)-Cys773	-4.41 -4.52	-6.5071
12	1	O(34)-Lys721	-4.70	-7.2109
Erlotinib	3	O(8)-Cys773 O(11)-Cys773 N(1)-Met769	-2.58 -3.00	-16.247
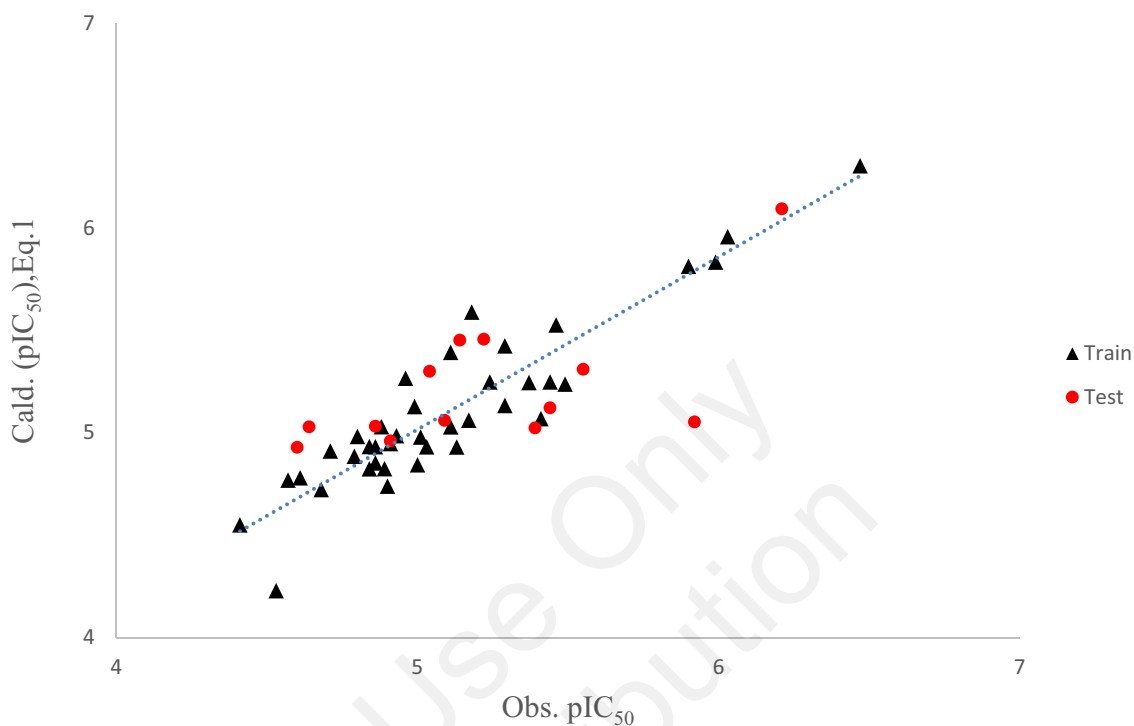
Fig. (1). A graph between the calculated and observed activities of compounds of Table **1**.
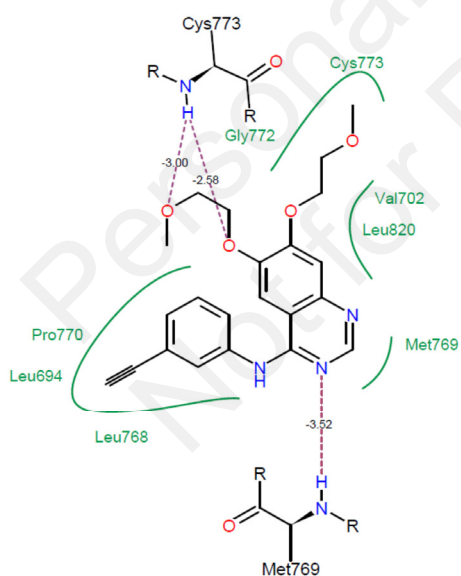
Fig. (2). A docked structure of Erlotinib in EGFR (PDB entry code 1M17).

Fig. (3). A docked structure of predicted compound 7 (Table **3**) in EGFR (PDB entry code 1M17).

CONCLUSION

All the cinnamic acid analogues studied as EGFR inhibitors are found to interact with the receptors involving electronic interactions. Their docking study shows that they also form hydrogen bonds and perform steric interactions with certain deep pockets available in the enzyme. Similar conclusions were also drawn regarding the binding of different categories of chemicals with EGFR, when QSAR and molecular modeling studies were performed on them by differ-

ent authors [18-20]. Electronic interactions and hydrogen bondings were found to be important in many other cases also such as in H^+/K^+-ATPase inhibition, recently studied by Agarwal *et al*. [21]. In the present case, the enzyme is also found to possess several deep pockets in which different portions of inhibitors are engulfed, *e.g.*, in Figure **2** a deep pocket of the enzyme formed by Pro770, Leu 694, and Leu768 residues engulfs Erlotinib's aryl ring with acetylenic moiety. Similarly in Figure **3,** which shows the interactions

of compound 7, one of the predicted compounds, a pocket formed by Asp831 and Phe699 engulfs a large portion of the pendent substituent at position 1 of the central ring and a pocket formed by Leu768, Leu694, Met769, and Leu820 engulfs the fluoro-substituted alkyl moiety of side chain. These are some of the major steric interactions shown in both the figures. Besides, there are some minor steric interactions which contribute to the activity of the compounds. It has been shown that compound 7 has better interactions with the enzyme than Erlotinib, and that is why its activity is higher than that of Erlotinib. It may be suggested that bulky compounds with better flexibility may possess good EGFR inhibition potency.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1]   Leslie, B. J.; Holaday, C. R.; Nguyen, T.; Hergenrother, P. J. Phenylcinnamides as novel antimitotic agents. *J. Med. Chem.,* **2010**, *53*, 3964–3972.

[2]   Cutillo, F.; D'Abrosca, B.; DellaGreca. M.; DiMarino, C.; Golino, A.; Previtera. L.; Zarrelli, A. Cinnamic acid amides from Chenopodium album: effects on seeds germination and plant growth. *Phyto. Chem.,* **2003**, *64*, 1381–1387.

[3]   Zhang, M.; Lu. X.; Zhang, H.-J.; Li, N.; Xiao, Y.; Zhu, H.-L.; Ye, Y.-H. Synthesis, structure, and biological assay of cinnamic amides as potential EGFR kinase inhibitors. *Med. Chem. Res.,* **2013**, *22*, 986–994.

[4]   Qian, Y.; Zhang, H.-J.; Hao, Zhang.; Xu, C.; Zhao, J.; Zhu, H.-L. Synthesis, molecular modeling, and biological evaluation of cinnamic acid metronidazole ester derivatives as novel anticancer agents. *Bioorg. Med. Chem.,* **2010**, *18*, 4991–4996.

[5]   Li, D.-D.; Lv, P.-C.; Zhang, H.; Zhang, H.-J.; Hou, Y.-P.; Liu. K.; Ye, Y.-H.; Zhu, H.-L. The combination of 4-anilinoquinazoline and cinnamic acid: A novel mode of binding to the epidermal growth factor receptor tyrosine kinase. *Bioorg. Med. Chem.,* **2011**, *19*, 5012–5022.

[6]   http://www.chemdoodle.com

[7]   Talete srl, Dragon (Software for Molecular Descriptor Calculation) Version 6.0 - 2013 - http://www.talete.mi.it/.

[8]   Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR Models. *J. Comput. Chem.,* **2013**, *34*, 2121-2132.

[9]   Gramatica, P.; Cassani, S.; Chirico, N. QSARINS-Chem: Insubria datasets and new QSAR/QSPR models for environmental pollutants in QSARINS. *J. Comput. Chem.,* **2014***, 35,* 1036-1044.

[10]   Shi, L. M.; Fang, H.; Tong, W.; Wu, J.; Perkins, R.; Blair, R. M.; Branham, W. S.; Dial, S. L.; Moland, C. L.; Sheehan, D. M. QSAR models using a large diverse set of estrogens. *J. Chem. Inf. Comput. Sci.,* **2001**, *41*, 186−195.

[11]   Schüürmann, G.; Ebert, R.; Chen, J.; Wang, B.; Kühne, R. External validation and prediction employing the predictive squared correlation coefficients test set activity mean vs training set activity mean. *J. Chem. Inf. Model.,* **2008**, *48*, 2140−2145.

[12]   Consonni, V.; Ballabio, D.; Todeschini, R. Comments on the definition of the Q^2 parameter for QSAR validation. *J. Chem. Inf. Model.,* **2009**, *49*,1669−1678.

[13]   Consonni, V, Ballabio, D.; Todeschini R. Evaluation of model predictive ability by external validation techniques. *J. Chemom.,* **2010**, *24*, 194−201.

[14]   Ojha, P. K.; Mitra, I.; Das, R. N.; Roy, K. Further exploring r^2_m metrics for validation of QSPR models. *Chemom. Intell. Lab. Syst.,* **2011**, *107*, 194−205.

[15]   Ojha, P. K.; Roy, K. Comparative QSARs for antimalarial endochins: Importance of descriptor-thinning and noise reduction prior to feature selection. *Chemom. Intell. Lab. Syst.,* **2011**, *109*, 146−161.

[16]   Chirico, N.; Gramatica, P. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *J. Chem. Inf. Model.,* **2011**, *51*, 2320−2335.

[17]   Lin, L. I. A concordance correlation coefficient to evaluate reproducibility. *Biometrics,* **1989**, *45*, 255−268.

[18]   Deeb, O.; Clare, B. W. QSAR of aromatic substances: EGFR inhibitory activity of quinazoline analogues. *J. Enzyme Inhib. Med. Chem.,* **2008**, *23*,763-75

[19]   Sun, M.; Chen, J.; Cai, J.; Cao, M.; Yin, S.; Min, Ji. Simultaneously optimized support vector regression combined with genetic algorithm for QSAR analysis of KDR/VEGFR-2 inhibitors. *Chem. Biol. Drug Des.,* **2010**, *75*, 494–505.

[20]   Zhang, W.-M.; Xing, M.; Zhao, T.- T.; Ren, Y.-J.; Yang, X.-H.; Yang, Y.-S.; Lv, P.-C.; Zhu, H.-L. Synthesis, molecular modeling and biological evaluation of cinnamic acid derivatives with pyrazole moieties as novel anticancer agents. *R.S.C. Adv.,* **2014**, *4*, 37197-37207.

[21]   Agarwal, N.; Bajpai, A.; Srivastava, V.; Gupta, S. P. A quantitative structure-activity relationship and molecular modeling study on a series of biaryl imidazole derivatives acting as H+/K+-ATPase inhibitors. *Struct. Biol*., **2013**, *2013,* 11.