

---

**2D, 3D QSAR AND MOLECULAR DOCKING STUDIES OF  
PYRAZOLYL-THIAZOLINONE DERIVATIVES AS EGFR  
INHIBITORS**

**BARAVKAR A<sup>\*1</sup>, SHAIKH A<sup>2</sup>, GANDHI S<sup>3</sup>, SHINDE N<sup>4</sup> AND AHER N<sup>5</sup>**

**1:** Agricultural Development of Trust's Shardabai Pawar Institute of Pharmaceutical Sciences  
and Research, Baramati, Maharashtra

**2:** Allana College of Pharmacy, Pune, Maharashtra

**3:** AISSMS College of Pharmacy, Pune, Maharashtra

**4:** Agricultural Development of Trust's, Shardabai Pawar Mahila Science College, Baramati,  
Maharashtra

**5:** Ashvin College of Pharmacy, Manchihill, Sangamner, Maharashtra

**\*Corresponding Author: Atul Baravkar: E Mail: [atul200678@gmail.com](mailto:atul200678@gmail.com)**

Received 19<sup>th</sup> April 2021; Revised 20<sup>th</sup> June 2021; Accepted 29<sup>th</sup> July 2021; Available online 1<sup>st</sup> Oct. 2021

<https://doi.org/10.31032/IJBPAS/2021/10.10.1016>

**ABSTRACT**

Epidermal growth factor (EGF) has been found to be expressed and altered in a variety of malignancies and it plays a significant role in tumor development and progression, including cell proliferation, regulation of apoptotic cell death, angiogenesis and metastatic spread. A series of para-substituted pyrazolyl-thiazolinone derivatives were chosen for molecular modelling studies to establish quantitative relationship between structure and biological activities of the compounds using software MDS version 4.2 (VlifeScience). Based on these studies new molecules were designed. Molecular docking study was performed to check the receptor and ligand interaction or binding affinity by using the receptor 1M17. It was found that there was only one sigma- $\pi$  bond in the binding pocket. The benzene ring of erlotinib formed one sigma- $\pi$  bond with PHE699A (distance: 3.97 Å<sup>0</sup>). The model between compound MCR 004 and the ATP binding site was similar to that with erlotinib. Comparing these models, it was found that the hydrophobic pockets of ATP binding were all nicely occupied

by these compounds, and were nicely occupied by these compounds, and the difference was the combination mode. This molecular docking result, along with the biological assay data, suggesting that the compound MCR 004 was a potential inhibitor of EGFR. MCR 004 was found to be best and was comparable with the standard drug erlotinib.

**Keywords:** Epidermal Growth Factor, Receptor Tyrosine Kinase, Pyrazolyl-thiazolinone derivatives, Erlotinib

## INTRODUCTION

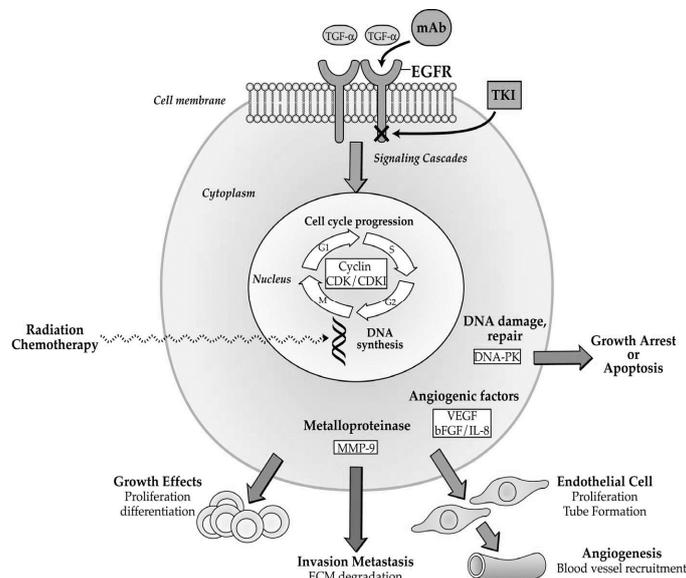
The epidermal growth factor (EGFR) belonging to the family of Receptor tyrosine kinase (RTKs) also called as phosphotyrosine kinase [1]. Receptor tyrosine kinase are single pass transmembrane protein important in intercellular signaling, by translating extracellular signals (ligands or growth factors) into activation of specific cell signaling cascades i.e. they plays an important role in signal transduction pathways that regulate cell division and differentiation [2] (**Figure 1**). The receptor tyrosine kinase family comprises of four members: EGFR (HER-1/ErbB-1), HER-2 (ErbB-2), HER-3 (ErbB-3) and HER-4 (ErbB-4) [3].

The Tyrosine kinase receptors are known to be activated by binding to different ligands (including EGF, TGFA) after a ligand binds to the extracellular domain of the receptor, the receptor forms functionally active dimers [EGFR-EGFR (homodimer) or EGFR-HER2, EGFR-HER3, EGFR-HER4 (heterodimer)] [4]. Dimerization induces the activation of the tyrosine kinase

domain, which leads to auto-phosphorylation of the receptor on multiple tyrosine residues. This leads to activation a series of intracellular signaling cascades to affect gene transcription, which in turn results in cancer cell proliferation, reduced apoptosis, invasion and metastasis and also stimulates tumor-induced angiogenesis [5]. EGF has been found to be expressed and altered in a variety of malignancies and clearly it plays a significant role in tumor development and progression, including cell proliferation, regulation of apoptotic cell death, angiogenesis and metastatic spread [6]. EGF tyrosine kinase-mediate cell growth signaling pathways plays an important role in the formation and development of many types of solid tumors, including head and neck, lung, breast, bladder, prostate and kidney cancers [7]. Therefore, EGF tyrosine kinase represents an attractive target for the development of novel anticancer agents. EGFR and HER-2 are the hottest targets in current cancer research and their over expression or abnormal activation often

cause cell malignant transformation. Gefitinib and Erlotinib are the representative drugs for this kind of

inhibitors and have been approved by US FDA for the treatment of patient with non-small-cell-lung cancer (NSCLC) [8].



### Figure 1: Growth Factors (Egfr) Signaling For Cell Growth And Division Inhibition of EGFR Signaling in Cancer Therapy:

Thiazolinone and their derivative have attracted continuing interest over the years because of their varied biological activities, such as anti-inflammatory, antimicrobial, antiproliferative, antiviral, anticonvulsant, antifungal and antibacterial [9]. Recent years thiazolinone derivative with their antitumor activity have become popular. Havrylyuk et al. reported that thiazolinones containing benzothiazole moiety has anticancer activity on leukemia, lung, colon, CNS, ovarian, renal, and prostate and breast cancers cell lines [10].

Many pyrazole derivatives are acknowledged to possess a wide range of bioactivities. The pyrazole motif makes up the core structure of numerous biologically

active like antiviral/antitumor, antibacterial, anti-inflammatory analgesic, fungistatic and antihypoglycemic activity [11]. Pyrazole derivative containing thiourea skeleton were recently reported as potent anticancer agents targeting EGFR. The pyrazole ring along with the thiazolinone ring the two combined substructures might exhibit synergistic anticancer effect. By integrating these two heterocyclic rings we can synthesize some novel pyrazolyl-thiazolinone derivative as a potential EGFR inhibitory agent.

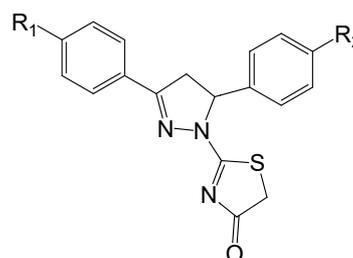


Figure 2: Chemical Structures Of 2-Substituted Pyrazolyl-Thiazolinone Derivative

Thus, a series of para-substituted pyrazolyl-thiazolinone derivatives were chosen for molecular modelling studies to establish quantitative structure relationship (QSAR) between structure and biological activities of the compounds using software MDS version 4.2 (VLifeScience). Also docking studies were carried out on newly designed molecules.

### METHODOLOGY

All computational experiments were performed on computer having genuine Intel Pentium Dual Core Processor and Windows XP operating system using the VLifeMDS software package of VLife Technologies Pvt. Ltd. Pune (www.vlifesciences.com) (version 4.3). The molecular structures were built in

VLifeMDS by 2D drawing builder and converted to 3D structures, and saved in .mol and .mds file format respectively and subjected to an optimization or energy minimization using MMFF method.

A dataset reported by Qiu *et al* (2012) [12] for a series of pyrazolyl-thiazolinone derivatives was utilized for 2D-QSAR and 3D-QSAR study for their EGFR inhibitory activity. EGFR inhibitory activities were converted into the corresponding pKi values to get the linear relationship in equation using the formula:  $pKi = -\log Ki$ , where Ki value represent inhibition activity against EGFR and HER-2 kinase in  $IC_{50}$  ( $\mu M$ ). The structure and biological activity of the selected series is shown in **Table 1**.

Table1:  $IC_{50}$  for EGFR Inhibitors

Compound	R <sub>1</sub>	R <sub>2</sub>	EGFR Activity	
			$IC_{50}(\mu g/ml)$	$pIC_{50}^*$
E1	H	F	4.86	-0.687
E2	H	Cl	3.49	-0.543
E3	H	Br	1.35	-0.130
E4	H	OMe	4.27	-0.63
E5	F	H	8.14	-0.911
E6	F	F	16.92	-1.228
E7	F	Cl	10.92	-1.038
E8	F	Br	4.79	-0.68
E9	F	Me	8.36	-0.922
E10	Cl	H	5.34	-0.728
E11	Cl	F	14.21	-1.153
E12	Cl	Cl	8.16	-0.912
E13	Cl	Me	6.67	-0.824
E14	Br	H	3.20	-0.505
E15	Br	F	6.48	-0.812
E16	Br	Cl	4.12	-0.615
E17	Br	Br	2.03	-0.307
E18	Br	Me	5.58	-0.747
E19	Me	H	1.08	-0.033
E20	Me	F	2.01	-0.303
E21	Me	Cl	1.66	-0.22
E22	Me	Br	0.24	0.62

\* $pIC_{50} = -\log IC_{50}$

**2D-QSAR:** It was studied on set of 23 molecules; selection of training and test set carried out using random data selection method. For the building of models partial least squares regression model building method using stepwise forward-backward variable selection method was used.

**3D-QSAR:** It was studied on same set of 23 molecules. Unsubstituted pyrazoly-thiazolinone was used as template for alignment of molecules using template alignment method. Selection of training and test set carried out using random data selection method. For the building of 3D-QSAR models k-NN MFA approach<sup>12</sup> using stepwise forward-backward variable selection method were used. Structure and biological activities of pyrazoly-thiazolinone derivatives are shown.

### **Molecular Docking studies**

**Software package:** Molecular docking studies were carried on newly designed molecules. For the present study, the software package QSARpro and MDS which were procured from VLife sciences Pvt. Ltd. Pune was used.

**Preparation of ligands and structure conformation generation:** Structures of compounds were sketched using the 2D structure draw application Vlife2Ddraw and converted to 3D structures. All the structures were minimized and optimized with the Merck Molecular Force Field

(MMFF) method taking the root mean square gradient (RMSD) and the iteration limit to 10,000. Conformers for each structure were generated using Monte Carlo by applying MMFF force field method and least energy conformer was selected for further study.

**Preparation of protein:** The PDB structure [1M17] [13] was downloaded and energy minimization of the protein complex was carried out. All the bound water molecules, ligands and cofactors were removed from the protein which were taken in pdb format. The tool neutralized the side chains that were not close to the binding cavity and did not participate in salt bridges. This step was then followed by restrained minimization of co-crystallized complex which reoriented side-chain hydroxyl groups and alleviated potential steric clashes. The complex obtained was minimized using Merck molecular force field. The minimization was terminated after either completion of 5,000 steps or after the energy gradient converged below 0.05 kcal/mol.

**Docking Methodology:** Docking study was performed on VlifeMDS version 4.3 on Lenovo computer, i3 processor with XP operating system. The GRIP-based ligand docking with genetic algorithm approximated a systematic search of positions, orientations, and conformations

of the ligand in the enzyme binding pocket via a series of hierarchical filters.

## RESULTS

**2D QSAR:** In 2D QSAR analysis, multiple linear regression analysis (MLR) coupled with stepwise forward backward variable selection method was applied to generate 2D models. Selection of training and test set was done by Random selection method (Test set: E5, E6, E7, E8, E9, E10). From these models, the one having good  $q^2$  and  $\text{pred}_r^2$  values was selected as best model. Statistically significant 2D QSAR models are shown in **Table 2**.

The selection of the best model is based on the values of  $r^2$  (squared correlation coefficient),  $q^2$  (cross-validated correlation coefficient),  $\text{pred}_r^2$  (predicted correlation coefficient for the external test set), F (Fisher ratio) reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F-test indicate that the model is statistically significant.  $r^2_{se}$ ,  $q^2_{se}$  and  $\text{pred}_r^2_{se}$  are the standard errors terms for  $r^2$ ,  $q^2$  and  $\text{pred}_r^2$  respectively. The statistically significant 2D-QSAR model is shown as follows.

### Model-1

$$\text{pIC}_{50} = 0.1070 (\text{T}_2\text{O}_7) - 0.0081 (\text{Quadrupole } 1) - 0.0001(\text{Mom Inertia X}) - 0.01664 (\text{T}_2\text{Cl}_7) + 9.9832$$

### Interpretation of the Model-1

From equation, model 1 explains 98.52% ( $r^2 = 0.9852$ ) of the total variance in the training set as well as it has internal ( $q^2$ ) and external ( $\text{pred}_r^2$ ) predictive ability of 75.23 % and 84.59 % respectively. The F test shows the statistical significance of 99.99 % of the model which means that probability of failure of the model is 1 in 10000. In addition, the randomization test shows confidence of 99.9999 (Alpha R and  $R^2 = 0.00000$ ) that the generated model is not random and hence chosen as the QSAR model. From QSAR model 1, positive coefficient value of  $\text{T}_2\text{O}_7$  [count of number of double bonded atoms (i.e. any double bonded atom,  $\text{T}_2$ ) separated from oxygen atom by seven bonds] on the biological activity indicated that higher values leads to good inhibitory activity while lower value leads to reduced inhibitory activity while negative coefficient value of  $\text{T}_2\text{Cl}_7$  [count of number of double bonded atoms (i.e. any double bonded atom,  $\text{T}_2$ ) separated from chlorine atom by seven bonds] on the inhibitory activity indicated that lower value leads to better inhibitory activity whereas higher value leads to decrease inhibitory activity. Mom Inertia X [the descriptor signifies moment of inertia at X axis] lower value or Zero value of Mom Inertia does not affect the activity. Negative coefficient value of quadrupole

1[Magnitude of first tensor of quadrupole moment] the inhibitory activity indicated that lower values lead to good inhibitory activity whereas higher value leads to decreased inhibitory activity.

Data fitness plot for model 1 is shown in **Figure 3**. The plot of observed vs predicted activity provides an idea about how well the model was trained and how well it predicts the activity of external test set.

The graph of observed vs predicted activity of training and test sets for model 1 is shown in **Figure 3** it reveals that the model is able to predict the activity of training set quite well as well as external test set, providing confidence of model. Result of the observed vs predicted inhibitory activity for the training and test compounds for the model 1 is shown in **Table 4**.

**3D-QSAR:** kNN-MFA samples the steric and electrostatic fields surrounding a set of ligands and constructs 3D-QSAR models by correlating these 3D fields with the corresponding biological activities. Molecular alignment was used to visualize the structural diversity in the given set of molecules.

The template structure i.e. para substituted pyrazolyl-thiazolinone derivatives was used for alignment by considering the common elements of the series as shown in **Figure 4**. A **Table 3** shows statistically significant

3D-QSAR models. (Test set- 12, 13, 14, 16, 17, 18)

The selection of the best model is based on the values of  $q^2$  (internal predictive ability of the model) and that of  $\text{pred}_r^2$  (the ability of the model to predict the activity of external test set). The statistically significant 3D-QSAR model for  $\text{pIC}_{50}$  (model-2) is given below.

#### Model-2

$$\text{pIC}_{50} = E_{935} (-2.2072-2.1229) + S_{807} (5.4542-5.5922)$$

#### Interpretation of the Model-2

The model 2 explains values of  $k$  (2),  $q^2$  (0.7870),  $\text{pred}_r^2$  (0.9194),  $q^2_{se}$  (0.1948), and  $\text{pred}_r^2 se$  (0.1997) prove that QSAR equation so obtained is statistically significant and shows the predictive power of the model is 96.03% (internal validation). **Table 4** represents the predicted inhibitory activity by the model-2 for training and test set.

The data fitness plot for model 2 is shown in **Figure 5**. The plot of observed vs predicted activity provides an idea about how well the model was trained and how well it predicts the activity of the external test set.

Result plot in which 3D-alignment of molecules with the important steric and electrostatic points contributing in the model-2 with ranges of values shown in the parenthesis represented in **Figure 6**.

It shows the relative position and ranges of the corresponding important steric and electrostatic fields in the model provides guideline for new molecule design as follows-

1. Electrostatic field, E<sub>38</sub> (0.324, 0.024) has positive range indicates that positive electrostatic potential is favorable for increase in the activity and hence less electronegative substituent group is preferred in that region.
2. Steric field, S<sub>447</sub> (-0.032, -0.032) has negative range indicates that negative steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.

### Design and activity prediction of newer derivatives

From the best models obtained, some newer compounds were designed claiming to possess better activity than the reported one. The structures were not reported earlier anywhere is confirmed by chemSpider.

**Molecular Docking studies:** To gain better understanding on the potency of the newly designed compounds (MCR 001-004) we proceeded to examine the interaction of it with EGFR (PDB code: 1M17) by molecular docking, which was performed by simulation of the compound into the ATP binding site in EGFR. The binding model of the compounds with EGFR is depicted in the **Figure 7**. The amino acid residues which had interaction with EGFR were labelled. In the binding model,

compound MCR 004 was nicely bound to the EGFR kinase with its carbonyl –group of A721 (LYS 721), forming a more optimal hydrogen bond interaction (distance: N-H...O = 2.58 Å<sup>0</sup>). Based on the favorable EGFR inhibitory activity of pyrazolyl-thiazolinone derivatives, it could be concluded that this H-bond played an important effect in the EGFR inhibitory. The end amino cation of PHE699A was also formed a  $\pi$ -cation interaction with the benzene ring of the compound MCR 004 (distance: 5.02 Å<sup>0</sup>), which enhanced the binding action between EGFR and the ligand compound MCR 004. The electron donating substituent amino-NH<sub>2</sub> strengthened the  $\pi$ -cation binding and the results indicated that it was primarily due to direct through-space interaction between the substituent and the cation.

Meanwhile, erlotinib was taken as a positive control in the docking procedure. It was found that there was only one sigma- $\pi$  bond in the binding pocket. The benzene ring of erlotinib formed one sigma- $\pi$  bond with PHE699A (distance: 3.97 Å<sup>0</sup>). The model between compound MCR 004 and the ATP binding site was similar to that with erlotinib. Comparing these models, it was found that the hydrophobic pockets of ATP binding were all nicely occupied by these compounds, and were nicely occupied by these compounds, and the difference

was the combination mode. This molecular docking result, along with the biological assay data, suggesting that the compound MCR 004 was a potential inhibitor of EGFR.

3D model shows interaction of Erlotinib, MCR 004 with their respective ATP binding site. The H-bond (green lines) is displayed as a dotted line, and the  $\pi$ -cation interaction is shown as yellow lines and the

hydrophobic interaction is shown as dotted blue lines (Figure 7).

2D model shows interaction between compound Erlotinib, MCR 004 with their ATP binding site. The H-bond (dark blue arrows) is displayed as dotted arrows & the  $\pi$ -cation interaction is shown as pink lines and hydrophobic bond is shown as faint blue lines.

Table 2: Statistical Evaluation Of 2d-Qsar Models Of Para-Substituted Pyrazolyl-Thiazolinone Derivatives (For Egfr Activity)

Trials	$r^2$	$q^2$	$r^2_{se}$	$q^2_{se}$	Pred_ $r^2$	Pred_ $r^2_{se}$	F test
(Model-1)	0.9852	0.7523	0.1715	0.2133	0.8459	0.2251	83.32
2	0.8350	0.7400	0.1084	0.2595	0.2380	0.5573	25.30
3	0.7816	0.6795	0.1951	0.2708	0.3208	0.3950	39.36
4	0.7392	0.6675	0.2339	0.6675	0.3605	0.3892	31.17

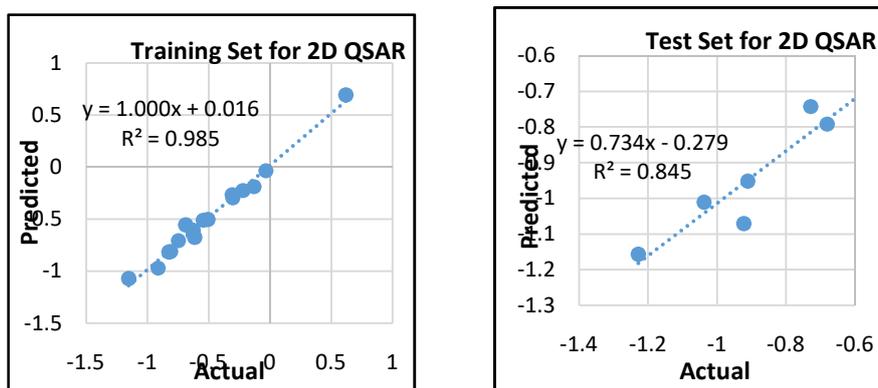


Figure 3: Data Fitness Plot For 2d Qsar (Model-1)

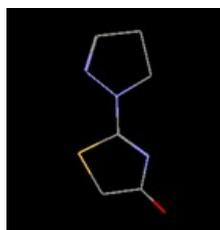


Figure 4: Para Substituted Pyrazolyl-Thiazolinone Derivatives (Template Structure)

Table 3: Statistical Evaluation Of 3d-Qsar Models Of Para-Substituted Pyrazolyl-Thiazolinone Derivatives (Egfr Activity)

Trials	kNN	DOF	$q^2$	$q^2_{se}$	pred $r^2$	pred $r^2_{se}$
1	2	21	0.7052	0.1831	0.4485	0.3547
(Model-2)	2	21	0.7870	0.1948	0.9194	0.1997
3	2	21	0.7044	0.1857	0.4104	0.3660
4	2	21	0.6293	0.2422	0.2580	0.2939

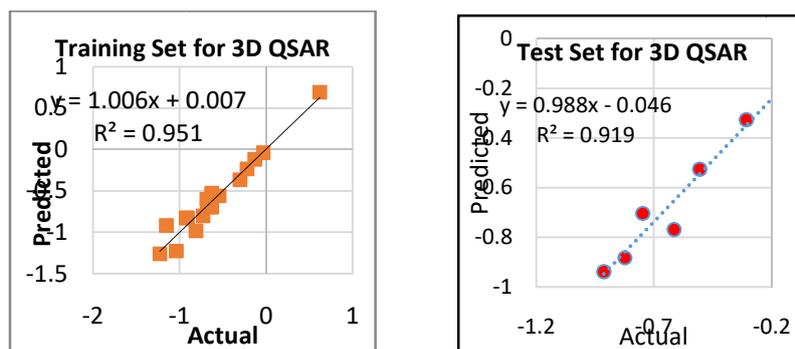


Figure 5: Data fitness plot for 3D QSAR (Model-2)

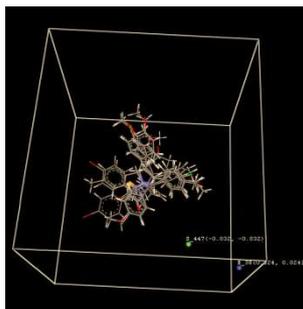


Figure 6: Plot of contribution chart for 3D QSAR (Model-2)

Table 4: Actual And Predicted Activities For 23 Compounds Based On The Best 2d/3d-Qsar Models Of Para-Substituted Pyrazolyl-Thiazolinone Derivatives

Compound code	Actual EGFR Activity $pIC_{50}$	2D QSAR Predicted EGFR Activity (Model 1) $pIC_{50}$	3D QSAR Predicted EGFR Activity (Model 2) $pIC_{50}$
E1	-0.687	-0.556	-0.597
E2	-0.543	-0.511	-0.559
E3	-0.130	-0.188	-0.117
E4	-0.63	-0.615	-0.524
E5	-0.911	-0.952	-0.828
E6	-1.228	-1.157	-1.26
E7	-1.038	-1.011	-1.227
E8	-0.68	-0.791	-0.693
E9	-0.922	-1.071	-0.824
E10	-0.728	-0.742	-0.799
E11	-1.153	-1.071	-0.917
E12	-0.912	-0.971	-0.939
E13	-0.824	-0.816	-0.882
E14	-0.505	-0.503	-0.526
E15	-0.812	-0.811	-0.982
E16	-0.615	-0.675	-0.769
E17	-0.307	-0.265	-0.327
E18	-0.747	-0.707	-0.703
E19	-0.033	-0.035	-0.036
E20	-0.303	-0.295	-0.364
E21	-0.22	-0.225	-0.234
E22	0.62	0.694	0.694

Table 5: Newly Designed Molecules Of Para-Substituted Pyrazolyl-Thiazolinone Derivatives

Compound	R <sub>1</sub>	R <sub>2</sub>	Predicted EGFR activity
MCR 001	NH <sub>2</sub>	OCH <sub>3</sub>	0.557
MCR 002	NO <sub>2</sub>	OH	0.173
MCR 003	NO <sub>2</sub>	F	0.195
MCR 004	NO <sub>2</sub>	OCH <sub>3</sub>	0.162



- stereoscopic vision. *Prog. Neurobiol*; 55: 191–224.
- [5] Fry GA, Kent PR (1944). The effects of base-in and base-out prisms on stereoacuity. *American Journal of Optometry and Archives of the American Academy of Optometry*; 21: 492–507
- [6] Trotter, Y. *et al* (1992). Modulation of neural stereoscopic processing in primate area V1 by the viewing distance. *Science*; 257: 1279–1281
- [7] Jiminez JR, Rubino M, Diaz SA, *et al* (2000). Changes in stereoscopic depth perception caused by decentration of spectacle lenses. *Optom Vis Sci*; 77: 421-7.
- [8] Kaye S.B, Siddiqui A, Ward A, *et al* (1999). Monocular and binocular depth discrimination threshold. *Optometry and Vision Science*; 76 (11), 770-782.
- [9] Fricke TR, Siderov J (1997). Stereopsis, stereotests, and their relation to vision screening and clinical practice. *Clinical and Experimental Optometry*; 80(5): 165-172.
- [10] Osuobeni E.P, Al- Musa K.A. (1993). Gender differences in interpupillary distance among Arabs. *Ophthalmology & Visual Science*; 70 (12): 1027- 1030.
- [11] Zaroff C.M, Knutelska M, Frumkes T.E. (2002). Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender. *Ophthalmology & Visual Science*; 44 (2).
- [12] Laframboise S, Deguise D, Faubert J (2006). Effect of aging on stereoscopic interocular correlation. *Optometry & Vision Science*; 83(8): 589-593.