

RESEARCH

Open Access



In-silico activity prediction, structure-based drug design, molecular docking and pharmacokinetic studies of selected quinazoline derivatives for their antiproliferative activity against triple negative breast cancer (MDA-MB231) cell line

Sagiru Hamza Abdullahi*, Adamu Uzairu, Gideon Adamu Shallangwa, Sani Uba and Abdullahi Bello Umar

Abstract

Background: Cancer is a major health threat especially in unindustrialized nations. It surpasses coronary diseases and takes the number one killer position as a result of different global wide influences. Among many breast cancer substrates, triple-negative breast cancer (TNBC) is particularly devastating because it rapidly metastasize to other parts of the body, with a high risk of earlier recession and mortality.

Result: In this research work, four (4) quantitative structure activity relationship (QSAR) models were developed using a series of quinazoline derivatives with activities against triple negative breast cancer cell line (MDA-MB231), model 1 was selected due to its statistical fitness with the following validation parameters: $R^2 = 0.875$, $Q^2 = 0.837$, $R^2 - Q^2 = 0.038$, $N_{\text{ext test set}} = 5$, and $R^2_{\text{ext}} = 0.655$. Molecular docking studies was performed for the quinazoline series as well as the reference drug (Gefitinib) and the active site of the epidermal growth factor receptor (EGFR) (pdb id = 3ug2). Eight compounds (6, 10, 13, 16, 17, 18, 19 and 20) were observed to have better docking score docking scores relative to Gefitinib. Compound number nineteen from the training set (pred $\text{pIC}_{50} = 5.67$, Residual = - 0.04 and MolDock score = - 123.238) was identified as the best compound since it has the best Moldock score and was excellently predicted by the selected model with least residual value, Hence was adopted as template for the design of Ten (10) new novel compounds with better activities and better docking scores. The inhibitive activities of the designed compounds were predicted by the selected model and most of them possess an improved activity relative to the template compound (19). The designed compounds were also redocked on to active pocket of the EGFR receptor and it was observed that they displayed better docking scores compared to the Template and the reference drug (Gefitinib) utilized in the design. Furthermore, the designed compounds were subjected to ADMET and drug-likeness studies using SWISSADME and pkCSM online web tools and they were observed to be pharmacologically active, easily synthesized and do not violate the Lipinski's rule of five.

Conclusion: Hence, the designed compounds can be employed as inhibitors of MDA-MB231 cell line after passing through in vivo and in vitro evaluation.

*Correspondence: sagirwasai@gmail.com
Department of Chemistry, Faculty of Physical Sciences, Ahmadu Bello
University, P.M.B.1045, Zaria, Kaduna State, Nigeria

Keywords: Density function theory, Quantitative structure activity relationship, Triple negative breast cancer, Molecular docking, Pharmacokinetic studies

Background

Cancer is a challenging problem for the global health community, and its increasing burden necessitates seeking novel and alternatives therapies (Rajabi et al. 2021). It takes the number one killer position as a result of different global wide influences. Although considerable progress were made in the chemotherapeutic remedy of some victims, the unrelenting obligation to the difficult task of detecting new anti-cancer drugs is still crucial. Breast cancer is the most predominant class of cancer diagnosed in females around the globe, with an incidence that intensifies vividly with age. Among many breast cancer substrates, triple-negative breast cancer (TNBC) is particularly devastating because it rapidly disperse to other parts of the body, with a high risk of earlier recession and mortality (Hu et al. 2012). Annually, at least one million females are identified with breast tumor and TNBC is accountable for close to 15–20% of the complete breast cancer identified (Jo et al. 2019).

The epidermal growth factor receptor (EGFR) plays a crucial part in the control of cell growth and is regarded as one of the most seriously evaluated tyrosine kinase's (TK) target inhibitors (El-Azab et al. 2010). Numerous TKs had a vital functions in cell propagation, division, metastasis and endurance, besides their uncontrolled triggering via processes such as point mutations leads to a substantial proportion of clinical cancers. EGFR is over expressed in numerous tumors, such as brain, lung, bladder, ovarian, colon, breast, head, and prostate tumors (Tiwari et al. 2015).

Components of the erbB class of EGFR-TKs, which comprise of erbB2 (HER2), erbB3 (HER3), and erbB4 (HER4), are overexpressed in a substantial ratio of human tumors, and this is attributed to the miserable prognosis of the malady (Chandregowda et al. 2009). Hence, inhibitors of erbB1 and erbB2 were acknowledged as possible anticancer drugs (Hynes and Lane 2005).

Extermination of cancer cells without causing damage on other normal tissues or cells is the main purpose of anti-cancer drugs. However, the fact that some of these drugs usually destroys some other normal cells and the resistance to these drugs experienced by some patients during early period of treatment necessitates the global search for identifying new higher quality drugs that are safe for the prevention and remedy of cancer (Al-Suwaitan et al. 2016). Immediate recognition, understanding of the cause and pathway of this disorder, and improvement in remedy have played a pivotal part in

curtailing breast cancer mortality rates over the past few years. Chemotherapy is still the central key to thorough therapy since it can exterminate tumor cells rapidly in the human system (Kaplan 2013).

In-silico approach of drug discovery have proven to accelerate the drug discovery process, as it lessen the time taken, resources and it enables the estimation of properties of new molecules such as toxicity and efficiency even before their synthesis. A mathematical relations that are able to establish a quantitative relationship between biological activities of a molecules and their molecular structures in form of linear equation is called Quantitative Structure–Activity Relationships (QSAR) (Abdullahi et al. 2021). Efficiency and safety of the drug to the system are the two major causes leading to drug failure. Therefore, it is compulsory to find potent molecules with better ADMET properties “drug-likeness” (Lawal et al. 2021).

This research is mainly purposed in utilizing QSAR approach to compute the inhibitory activities of a series of quinazoline derivatives against MDA-MB231 breast cancer cell line, perform molecular docking studies to understand the nature of interaction between the compounds and the EGFR protein receptor, design new potent compounds based on their docking scores and examine their ADMET and drug likeness properties.

Methods

Data sets retrieval

A series of 23 quinazoline derivatives with inhibitory activities (IC_{50} in $\mu\text{g/ml}$) against Triple Negative Breast cancer cell line MDA-MB231 are retrieved from Abuelizz et al. (2017). The inhibitory activities were linearized by taking their negative logarithm to base 10 as shown in Eq. 1.

$$pIC_{50} = -\log_{10}(IC_{50} \times 10^{-6}) \quad (1)$$

Chemical structure of the quinazoline analogs as well as their respective inhibitive capacity at 50% concentration (pIC_{50}) are presented in Table 1.

Calculation of molecular descriptors

2D structures of the quinazoline analogs were sketched by utilizing Chemdraw version 16.0 and they were transformed to 3D format using Spartan 14 software. Molecular mechanics force field were employed to clean the 3D structures to eliminate all strain from the structure of

Table 1 Chemical structure, experimental and predicted pIC₅₀ of Quinazoline derivatives

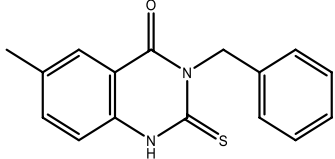
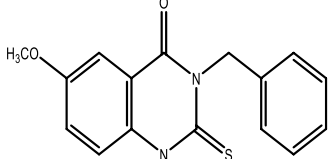
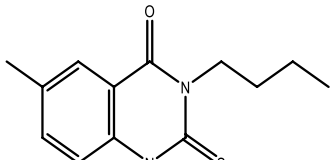
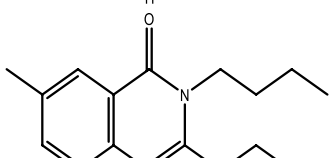
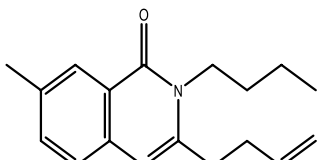
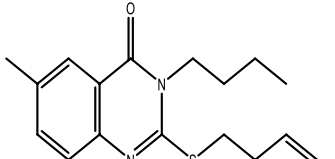
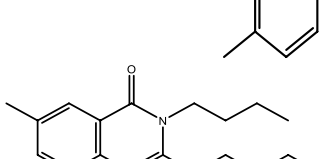
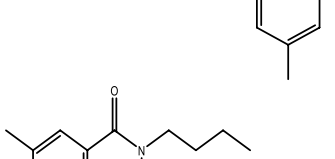
S/no	Structure	Exp Pic ₅₀	Pred pIC ₅₀	Residual	MolDock score
1 ^b		5.49	5.49	0.00	-77.1896
2 ^a		5.31	5.35	-0.04	-83.5156
3 ^a		5.56	5.57	-0.01	-69.7382
4 ^a		5.42	5.48	-0.06	-90.2325
5 ^b		5.42	5.43	-0.01	-97.2109
6 ^a		5.35	5.30	-0.05	-110.38
7 ^a		5.30	5.43	-0.13	-103.146
8 ^b		5.38	5.29	0.09	-99.0752

Table 1 (continued)

9 ^a		5.39	5.33	0.06	-101.162
10 ^a		5.37	5.28	0.09	-112.464
11 ^a		4.02	4.21	-0.19	-113.842
12 ^a		5.29	5.25	0.04	-106.971
13 ^a		5.29	4.98	0.31	-112.695
14 ^a		5.30	5.33	-0.03	-105.149
15 ^a		5.53	5.29	0.24	-102.592
16 ^a		5.30	5.40	-0.10	-114.001
17 ^b		5.49	5.32	0.17	-110.495

Table 1 (continued)

18 ^a		5.46	5.40	0.06	-119.287
19 ^a		5.63	5.67	-0.04	-123.238
20 ^a		5.59	5.34	0.25	-123.253
21 ^a		5.55	5.64	-0.09	-106.836
22 ^b		5.56	5.50	0.06	-77.2584
23 ^a		5.25	5.34	-0.09	-88.4194
Gefitinib		4.54			-108.022

^a Model building set^b External validation set

the molecule as well as guaranteeing a well-defined conformer relationship within the compounds (Viswanadhan et al. 1989). Density Functional Theory (DFT) quantum mechanical calculation was employed for the geometry optimization using B3LYP/631G^{*} basis set. The optimized structures were saved in Spatial Document File (SDF) format and then exported to PADEL descriptor calculation software to compute the molecular descriptors (Amin and Gayen 2016).

Data set partitioning

The data set were partitioned into two separate set: Modeling set (training set) and external validation set (test set). The modeling set consist of eighteen (18) compounds while the external validation set is made up of five (5) compounds. Models are built using the modeling set while the predictive ability of the built model

was ascertained using the external validation set (Tropsha et al. 2003). This splitting certifies that an analogous standard can be engaged to predict the activity of the test set. Kennard–Stone Algorithm was applied for dividing dataset into a modeling and test set (Kennard and Stone 1969).

QSAR model building and external validation

The most important aspect of QSAR studies is the designation and sampling of descriptors that offers an ultimate information in activity disparities and have minimal co-linearity. Hence, genetic function algorithm (GFA) progresses the model accurateness while selecting relevant molecular descriptors (Learidi 1996). Multi linear regression (MLR) was utilized on the model building set to express the mathematical relations between the depending variable A (pIC₅₀) and independent variable

B (molecular descriptors). An exceptional feature of GFA algorithm is that it is able to generate multiple models rather than single model. Validation parameters that provides a guide in selecting the best QSAR model include correlation coefficient (R^2), adjusted R^2 (R^2_{adj}), cross-validation coefficient (Q^2_{cv}) and correlation coefficient of the external validation set (R^2_{ext}), all are expressed in Eqs. (2, 3, 4 and 5) respectively.

$$R^2 = 1 - \frac{\sum(Y_{exp} - Y_{pred})^2}{\sum(Y_{exp} - Y_{mtraining})^2} \quad (2)$$

$$R^2_{adj} = 1 - \left(1 - R^2\right) \frac{N - 1}{N - P - 1} = \frac{(N - 1)R^2 - P}{N - P + 1} \quad (3)$$

$$Q^2_{cv} = 1 - \frac{\sum(Y_{pred} - Y_{exp})^2}{\sum(Y_{exp} - Y_{mtraining})^2} \quad (4)$$

$$R^2_{ext} = 1 - \frac{\sum(Y_{pred} - Y_{exp})^2}{\sum(Y_{exp} - Y_{mtraining})^2} \quad (5)$$

where P is the number of independent variables in the model and N is the sample size. Y_{exp} , Y_{pred} , and $Y_{mtraining}$ are the experimental activity, the predicted activity, and the mean experimental activity of the compounds in the modeling set, respectively (Tropsha et al. 2003). The least recommended values for these parameters are shown in Table 2.

Y-randomization test

In order to assess the robustness of the model and to affirm that the model was not obtained by chance correlation Y -randomization was performed on the model building set data (Tropsha et al. 2003). A new QSAR model was generated using the descriptor matrix by shuffling the activity matrix randomly. A built QSAR model is robust and reliable only when it has low values of R^2 and Q^2 for numerous trials. Another validation parameter is the coefficient of determination for Y -randomization cR^2_p , and it should exceed 0.5 for passing this test as in Eq. 6

$$cR^2_p = R \times [R^2 - R_r^2] \quad (6)$$

cR^2_p is coefficient of determination for Y -randomization, R is the coefficient of determination for Y -randomization and R_r is average 'R' of random models.

Molecular docking studies

Ligand-Protein molecular docking studies was performed on all the quinazoline derivatives to study the nature of interactions between active pocket of the EGFR protein receptor and the ligands on HP laptop equipped with a dual-core Intel (R) PENTIUM (R) B940 CPU processor running at 2.0 GHz and 4.0 GB of RAM running on Windows 8 using Molegro Virtual Docker (MVD) software and Discovery studio.

Ligand preparation

The least energy optimized structures was saved in pdb file format prior to docking studies (Abdullahi et al. 2021).

Protein retrieval and Preparation

3D X-ray crystallized structure of the EGFR protein receptor (pdb id = 3ug2) was obtained from the protein data bank (<https://www.rcsb.org/>), and was prepared on the MVD workspace by eliminating water molecules and co-crystallized ligand enclosed in the crystal structure. The amino acid residues with structural error were repaired/rebuilt. The Fully prepared protein structure was also save in pdb format prior to docking process.

Docking of the ligands and receptor using molegro virtual docker (MVD)

Due to its ability to produce a better and consistent results relative to other docking softwares, Molegro Virtual Docker software program was utilized for the docking study in this research. Before the start of the process, the prepared protein was exported from its folder to the MVD work space, cavities were detected and surface was created. The active pocket of the EGFR receptor was anticipated and was set inside a regulated sphere of X : 0.91 Y : 50.08 and Z : 22.54 with 15 Å radius respectively. The prepared ligands were imported into the MVD work

Table 2 Least approved values for predicting an acceptable QSAR model and values for the selected model

Symbol	Name	Approved Value	Selected model
R^2	Correlation coefficient	≥ 0.6	0.875
Q^2	Cross validation coefficient	≥ 0.5	0.837
$R^2 - Q^2$	Difference between R^2 and Q^2	≤ 0.3	0.038
$N_{ext\ test\ set}$	Minimal number of test set	≥ 5	5.000
R^2_{ext}	Correlation coefficient for the test set	≥ 0.6	0.655

space and docking process was executed using a grid resolution of 0.30 Å. The Root Mean Square Deviation (RMSD) threshold was set as 2.00 Å for the multiple clusters poses with 100.00 energy penalty values. The docking algorithm was set for a maximum of 1500 iteration with a maximum population size of 50. The docking simulation was run for a minimum of 50 times for the 10 poses, and the best poses were determined based on the MolDock score scoring function (Thomsen and Christensen 2006).

Hydrogen bond, Hydrophobic, alkyl, pi-alkyl, Halo bonds and aryl intermolecular interactions were viewed with the aid of Discovery studio software.

ADMET and drug-likeness properties

pkCSM ([http:// structure. bioc. cam. ac. uk/ pkcsml](http://structure.bioc.cam.ac.uk/pkcsml)), and SwissADME (<http:// www. swiss adme. ch/ index. php>) are free and easily accessible online web sites that are utilized to explore the ADMET and medicinal properties of tiny molecules. The ADMET and Drug-resemblance properties of the compounds are obtained from these sites in the current research. At the pre-clinical phase of drug invention, the most crucial parameter is the Lipinski's rule of five (Abdullahi et al. 2021), and it states that any molecule that violates more than two (2) of the following criteria is not easily permeable or readily absorbed into the body system. The criteria are; $MW < 500$, $HBD < 5$, $HBA < 10$, $\text{Log } p < 5$ and $PSA (PSA) < 140 \text{ \AA}^2$.

Structure based drug design

The design of drugs based on structures is also called the direct drug design, a very important, powerful and useful method of drug discovery. This includes the collection of information on the three-dimensional structure of the target receptor (protein) via approaches such as X-ray crystallography, NMR spectroscopy, or homology modeling, followed by the design of promising drug candidates based on binding and selective efficiencies for their target groups. The method involves several steps, such as retrieval and preparation of protein structure, preparation of ligand archives and manual design of new novel compounds (Batool et al. 2019).

Results

In this research work Genetic function algorithm was employed to generate the QSAR models due to its ability to produce a vast population of model instead of just a single model. Four models were generated from the model building set and the first one was chosen because of its statistical significance.

Model 1

$$Y = 610.603330617 * \text{AATSC8c} \\ - 5.196120723 * \text{MATSC8c} \\ - 0.030888015 * \text{SpMAD_Dzs} \\ + 3.035899535 * \text{BCUTc} - 11 \\ + 6.602683196.$$

Model 2

$$\text{pIC}_{50} = 623.468540876 * \text{AATSC8c} \\ - 5.213000244 * \text{MATSC8c} \\ - 0.000914275 * \text{SpDiam_Dzs} \\ + 3.132477307 * \text{BCUTc} - 11 \\ + 6.488595705$$

Model 3

$$\text{pIC}_{50} = 627.526620329 * \text{AATSC8c} \\ - 5.260233300 * \text{MATSC8c} \\ - 0.001549236 * \text{SpMax_Dzs} \\ + 3.125046129 * \text{BCUTc} - 11 \\ + 6.501710152$$

Model 4

$$\text{pIC}_{50} = 627.526620329 * \text{AATSC8c} \\ - 5.260233300 * \text{MATSC8c} \\ - 0.001549236 * \text{EE_Dzs} \\ + 3.125046129 * \text{BCUTc} - 11 \\ + 6.501710152$$

Discussion

The results of various statistical parameters of the selected model are shown in Table 1, therefore, the model satisfies the least required values for the evaluation of a robust QSAR model. Additionally, the selected model was utilized to estimate the inhibitive capacity of the external validation set and it was found to have passed the external validation test with $R^2_{\text{ext}} = 0.655$ (Table 2). The structures, experimental as well as predicted inhibitive activities of the compounds in this research work are placed in Table 1 respectively.

A plot of experimental pIC_{50} against the predicted pIC_{50} of both the model building as well as the external validation set is shown in Fig. 1, while Fig. 2 represent a plot of experimental Pic_{50} of both the model building

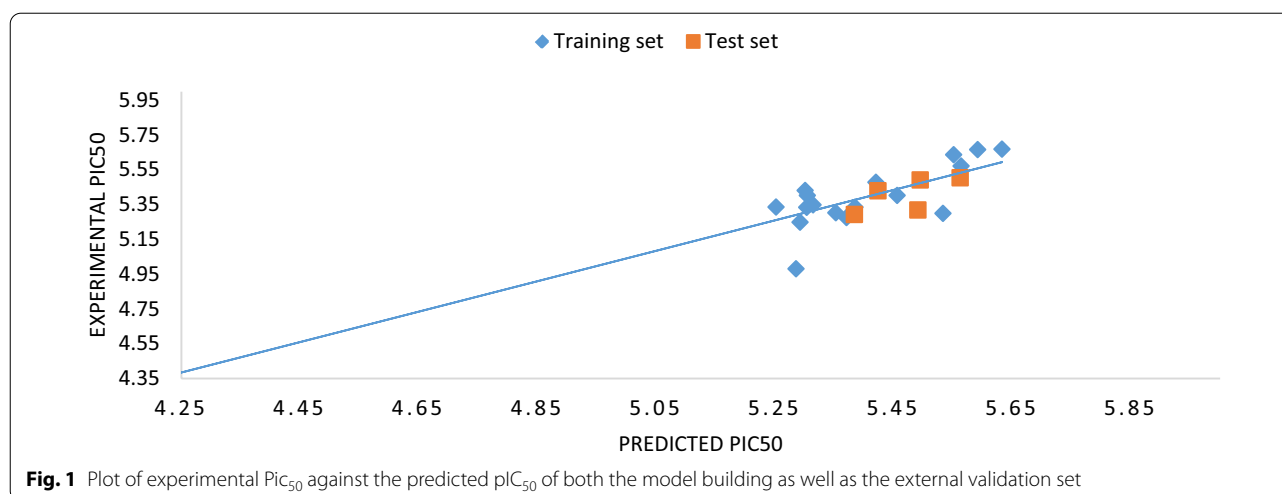


Fig. 1 Plot of experimental Pic_{50} against the predicted pIC_{50} of both the model building as well as the external validation set

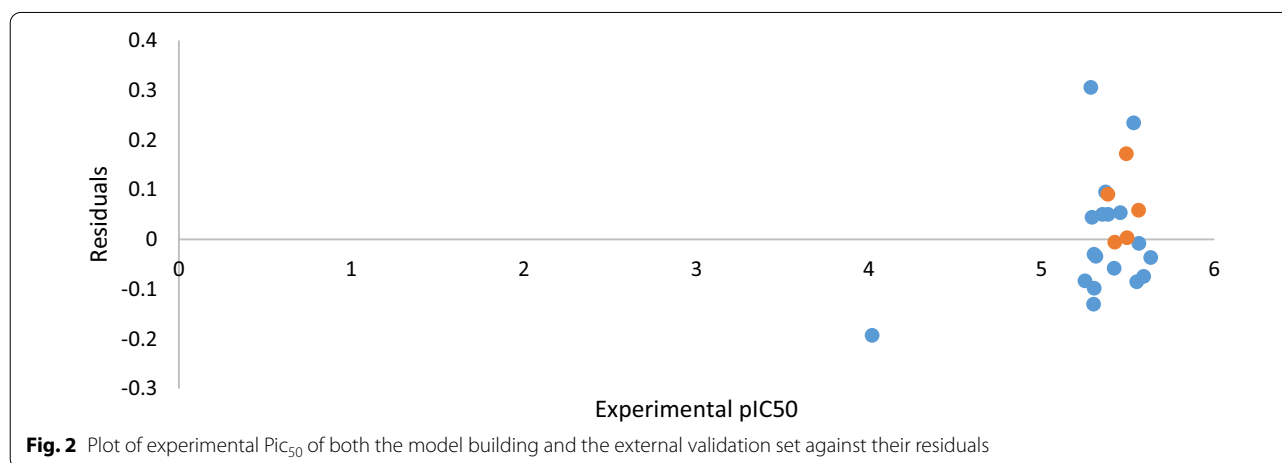


Fig. 2 Plot of experimental Pic_{50} of both the model building and the external validation set against their residuals

and the external validation set against their residuals. The selected model does not exhibit systematic error since the standardized residuals are randomly dispersed on either side of zero (Ibrahim et al. 2018).

Pearson's correlation matrix of the selected model indicates that descriptors are not correlated to each other, this illustrates that they are very good (Table 3).

Y-randomization test

Result of Y -randomization test is shown in Table 4 and the test is used to confirm that a model was not obtained by coincidental correlation. The coefficient of determination for Y -randomization cR^2p is the most crucial parameter for this test and for a robust QSAR model its value must exceed 0.5. Its value in this research work is 0.74, this illustrates that the model is powerful enough and was not purely due to chance and has satisfied the minimal requirement for robustness.

Table 3 Pearson's correlation matrix of the descriptors in the selected model

	AATSC8c	MATS8c	SpMAD_Dzs	BCUTc-11
AATSC8c	1			
MATS8c	0.989573	1		
SpMAD_Dzs	0.194418	0.168297	1	
BCUTc-11	-0.1423	-0.12496	0.000392	1

Mean effect of the descriptors in the selected model

The mean effect designate the individual function and the influence of each descriptor in a model, and it is computed for each of the molecular descriptors using the below equation:

Table 4 Result of Y-randomization test of the selected model

Model	R	R ²	Q ²
Original	0.935611	0.875367	0.472886
Random 1	0.673207	0.453208	-0.49528
Random 2	0.339829	0.115484	-1.11468
Random 3	0.494563	0.244592	-0.21047
Random 4	0.450413	0.202872	-0.85992
Random 5	0.556608	0.309812	-1.14735
Random 6	0.5238	0.274366	-0.46273
Random 7	0.551894	0.304587	-0.41126
Random 8	0.293835	0.086339	-0.72515
Random 9	0.529202	0.280055	-1.12015
Random 10	0.567466	0.322018	-0.36167
Random models parameters			
Average R	0.498082		
Average R ²	0.259333		
Average Q ²	-0.69087		
cRp ²	0.741014		

$$MF_j = \frac{B_j \sum_{j=1}^{i=n} d_{ij}}{\sum_j^m B_j \sum_i^n d_{ij}} \quad (7)$$

MF_j denotes the mean effect for the indicated descriptor *j*, the coefficient of the descriptor *j* is denoted by β_j, *d*_{ij} is the value of the Target descriptor for each molecule and *m* population of descriptors in the model (Dimić et al. 2015).

The MF value offers essential information on the effect of each molecular descriptors in the picked model; the size and signals of these descriptors combined with their mean effects reveal their stability and path in inducing the activity of a molecule. The mean effect values are presented in Table 5. BCUTc-11, MATS8c and SpMAD_Dzs were found to possess the most pronounced influence on the model performance due to their large and positive mean effect values. Their positive sign indicated that increase in their value increases the inhibitive activities of a compound against MDA-MB231 breast cancer cell line. The other descriptor, AATSC8c is negatively correlated

with the inhibitive activities of the compounds against the breast cancer cell line, higher value of this descriptor will be responsible for hindering the potency of these compound.

Variance inflation factor (VIF)

The inter-correlation amongst molecular descriptors in a model is detected using their variation inflation factors (VIF), to check whether the descriptors are highly correlated with one another or not. computed VIF values less than 1 illustrates that there is no inter-correlation between the descriptors between 1 and 5, the model can be accepted; and if it is higher than 10, the model cannot be accepted. It can be calculated using the Eq. 8 below. In this research work VIF values for all the descriptors are less than 10, this demonstrates the fitness of the selected model and the descriptors were independent of one another (Table 3).

Table 5 Descriptors in the selected, their definition, class and Mean effect values

Descriptors	Definition	Class	VIF	MF
AATSC8c	Average centered Broto-Moreau autocorrelation – lag 8/weighted by charges	2D	5.11012	-0.20071
MATS8c	Moran autocorrelation – lag 8/weighted by charges	2D	5.034677	0.1154
SpMAD_Dzs	Spectral mean absolute deviation from Barysz matrix/weighted by I-state	2D	1.072967	0.305655
BCUTc-11	nhigh lowest partial charge weighted BCUTS	2D	1.03593	0.779653

VIF VARIANCE INflation factor, MF mean Effect

$$\text{VIF} = \frac{1}{1 - R^2} \quad (8)$$

Williams' plot of the selected model

A plot of leverage values against standardized residuals of a particular set of compounds is called Williams plot. It enables the detection of a sample that is outside the defined domain of applicability of a particular model. A compound having a leverage value that exceeds the cut-off leverage is tagged as influential as it may affect the performance of the model. In this research, the Williams plot for model 1 is shown in Fig. 3, the cut-off leverage is 0.833 hence, and only the external validation set compounds lies beyond the defined domain of applicability (leverage values > 0.833). These compounds affects the performance of the model but cannot be tagged as outliers since their standardized residual values lies within ± 3 region.

Molecular docking studies

Molecular docking studies is performed to have the knowledge on the nature of binding interactions and the amino acid residues that are accounted for inducing the biological activity of a molecule. In this research work docking simulation study was performed between all the studied quinazoline derivatives and the binding pocket of the EGFR protein receptor (pdb id = 3ug2) and the results are placed in Table 2. The reference drug (Gefitinib) was also redocked into the same binding pocket to revalidate the docking results. Eight compounds (6, 10, 13, 16, 17, 18, 19 and 20) were observed to have better docking score as well as experimental and predicted activities than Gefitinib as such they were tagged as potential hit compounds. Various types of Amino acid interactions between the potential hit compounds and the active site of the EGFR receptor are presented in Table 6.

Compound 6 is observed to have interacted with the binding site of the EGFR receptor via one (1) conventional Hydrogen bond, two (2) Carbon-Hydrogen bond, one (1) Pi-sulfur interaction and several Alkyl and Pi-Alkyl Interactions. The Carbonyl Oxygen of the quinazoline ring forms a conventional Hydrogen bond with MET793 at a distance 1.86 Å, and a Carbon-Hydrogen bond with LEU792 at a distance 2.46 Å. Other Carbon-Hydrogen bond is between the Hydrogen atom H8 and GLN791 amino acid residue at a distance 2.85 Å. The benzene ring intercalated in space and forms a π -Sulfur interaction with MET790 at a distance 3.46 Å. ALA743, LEU718, LEU792, CYS775, MET790, MET793, LEU844, VAL726 and LYS745 residues forms Alkyl interactions with the compound and ALA743, LEU844, LEU718, LYS745 and LEU788 amino acid residues forms π -Alkyl interactions. 2D and 3D binding nature of compound 6 in the binding pocket of the EGFR receptor is shown Fig. 4.

Compound 10 interacted with the binding pocket of the EGFR receptor through a single Conventional Hydrogen bond, double Carbon-Hydrogen bond, single electrostatic π -Cation Hydrogen bond, Hydrophobic π -Sigma, one π -Sulfur, and several Alkyl and π -Alkyl interactions. The quinazoline group carbonyl oxygen forms a conventional Hydrogen bond with MET793 at a bond distance 2.15 Å, and forms a Carbon-Hydrogen bond with LEU792 at a bond distance 2.79 Å, ALA743 forms the other Carbon-Hydrogen bond with the cyanide carbon at a bond distance 3.14 Å. The phenyl ring intercalated in space and forms a π -cation Hydrogen bond with LYS745 at a bond distance 3.10 Å. LYS745 also forms a single electrostatic π -Sigma with the benzene ring, CYS775 forms a π -Sulfur interaction with the benzene ring of the quinazoline scaffold at a distance 5.58 Å. CYS775, MET790, MET793 and LEU718 residues forms an Alkyl interactions while ALA743, MET790, LEU844, MET793 and LEU844 amino acid residues forms a π -Alkyl interactions with the compound. 2D and 3D binding mode of

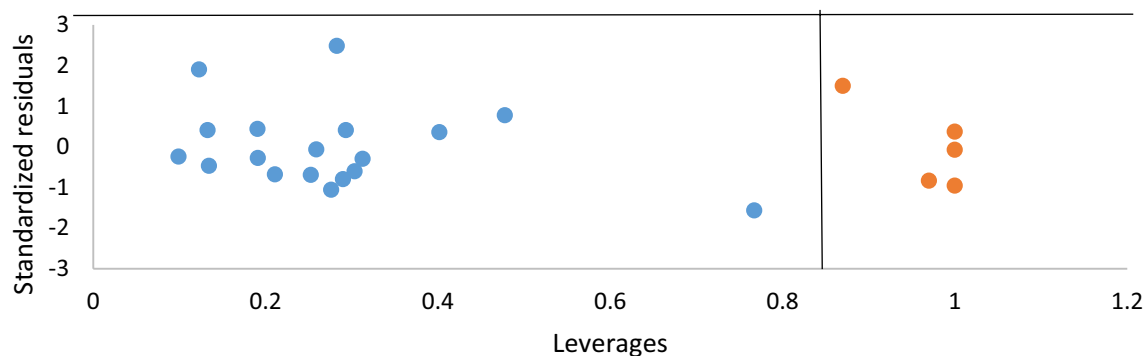


Fig. 3 Williams plot of the selected model

Table 6 various types of Amino acid interactions between the potential hit compounds and the active site of the EGFR receptor

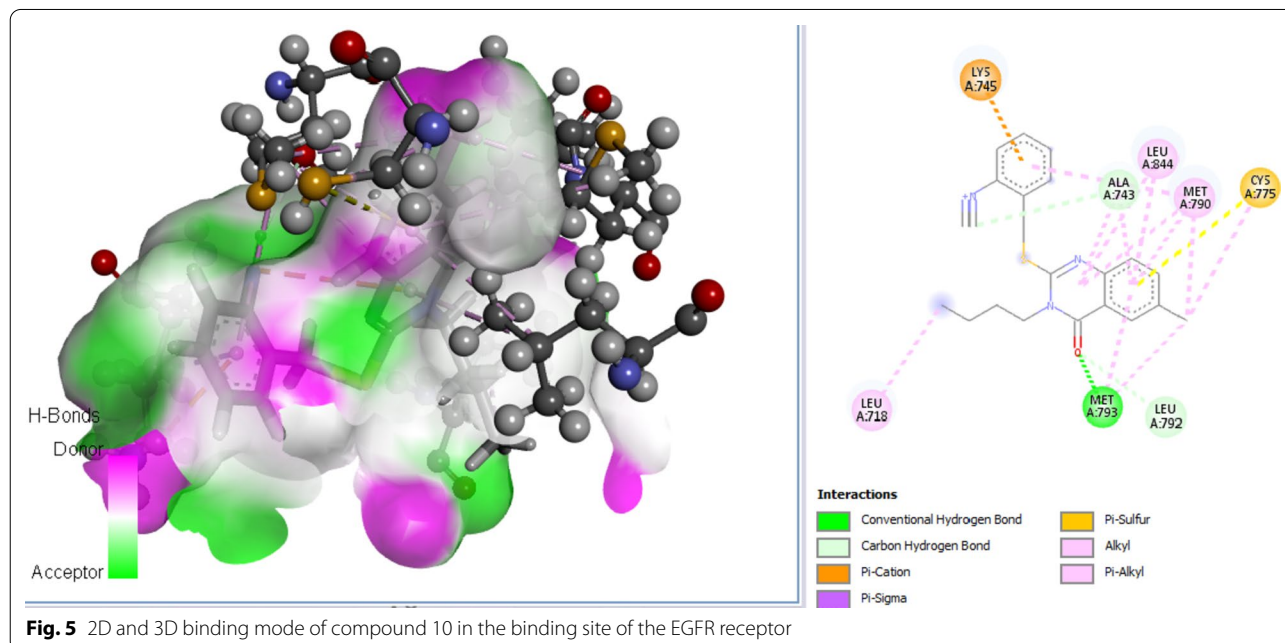
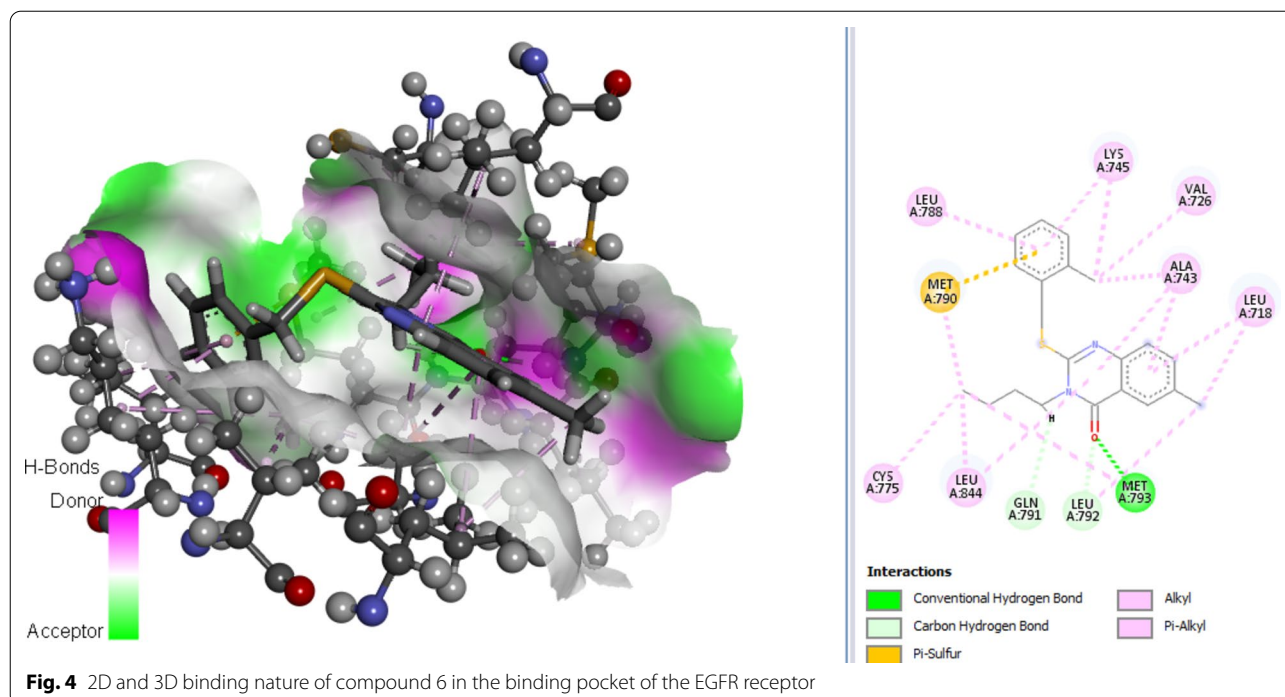
Complex	Hydrogen bond interactions	Distance (Å)	Electrostatic and Hydrophobic interactions
6	MET793 conventional hydrogen bond	1.86	MET790 Pi-Sulfur
	LEU792 and GLN791 carbon-hydrogen	2.46 and 2.85	ALA743, LEU718, LEU792, CYS775, MET790, MET793, LEU844, VAL726 and LYS745 Alkyl ALA743, LEU844, LEU718, LYS745 and LEU788 Pi-Alkyl
10	MET793 conventional hydrogen bond	2.14954	LYS745 Pi-cation
	LEU792 and ALA743 carbon-hydrogen bond	2.79 and 3.14	LYS745 Pi-Sigma CYS775 Pi-Sulfur CYS775, MET790, MET793 and LEU718 Alkyl ALA743, MET790, LEU844, MET793 and LEU844 Pi-Alkyl
13	MET793 conventional hydrogen bond	1.61	LEU718 Pi-Sigma
	LEU792, GLN791 and GLU762 carbon-hydrogen bond	2.45, 2.84 and 2.89	MET790 Pi-Sulfur LEU718, CYS775, MET790, MET793 and LEU844 Alkyl LEU718, VAL726, ALA743, LEU844, LYS745 and LEU788 Pi-Alkyl
16	MET793 conventional hydrogen bond	1.65	CYS775 Pi-Sulfur
	LEU792 carbon-hydrogen bond	2.55	CYS775, MET793, LYS745, LEU788 and MET790 Alkyl VAL726, ALA743, LYS745, MET790, LEU844, ALA743, MET793 and LEU718 Pi-Alkyl
17	MET793 conventional hydrogen bond	2.35	MET790 Pi-Sulfur
	LEU792, ASN842 and GLN791 carbon-hydrogen bond	2.85, 2.73 and 2.69	LEU718 and LEU792 Alkyl VAL726, ALA743, LEU844, LEU718, VAL726, CYS775, MET790, MET793 and LEU844 Pi-Alkyl
18	SER719 conventional hydrogen bond	2.88	ASP855 Pi-anion LEU844, VAL726, ALA743, LYS745 and MET790 Pi-Alkyl
19	LYS745 and MET793 conventional hydrogen bond	2.92 and 2.96	LYS745 Pi-Cation
			LYS745 Pi-Sigma MET790 Pi-Sulfur VAL726, ALA743, LYS745, MET790, LEU788, LEU718 And LEU792 Pi-Alkyl
20	LYS745 and MET793 conventional hydrogen bond	1.98 and 2.19	ASP855 Pi-Anion MET790 Pi-Sulfur
	GLN791	1.83	LEU844, ALA743, CYS775, MET790 and MET793 Pi-Alkyl
Gefitinib	GLN791 conventional hydrogen bond	2.81	MET790 Pi-Sulfur
	ALA743, GLU762, ASP855, ARG841, ASN842 and ASP855	2.89, 3.01, 2.75, 2.77, 2.61, 2.95, and 2.25	ARG841, LEU718 and VAL726 Alkyl LYS745, MET790, ALA743 and LEU844 Pi-Alkyl

compound 10 in the binding site of the EGFR receptor is presented in Fig. 5 respectively.

Compound 13 was observed to have interacted with the EGFR receptor via a single conventional Hydrogen bond, three (3) Carbon-Hydrogen bonds, a single π -sigma and π -Sulfur interaction and several hydrophobic alkyl and π -Alkyl interactions. The carbonyl oxygen of the quinazoline group forms a conventional and Carbon-Hydrogen bonds with MET793 and LEU792 at a bond distances 1.61 and 2.45 Å, other Carbon-Hydrogen bond interactions are observed between Methoxy Hydrogen atom and GLU762 at a distance 2.89 Å, Hydrogen atom

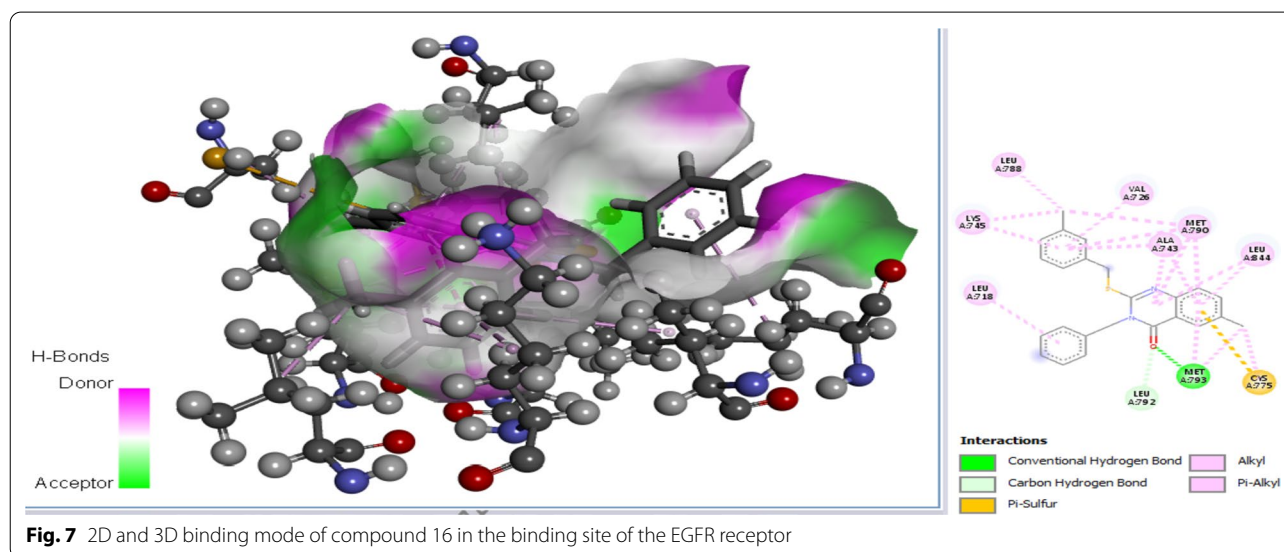
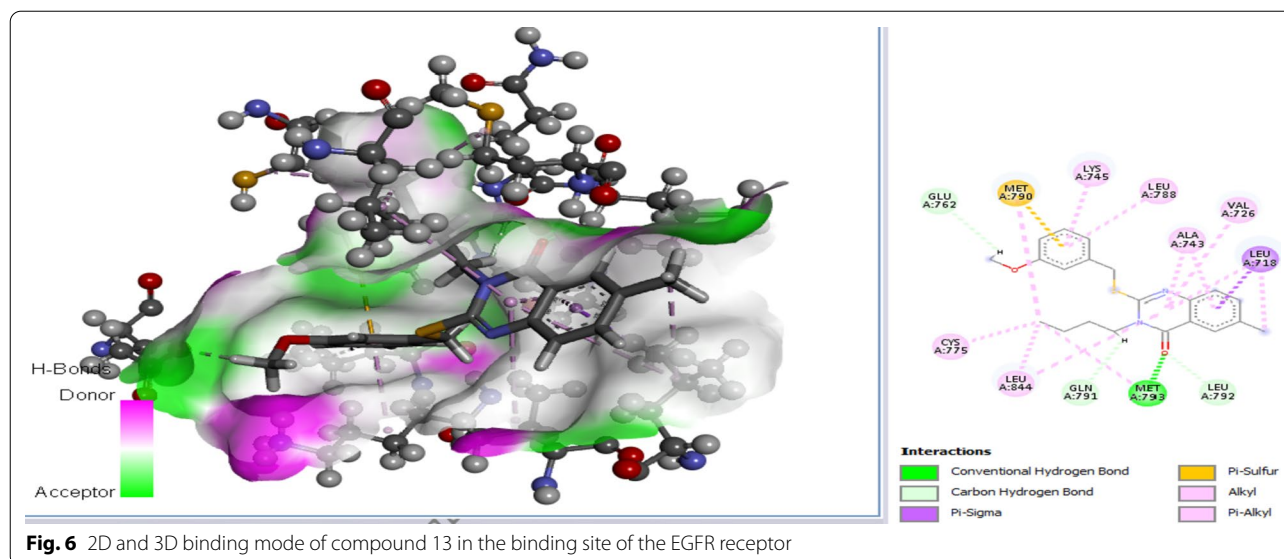
H8 and GLN791 at a bond distance 2.84 Å. The phenyl ring of the quinazoline scaffold intercalated in space to form a π -sigma hydrophobic interaction with LEU718, a π -Sulfur interaction is observed between the other benzene ring and MET790. LEU718, CYS775, MET790, MET793 and LEU844 residues forms Alkyl interactions while LEU718, VAL726, ALA743, LEU844, LYS745 and LEU788 forms a Pi-Alkyl hydrophobic interactions. 2D and 3D binding mode of compound 13 in the binding site of the EGFR receptor (pdb id = 3ug2) is pictured in Fig. 6.

The binding interactions of compound 16 is through a single conventional and Carbon-Hydrogen



bond, one (1) π -Sulfur, and numerous alkyl as well as π -alkyl interactions. The carbonyl oxygen atom of the quinazoline group forms a conventional and Carbon-Hydrogen bond with MET793 and LEU792 at a bond distance 1.65 and 2.55 Å. The phenyl ring of the quinazoline scaffold forms a π -Sulfur interaction with CYS775 at a bond distance 4.35 Å. CYS775, MET793,

LYS745, LEU788 and MET790 forms an alkyl interactions while VAL726, ALA743, LYS745, MET790, LEU844, ALA743, MET793 and LEU718 amino acid residues forms a Pi-Alkyl hydrophobic interactions. Figure 7 represent the 2D and 3D binding mode of

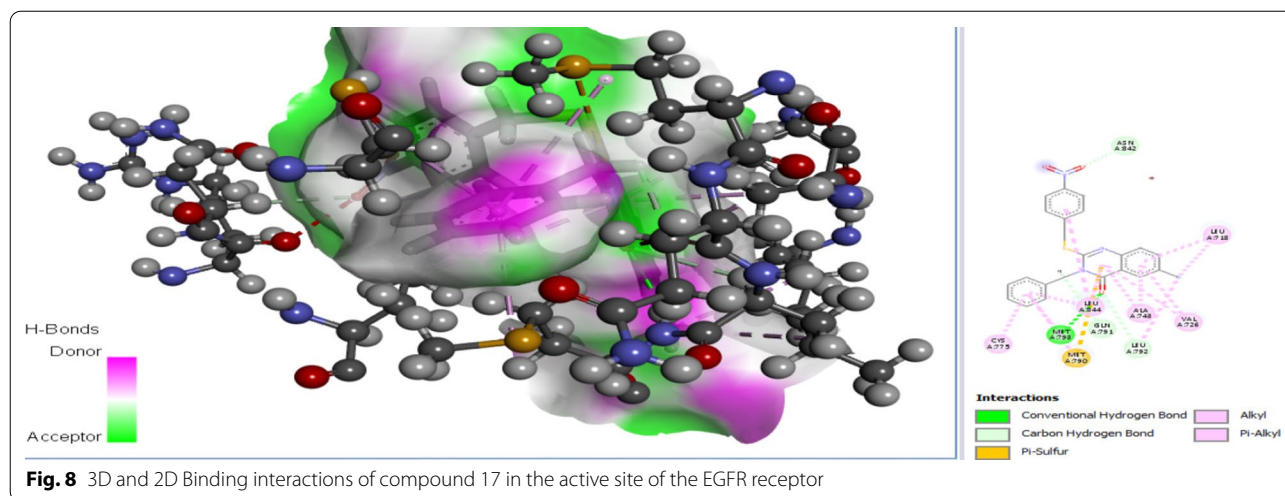


compound 16 in the binding site of the EGFR receptor (pdb id = 3ug2).

The binding mode of compound 17 is through single conventional Hydrogen bond, triple Carbon-Hydrogen bonds, single Pi-Sulfur interactions and many Alkyl as well as Pi-Alkyl Hydrophobic interactions. The quinazoline carbonyl oxygen interacted with MET793 to form a conventional Hydrogen bond at a distance 2.35 Å, and also forms a carbon-hydrogen bond when interacted with LEU792 at a distance 2.85 Å, the Nitro group oxygen forms another Carbon-Hydrogen bond with ASN842 at a bond distance 2.73 Å, Hydrogen atom of the methyl group that connects the quinazoline scaffold forms the

other Carbon-Hydrogen bond with GLN791 at distance 2.69 Å. MET790 forms a Pi-Sulfur interaction, LEU718 and LEU792 forms an Alkyl interaction while LEU844, VAL726, ALA743, LEU718, CYS775, MET790 and MET793 forms a hydrophobic Pi-Alkyl interactions. 3D and 2D Binding interactions of compound 17 in the active site of the EGFR receptor is placed in Fig. 8.

The binding mode of compound 18 in the active site of the EGFR receptor is through a single conventional Hydrogen bond, an electrostatic π -Anion and π -alkyl interactions. Oxygen atom of the oxadiazole group forms a conventional Hydrogen bond with SER719 at a bond distance 2.88 Å, the quinazoline benzene ring moiety

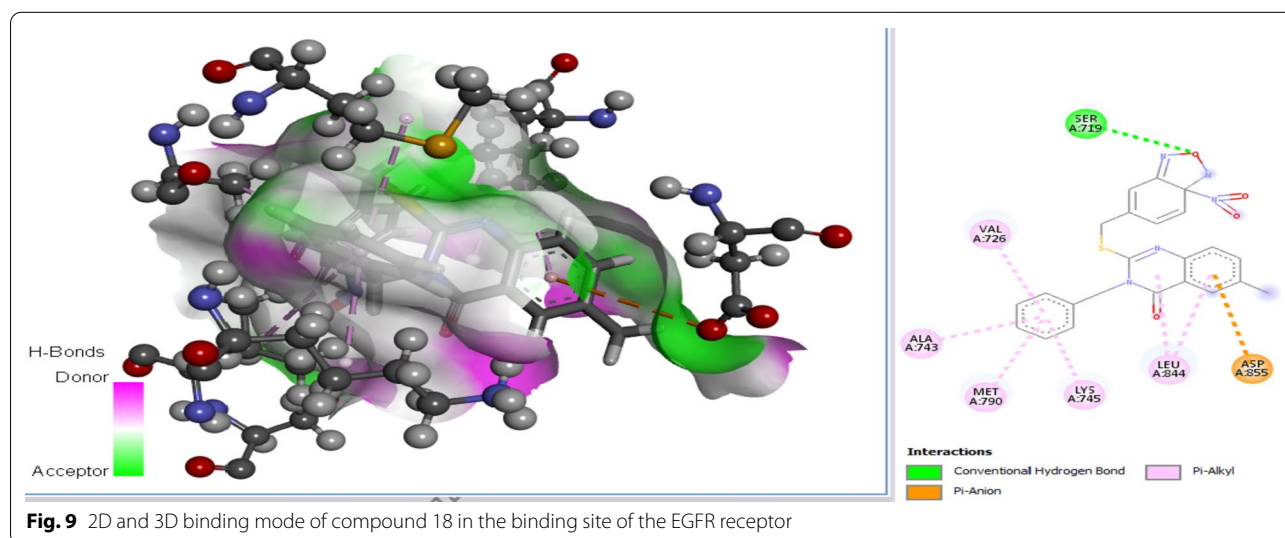


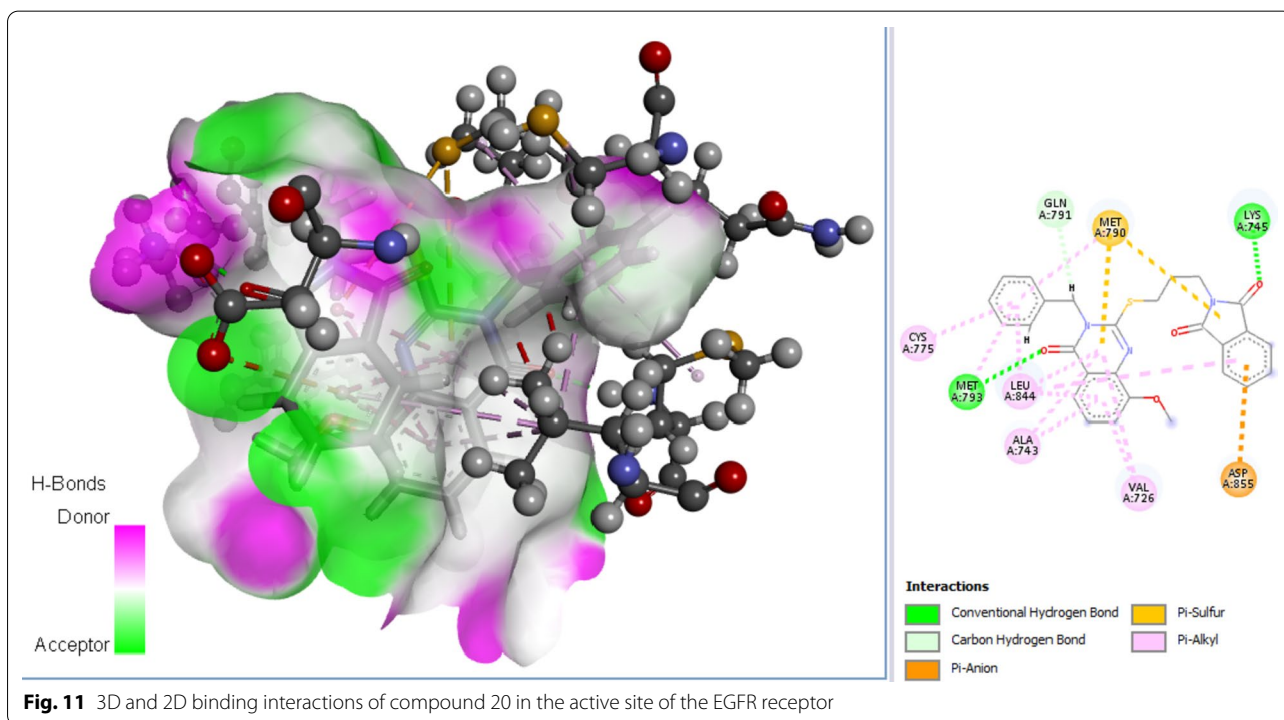
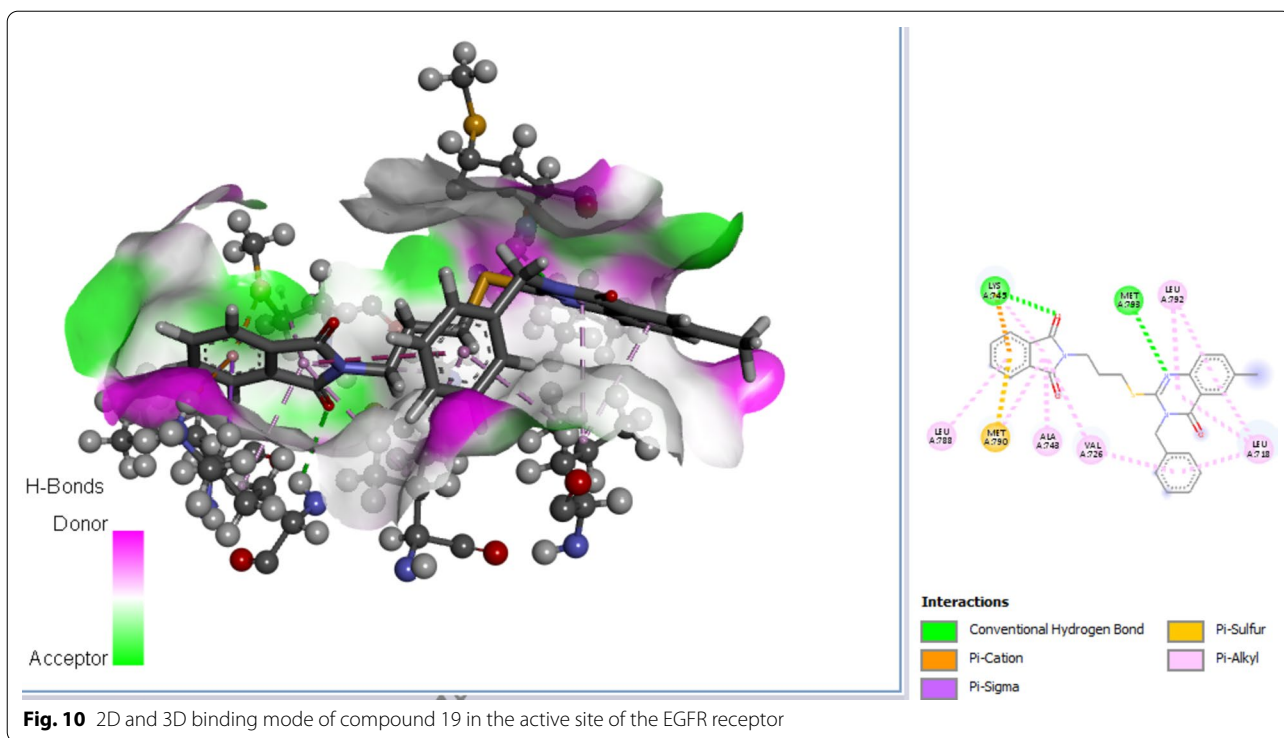
intercalated in space and forms an electrostatic π -Anion interaction with ASP855. LEU844, VAL726, ALA743, LYS745 and MET790 amino acid residues forms Pi- Alkyl hydrophobic interactions. 2D and 3D binding mode of compound 18 in the binding site of the EGFR receptor (pdb id = 3ug2) are shown in Fig. 9.

Compound 19 is bound to the active site of the EGFR receptor via two (2) conventional Hydrogen bonds, a single Pi-cation Hydrogen bond, a hydrophobic Pi-Sigma interaction, a Pi-sulfur and numerous Pi-Alkyl hydrophobic interactions. LYS745 and MET793 forms conventional Hydrogen bonds at bond distances 2.92 and 2.96 Å, LYS745 forms a Pi-Cation Hydrogen bond and Hydrophobic Pi-Sigma interactions at distances 3.33 and 2.57 Å, MET790 forms a Pi-sulfur interaction. VAL726, ALA743, LYS745, MET790, LEU788, LEU718 and LEU792 amino acid residues formed Pi- Alkyl hydrophobic interactions.

2D and 3D binding mode of compound 21 in the active site of the EGFR receptor (pdb id = 3ug2) are placed in Fig. 10 respectively.

Binding mode of compound 20 is through two (2) conventional Hydrogen bonds, single Carbon-Hydrogen bond, single electrostatic Pi-anion, single Pi-Sulfur and hydrophobic Pi-Alkyl Interactions. Carbonyl Oxygen atoms of the quinazoline scaffold and Isoindoline-1,3-dione group forms conventional Hydrogen bonds with MET793 and LYS745 at distances 2.16 and 1.98 Å, Hydrogen atom of the methyl group that connects the quinazoline scaffold to the phenyl ring forms a Carbon-Hydrogen bond with GLN791 at distance 1.83 Å. ASP855 forms an electrostatic Pi-Anion interaction, MET790 forms a Pi-Sulfur Interaction, while VAL726, ALA743, LEU844, VAL726, CYS775, MET790 and MET793 forms a hydrophobic Alkyl interactions. 3D and 2D binding





interactions of compound 20 in the active site of the EGFR receptor is shown in Fig. 11 respectively.

To validate the docking approach the reference drug, Gefitinib was also docked onto the binding pocket of

the EGFR receptor and was observed to interact with the protein kinase via a single conventional Hydrogen bond, eight (8) Carbon-Hydrogen bond, single Pi-Sulfur

interaction, several Alkyl and Pi-Alkyl hydrophobic interaction. Hydrogen atom attached to the Amino group formed a conventional Hydrogen bond with GLN791 at a bond distance 2.81 Å, Hydrogen atom H3 of the quinazoline group formed a Carbon-Hydrogen bond with ALA743 at distance 2.89 Å, Hydrogen atom of the Hydroxyl group attached to the Quinazoline frame forms another Carbon-Hydrogen bond with GLU762 at a distance 3.00 Å, Additionally, Hydrogen atom of the Morpholine group and that of a methyl group adjacent to it forms double Carbon-Hydrogen bonds with ASP855 at a bond distances 2.75 and 2.25 Å respectively, ARG841 forms double Carbon-Hydrogen bonds with two morpholine group Hydrogen atoms at distances 2.76 and 2.61 Å, the last one is with ASN842 at distance 2.95 Å. MET790 forms a Pi-Sulfur interactions, ARG841, LEU718, VAL726, LYS745, MET790, ALA743 and LEU844 forms Alkyl and Pi-Alkyl Interactions. Figure 12 represent the 3D and 2D binding mode of Gefitinib in the active site of EGFR receptor.

Structure-based drug design

In this research work, all the quinazoline derivatives were docked on to the binding site of the EGFR (pdb id=3ug2). Compound 19 (pred $pIC_{50}=5.67$, Residual = -0.04 and MolDock score = -123.238) was identified as the best compound since it has the best MolDock score and was excellently predicted by the model selected with least residual value and was within the defined

applicability domain, hence, it is adopted as template for the design. Ten (10) novel compounds were designed by addition of various groups on the Meta, Para and Ortho positions of the isoindoline-1, 3-dione phenyl ring. The inhibitive activities of the designed compounds were predicted by the selected model and most of them possess an improved activity relative to the template compound. The structure, predicted activity and MolDock score of the designed compounds are presented in Table 7.

Molecular docking studies of the designed compounds

Molecular docking investigation was also performed for the designed compounds and the binding site of the EGFR receptor (pdb id=3ug2) using Molegro Virtual Docker (MVD) software. The designed compounds were optimized to obtain the most stable and least energy conformer using DFT calculations utilizing B3LYP 631G* basis set and the optimized molecules were saved in pdb format. All the designed compounds displayed better docking scores compared to the Template and the reference drug (Gefitinib) utilized in the design. Types of amino acid interactions of the designed compounds and the active site of the EGFR receptor are presented in Table 8. The results of three (3) compounds with best docking scores are discussed in this research.

Designed compound number three (3) has the best docking score (MolDock score = -159.63) and it is found to interact with the EGFR receptor via three (3) Carbon-Hydrogen bonds, two (2) Alkyl and many Pi-Alkyl

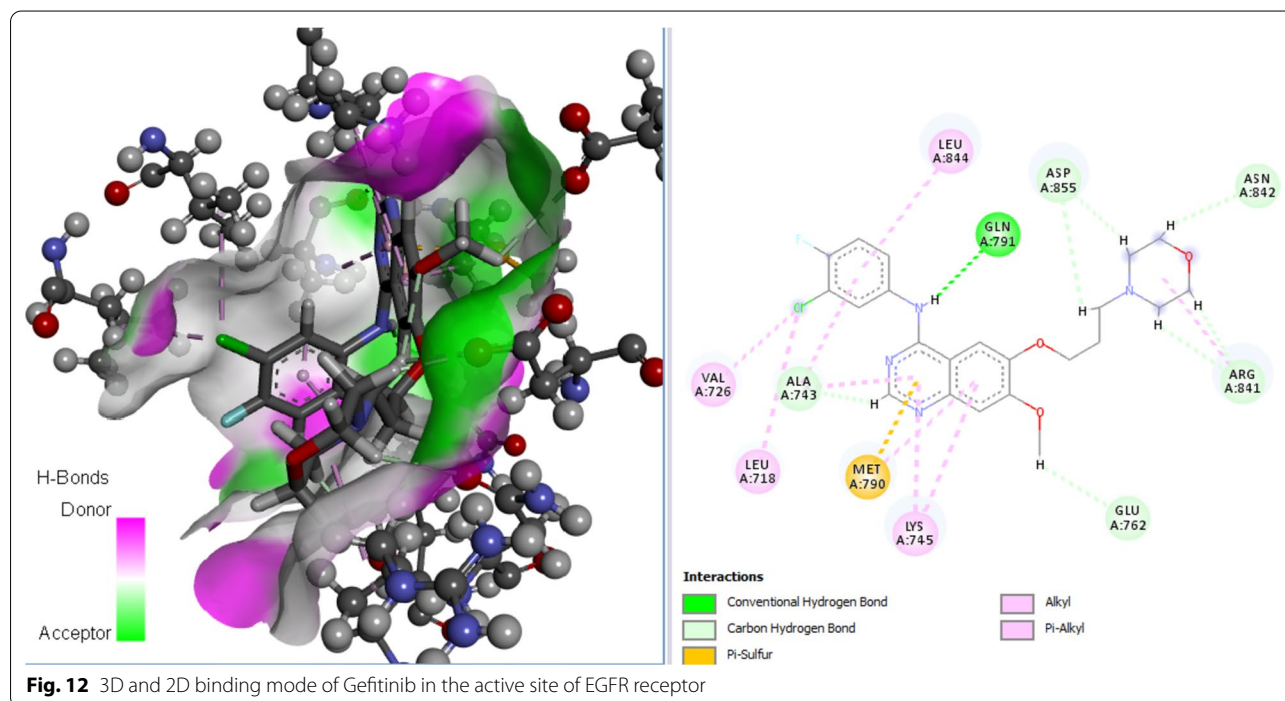
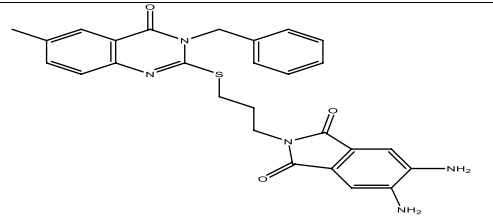
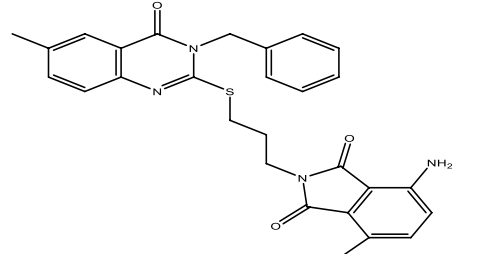
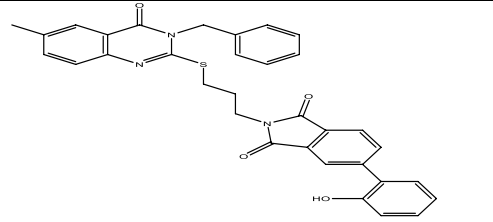


Table 7 Structure, predicted pIC₅₀ and MolDock scores of the designed compounds

S/no	Structure	PredpIC ₅₀	MolDock score
1		5.487	-143.89
2		5.564	-152.085
3		5.329	-159.63
4		5.844	-135.074
5		5.931	-138.734
6		5.657	-146.947
7		5.844	-136.569

Table 7 (continued)

8		5.905	-138.91
9		6.04	-140.05
10		5.627	-146.806

hydrophobic interactions. GLU762 forms two (2) Carbon-Hydrogen bonds with dimethyl amine Hydrogen atoms at a distance 2.64 and 2.98 Å, THR854 residue forms the other Carbon-Hydrogen bond with the Hydrogen atom of the methyl benzene that is directly attached to the Quinazoline scaffold at a distance 2.51 Å. LYS745 and LEU788 forms the alkyl interactions while ALA743, LYS745, LEU788, MET790, CYS775, MET793, LEU844 and LYS852 residues forms the Pi-Alkyl interactions. The 3D and 2D binding mode of designed compound 3 in the active pocket of the EGFR receptor is shown in Fig. 13.

Designed compound 2 has the second best docking score (-152.085) and it is found to have interacted with the EGFR receptor through three (3) conventional Hydrogen bonds, three (3) Carbon-Hydrogen bonds, single electrostatic Pi-Cation, Pi-Anion and Hydrophobic Pi-Sigma, two (2) Pi-Sulfur, single Alkyl and many Pi-Alkyl hydrophobic interactions. The Oxygen atom of the Isoindoline-1,3-dione forms a conventional Hydrogen bond with MET793 at a bond distance 1.66 Å, ASP855 forms another conventional Hydrogen bond with Hydrogen atom of the OH- group attached to the amino benzene group at a bond distance 1.87 Å, Carbonyl Oxygen of the quinazoline scaffold forms the other Conventional Hydrogen bond with CYS797 at a distance 2.26 Å. Similarly, Oxygen atom of the Isoindoline-1, 3-dione forms a Carbon-Hydrogen bond with LEU792 at a distance

2.77 Å, Methyl group Hydrogen atom attached to the Isoindoline-1,3-dione forms another Carbon-Hydrogen bond with MET793 at a distance 2.92 Å and the last one is between ASP800 and the Hydrogen atom of the methyl group that connects the phenyl ring to the main Quinazoline scaffold at a distance 2.87 Å. Additionally, the para-Hydroxy amino benzene intercalated in space to form an electrostatic Pi-Anion interaction with LYS745, ASP800 forms a Pi-Anion electrostatic interaction, GLY796 forms a Hydrophobic Pi-Sigma, MET790 and CYS797 forms a Pi-Sulfur interaction, LEU844 forms Alkyl and VAL726, ALA743 and LEU844 residues formed a hydrophobic Pi-Alkyl interactions. 3D and 2D binding pattern of designed compound 2 in the active site of the EGFR receptor (pdb id = 3ug2) is presented in Fig. 14.

Designed compound 6 also having a promising docking score (MolDock score = -146.947) interacted with the active pocket of the EGFR receptor via two (2) conventional Hydrogen bonds, four (4) Carbon-Hydrogen bonds, two (2) Pi-Anion electrostatic interactions, and Pi-Alkyl hydrophobic interactions. LYS745 and ASP855 forms a conventional Hydrogen bonds with the Oxygen atoms of the Isoindoline-1,3-dione at a distances 1.54 Å and 2.68 Å, GLU762 forms a Carbon-Hydrogen bond with methyl group Hydrogen atoms attached directly to the Isoindoline group at distances 2.63 Å and 2.73 Å,

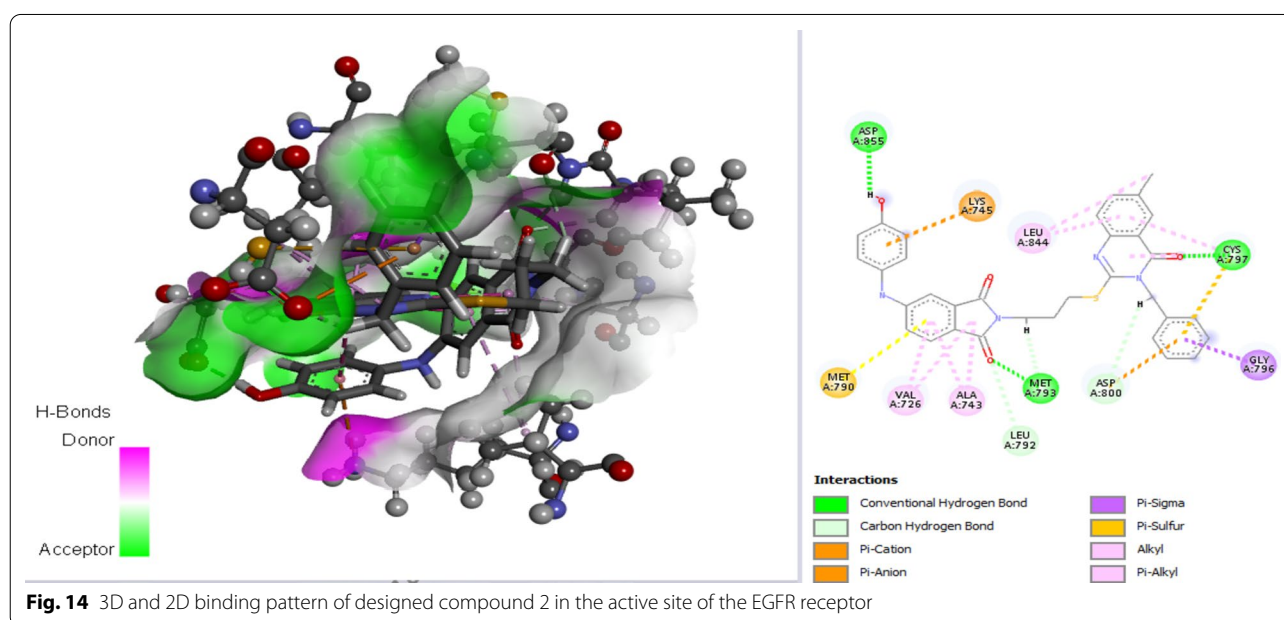
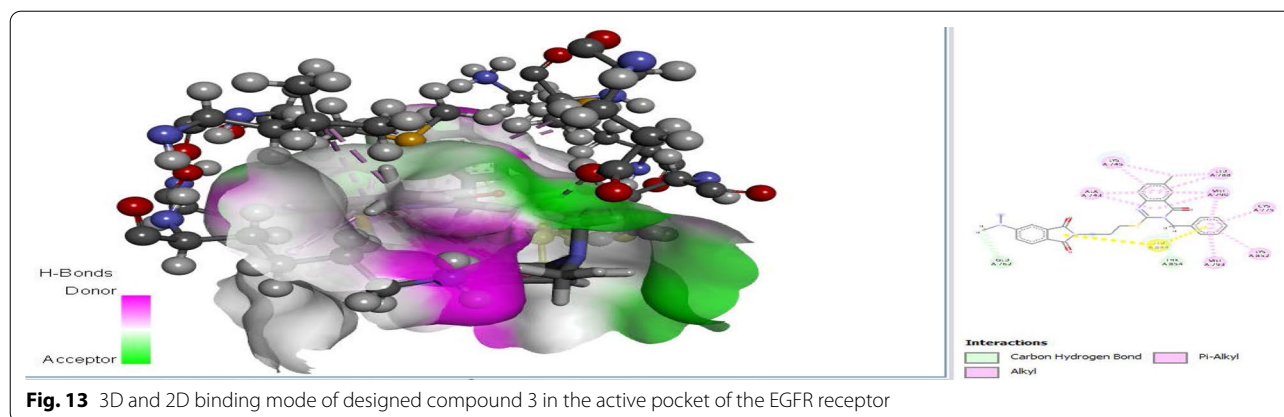
Table 8 Types of amino acid interactions of the designed compounds and the active site of the EGFR receptor

S/no	Hydrogen bond interactions	Distance (Å)	Electrostatic and hydrophobic interactions
1	LYS745, ASP855 and ARG841 conventional hydrogen bond	2.32, 2.28 and 2.49	ASP855 Pi-Anion VAL726, LYS745, ALA743, LEU844 and ARG841 Pi-Alkyl
	GLU762 carbon–hydrogen bond	2.24	
2	MET793, CYS797 and ASP855 conventional hydrogen bond	1.66, 2.26 and 1.87	LYS745 Pi-Cation
	LEU792, MET793 and ASP800 carbon–hydrogen bond	2.77, 2.92 and 2.87	ASP800 Pi-anion GLY796 Pi-Sigma MET790 and CYS797 Pi-Sulfur LEU844 Alkyl VAL726, ALA743 and LEU844 Pi-Alkyl
3	THR854 and GLU762 carbon–hydrogen bond	2.51, 2.64 and 2.98	LYS745 and LEU788 Alkyl ALA743, LYS745, LEU788, MET790, CYS775, MET793, LEU844 and LYS852 Pi-Alkyl
4	ALA743 conventional hydrogen bond	1.67	LYS745 Pi-Sigma MET790 and CYS797 Pi-Sulfur
	LYS745 Pi-cation hydrogen bond	3.03	LYS745, VAL726, ALA743, MET790 and LEU844 Pi-Alkyl
5	MET793 conventional hydrogen bond	2.54	LEU718 and LEU792 Alkyl CYS797, MET790, LEU844, LEU718, ALA743, CYS775 and LYS852 Pi-Alkyl
	LEU792, GLN791 and ARG776 carbon–hydrogen bond	2.98, 2.90 and 3.17	
6	LYS745 and ASP855 conventional hydrogen bond	1.54 and 2.68	ASP855 Pi-anion VAL726, LYS745 and ALA743 Pi-Alkyl
	GLU762, ASN842 and ARG841 carbon–hydrogen bond	2.63, 2.73, 2.97 and 2.49	
7	LYS745 and ARG841 conventional hydrogen bond	3.04 and 2.23	CYS775 Pi-Sulfur CYS797 Alkyl LEU844, VAL726, ALA743 and MET790 Pi-Alkyl
8	LYS745 and ARG841 conventional hydrogen bond	1.77 and 2.37	ASP855 Pi-anion MET790 Pi-Sulfur VAL726 Band ALA743 Pi-Alkyl
	GLU762 carbon–hydrogen bond	2.37	
9	LYS745 and ASP855 conventional hydrogen bond	2.17 and 3.01	ASP855 Pi-anion MET790 and GLN791 Amide Pi-Stacked
	LYS745 and GLU762 carbon–hydrogen bond	3.02 and 2.20	LYS745, VAL726, ALA743, MET790 and LEU844 Pi-Alkyl
10	MET793, CYS797 and THR854 conventional hydrogen bond	1.48, 2.28 and 1.83	ASP800 Pi-anion
	LEU792 and ASP800 carbon–hydrogen bond	2.52 and 2.85	GLY796 Pi-Sigma CYS797 Pi-Sulfur LEU844 Alkyl VAL726, ALA743, LEU844, MET790, LEU844, CYS797 and LYS745 Pi-Alkyl

Oxazolidine Hydrogen atoms forms another Carbon-Hydrogen bonds with ASN842 and ARG841 at distances 2.97 Å and 2.49 Å respectively. ASP855 forms double Pi-Anion electrostatic interactions, and lastly, VAL726, LYS745 and ALA743 forms a hydrophobic Pi-Alkyl interactions. 3D and 2D binding pattern of designed compound 6 in the active site of the EGFR receptor (pdb id = 3ug2) is presented in Fig. 15.

ADMET and pharmacokinetics studies of the designed compounds

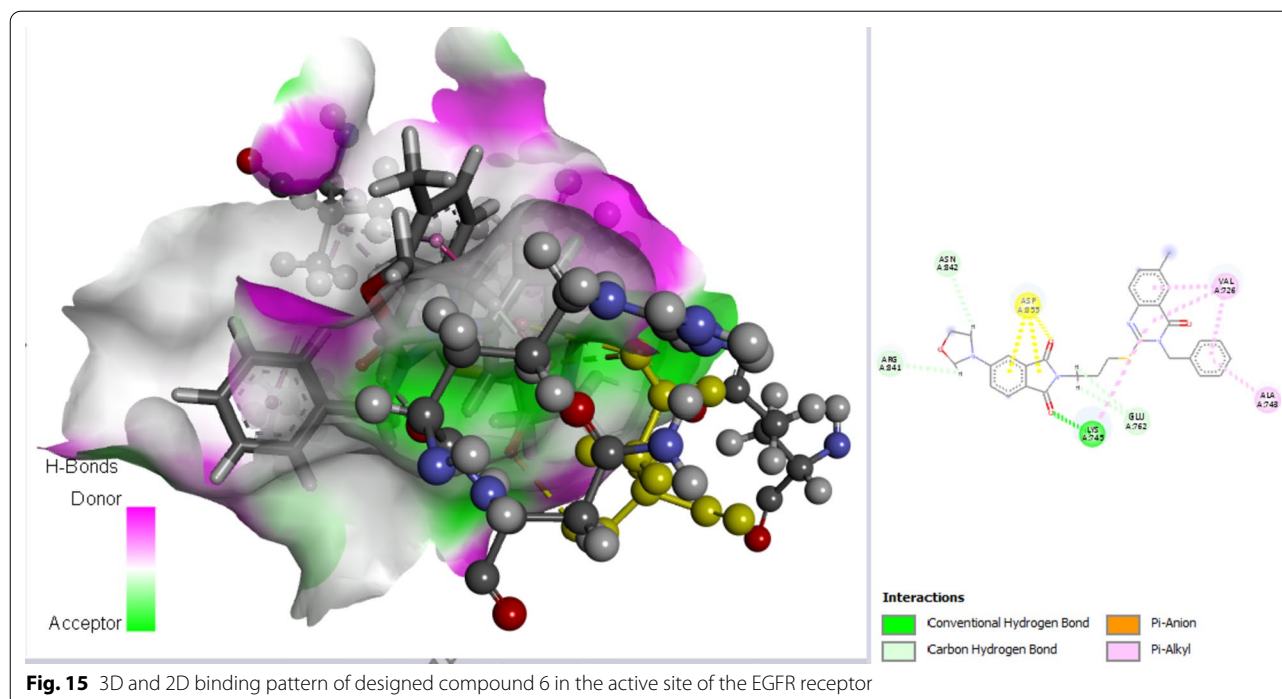
The results of ADMET and Pharmacokinetics of the designed compounds are depicted in Tables 9 and 10 respectively. None of the compounds designed violate greater than two of the acceptable thresholds established by Lipinski's rule of five filters for tiny molecules. Accordingly, they were expected to be permeable across the cell membrane, easily absorbed, transported and diffused (Ibrahim et al. 2020). In addition, the designed



compounds were further evaluated for their synthetic accessibility, testing on a scale between 1 (simply synthesizable) to 10 (tediously synthesizable). The designed compound's synthetic accessibility values range from 3–4 (Table 10), consequently, they are easily synthesizable. Furthermore, all the designed compounds displayed excellent intestinal (human) absorbance values from 81 to 100% (Table 9). The results of ADMET studies of the designed compounds illustrate that they have the potential of crossing blood brain barrier and central nervous system since for a molecule, blood–brain barrier (BBB), and the penetrability of central nervous system (CNS) approved values are > 0.3 to < -1 log BB and > -2 to < -3 log PS respectively (Clark 2003).

The class of super enzymes Cytochrome P450 (CYP450) have a vigorous role in drug metabolism

since it is the key liver protein system responsible for phase-1 metabolism (oxidation). Up to today, only seventeen CYP families were recognized in humans, among which only (CYP1, CYP2, CYP3, and CYP4, respectively) are involved in the drug metabolism, with a CYP (1A2, 2C9, 2C19, 2D6 and 3A4, respectively) were known to be accountable for the biotransformation of at least 90% of the drugs that undertake phase-I metabolism (Šrejber et al. 2018), and have been predicted and presented in Table 9. Moreover, cytochrome CYP3A4 inhibition is the most important principle in the current study. The results of this research indicated that all the designed compounds are the substrates of CYP3A4 as well as inhibitors of CYP3A4, respectively (Table 9). Clearance designates the correlation between the body's drug levels to its

**Table 9** Predicted ADMET properties of the designed compounds

S/no	Absorption Intestinal (human) absorption	Distribution		Metabolism							Excretion Total clearance	Toxicity AMES toxicity
		LogBB	LogPS	Substrate				Inhibitors				
				2D6	3A4	1A2	2C19	2C9	2D6	3A4		
1	93.058	-0.287	-1.871	No	Yes	No	Yes	Yes	No	Yes	0.093	No
2	95.176	-1.257	-2.054	Yes	Yes	No	Yes	Yes	No	Yes	-0.021	No
3	96.696	-0.864	-2.264	No	Yes	No	Yes	Yes	No	Yes	0.325	No
4	91.265	-0.512	-2.248	No	Yes	No	Yes	Yes	No	Yes	0.103	No
5	91.524	-0.719	-1.896	No	Yes	No	Yes	Yes	No	Yes	0.203	No
6	96.662	-0.707	-2.343	No	Yes	No	Yes	Yes	No	Yes	0.392	No
7	91.265	-0.512	-2.248	No	Yes	No	Yes	Yes	No	Yes	0.103	No
8	84.163	-1.094	-2.492	No	Yes	Yes	Yes	Yes	No	Yes	0.189	No
9	81.240	-1.011	-2.460	No	Yes	No	Yes	Yes	No	Yes	0.156	No
10	100.00	-0.398	-1.972	No	Yes	No	Yes	Yes	No	Yes	0.245	No

elimination per unit time. A minimal value of the total clearance suggests good perseverance of the drugs in the human body, and all the designed compounds displayed a promising persistence in the body for the drug. Additionally, it is advisable to explore the toxicity level of the designed compounds as it plays an important part in screening of drugs candidates. Based on the results of this research, all the designed compounds are found to be non-toxic (Table 9).

Conclusions

In this research work four (4) QSAR models were developed for a series of quinazoline derivatives with inhibitory activities against Triple Negative Breast Cancer cell line (MDA-MB231). Model 1 was chosen due to its statistical fitness with the following validation parameters $R^2 = 0.875$, $Q^2 = 0.837$, $R^2 - Q^2 = 0.038$, $N_{\text{ext test set}} = 5$, and $R^2_{\text{ext}} = 0.655$. The selected model was utilized for the prediction of the activities of both

Table 10 Predicted Drug-likeness properties of the designed compounds

S/no	MW	TPSA	WLogP	HBA	HBD	SA	RO5 Violation	Drug-Likeness
1	560.67	109.60	5.89	4	1	3.98	2	Yes
2	576.66	129.83	5.6	5	2	3.99	2	Yes
3	512.62	100.81	4.22	4	0	3.73	2	Yes
4	484.57	123.59	3.74	4	1	3.47	0	Yes
5	552.47	97.57	5.3	4	0	3.57	2	Yes
6	540.63	110.04	3.56	5	0	3.95	1	Yes
7	484.57	123.59	3.74	4	1	3.47	0	Yes
8	499.58	149.61	3.33	4	2	3.56	0	Yes
9	500.57	143.82	3.45	5	2	3.57	1	Yes
10	561.65	117.8	5.52	5	1	3.9	2	Yes

MW molecular weight, TPSA topological polar surface area, HBA hydrogen bond acceptors, HBD: hydrogen bond donors, SA synthetic accessibility, RO5 rule of five

the model building as well as external validation set of quinazoline derivatives. Molecular docking studies was performed on all the quinazoline series and the reference drug (Gefitinib) and the active site of the EGFR receptor (pdb id=3ug2). Eight compounds (6, 10, 13, 16, 17, 18, 19 and 20) were observed to have better docking score docking scores relative to Gefitinib. Compound 19 (pred $pIC_{50} = 5.67$, Residual = -0.04 and MolDock score = -123.238) was identified as the best compound since it has the best Moldock score and was excellently predicted by the model selected with least residual value and lies within the defined applicability domain, Hence, it was adopted as template for the design of Ten (10) new novel compounds with better activities and better docking scores. The inhibitive activities of the designed compounds were predicted by the selected model and most of them possess an improved activity relative to the template compound (19). The designed compounds were also redocked on to active pocket of the EGFR receptor and it was observed that they displayed better docking scores compared to the Template and the reference drug (Gefitinib) utilized in the design. Furthermore, the designed compounds were subjected to ADMET and drug-likeness studies using SWISSADME and pkCSM online web tools and they were observed to be pharmacologically active, easily synthesized and don not violate the Lipinski's rule of five. The designed compounds are said to be well absorbed by the Human intestine values due to their high absorbance values from 81 to 100%, they have the potential of crossing blood-brain barrier and central nervous system, they are all substrate as well as inhibitors of the most vital CYP families 3A4, displayed a promising persistence in the body and are all non-toxic. Hence, the designed

compounds can be employed as inhibitors of triple negative breast cancer cell line (MDA-MB231) after passing through in vivo and in vitro evaluation.

Abbreviations

EGFR: Epidermal Growth Factor Receptor; TNBC: Triple Negative Breast Cancer; QSAR: Quantitative structure activity relationship; DFT: Density functional theory; PADEL: Pharmaceutical Data Exploration Laboratory; B3LYP: Bee -3- Lee Yang Par; GFA-MLR: Genetic function algorithm-Multi linear regression; ME: Mean effect; ADMET: Absorption, Distribution, Metabolism, Excretion and Toxicity.

Acknowledgements

The author sincerely acknowledges all the contributors of this exploratory group for their guidance and motivation at some stages of this research work and Ahmadu Bello University for providing the softwares and conducive environment utilized for this analysis.

Authors' contributions

All authors contribute accordingly, SHA performed all the computational analysis and wrote the manuscript, AU provided all the softwares for this research and edited the manuscript to ensure that error was minimized before final submission, ABU analyzed the results obtained from the softwares, GAS and SU assessed the manuscript by using Plagiarism checker as well as arranged the manuscript according to the journal format. All authors read and approved the final manuscript.

Funding

Funds allocation was not received by the authors for this research.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable, because this article does not contain any studies with animal or human subjects.

Consent of publication

Not applicable.

Competing interests

The correspondents did not acknowledge competing of interest.

Received: 28 November 2021 Accepted: 18 December 2021
Published online: 04 January 2022

References

- Abdullahi SH, Uzairu A, Ibrahim MT, Umar AB (2021) Chemo-informatics activity prediction, ligand based drug design, Molecular docking and pharmacokinetics studies of some series of 4, 6-diaryl-2-pyrimidinamine derivatives as anti-cancer agents. *Bull Natl Res Cent* 45:167. <https://doi.org/10.1186/s42269-021-00631-w>
- Abuelizz HA, Marzouk M, Ghabbour H, Al-Salahi R (2017) Synthesis and anticancer activity of new quinazoline derivatives. *Saudi Pharma J* 25(2017):1047–1054
- Al-Suwaidan IA, Abdel-Aziz AA-M, Shower TZ, Ayyad RR, Alanazi AM, El-Morsy AM, Mohamed MA, Abdel-Aziz NI, El-Sayed MA-A, El-Azab AS (2016) Synthesis, antitumor activity and molecular docking study of some novel 3-benzyl-4 (3H) quinazolinone analogues. *J Enzyme Inhib Med Chem* 31(1):78–89
- Amin SA, Gayen S (2016) Modelling the cytotoxic activity of pyrazolo-triazole hybrids using descriptors calculated from the open source tool "PaDEL-descriptor". *J Taibah Univ Sci* 10(6):896–905
- Batool M, Ahmad B, Choi S (2019) A structure-based drug discovery paradigm. *Int J Mol Sci* 20(11):2783
- Chandregowda V, Kush AK, Reddy GC (2009) Synthesis and in vitro antitumor activities of novel 4-anilinoquinazoline derivatives. *Eur J Med Chem* 44:3046–3055
- Clark DE (2003) In silico prediction of blood-brain barrier permeation. *Drug Discov Today* 8(20):927–933
- Dimić D, Mercader AG, Castro EA (2015) Chalcone derivative cytotoxicity activity against MCF-7 human breast cancer cell QSAR study. *Chemom Intell Lab Syst* 146:378–384. <https://doi.org/10.1016/j.chemolab.2015.06.011>
- El-Azab AS, Al-Omar MA, Abdel-Aziz AAM, Abdel-Aziz NI, El-Sayed MAA, Aleisa AM, Sayed-Ahmed MM, Abdel-Hamide SG (2010) Design, synthesis and biological evaluation of novel quinazoline derivatives as potential antitumor agents: molecular docking study. *Eur J Med Chem* 45:4188–4198
- Hu K, Law JH, Fotovati A, Dunn SE (2012) Small interfering RNA library screen identified polo-like kinase-1 (PLK1) as a potential therapeutic target for breast cancer that uniquely eliminates tumor-initiating cells. *Breast Cancer Res* 14(1):1–5
- Hynes NE, Lane HA (2005) ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer* 5:341–354
- Ibrahim MT, Uzairu A, Shallangwa GA, Ibrahim A (2018) In-silico studies of some oxadiazoles derivatives as anti-diabetic compounds. *J King Saud Univ Sci*. <https://doi.org/10.1016/j.jksus.2018.06.006>
- Ibrahim MT, Uzairu A, Shallangwa GA, Uba S (2020) Structure-based design of some quinazoline derivatives as epidermal growth factor receptor inhibitors. *Egypt J Med Hum Genet* 21:63. <https://doi.org/10.1186/s43042-020-00107-y>
- Jo J, Kim SH, Kim H, Jeong M, Kwak JH, Han YT, Jeong JY, Jung YS, Yun H (2019) Discovery and SAR studies of novel 2-anilinopyrimidine-based selective inhibitors against triple-negative breast cancer cell line MDA-MB-468. *Bioorg Med Chem Lett* 29(1):62–65
- Kaplan W (2013) Background paper 6.5 cancer and cancer therapeutics. In: World Health Organization (ed) *Priority medicines for Europe and the world: update*, pp 5–6
- Kennard RW, Stone LA (1969) Computer aided design of experiments. *Technometrics* 11(1):137–148
- Lawal HA, Uzairu A, Uba S (2021) QSAR, molecular docking studies, ligand-based design and pharmacokinetic analysis on Maternal Embryonic Leucine Zipper Kinase (MELK) inhibitors as potential anti-triple-negative breast cancer (MDA-MB-231 cell line) drug compounds. *Bull Natl Res Cent* 45:90. <https://doi.org/10.1186/s42269-021-00541-x>
- Leardi R (1996) Genetic algorithms in molecular modeling. Elsevier, pp 67–86
- Rajabi S, Maresca M, Yumashev AV, Choopani R, Hajimehdipoor H (2021) The most competent plant-derived natural products for targeting Apoptosis in cancer therapy. *Biomolecules* 11(4):534. <https://doi.org/10.3390/biom1104053>
- Šrejber M, Navrátilová V, Paloncýová M, Bazgier V, Berka K, Anzenbacher P, Otyepka M (2018) Membrane-attached mammalian cytochromes P450: an overview of the membrane's effects on structure, drug binding, and interactions with redox partners. *J Inorg Biochem* 183:117–136
- Thomsen R, Christensen MH (2006) MolDock: a new technique for high accuracy molecular docking. *J Med Chem* 49(11):3315–3321
- Tiwari SK, Sachan S, Mishra A, Tiwari S, Pandey V (2015) The anticancer activity of some novel 4-anilino quinazoline derivatives as tyrosine kinase (EGFR) inhibitor and the quantitative structure activity relationships. *Int J Pharm Life Sci* 6:4819–4828
- Tropsha A, Gramatica P, Gombar VK (2003) The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *Mol Inf* 22(1):69–77
- Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. *J Chem Inf Comput Sci* 29(3):163

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)