Structure-based design of some quinazoline derivatives as epidermal growth factor receptor inhibitors

Muhammad Tukur Ibrahim, Adamu Uzairu, Gideon Adamu Shallangwa and Sani Uba

Abstract

Background: The discovery of epidermal growth factor receptor (EGFR) inhibitors for the treatment of lung cancer, most especially non-small cell lung cancer (NSCLC), was one of the major challenges encountered by the medicinal chemist in the world. The treatment of EGFR tyrosine kinase to manage NSCLCs becomes an urgent therapeutic necessity. NSCLC was the foremost cause of cancer mortality worldwide. Therefore, there is a need to develop more EGFR inhibitors due to the development of drug resistance by the mutation. This research is aimed at designing new EGFR inhibitors using a structure-based design approach. Structure-based drug design comprises several steps such as protein structure retrieval and preparation, ligand library preparation, docking, and structural modification on the best hit compound to design new ones.

Result: Molecular docking virtual screening on fifty sets of quinazoline derivatives/epidermal growth factor receptor inhibitors against their target protein (EGFR tyrosine kinase receptor PDB entry: 3IKA) and pharmacokinetic profile predictions were performed to identify hit compounds with promising affinities toward their target and good pharmacokinetic profiles. The hit compounds identified were compound 6 with a binding affinity of $-9.3 \text{ kcal/mol}$, compounds 5 and 8, each with a binding affinity of $-9.1 \text{ kcal/mol}$, respectively. The three hit compounds bound to EGFR tyrosine kinase receptor via four different types of interactions which include conventional hydrogen bond, carbon-hydrogen bond, electrostatic, and hydrophobic interactions, respectively. The best hit (compound 6) among the 3 hit compounds was retained as a template and used to design sixteen new EGFR inhibitors. The sixteen newly designed compounds were also docked into the active site of EGFR tyrosine kinase receptor to study their mode of interactions with the receptor. The binding affinities of these newly designed compounds range from $-9.5 \text{ kcal/mol}$ to $-10.2 \text{ kcal/mol}$. The pharmacokinetic profile predictions of these newly designed compounds were further examined and found to be orally bioavailable with good absorption, low toxicity level, and permeable properties.

Conclusion: The sixteen newly designed EGFR inhibitors were found to have better binding affinities than the template used in the designing process and afatinib the positive control (an FDA approved EGFR inhibitor). None of these designed compounds was found to violate more than the permissible limit set by RO5. More so, the newly designed compounds were found to have good synthetic accessibility which indicates that these newly designed compounds can be easily synthesized in the laboratory.

Keywords: Structure-based, Design, EGFR, Inhibitors, Pharmacokinetic, Prediction
Background
One of the major challenges faced by the medicinal chemist was discovering epidermal growth factor (EGFR) receptor inhibitors for treating lung cancer, especially, non-small cell lung cancer (NSCLC) or EGFR tyrosine kinase mutations [1]. Treatment of EGFR tyrosine kinase to control NSCLCs has become an urgent therapeutic requirement due to the development of drug resistance by the mutation [2].

Lung cancer was one of the world’s leading cancer problems. A lot of deaths are recorded every year, estimated to be around one third of all cancer deaths. Non-small cell lung cancer (NSCLC) was the main subset of lung cancers that accounted for about 85% of cancer-related problems [3]. Overexpression of epidermal growth factor receptor kinase was identified to be the common cause of NSCLCs. Based on the report on the population of patients with NSCLCs in Caucasias, the figure is approximately 10–15% and in Asia is approximately 30–40%, respectively [3].

NSCLC therapeutic agents showed a high response rate in patients with increase modifications of EGFR. NSCLC therapeutic agents or EGFR inhibitors were classified into reversible EGFR inhibitors (first-generation EGFR inhibitors) and irreversible EGFR inhibitors (second- and third-generation EGFR inhibitors). Unfortunately, the potency period of these first-generation EGFR inhibitors (Gefitinib and Erlotinib) is shortened due to the development of drug resistance by the secondary mutation T790M [4]. The second-generation irreversible EGFR inhibitors, such as afatinib and canertinib, were subsequently developed for the treatment of the NSCLC EGFR*<sup>T790M</sup> mutations [5]. However, due to severe side effects, such as skin rashes and diarrhea, the second-generation EGFR inhibitors could not show any significant advantage over the first-generation reversible EGFR inhibitors. It is believed that the activities upon wild-type EGFR will narrow the possible activities on the patients with the T790M mutation [1, 6–8].

To address the unmet clinical demands, many third-generation irreversible EGFR inhibitors such as WZ4002, Rociletinib, Olmutinib, and Osimertinib were designed to inhibit the T790M resistance mutation while being more selective for wild type EGFR [1, 9–12].

In structure-based design, molecular docking is used to screen a library of compounds to identify compounds with higher affinities toward their target protein [13]. ADME and drug-likeness properties prediction played a vital role in structure-based design in the determination of the pharmacokinetic profiles of drug-like compounds in the early stage of drug development [14].

The purpose of this work is to apply the concept of structure-based design on fifty sets of quinazoline derivatives (EGFR inhibitors) previously reported in literature and design new quinazoline derivatives (new EGFR inhibitors) that may have better binding affinities than the previous quinazoline derivatives and further evaluate their pharmacokinetic properties.

Method
Computational environment and tools
This computational work was carried out on a Dell personal computer laptop, with the following specifications: Intel * Core*™ i7 Dual CPU, M330 @2.75 GHz 2.75GHz, and 8GB of RAM. The following software was utilized to achieve the success of this research: Pyrex virtual screening software, UCSF Chimera, Discovery studio, and SWISSADME an online web tool.

Compounds under investigation
Fifty sets of quinazoline derivatives as EGFR inhibitors previously reported by Zhang et al. were collected and used in this research work [15]. The compounds under investigation were synthesized under the same condition. After the retrieval of the compounds under investigation, the 2D structures of all the compounds were drawn with the help of Chemdraw software and presented in the Supplementary Table 1 [16, 17].

Ligand data preparation
Ligand preparation is very vital and also an important step in molecular docking study. As such, the preparation of the ligands in this work was done by finding the most optimum geometries of all the compounds under investigation, at B3LYP/6-311G* level of theory, with density functional theory (DFT) method using Spartan 14 wave software [18]. The optimum conformations of all the compounds under investigation were then saved in protein data bank file format for the next step. A prepared 3D conformation of one of the compounds under investigation/ EGFR inhibitor is shown in Fig. 1.

EGFR enzyme preparation
The crystal structure of the epidermal growth factor receptor tyrosine kinase enzyme covalently binding to WZ4002 with protein data bank entry: 3IKA was retrieved from the RCSB protein data bank database. After successful retrieval of the enzyme, the preparation of the enzyme for the molecular docking simulation was done using a discovery studio visualizer. In the process of its preparation, the co-crystallized ligand (WZ4002) and molecule of water present on the structure were deleted. Before that, polar hydrogen was added to the crystal structure. The 3D structure of the prepared EGFR tyrosine kinase enzyme is shown in Fig. 2.
Molecular docking execution
The docking of the compounds/EGFR inhibitors/ligands under investigation in the active site of the EGFR tyrosine kinase enzyme was achieved with the help of Autodock vina of Pyrex virtual screening software [19]. UCSF Chimera software was used for the re-coupling of the docked ligand with the receptor. The discovery studio visualizer was used to visualize the re-coupled complexes so as to view the nature of interactions between the ligands and the enzyme.

Pharmacokinetic profile prediction
The drug-likeness evaluation and ADMET screening were further performed using the SwissADME free web tool, developed by the Swiss Institute of bioinformatics, and freely available at http://www.swissadme.ch [20]. All the compounds under investigation were subjected to this part of in silico screening as filtering criteria in compliance with Lipinski’s rule of five (RO5) [21, 22].

Structure-based design
Structure-based drug design is a very important, robust, and useful technique in the drug discovery arena. It is also called direct drug design which involves the acquisition of the information regarding the three-dimensional structure of the molecular target (protein) through methods such as X-ray crystallography, NMR spectroscopy, or homology modeling, followed by the design of suitable drug candidates based on the binding affinities and selectivity for their target molecules. Structure-based drug design comprises several steps such as protein structure retrieval and preparation, ligand library preparation, docking, and manual design of new compounds [23].

Results
Molecular docking simulation
The results of the molecular docking simulation of the best top ten compounds are presented in Table 1 and Fig. 3a, b, respectively.

Drug-likeness and ADME property prediction
The results of the drug-likeness and ADME properties prediction for the best top ten compounds are presented in Tables 2 and 3, respectively.

Design
The structures and binding affinities of the template and the newly designed compounds are given in Table 4.
| Ligand-receptor (3IKA) | Binding affinity (kcal/mol) | Hydrogen bond | Bond distance (Å) | Hydrophobic and electrostatic interactions |
|-------------------------|-----------------------------|---------------|------------------|------------------------------------------|
| 3                       | − 8.4                       | GLN791, UNK1, LYS745, CYS797, & GLY796 | 2.76935, 2.67256, 2.19566, 2.86768, & 3.27741 | UNK1, LYS745, VAL726 (2), LYS745, & PHE723 |
| 5                       | − 9.1                       | GLN791, LYS745, CYS797, ASN842, PRO794, & GLY796 | 2.8768, 2.83195, 2.80635, 3.65564, 3.22827, & 3.2556 | ASP855 (2), LEU718, VAL726 (2), & PHE723 |
| 6                       | − 9.3                       | GLN791, PRO794, & GLY796 | 2.87813, 3.44185, & 3.37399 | LYS745, ASP855, PHE723, LEU718, & VAL726 |
| 7                       | − 9                         | GLN791, ASP855, CYS797, ASN842, PRO794, & GLY796 | 2.98981, 2.09941, 2.78308, 3.61807, 3.31544, & 3.23449 | ASP855, VAL726 (2), LEU718, & PHE723 |
| 8                       | − 9.1                       | GLN791, ASP855, PRO794, & GLY796 | 2.92147, 2.69743, 3.3526, & 3.30683 | LYS745, LEU718, & VAL726 |
| 9                       | − 8.3                       | UNK1          | 3.48721          | LEU718 (2), LYS728, LEU792, VAL726, LEU844, & PHE723 |
| 12                      | − 8.5                       | GLN791, PRO794, & GLY796 | 2.91912, 3.33444, & 3.32438 | ASP855, LYS745 (2), VAL726 (2), LEU718, & PHE723 |
| 22                      | − 8.8                       | GLN791, GLN791, ARG841, PRO794, & GLY796 | 3.04006, 3.10979, 3.52195, 3.44184, & 3.31464 | LYS745, ASP855, PHE723, LEU718, VAL726, & LYS745 |
| 29                      | − 8.7                       | GLN791, PRO794, & GLY796 | 2.93488, 3.37212, & 3.30043 | ASP855, LYS745 (2), LEU718, ARG841, VAL726 (2), & PHE723 |
| 34                      | − 8.6                       | GLN791, PRO794, GLY796, & ARG841 | 2.89571, 3.36852, 3.34267, & 3.44426 | ASP855, LYS745 (2), LEU718, VAL726 (2), & PHE723 |
Molecular docking investigation of the newly designed compounds

The results of the molecular docking of the newly designed compounds are presented in Table 5 and Fig. 4a, b, respectively.

Drug-likeness and ADME property prediction of the newly designed compounds

The results of the drug-likeness and ADME properties prediction for the designed compounds are presented in Tables 6 and 7, respectively.
Molecular docking

Molecular docking virtual screening, one of the methods applied in structure-based drug design, was used to screen fifty sets of EGFR inhibitors in order to identify hit compound that could be used to design new EGFR inhibitors by investigating their binding interactions in the active site of EGFR tyrosine kinase enzyme (3IKA) (Supplementary Table 2). The binding affinities of this docking study presented in Supplementary Table 2 in kcal/mol range from −7.1 kcal/mol to −9.3 kcal/mol. Also, all possible interactions (hydrogen bond, hydrophobic, and electrostatic interactions) between the ligands and the target protein were shown in the same table. The results of the best top 10 compounds under investigation with higher binding affinities/lower dockings represent in Table 1 out of which the best three are discussed.

The best three hit compounds identified were compound 6 with the highest binding affinity of −9.3 kcal/mol, followed by compounds 5 and 8, each with the binding affinity of −9.1 kcal/mol, respectively. The best hit compound 6 bound to EGFR tyrosine kinase receptor via four different types of interactions including conventional hydrogen, carbon-hydrogen, electrostatic, and hydrophobic bond interactions, respectively. Nitrogen one (N1) of the quinazoline ring of compound 6 binds to EGFR tyrosine kinase receptor via conventional hydrogen bond interaction with GLN791 amino acid residue with a bond distance of 2.87813 Å. PRO794 and GLY796 amino acid residues formed carbon-hydrogen bonds with the active site of the receptor each with a bond distance of 3.44185 Å and 3.37399 Å, respectively. EGFR tyrosine kinase receptor interacted also via three hydrophobic interactions with PHE723, LEU718, and VAL726 amino acid residues with different parts of the compound, respectively. Furthermore, electrostatic interaction was observed between compound 6 and the active site of the receptor with LYS745 and ASP855 amino acid residues, respectively.

Compound 5 was found to bind with the active site of EGFR tyrosine kinase receptor via two hydrogen bonds (conventional and carbon-hydrogen), four hydrophobic, and two electrostatic interactions. The conventional hydrogen bonds were between the hydrogen attached to one of the nitrogen of the quinazoline ring, oxygen of the acrylamide moiety, and fluorine on the benzyl ring.

Table 2 Drug-likeness properties of ten top compounds

| Molecule  | MW    | TPSA  | WLOGP | No. of H-bond donors | No of H-bond acceptors | ROS violations |
|----------|-------|-------|-------|----------------------|------------------------|---------------|
| Molecule 3 | 501.46 | 114.47 | 4.44  | 3                    | 9                      | 1             |
| Molecule 5 | 534.97 | 94.6   | 4.11  | 2                    | 7                      | 1             |
| Molecule 6 | 500.52 | 94.6   | 4.46  | 2                    | 7                      | 1             |
| Molecule 7 | 517.96 | 107.49 | 4.95  | 2                    | 7                      | 1             |
| Molecule 9 | 400.43 | 85.37  | 3.54  | 2                    | 5                      | 0             |
| Molecule 12 | 428.84 | 85.37  | 4.69  | 2                    | 6                      | 0             |
| Molecule 22 | 557.62 | 97.84  | 4.39  | 2                    | 8                      | 1             |
| Molecule 29 | 526    | 88.61  | 4.16  | 2                    | 7                      | 1             |
| Molecule 34 | 560.04 | 124.89 | 4.61  | 2                    | 8                      | 1             |
| Molecule 43 | 441.52 | 79.38  | 3.85  | 2                    | 5                      | 0             |

Table 3 ADME properties of ten top compounds

| Molecule  | GI absorption | BBB permeant | Pgp substrate | Bioavailability score | Synthetic accessibility |
|----------|---------------|--------------|---------------|-----------------------|-------------------------|
| Molecule 3 | Low           | No           | No            | 0.55                  | 3.54                    |
| Molecule 5 | Low           | No           | Yes           | 0.55                  | 4.18                    |
| Molecule 6 | High          | No           | Yes           | 0.55                  | 4.11                    |
| Molecule 7 | High          | No           | Yes           | 0.55                  | 4.16                    |
| Molecule 9 | High          | No           | No            | 0.55                  | 3.67                    |
| Molecule 12 | High         | No           | No            | 0.55                  | 3.58                    |
| Molecule 22 | High         | No           | Yes           | 0.55                  | 4.63                    |
| Molecule 29 | High         | No           | Yes           | 0.55                  | 4.32                    |
| Molecule 34 | Low           | No           | Yes           | 0.55                  | 4.67                    |
| Molecule 43 | High          | No           | Yes           | 0.55                  | 3.48                    |
of compound 5 with ASP855 (2.8768 Å), LYS745 (2.83195 Å), and CYS797 (2.80635 Å) amino acid residue of the receptor, respectively. The carbon-hydrogen bond was between the carbon and oxygen attached to the tetrahydrofuran-3-yl moiety and also carbon of the fluoro benzyl ring with ASN842, PRO794, and GLY796 amino acid residues of the receptor each with a bond distance of 3.65564 Å, 3.22827 Å, and 3.2556 Å, respectively. A different part of the ligand was also observed to bind with the binding pocket of the receptor via two electrostatic and four hydrophobic bonds with ASP855 (2), LEU718, VAL726 (2), and PHE723 amino acids.

Compound 8 among the hit compounds formed two conventional hydrogen bonds with GLN791 and ASP855 amino acid residues, with a bond distance of 2.92147 Å and 2.69743 Å, respectively. Carbon-hydrogen bonds were also observed between the compound and the active site of the receptor with PRO794 (3.3526 Å) and GLY796 (3.30683 Å) residues, respectively. The compound was further observed to bind with the active site of the receptor protein via hydrophobic and electrostatic interactions with LYS745, LEU718, and VAL726 amino acid residues, respectively.

The common amino acid residues to these hit compounds were GLN791, LYS745, PRO794, GLY796, LEU718, and VAL726, respectively. These amino acid residues might be responsible for their higher binding affinities. Furthermore, Fig. 3a, b shows the 2D structures of compounds 6 and 8 using discovery studio visualizer.

**Drug-likeness and ADME property prediction of the studied compounds**

The drug-likeness and ADME properties of all the compounds under investigation were predicted using SWISSA DME online web tool. The drug-likeness and ADME properties of all the compounds are presented in Supplementary Tables 3 and 4 following the notable Lipinski’s rule of five. It states that the permeation of an orally administered compound is more likely to be better if the molecule satisfies the following conditions: (i) hydrogen bond donors ≤ 5 (OH and NH groups), (ii) hydrogen bond acceptors ≤ 10 (N and O atoms), (iii) molecular weight < 500, (iv) calculated WlogP < 5 (v), and TPSA ≤ 140, respectively. Any compound that violates more than two of these conditions may have bioavailability-related problems. The results of the best top 10 compounds are presented in Tables 2 and 3.

All the compounds satisfied the Lipinski’s rule of five without violating more than one of the conditions stated except compound 16 which has two violations. Thus, predicting their good permeability properties, easy transportation, absorption, and diffusion. The number of hydrogen bond acceptors and donors for all the compounds under investigation was less than 5 and 10, respectively, per the
| Ligand receptor (3IKA)         | Binding energy (kcal/mol) | Hydrogen bond                      | Bond distance (Å)                                                                 | Hydrophobic and other interactions                                                                 |
|-------------------------------|---------------------------|-----------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| 1                             | −10                       | ASP855, GLN791, & GLY796          | 2.79452, 3.28946, & 3.4809                                                      | LYS745 (3), ASP855, LEU718 (2), ALA743, LEU844, VAL726, & LEU858                                                                 |
| 2                             | −9.9                      | GLN791 & GLY796                  | 3.11371 & 3.36776                                                               | LYS745, ASP855, LEU844, PHE723, LEU718 (2), LEU747, ALA755, VAL726, & ALA743                                                                 |
| 3                             | −9.7                      | GLN791 & GLY796                  | 3.09658 & 3.39629                                                               | LYS745 (2), ASP855, LEU844, PHE723, LEU718 (2), ALA743, VAL726, & LEU858                                                                 |
| 4                             | −9.5                      | GLN791 & GLY796                  | 3.18431 & 3.36247                                                               | LYS745 (2), ASP855, PHE723, LEU718 (2), LYS875, ALA743, LEU844, VAL726, & LEU858                                                                 |
| 5                             | −9.8                      | LYS875, GLN791, PRO794, & GLY796 | 2.42771, 3.15611, 3.72875, & 3.37413                                             | LYS745 (2), ASP855, PHE723, LEU718 (2), ALA743, LEU844, VAL726, & LYS875                                                                 |
| 6                             | −10                       | ASP855, GLN791 & GLY796          | 2.65139, 3.25938, & 3.49579                                                      | LYS745 (3), ASP855, LEU844, LYS875, LEU718 (2), PRO877, LEU858 (2), ALA743, & VAL726                                                                 |
| 7                             | −10                       | ASP855, GLN791, & GLY796         | 2.67073, 3.26813, & 3.45576                                                      | LYS745 (3), ASP855, LEU718 (2), ALA743, LEU844, VAL726, & LEU858                                                                 |
| 8                             | −9.6                      | ASP855 & MET793                  | 2.50235 & 2.32716                                                               | ASP855, LEU718, LEU844, MET790, PHE723, VAL726, CY797, & ALA743                                                                 |
| 9                             | −9.7                      | ASN842, ASP837, MET793, ARG841, THR854, & PHE723 | 2.30969, 2.79936, 2.57049, 3.05655, 2.9311, & 2.86845 | ASP855, LEU718, GLY796, LEU844 (2), VAL726, LYS745, MET790 (2), LYS875, CY797, & ALA743                                                                 |
| 10                            | −10.2                     | ASP855, ARG841, LYS875, GLN791, & GLY796 | 2.25702, 2.53282, 2.46915, 3.18433, & 3.47632                                  | LYS745 (2), ASP855, LEU844, LEU718 (2), PRO877 (2), LYS875, VAL726, ALA743, & LEU858                                                                 |
| 11                            | −10                       | ASP855, GLN791, PRO794, & GLY796 | 2.74058, 3.225, 3.72385, & 3.34475                                               | LYS745 (3), ASP837, ASP855, LEU718 (2), ALA743, LEU844, VAL726, & LEU858                                                                 |
| 12                            | −9.5                      | ARG841, GLN791, PRO794, & GLY796 | 2.67415, 3.15579, 3.7195, & 3.31756                                               | LYS745, ASP855, PHE723, LEU718 (2), ALA743, LEU844, VAL726, & LYS745                                                                 |
| 13                            | −9.5                      | MET793, MET793, & LEU718         | 2.7385, 2.00214, & 3.73657                                                      | LYS745, ASP855, LEU718 (3), CY797 (2), ARG841, ALA743, LEU844, VAL726, & ARG481                                                                 |
| 14                            | −10.1                     | GLN791, GLY796, & PRO877         | 3.09152, 3.40727, & 3.29649                                                      | ASP837 (2), LYS745 (3), ASP855, LEU844, PHE723, PRO877, LEU718 (2), ALA743, VAL726, & LEU858                                                                 |
| 15                            | −9.6                      | GLN791 & GLY796                  | 3.07325 & 3.60625                                                                | LYS745 (2), ASP85, LEU844, PHE723, LEU718 (2), VAL726, ALA743, LEU844, & LEU858                                                                 |
| 16                            | −10.1                     | ARG841, ARG841, GLN791, & GLY796 | 2.90956, 2.69717, 3.21588, & 3.44122                                              | LYS745 (3), ASP837, ASP855, LEU718 (2), LYS875, PRO877, VAL726, ALA743, LEU844, & PHE723                                                                 |
| Afatinib                      | −7.9                      | GLN791, LEU718, ASP800, LEU718, MET793, & GLY796 | 3.01018, 3.79665, 3.63169, 3.54526, 3.11324, & 3.34059                             | ASP855, LYS745, LEU718 (2), VAL726 (2), ALA743, LEU844, & PHE723                                                                 |
notable RO5. From all these predicted parameters, it can be predicted that all the compounds under investigation including the three hit compounds can be orally bioavailable and also orally active as they obeyed the notable RO5.

For a drug to be orally active, it is expected to have high gastrointestinal absorption and all the compounds under investigation including the three hit compounds exhibited high gastrointestinal absorption except compounds 2, 16, 21, 34, and 39 with low gastrointestinal absorption. None of them was seen to permeant through the blood-brain barrier indicating lower toxicity. The most vital factor indicating good absorption was recognized to be the bioavailability score (which give the amount of drug present in the plasma). All the compounds under investigation including the three hit compounds were found to have high bioavailability scores of
except compound 16 which has a lower bioavailability score of 0.17. P-gp substrate served as a defender to the central nervous system (CNS) from xenobiotics. Also, the P-gp substrate of all the compounds under investigation including the 3 hit compounds was predicted. Some were found to be a substrate to P-gp while some were not. Moreover, synthetic accessibility scores refer to how easy a compound can be synthesized in a laboratory and scales of easy to hard range between 0 and 10. Synthetic accessibility score of all the compounds under investigation including the three (3) hit compounds was less than 5 showing that they can be easily synthesized in the laboratory. The compounds under investigation including the three hit compounds were predicted to have good pharmacokinetic profile and drug-likeness except compound 16 [24, 25].

| Table 6 Drug-likeness properties of the newly designed compounds |
|---------------------------------------------------------------|
| **Molecule** | **MW** | **TPSA** | **WLOGP** | **No. of H-bond donors** | **No of H-bond acceptors** | **ROS violations** |
|----------------|-------|---------|-----------|------------------------|--------------------------|------------------|
| Molecule 1     | 591.63| 106.63  | 7.2       | 3                      | 7                        | 2                |
| Molecule 2     | 674.55| 106.63  | 8.36      | 3                      | 7                        | 2                |
| Molecule 3     | 591.63| 106.63  | 7.2       | 3                      | 7                        | 2                |
| Molecule 4     | 626.08| 106.63  | 7.86      | 3                      | 7                        | 2                |
| Molecule 5     | 609.62| 106.63  | 7.76      | 3                      | 8                        | 2                |
| Molecule 6     | 626.08| 106.63  | 7.86      | 3                      | 7                        | 2                |
| Molecule 7     | 609.62| 106.63  | 7.76      | 3                      | 8                        | 2                |
| Molecule 8     | 674.55| 106.63  | 8.36      | 3                      | 7                        | 2                |
| Molecule 9     | 658.09| 106.63  | 8.51      | 3                      | 8                        | 2                |
| Molecule 10    | 641.64| 106.63  | 8.66      | 3                      | 9                        | 2                |
| Molecule 11    | 621.66| 115.86  | 7.21      | 3                      | 8                        | 2                |
| Molecule 12    | 621.66| 115.86  | 7.21      | 3                      | 8                        | 2                |
| Molecule 13    | 644.07| 106.63  | 8.42      | 3                      | 8                        | 2                |
| Molecule 14    | 644.07| 106.63  | 8.42      | 3                      | 8                        | 2                |
| Molecule 15    | 639.65| 115.86  | 7.77      | 3                      | 9                        | 2                |
| Molecule 16    | 656.1 | 115.86  | 7.87      | 3                      | 8                        | 2                |

| Table 7 ADME properties of the newly designed compounds |
|--------------------------------------------------------|
| **Molecule** | **GI absorption** | **BBB permeant** | **Pgp substrate** | **Bioavailability score** | **Synthetic accessibility** |
|----------------|-------------------|------------------|-------------------|--------------------------|-----------------------------|
| Molecule 1     | Low               | No               | No                | 0.55                     | 4.67                        |
| Molecule 2     | Low               | No               | No                | 0.55                     | 4.88                        |
| Molecule 3     | Low               | No               | No                | 0.55                     | 4.67                        |
| Molecule 4     | Low               | No               | No                | 0.55                     | 4.66                        |
| Molecule 5     | Low               | No               | No                | 0.55                     | 4.67                        |
| Molecule 6     | Low               | No               | No                | 0.55                     | 4.67                        |
| Molecule 7     | Low               | No               | No                | 0.55                     | 4.67                        |
| Molecule 8     | Low               | No               | No                | 0.55                     | 4.87                        |
| Molecule 9     | Low               | No               | No                | 0.17                     | 5.16                        |
| Molecule 10    | Low               | No               | No                | 0.55                     | 4.84                        |
| Molecule 11    | Low               | No               | No                | 0.55                     | 4.86                        |
| Molecule 12    | Low               | No               | No                | 0.55                     | 4.83                        |
| Molecule 13    | Low               | No               | No                | 0.17                     | 4.68                        |
| Molecule 14    | Low               | No               | No                | 0.17                     | 4.69                        |
| Molecule 15    | Low               | No               | No                | 0.55                     | 4.87                        |
| Molecule 16    | Low               | No               | No                | 0.55                     | 4.87                        |
**Designed compounds**

Based on the virtual screening carried out using molecular docking and pharmacokinetic studies on quinazoline derivatives, compound 6 with the highest binding affinity of \(-9.3\ \text{kcal/mol}\), good pharmacokinetic profile, and drug-likeness property was identified as the best hit compound in the analogs. Compound 6 being the best hit was used as the template for designing new compounds. Sixteen new compounds (Table 4) were designed by carrying out structural modification on the meta position of the fluorobenzyl ring of compound 6 (the template). Studying the designed compounds, the addition of phenyl-amino rings and halo substituted phenyl-amino rings on the meta position of the fluorobenzyl ring attached to "oxy-phenyl amino ring" moiety of the template significantly increase the binding affinities of the designed compounds.

**Molecular docking investigation of the newly designed compounds**

The newly developed compounds were also docked at the active site of EGFR tyrosine kinase receptor (PDB code: 3IKA). Table 5 shows the docking results of all the newly designed compounds in the active site of the target protein (EGFR tyrosine kinase receptor). The binding affinities of these newly designed compounds range from \(-9.5\ \text{kcal/mol}\) to \(-10.2\ \text{kcal/mol}\). When compared with the template and the control afatinib, the docked-designed compounds were seen to have better binding affinities than the template (\(f = 9.3\ \text{kcal/mol}\)) and the control (\(-7.9\ \text{kcal/mol}\)).

Designed compound SED10 was found to be the best among all designed compounds with a binding affinity of \(-10.2\ \text{kcal/mol}\). It interacted with the binding pocket of EGFR tyrosine kinase receptor via three conventional hydrogen bonds, two carbon-hydrogen bonds, four halogen bonds, three electrostatic bonds, and nine hydrophobic bonds with the following amino acid residues: ARG841, GLN791, GLY796, LYS745 (2), ASP855, LEU844, LEU718 (2), PRO877 (2), LYS875, VAL726, ALA743, and LEU858, respectively (Table 5).

The second-best among the designed compounds was SAD14. It bound to the binding pocket of its target via three carbon-hydrogen bonds, four halogen bonds, three electrostatic bonds, and nine hydrophobic bonds with the following amino acids GLN791, GLY796, ASP837 (2), LYS745 (3), ASP855, LEU844, PHE723, PRO877, LEU718 (2), ALA743, VAL726, and LEU858, respectively.

The following amino acid residues GLN791, GLY796, LYS745, ASP855, LEU844, LEU718, PRO877, VAL726, ALA743, and LEU858 were common to the best two designed compounds. This may be the reason they have higher binding affinities. Furthermore, Afatinib, the positive control, was used to validate the docking procedure in this study, then compared with the designed compounds. The designed compounds were observed to have better binding affinities than Afatinib with a binding affinity of \(-7.9\ \text{kcal/mol}\). The 2D-structures of designed compound SED10 and SED14 in complex with the receptor were presented in Fig. 4a, b.

**Drug-likeness and ADME prediction of the newly designed compounds**

The drug-likeness of the newly designed compounds was also predicted following Lipinski’s rule of five (Table 6). None of the designed compounds was found to violate more than two of the permissible limit set by the Lipinski’s rule of five filters for small molecules. Based on that, their permeability across the cell membrane, easy transportation, absorption, and diffusion was predicted [24, 26].

ADME properties of these newly designed compounds were also predicted (Table 7). All were observed to have low gastrointestinal absorption. But none was observed to permeate through the brain indicating lower toxicity. All designed compounds have higher bioavailability score of 0.55 except compounds 9, 13, and 14 with lower bioavailability scores of 0.17. Interestingly, all the designed compounds have a good synthetic accessibility score of < 5 except molecule 9 with the synthetic accessibility score > 5 (5.16) which indicates that these designed compounds can be easily synthesized in the laboratory [25, 27].

**Conclusion**

Molecular docking virtual screening carried out on fifty quinazoline derivatives/EGFR inhibitors reveals that compound 6 was the best hit compound among the investigated ones. Compound 6 is the best and was retained as the template for the structural modification for the design of new EGFR inhibitors.

The pharmacokinetic profile predictions of these hit compounds and the rest were further examined and found to be orally bioavailable with good absorption, low toxicity level, and permeable properties except compound 16.

Sixteen new EGFR inhibitors with better binding affinities than the template and Afatinib (the positive control) were designed using the best hit compound 6 as a template. The structural modification was made by adding a phenyl-amino ring, and halo substituted phenyl-amino ring on the meta position of the fluorobenzyl ring attached to “oxy-phenyl amino ring” moiety of the template might be responsible for the significant increases in the binding affinity of the designed compounds.

None of the designed compounds was found to violate more than the permissible limit set by RO5, thereby predicting their transportation, absorption, and diffusion. More so, the designed compounds were found to have good synthetic accessibility which indicates that these designed compounds can be synthesized in the laboratory.
Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43042-020-00107-y.

Additional file 1: Supplementary Table 1. The structures of the compounds under investigation. Supplementary Table 2. The binding affinity and different interactions of the ligands under investigation with EGFR tyrosine kinase receptor. Supplementary Table 3. Drug-likeness of properties of compounds under investigation. Supplementary Table 4. ADME properties of compounds under investigation.

Abbreviations

DFT: Density function theory; B3LYP: Becke’s three-parameter read-Yang-Parr hybrid; PDB: Protein data bank; NSCLC: Non-small cell lung cancer agents; EGFR: Epidermal growth factor receptor

Acknowledgements

The authors acknowledge the technical effort of Ahmadu Bello University, Zaria, Nigeria.

Authors’ contributions

MTI: contributed throughout the research work. AU: gives directives and technical advice. GAS: partake in the technical activities. SU: also partake in the technical activities. All authors have read and approved the manuscript.

Funding

This research did not receive any funding from anybody.

Availability of data and materials

All data and materials are available upon request.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

All authors have declared that there is no conflict of interest regarding this submission.

Received: 9 July 2020 Accepted: 30 October 2020

Published online: 19 November 2020

References

1. Song J, Jang S, Lee JW, Jung D, Lee S, Min KH (2019) Click chemistry for improvement in selectivity of quinazoline-based kinase inhibitors for mutant epidermal growth factor receptors. Bioorg Med Chem Lett 29(3):477–480
2. Hanai E, Baumgardner M, Bryan MC, Chen Y, Eigenbrot C, Fan P et al (2016) 4-Aminoindazolyl-dihydrofuro [3, 4-d] pyrimidines as non-covalent inhibitors of mutant epidermal growth factor receptor tyrosine kinase. Bioorg Med Chem Lett 26(2):534–539
3. Kong L-L, Ma R, Yao M-Y, Yan X-E, Zhu S-J, Zhao P et al (2017) Structural pharmacological studies on EGFR T790M/C797S. Biochem Biophys Res Commun 488(2):266–272
4. Balak MN, Gong Y, Rely GJ, Somwar R, Li AR, Zakowski MF et al (2006) Novel D761Y and common secondary T790M mutations in epidermal growth factor receptors. Bioorg Med Chem Lett 29(3):477–480
5. Dungo RT, Keating GM (2013) Afatinib: first global approval. Drugs. 73(13):1503–1515
6. Solca F, Dahl G, Zoepfle A, Bader G, Sanderson M, Klein C et al (2012) Target binding properties and cellular activity of afatinib (BIBW 2992), an irreversible Erbb family blocker. J Pharmacol Exp Ther 342(2):342–350
7. Tsao M-S, Sakurada A, Cutz J-C, Zhu C-Q, Karmel-Reid S, Squie J et al (2005) Erlotinib in lung cancer—molecular and clinical predictors of outcome. N Engl J Med 353(2):133–144
8. Cross DA, Ashton SE, Ghiorghiu S, Eberlein C, Nebhan CA, Spitzler PJ et al (2014) AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. Cancer Discov 4(9):1046–1061
9. Walter AO, Sjøn RTT, Haringsma HJ, Oishi K, Sun J, Lee K et al (2013) Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC. Cancer Discov 3(12):1404–1415
10. Park K, Lee J-S, Lee KH, Kim J-H, Cho BC, Min YJ et al (2016) Olmutinib (BI 1487294, HM61713), an EGFR mutant-specific inhibitor, in T790M+ NSCLC: efficacy and safety at the RP2D. J Clin Oncol 34(suppl):abstr 9055
11. Paz-Ares L, Tan E-H, Obinaya K, Zhang L, Hirsh V, Boyer M et al (2017) Afatinib versus gefitinib in patients with EGFR mutation-positive advanced non-small-cell lung cancer: overall survival data from the phase IIb LUX-Lung 7 trial. Ann Oncol 28(2):270–277
12. Song Z, Ge Y, Wang C, Huang S, Zhu X, Liu K et al (2016) Challenges and perspectives on the development of small-molecule EGFR inhibitors against T790M-mediated resistance in non-small-cell lung cancer: miniperspective. J Med Chem 59(14):6580–6594
13. Beheshti A, Pourbasheer E, Nekoei M, Vahdani S (2016) QSAR modeling of antimalarial activity of urea derivatives using genetic algorithm—multiple linear regressions. J Saudi Chem Soc 20(3):282–290
14. Khan MF, Verma G, Akhtar W, Shakiquezzaman M, Akhter M, Rizvi MA et al (2019) Pharmacophore modeling, 3D-QSAR, docking study and ADME prediction of acyl 1, 3-4-thiazolide amides and sulfonamides as antitubulin agents. Arab J Chem. 12:5000–18.
15. Zhang L, Yang Y, Zhou H, Zheng Q, Li Y, Zheng S et al (2015) Structure-activity study of quinazoline derivatives leading to the discovery of potent EGFR-T790M inhibitors. Eur J Med Chem 102:445–463
16. Ibrahim MT, Uzairu A, Shallangwa GA, Ibrahim A (2018) Computational studies of some biscoumarin and biscoumarin thiourea derivatives asso-glucosidase inhibitors. J Eng Exact Sci 4(2):0276–0285
17. Mills N. ChemDraw Ultra 10.0 CambridgeSoft, 100 CambridgePark Drive, Cambridge, MA 02140. www.cambridgesoft.com. Commercial Price: 1910 fordordownload, 2150 for CD-ROM; Academic Price: 710 fordordownload, 800 for CD-ROM. ACS Publications; 2006.
18. Kohn W, Becke AD, Parr RG (1996) Density functional theory of electronic structure. J Phys Chem 100(31):12974–12980
19. Ibrahim MT, Uzairu A, Shallangwa GA, Ibrahim A (2020) In-silico studies of some oxadiazoles derivatives as anti-diabetic compounds. J King Saud Univ Sci. 32:423-432.
20. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 7:42717
21. Ismail SY, UZAIR A, Sagagi B (2019) In silico studies of sulfur-containing shikonin oxime derivatives as inhibitors of mgc 803 gastric cancer cell line. J Eng Exact Sci 5(1):0079–0089
22. Ibrahim MT, Uzairu A, Uba S, Shallangwa GA (2020) Computational modeling of novel quinazoline derivatives as potent epidermal growth factor receptor inhibitors. Heliyon. 6(2):e03289
23. Batool M, Ahmad B, Choi S (2019) A structure-based drug discovery paradigm. Int J Mol Sci 11(2):2783
24. Khan I, Gankapati KR, Setti A, Shaiq AB, Makani VK, Shareef MA et al (2019) Design, synthesis, in silico pharmacokinetics prediction and biological evaluation of 1, 4-dihydroindeno [1, 2-c]-pyrazole chalcone as EGFR/Act pathway inhibitors. Eur J Med Chem 163:636–648.
25. Hosen S, Dash R, Khatun M, Akter R, Bhuiyan MHR, Rezaul M et al (2017) Pharmacophore modeling, 3D-QSAR, docking study and ADME prediction of acylamide derivatives and sulfonamides as anti-tubulin agents. J Mol Liq 237:385–395
26. Ibrahim MT, Uzairu A, Shallangwa GA, Ibrahim A (2020) In-silico studies of some oxadiazoles derivatives as anti-diabetic compounds. J King Saud Univ Sci. 32:423-432.
27. Patool M, Ahmad B, Choi S (2019) A structure-based drug discovery paradigm. Int J Mol Sci 11(2):2783
28. Khan I, Gankapati KR, Setti A, Shaiq AB, Makani VK, Shareef MA et al (2019) Design, synthesis, in silico pharmacokinetics prediction and biological evaluation of 1, 4-dihydroindeno [1, 2-c]-pyrazole chalcone as EGFR/Act pathway inhibitors. Eur J Med Chem 163:636–648.
29. Hosen S, Dash R, Khatun M, Akter R, Bhuiyan MHR, Rezaul M et al (2017) In silico ADME/T and 3D QSAR analysis of KDR inhibitors. J Appl Pharm Sci 7(1):120–128
30. Ibrahim MT, Uzairu A, Shallangwa GA, Uba S (2020) Computer-aided molecular modeling studies of some 2, 3-dihydropyridin-1(2H) diazo [2, 3-4] quinazoline derivatives as EGFR WT inhibitors. Beni-Suef Univ J Basic Appl Sci 9:1–10
31. Ibrahim MT, Uzairu A, Shallangwa GA, Uba S (2020) In-silico activity prediction and docking studies of some 2, 3-disubstituted 8-phenylthio/phenylsulfinyl-9 h-purine derivatives as anti-proliferative agents. Heliyon. 6(1):e03158

Publisher's Note

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