Computational virtual screening and structure-based design of some epidermal growth factor receptor inhibitors

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

Abstract

Background: The foremost cause of cancer mortality worldwide was lung cancer. Lung cancer is divided into small cell lung cancer and non-small cell lung cancer (NSCLC). The latter is the main type of lung cancer that account for about 90% of the cancer issues and estimate about 25% of the cancer mortality each year in the world. Among the types of lung cancer with about 1.5 million patients and less than 20% survival rate is NSCLC. Overexpression of EGFR tyrosine kinase was recognized to be the cause of NSCLC. Therefore, there is a need to develop more EGFR inhibitors due to drug-resistance development by the mutation.

Result: Computational virtual screening on some epidermal growth factor receptor inhibitors (EGFRL858R/T790M inhibitors or NSCLC therapeutic agents) against their target protein (EGFR tyrosine kinase receptor pdb entry 3IKA) was performed via molecular docking simulation and pharmacokinetics to identify hit compounds with a promising affinity toward their target. The hit compounds discovered were compound 22 with $-9.8$ kcal/mol, 24 with $-9.7$ kcal/mol, 17 with $-9.7$ kcal/mol, and 19 with $-9.5$ kcal/mol respectively. These lead compounds were further subjected to drug-likeness and ADME prediction and found to be orally bioavailable. Six (6) new EGFRL858R/T790M inhibitors using compound 22 with the highest binding affinity as a template were designed.

Conclusion: The six newly EGFRL858R/T790M inhibitors were found to have a better binding affinity than the template used in the designing process and AZD9291 (the positive control). None of the designed compounds was found to violate more than the permissible limit set by ROS thereby predicting their easy transportation, absorption, and diffusion. More so, the designed compounds were found to have good synthetic accessibility which indicates that these designed compounds can be easily synthesized in the laboratory.

Keywords: Computational, Docking, EGFRL858R/T790M, Inhibitors, ADME, Drug-likeness

Background

Lung cancer is one of the leading cancer problems in the globe. It was reported to cause a lot of death every year (estimated to take about one-third of the entire cancer deaths). Non-small cell lung cancer (NSCLC) is the main subset of lung cancers that accounts for about 85% of the cancer problems [1]. Overexpression of epidermal growth factor receptor kinase was identified to be the common cause of NSCLCs. Report on NSCLCs on the population of patients in the Caucasia rise to about 10–15% and 30–40% in Asia [1]. The discovery of NSCLC therapeutic agents for the treatment of EGFR tyrosine kinase is one of the major challenges encountered by the medicinal chemist [2]. The treatment of EGFR tyrosine kinase to managed NSCLCs became a very urgent therapeutic necessity [3].

NSCLC therapeutic agents show a very high response rate in patients with arousing modifications of EGFR. NSCLC therapeutic agents or EGFR inhibitors are classified into reversible EGFR inhibitors (first-generation EGFR inhibitors); gefitinib and erlotinib are the example of this class of EGFR inhibitors.
Table 1 The Structures of the data set

| S/No | Structure |
|------|-----------|
| 1    | ![Structure 1](image1) | 15   | ![Structure 15](image15) |
| 2    | ![Structure 2](image2)  | 16   | ![Structure 16](image16) |
| 3    | ![Structure 3](image3)  | 17   | ![Structure 17](image17) |
| 4    | ![Structure 4](image4)  | 18   | ![Structure 18](image18) |
| 5    | ![Structure 5](image5)  | 19   | ![Structure 19](image19) |
| 6    | ![Structure 6](image6)  | 20   | ![Structure 20](image20) |
| 7    | ![Structure 7](image7)  | 21   | ![Structure 21](image21) |
| 8    | ![Structure 8](image8)  | 22   | ![Structure 22](image22) |
| 9    | ![Structure 9](image9)  | 23   | ![Structure 23](image23) |
| 10   | ![Structure 10](image10)| 24   | ![Structure 24](image24) |
inhibitors include. Unluckily, the span of the potency of these first-generation EGFR inhibitors is narrowed due to the development of drug resistance by the secondary mutation T790M [4]. And the irreversible EGFR inhibitors (second- and third-generation EGFR inhibitors), afatinib and osimertinib, are the examples of these EGFR inhibitors. In line to defeat the resistance to the first-generation of EGFR inhibitors, the second-generation irreversible EGFR inhibitors, such as afatinib and canertinib, were afterward devised to treat NSCLC EGFR T790M mutation [5]. Yet, due to severe side effects, such as skin rash and diarrhea, these second-generation inhibitors cannot attain satisfaction over the first-generation reversible inhibitors. It is believed that the activities upon wild-type EGFR will narrow the possible activities on the patients with the T790M mutation [2, 6−8].

To approach the unmet clinical demands, many third-generation irreversible EGFR inhibitors, such as WZ-4002, rociletinib, olmutinib, and osimertinib were designed to inhibit the T790M resistance mutation while being more selective for wild type EGFR [2, 9−12].

Molecular docking is a molecular modeling technique used in structure-based design to screen a library of compounds to identify compounds with a higher affinity toward their target protein by elucidating their mode of interaction with their target utilizing their 3D structures [13]. Pharmacokinetics and drug-likeness properties prediction of hit compounds play a vital role in structure-based design in the determination of the pharmacokinetic profile of the hit compounds under investigation in the early stage of the drug pipeline [14].

This work is aimed at carrying out computational virtual screening on some EGFR L858R/T790M inhibitors using molecular docking to identify hit compounds with a promising affinity against their target receptor (EGFR tyrosine kinase receptor), confirm their bioavailability via their pharmacokinetics and drug-likeness properties, and design new potent EGFR L858R/T790M inhibitors that have better binding affinity than the template.

Method
This computational work was done on a Dell personal computer laptop, with these specifications: Intel “Core”? i7 Dual CPU, M330 @2.75 GHz 2.75 GHz, and 8 GB of RAM. The following software was utilized to achieve the success of this research: Pyrex virtual screening software,
Table 2 The interactions of the molecules under investigation in the active site of EGFR receptor Complex Binding affinity (Kcal/mol) H Bond Bond distance (Å) Hydrophobic, electrostatic and Other Interactions
Complex 1 −8.2 LYS745 LYS875 GLY857 PHE723 2.4833 3.59309 3.72533 2.6702 ASP855, LEU718, PHE723 and LEU844
Complex 2 −8.0 ASP855 MET790 LYS745 ASP855 MET793 GLY857 PHE723 2.98309 2.54859 2.53945 3.28835 3.63827 3.69781 2.74564 ASP855, LEU718, PHE723 and LEU844
Complex 3 −8.3 LYS745 ASN842 ASP855 ASP837 GLY857 PHE723 2.55654 3.60971 3.46169 3.79142 3.5789 2.70142 ASP855, LEU718, PHE723, LEU718 and LEU844
Complex 4 −8.6 LYS745 ASN842 GLY857 PHE723 2.57931 3.47241 3.71569 2.53735 ASP855, MET790, PHE723, LEU718(3), VAL726, ALA743 and LEU844 (2)
Complex 5 −8.4 MET790 LYS745 ASN842 ASP855 GLY857 PHE723 2.63263 2.57166 3.5002 3.48749 3.7168 3.61079 2.69448 ASP855, LEU718 (2), PHE723, LEU792 and LEU844
Complex 6 −8.1 MET790 ASP855 LYS745 GLY857 PHE723 2.63344 2.45808 2.51691 2.75431 3.42724 3.68209 ASP855, LEU718, PHE723 and LEU844
Complex 7 −8.2 MET790 LYS745 ASN842 GLY857 PHE723 2.66828 2.50517 3.64733 3.71982 2.63929 ASP855, LEU718 (2), PHE723 and LEU844
Complex 8 −8.2 MET790 LYS745 ASP855 GLY857 PHE723 2.69124 2.49593 3.31777 3.56354 3.70766 2.66609 ASP855, PHE723, LEU718 (2) and LEU844
Complex 9 −8.1 LYS745 LYS745 GLY857 PHE723 2.47818 2.56406 3.79409 3.76916 2.6504 PHE723, LEU718 (2) and LEU844
Complex 10 −8.2 LYS745 LYS745 ASP855 GLY857 2.499 2.82934 3.37238 3.46133 3.67217 ASP855, LEU844, PHE723, and ALA743
Complex 11 −8.2 LYS745 LYS745 ASN842 GLY857 2.49554 2.73608 3.6529 3.53467 3.69398 ASP855, LEU844, PHE723, ALA743, MET790, LEU844, LEU718, and ALA743
Complex 12 −8.3 LYS745 ASP855 GLY857 PHE723 2.57602 ASP855, PHE723, LEU718 (2), and LEU844
Table 2 The interactions of the molecules under investigation in the active site of EGFR receptor (Continued)

| Complex | Binding affinity (Kcal/mol) | H Bond | Bond distance (Å) | Hydrophobic, electrostatic and Other Interactions |
|---------|----------------------------|--------|------------------|-----------------------------------------------|
|         |                            |        |                  |                                               |
| 12      | −8.0                       | LYS745LYS745ASP855GLY857 | 2.51678 | ASP855, LEU718, PHE723, LEU718, LEU792, MET793, LEU844, LEU844 |
|         |                            |        |                  |                                               |
| 13      | −8.3                       | ASP855LYS745GLY857      | 2.56789 | ASP855, LEU718 (2), PHE723 and LEU844         |
|         |                            |        |                  |                                               |
| 14      | −8.3                       | MET790LYS745ASN842ASP855GLY857PHE723 | 2.57269 | ASP855, PHE723, LEU718 (2), and LEU844       |
|         |                            |        |                  |                                               |
| 15      | −9.0                       | MET790LYS745ASP837ASP855GLY857PHE723 | 2.59334 | ASP855, A:LEU718 (3), PHE723, LEU792, and LEU844 |
|         |                            |        |                  |                                               |
| 16      | −9.0                       | ASP855LYS745GLY857      | 2.48189 | ASP855, PHE723, LEU718 (2), and LEU844       |
|         |                            |        |                  |                                               |
| 17      | −9.5                       | MET790LYS745ASP837ASP855GLY857PHE723 | 2.59334 | ASP855, A:LEU718 (3), PHE723, LEU792, and LEU844 |
|         |                            |        |                  |                                               |
| 18      | −8.8                       | LYS745GLY796GLY857PHE723 | 2.56789 | ASP855, LEU844, MET790, PHE723, LEU718 (3), ALA743, and LEU844 (2) |
|         |                            |        |                  |                                               |
| 19      | −8.8                       | LYS745GLY796ASP800      | 2.48189 | ASP855, MET790, PHE723, LEU718 (2), ALA743, and LEU844 (2) |
|         |                            |        |                  |                                               |
| 20      | −9.0                       | LYS745GLY796ASP800      | 2.48189 | ASP855, MET790, PHE723, LEU718 (2), ALA743, and LEU844 (2) |
|         |                            |        |                  |                                               |
| 21      | −9.0                       | LYS745GLY796ASP800      | 2.48189 | ASP855, MET790, PHE723, LEU718 (2), ALA743, and LEU844 (2) |
|         |                            |        |                  |                                               |
| 22      | −8.8                       | LYS745GLY796ASP800      | 2.48189 | ASP855, MET790, PHE723, LEU718 (2), ALA743, and LEU844 (2) |
|         |                            |        |                  |                                               |
| 23      | −8.8                       | LYS745GLY796ASP800      | 2.48189 | ASP855, MET790, PHE723, LEU718 (2), ALA743, and LEU844 (2) |
|         |                            |        |                  |                                               |
| 24      | −9.0                       | LYS745GLY796ASP800      | 2.48189 | ASP855, MET790, PHE723, LEU718 (2), ALA743, and LEU844 (2) |
Table 2 The interactions of the molecules under investigation in the active site of EGFR receptor (Continued)

| Complex | Binding affinity (Kcal/mol) | H Bond | Bond distance (Å) | Hydrophobic, electrostatic and Other Interactions |
|---------|----------------------------|--------|-------------------|---------------------------------------------------|
| Complex 25 | -9.3 | LYS745ASP837GLY857PHE723 | 2.53735 | ASP855, LEU718(2), LEU844 (2), MET790, PHE723, VAL726, ALA743 |
| Complex 26 | -7.9 | LYS745ARG841ASP837GLY857 | 2.72291 | ASP855, LEU718, VAL726, ALA743, and LEU844 |
| Complex 27 | -7.6 | ASP855MET793 | 2.60582 | PHE723 and LEU858 |
| Complex 28 | -8.1 | ASP855 | 2.80197 | ASP855, PHE723, and LEU858 |

Fig. 3 2D and 3D structures of (a) Complex 22, (b) Complex 24, (c) Complex 17, and (d) Complex 19 using Discovery studio visualizer
Source and sketching of dataset under investigation
Twenty-eight (28) sets of EGFR$_{L858R/T790M}^{-}$ inhibitors were gotten from the work of Hu et al. [15] and used in this research. Immediately after the retrieval of the data, the next thing is drawing of all the molecules under investigation in 2D format. Chemdraw software was then used to draw the 2D structures of all the molecules under investigation [16]. Table 1 presents the structures of all the data set under investigation.

Determination of the optimum structures under investigation
Determination of the most stable/optimum geometry of all the molecules on potential energy surface (PES) was achieved by the use of Spartan 14 wave software in this research. B3LYP/6-311G* level of theory with density functional theory (DFT) was used to achieve the searching for the optimum structures under investigation [17].

Ligands, EGFR enzyme preparation and execution of the molecular docking simulation
Ligands preparation is very vital in any molecular docking studies. As such, the preparation of the ligands in this work was done using the optimum geometry of each of the ligands obtained in 2.1 above before the elucidation of their binding interactions and the binding pose of the EGFR enzyme [18]. Figure 1 shows the 3D geometry of a prepared EGFR$_{L858R/T790M}^{-}$ inhibitor (ligand) under investigation.

Fig. 4 2D and 3D structures of (a) Complex 22, (b) Complex 24, (c) Complex 17, and (d) Complex 19 using PyMOL
The EGFR enzyme with protein data bank code: 3IKA was retrieved from the RCSB protein data bank database. After successful retrieval of the enzyme, the preparation of the EGFR enzyme for the molecular docking simulation was done using discovery studio visualizer, in the process of the preparation of the enzyme, the co-crystalline ligand and molecule of water present on the structure were deleted. Before that, polar hydrogen was added. Figure 2 shows the 3D structure of the prepared EGFR enzyme.

The docking of the ligands to the binding pose of the EGFR enzyme was achieved with the help of Autodock vina of Pyrex virtual screening software [19]. After a successful docking procedure, since Pyrex was used there is a need to re-couple the docked ligand and the receptor for further investigation. UCSF Chimera software was used for the re-coupling of the docked ligand and the receptor. PyMOL and Discovery studio were used to achieve the visualization of recoupled complexes in order to view the nature of the interaction between the ligand and the receptor.

**Drug-likeness and ADME properties prediction**

Pharmacokinetics and drug-likeness of the EGFR(R858R/T790M) inhibitors under investigation were determined using SwissADME, a free online web tool [20]. Lipinski’s rule of five was the criteria used in the determination of the drug-likeness of the molecules under investigation which states that if any small molecule violates more than 2 of these criteria, the molecules might not be orally bioavailable [21].

**Design**

Structure-based drug design is a very robust and useful technique. Structure-based drug design is also called direct design which involves the acquisition of the information regarding the three-dimensional structure of the

### Table 3 Pharmacokinetics properties

| Molecule   | MW      | No. of H-bond acceptors | No. of H-bond donors | TPSA | WLOGP | No. of Lipinski’s rule violations |
|------------|---------|-------------------------|----------------------|------|-------|----------------------------------|
| Molecule 1 | 486.57  | 6                       | 2                    | 100.44 | 3.5   | 0                                |
| Molecule 2 | 516.59  | 7                       | 2                    | 109.67 | 3.51  | 1                                |
| Molecule 3 | 500.6   | 6                       | 2                    | 100.44 | 3.81  | 0                                |
| Molecule 4 | 528.65  | 6                       | 2                    | 100.44 | 4.62  | 1                                |
| Molecule 5 | 514.62  | 6                       | 2                    | 100.44 | 4.06  | 1                                |
| Molecule 6 | 504.56  | 7                       | 2                    | 100.44 | 4.06  | 1                                |
| Molecule 7 | 521.01  | 6                       | 2                    | 100.44 | 4.15  | 1                                |
| Molecule 8 | 565.46  | 6                       | 2                    | 100.44 | 4.26  | 1                                |
| Molecule 9 | 612.47  | 6                       | 2                    | 100.44 | 4.1   | 1                                |
| Molecule 10| 504.56  | 7                       | 2                    | 100.44 | 4.06  | 1                                |
| Molecule 11| 521.01  | 6                       | 2                    | 100.44 | 4.15  | 1                                |
| Molecule 12| 565.46  | 6                       | 2                    | 100.44 | 4.26  | 1                                |
| Molecule 13| 555.46  | 6                       | 2                    | 100.44 | 4.81  | 1                                |
| Molecule 14| 583.46  | 7                       | 2                    | 100.44 | 4.82  | 1                                |
| Molecule 15| 539     | 7                       | 2                    | 100.44 | 4.71  | 1                                |
| Molecule 16| 480.61  | 6                       | 2                    | 100.44 | 3.7   | 0                                |
| Molecule 17| 526.63  | 6                       | 2                    | 100.44 | 3.62  | 1                                |
| Molecule 18| 610.45  | 6                       | 2                    | 100.44 | 3.1   | 1                                |
| Molecule 19| 513.59  | 6                       | 2                    | 106.43 | 4.08  | 1                                |
| Molecule 20| 597.41  | 6                       | 2                    | 106.43 | 3.56  | 1                                |
| Molecule 21| 628.46  | 6                       | 3                    | 120.41 | 3.4   | 1                                |
| Molecule 22| 542.63  | 6                       | 3                    | 120.41 | 2.91  | 1                                |
| Molecule 23| 626.45  | 6                       | 3                    | 120.41 | 2.39  | 1                                |
| Molecule 24| 529.59  | 6                       | 3                    | 126.4  | 3.37  | 1                                |
| Molecule 25| 613.41  | 6                       | 3                    | 126.4  | 2.86  | 1                                |
| Molecule 26| 515.61  | 6                       | 3                    | 112.47 | 3.79  | 1                                |
| Molecule 27| 543.66  | 6                       | 3                    | 112.47 | 4.83  | 1                                |
| Molecule 28| 569.7   | 6                       | 3                    | 112.47 | 5.37  | 1                                |
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 affinity 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 [22].

Results

Molecular docking simulation

The results of the molecular docking simulation are presented in Table 2 and Figs. 3 and 4 respectively.

Drug-likeness and ADME properties prediction

The results of the drug-likeness and ADME properties prediction are presented in Tables 3 and 4, Figs. 5 and 6 respectively.

Table 4 Boiled-egg and CYP isoforms inhibition

| Molecule | GI absorption | BBB permeant | Pgp substrate | CYP1A2 inhibitor | CYP2C19 inhibitor | CYP2C9 inhibitor | CYP2D6 inhibitor | CYP3A4 inhibitor |
|----------|----------------|--------------|---------------|------------------|------------------|------------------|------------------|------------------|
| Molecule 1 | High           | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 2 | High           | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 3 | High           | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 4 | High           | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 5 | High           | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 6 | High           | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 7 | High           | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 8 | High           | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 9 | High           | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 10 | High         | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 11 | High          | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 12 | High          | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 13 | High          | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 14 | High          | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 15 | High          | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 16 | High          | No           | No            | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 17 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 18 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 19 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 20 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 21 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 22 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 23 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 24 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 25 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 26 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 27 | High          | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |
| Molecule 28 | Low           | No           | Yes           | No               | Yes              | Yes              | Yes              | Yes              |

Molecular docking of designed compounds

The results of the molecular docking of designed compounds are presented in Tables 5 and 6 and Fig. 7 respectively.

Drug-likeness and ADME properties prediction

The result of the drug-likeness and ADME properties prediction is presented in Tables 7 and 8, respectively.

Discussion

Molecular docking simulation

Molecular docking simulation was used to screen twenty-eight (28) sets of EGFR L858R/T790M inhibitors in order to identify hit compounds that could be used to design new EGFR L858R/T790M inhibitors by investigating their binding interactions in the binding pose of EGFR receptor (3IKA) (Table 2). The result of the four best hit
compounds with the lowest docking scores/highest binding affinity will be discussed.

Compound 22 was the best among the four selected compounds that have the lowest docking score of $-9.8$ kcal/mol due to the major number of interactions in the binding pocket of the enzyme. Discovery studio visualizer was used to investigate its interaction in the binding pose of the enzyme, it was seen to interact with MET790 (2.65 Å), LYS745 (2.67 Å), ASP855 (3.21 Å), GLY857 (3.69 Å) and PHE723 (2.63 Å) amino acid residues in the active site of EGFR receptor via both conventional and carbon-hydrogen bond interactions. Beside conventional and carbon-hydrogen bond interactions, it also bound to LEU844, PHE723, LEU718 (3), ALA743, and LEU844 residues via Pi-Sigma, Pi-Sulfur, Pi-Pi Stacked, Alkyl, and Pi-Alkyl hydrophobic interactions. Pi-Anion electrostatic interaction with ASP855 and Pi-Sulfur interaction with MET790 were also observed.

The second best with a binding affinity of $-9.7$ kcal/mol was compound 24. It bound with LYS745 (2.58 Å), ASN842 (3.47 Å), GLY857 (3.72 Å), and PHE723 (2.54 Å) residues in the binding pose of the receptor through conventional and carbon-hydrogen bond interactions. Pi-Sigma, Pi-Pi Stacked, Alkyl, and Pi-Alkyl hydrophobic interactions were also observed with ASP855, LEU844 (2), PHE723, LEU718 (3), VAL726, and ALA743 residues. Apart from the interaction mentioned, Pi-Anion electrostatic interaction with ASP855 and Pi-Sulfur interaction with MET790 amino acid residues respectively were also seen.

The third best in the trend is compound 17 which also found to bound via conventional and carbon-hydrogen bond interactions with MET790 (2.57 Å), LYS745 (2.58 Å), ASN842 (3.57 Å), ASP855 (3.31 Å), GLY857 (3.71 Å), and PHE723 (2.69 Å) amino acid residues respectively. Apart from conventional and carbon-hydrogen bond interactions, it interacted via...
Pi-Sigma, Pi-Pi Stacked, Alkyl, and Pi-Alkyl hydrophobic interactions with PHE723, LEU718 (3), LEU792, and LEU844 amino acid residues and also via Pi-Sulfur with MET790 amino acid residue in the binding pose of the receptor. The last one in the trend is compound 19 which also bound with the active site of the receptor via conventional and carbon-hydrogen bond interactions, Pi-Sigma, Pi-Pi Stacked, Alkyl, Pi-Alkyl hydrophobic interactions, and Pi-Anion electrostatic interactions as shown in Table 2. Figures 3 and 4 showed the 2D and 3D structures of the four lead compounds investigated using discovery studio visualizer and Pymol.

Drug-likeness and ADME properties prediction of the studied compounds

Table 3 presents the computed drug-likeness of the compounds under investigation. It was observed in the table that none of the molecules under investigation violated more than the maximum permissible limit of the criteria stated by Lipinski’s filters, it therefore means that there is a high tendency that all of these molecules might be pharmacologically very active. In fact, all these molecules under investigation are said to have good absorption, low toxicity level, orally bioavailable, and permeable properties except molecule 28 which has WlogP value (it predicts whether a molecule has low toxicity level or not) greater than 5. The Bioavailability Radar of the four selected molecules under investigation was shown to further confirm their drug-likeness properties (Fig. 5). The compounds under investigation could be said to be orally bioavailable.

Table 4 presents the gastrointestinal (GI) absorption, blood-brain barrier (BBB) permeant, Pgp substrate, and CYP isoforms inhibition properties of all the molecules under investigation. From the table, all the molecules under investigation have high GI absorption, none has BBB permeant, some were found to be able to permeate through the skin and some cannot, also all were observed to inhibit the CYP isoforms except CYP1A2. The boiled-egg plot was performed to further confirm the GI absorption and BBB permeant properties of the four hit compounds (Fig. 6). It is further confirmed from the plot that none of them passed through the BBB but they were within the GI absorption region.

Molecular docking of designed compounds

Six new EGFR\textsuperscript{L858R/T790M} inhibitors were designed using compound 22 with the highest binding affinity of $-9.8$ kcal/mol as the template (Table 5). Based on the interaction of compound 22 with the EGFR receptor, structural modifications were carried out on the template by the addition of substituents on the piperazin-1-yl moiety and isopropyl phenyl ring of the template.

The addition of acetyl group on the piperazin-1-yl moiety and 2 chlorine molecules at the meta position of the isopropyl phenyl ring of the template showed a significant increase in the interaction of the designed compound (D3) with the EGFR receptor with $-10.2$ kcal/mol binding energy. It was found to bind with the EGFR receptor through conventional and carbon-hydrogen bonds, hydrophobic, electrostatic, and other interactions (Table 6). Four amino acid residues (ASP855, MET790, LYS745, and LYS745) of the enzyme with bond distance 2.9622 Å, 2.49526 Å, 2.61911 Å, and 2.38759 Å were
### Table 5 The designed compounds with their binding affinities

| S/NO | Structures | Docking scores (kcal/mol) |
|------|------------|---------------------------|
| D1   | ![D1 Structure](image1) | -10                        |
| D2   | ![D2 Structure](image2) | -10                        |
| D3   | ![D3 Structure](image3) | -10                        |
| D4   | ![D4 Structure](image4) | -10.2                      |
| D5   | ![D5 Structure](image5) | -9.9                       |
| D6   | ![D6 Structure](image6) | -10.1                      |
| AZD9291 | ![AZD9291 Structure](image7) | -10                        |

Template Compound 6 (Binding energy = -9.8)
observed to form a conventional hydrogen bond with a different part of the ligand as depicted in Fig. 7a. Carbon-hydrogen bond was also observed in the binding pocket of the enzyme between these two amino acid residues ASP855 (3.24379 Å) and PHE723 (2.57647 Å) and the ligand. The ten (10) amino acid residues in the binding pocket of the enzyme who interacted with the ligands via hydrophobic interaction were LEU844 (2), MET790, PHE723, LEU718 (3), LEU792, CYS797, and ALA743 (2) respectively. Besides the mentioned interactions, electrostatic interaction was also observed between the ligand (D3) and ASP855 residue in the binding pocket of the receptor. The only amino acid who interacted via Pi-Sulfur interaction (other) was MET790. The addition of only the acetyl group on the piperazin-1-yl moiety of the template yielded significant change also in the interaction of the designed compound (D5) with the EGFR receptor with a very good binding affinity of −10.1 kcal/mol (Table 5). Designed compound D5 bounded to EGFR receptor via a hydrogen bond, hydrophobic interactions, and other interaction as shown in Table 6. The same number of the conventional

Table 6 The interactions of the designed compounds in the active site of the EGFR receptor

| Ligand receptor (3IKA) | Binding energy (kcal/mol) | Hydrogen bond | Bond distance (Å) | Hydrophobic and other interactions |
|------------------------|---------------------------|---------------|------------------|-----------------------------------|
| 1                      | −10                       | ASP855, MET790, LYS745, LYS745, ASP855 & PHE723 | 2.91264, 2.46131, 2.6733, 2.46143, 3.29523 & 2.58927 | ASP855, LEU718 (3), MET790, PHE723, LEU792, ALA743 & LEU844 (2) |
| 2                      | −10                       | ASP855, MET790, LYS745, ASP855 & PHE723 | 2.81294, 2.45287, 2.74079, 3.36729 & 2.75977 | ASP855, LEU844 (3), MET790, PHE723, LEU718 (4), LEU792, CYS797 & ALA743 (2) |
| 3                      | −10.2                     | ASP855, MET790, LYS745, LYS745, ASP855 & PHE723 | 2.9622, 2.49526, 2.61911, 2.38759, 3.24379 & 2.57647 | ASP855, LEU844 (2), MET790, PHE723, LEU718 (3), LEU792, CYS797 & ALA743 (2) |
| 4                      | −9.9                      | ASP855, THR854 & PHE723 | 2.6829, 2.50494 & 3.54003 | ASP855, LEU718, GLY796, PHE723, GLY857, ALA755, LEU747, ILE759, VAL726 & LEU858 (2) |
| 5                      | −10.1                     | ASP855, MET790, LYS745, LYS745, ASP855 & PHE723 | 2.83123, 2.35804, 2.58025, 2.43397, 3.31176 & 2.70789 | ASP855, LEU718 (3), MET790, PHE723, LEU792, ALA743 & LEU844 (2) |
| 6                      | −10                       | MET790, LYS745, ASP855 & PHE723 | 2.45034, 2.68227, 3.3187 & 2.46677 | ASP855, LEU718 (3), MET790, PHE723, LEU792, LEU858 (2), LYS875 (2), ALA743 & LEU844 |
| AZD9291                | −8.1                      | ASP855, ASP837 & GLY857 | 2.39196, 3.61636 & 3.55164 | MET790, LEU718, VAL726 LEU844 (2) & ALA743 (2) |

Template: Compound 22 (binding energy = −9.8)

Table 7 Drug-likeness properties of the designed compounds

| Molecule | MW     | TPSA | WLOGP | No. of H-bond donors | No. of H-bond acceptors | ROS violations |
|----------|--------|------|-------|----------------------|-------------------------|----------------|
| Molecule 1 | 588.63 | 137.48 | 3.38 | 3 | 10 | 2 |
| Molecule 2 | 637.1 | 137.48 | 4.13 | 3 | 11 | 2 |
| Molecule 3 | 653.56 | 137.48 | 3.98 | 3 | 11 | 2 |
| Molecule 4 | 655.7 | 132.44 | 4.23 | 4 | 11 | 2 |
| Molecule 5 | 570.64 | 137.48 | 2.83 | 3 | 6 | 1 |
| Molecule 6 | 598.7 | 137.48 | 3.46 | 3 | 6 | 1 |
hydrogen bond, carbon-hydrogen bond, electrostatic, and Pi-Sulfur (other) interactions were observed between D5 and the receptor except in the hydrophobic interaction where there were eight amino acids which interacted with the ligand. The four amino acid residues with the bond distance that interacted via conventional hydrogen bond with a different part of the ligand as shown in Fig. 7b were ASP855 (2.83123 Å), MET790 (2.35804 Å), LYS745 (2.58025 Å), and LYS745 (2.43397 Å) respectively. The two amino acids that were observed to the carbon-hydrogen bond in the binding pocket of the enzyme and the ligand were ASP855 (3.31176 Å) and PHE723 (2.70789 Å). The eight (8) amino acid residues in the binding pocket of the enzyme which interacted with the ligands via hydrophobic interaction were LEU718 (3), PHE723, LEU792, ALA743, and LEU844 (2) respectively. Besides the mentioned interactions, ASP855 residue was the only that form electrostatic interaction between the ligand and in the binding pocket of the receptor and MET790 was the only residue who interacted via Pi-Sulfur (other) interaction. This might be possible as the result of not having halogens in the designed compound 5 (D5) which is why the number of hydrophobic interactions were less than that of D3.

The other designed compounds (D1, D2, D4, and D6) showed good interactions with higher binding affinity in the binding pocket of the EGFR tyrosine kinase receptor (Table 6). They were observed to have interacted with the binding pocket of the enzyme via the same conventional hydrogen, carbon-hydrogen bond, hydrophobic, electrostatic, and Pi-Sulfur (other) interactions except D4 which has not interacted via

**Table 8 ADME Properties of the designed compounds**

| Molecule | GI absorption | BBB permeant | Pgp substrate | Bioavailability Score | Synthetic Accessibility |
|----------|---------------|--------------|---------------|----------------------|------------------------|
| Molecule 1 | Low | No | Yes | 0.17 | 4.22 |
| Molecule 2 | Low | No | Yes | 0.17 | 4.88 |
| Molecule 3 | Low | No | No | 0.17 | 4.37 |
| Molecule 4 | Low | No | No | 0.17 | 4.68 |
| Molecule 5 | Low | No | Yes | 0.17 | 4.17 |
| Molecule 6 | Low | No | Yes | 0.17 | 4.4 |

![Fig. 7 2D-structure of designed compound (a) D3 and (b) D5 in complex with EGFR enzyme](image)
Drug-likeness and ADME prediction of designed compounds

Using the Lipinski’s rule of five as a standard filter for small molecule, the drug-likeness of the designed compounds were also predicted as presented in Table 7. From the table, no any designed compound was found to violate more than the permissible limit set by Lipinski’s rule of five filters and therefore predicting their easy transportation, absorption, and diffusion [23, 24].

ADME properties of these designed compounds were also predicted and presented in Table 8. All were observed to have low gastrointestinal absorption. But none was observed to permeant through the brain. All designed compounds have a lower bioavailability score of 0.17. Based on the synthetic accessibility score (Table 8), they can all be synthesized in the laboratory [25, 26].

Conclusion

In conclusion, molecular docking simulation carried out on the twenty-eight (28) EGFRL858R/T790M inhibitors has identified four hit compounds with a higher binding affinity toward their target. The hit compounds discovered were compound 22 with −9.8 kcal/mol, 24 with −9.7 kcal/mol, 17 with −9.7 kcal/mol, and 19 with −9.5 kcal/mol respectively. These lead compounds were further subjected to drug-likeness and ADME prediction and found to be orally bioavailable with good absorption, low toxicity level, and permeable properties. The best among the hit compounds was retained as a template and used to design six new EGFRL858R/T790M inhibitors with better binding affinity than the template and AZD9291 (the positive control). None of the designed compounds was found to violate more than the permissible limit set by RO5 thereby predicting their easy 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.

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 technical activities. SU also partake in 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: 5 June 2020 Accepted: 6 August 2020

Published online: 18 August 2020

References

1. Kong L-L, Ma R, Yao M-Y, Yan X-E, Zhu S-L, Zhao P, Yun C-H (2017) Structural pharmacological studies on EGFR T790M/C797S. Biochem Biophys Res Commun 488(2):266–272
2. 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
3. Hanan EJ, Baumgardner M, Bryan MC, Chen Y, Eigenbrot C, Fan P, Gu X-H, La H, Malek S, Purkey KE (2016) 4-Aminindazolyl-dihydrofuro [3, 4-d] pyrimidines as non-covalent inhibitors of mutant epidermal growth factor receptor tyrosine kinase. Bioorg Med Chem Lett 26(5):534–539
4. Balak MN, Gong Y, Riely GJ, Somwar R, Li AR, Zakowski MF, Chiang A, Yang G, Ouefelli O, Kris MG (2006) Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor–mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. Clin Cancer Res 12(21):6494–6501
5. Dungo RT, Keating GM (2013) Aftatinib: first global approval. Drugs 73(13):1503–1515
6. Solca F, Dahl G, Zoepfel A, Bader G, Sanderson M, Klein C, Kraemer O, Himmelbach F, Haaksma E, Adolph GR (2012) Target binding properties and cellular activity of aflatinib (BBBW 2992), an irreversible ErbB family blocker. J Pharmcol Exp Ther 343(2):342–350
7. Tsao M-S, Sakurada A, Cutz J-C, Zhu C-Q, Kernl-Reid S, Squire J, Lorimer I, Zhang T, Liu N, Daneshmand M (2005) Erlotinib in lung cancer—molecular and clinical predictors of outcome. N Engl J Med 353(2):133–144
8. Cross DA, Ashton SE, Ghorgiu S, Eberlein C, Nebhan CA, Spitzler PJ, Orme JP, Finlay MRV, Ward RA, Mellor MJ (2014) AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. Cancer discovery 4(9):1046–1061
9. Walter AO, Sijin RTT, Haringsma HJ, Ohashi K, Sun J, Lee K, Dubrovskiy A, Labinski M, Zhu Z, Wang Z (2013) Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC. Cancer discovery 3(12):1404–1415
10. Park, K., J.-S. Lee, K.H. Lee, J.-H. Kim, B.C. Cho, Y.J. Min, J.Y. Cho, J.-Y. Han, B.-S. Kim, and J.-S. Kim, (2016), Olmutinib (BI 1482694; 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, Obyrne K, Zhang L, Hirsch V, Boyer M, Yang JH, Mck T, Lee K, Lu S (2017) Afatinib versus gefitinib in patients with EGFR mutation-positive advanced non-small-cell lung cancer: overall survival data from the phase IIIb LUX-Lung 7 trial. Ann Oncol 28(2):270–277
12. Song Z, Ge Y, Wang C, Huang S, Shu X, Liu K, Zhou Y, Ma X (2016) Challenges and perspectives on the development of small-molecule EGFR inhibitors.

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.
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. Journal of Saudi Chemical Society 20(3):282–290

14. Khan, M.F., G. Verma, W. Akhtar, M. Shauquzzaman, M. Akhter, M.A. Rizvi, and M.M. Alam, (2016), Pharmacophore modeling, 3D-QSAR, docking study and ADME prediction of acyl 1, 3, 4-thiadiazole amides and sulfonylamides as antitubulin agents, Arabian Journal of Chemistry.

15. Hu J, Han Y, Wang J, Liu Y, Zhao Y, Liu Y, Gong P (2018) Discovery of selective EGFR modulator to inhibit L858R/T790M double mutants bearing a N-9-Diphenyl-9H-purin-2-amine scaffold. Bioorg Med Chem 26(8):1810–1822

16. Ibrahim MT, Uzairu A, Shallowanga GA, Ibrahim A (2018) Computational studies of some biscoumarin and bisoumarin thiourea derivatives ASo-GLUCOSIDASE INHIBITORS. The Journal of Engineering and Exact Sciences 4(2):0276–0285

17. Kohn W, Becke AD, Parr RG (1996) Density functional theory of electronic structure. J Phys Chem 100(31):12974–12980

18. Adeniji, S.E., S. Uba, and A. Uzairu, (2018), Quantitative structure-activity relationship and molecular docking of 4-Alkoxy-Cinnamic analogues as anti-mycobacterium tuberculosis, Journal of King Saud University-Science.

19. Ibrahim, M.T., A. Uzairu, G.A. Shallowanga, and A. Ibrahim, (2018), In-silico studies of some oxadiazoles derivatives as anti-diabetic compounds, Journal of King Saud University-Science.

20. Batool M, Ahmad B, Choi S (2019) A structure-based drug discovery paradigm. Int J Mol Sci 20(11):2783

21. Ibrahim MT, Uzairu A, Shallowanga GA, Uba S (2020) Computer-aided molecular modeling studies of some 2, 3-dihydro-1, 4 dioxino [2, 3-f] quinazoline derivatives as EGFR WT inhibitors. Beni-Suef University Journal of Basic and Applied Sciences 9:1–10

22. Khan I, Sarikapati KR, Setti A, Shaik AB, Makani VKK, Shareef MA, Rajpurohit H, Vangara N, Pal-Bhadra M, Karnal A (2019) Design, synthesis, in silico pharmacokinetics prediction and biological evaluation of 1, 4-dihydropyridino [1, 2-c] pyrazole chalcone as EGFR/Akt pathway inhibitors. Eur J Med Chem 163:636–648

23. Ibrahim MT, Uzairu A, Shallowanga GA, Uba S (2020) In-silico activity prediction and docking studies of some 9-disubstituted 8-phenylthio/phenylsulfanyl-9-h-purine derivatives as Anti-proliferative agents. Heliyon 6(1):e03158

24. Ismail, S.Y., A. Uzairu, B. Sagagi, and M. Sabiu, (2018), In silico molecular docking and pharmacokinetic study of selected phytochemicals with estrogen and progesterone receptors as anticancer agent for breast cancer, 5 (3), 1337-1350.

25. Ibrahim MT, Uzairu A, Shallangwa GA, Uba S (2020) Computer-aided molecular modeling studies of some 2, 9-disubstituted 8-phenylthio/phenylsulfanyl-9-h-purine derivatives as Anti-proliferative agents. Journal of Applied Pharmaceutical Science 7(01):120–128

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