In-silico screening of inhibitor on protein epidermal growth factor receptor (EGFR)

To cite this article: Jeremi Ongko *et al* 2022 *IOP Conf. Ser.: Earth Environ. Sci.* **1041** 012075

In-silico screening of inhibitor on protein epidermal growth factor receptor (EGFR)

Jeremi Ongko, Jesica Viona Setiawan, Alfina Gracia Feronytha, Agnes Juliana, Andy Effraim, Mariana Wahjudi, Yulanda Antonius*

Universitas Surabaya, Surabaya, Indonesia

*yulandaantonius@staff.ubaya.ac.id

Abstract. The screening process to detect early-stage lung cancer is injurious to a patient's survival. Fortunately, there are natural compounds that have been acknowledged to possess anticancer properties, work as the protein binding inhibitors of lung cancer promotors: EGF and EGFR. The study aims to identify inhibitors of EGFR protein binding. Assessments were accomplished based on several parameters related to EGFR proteins, such as pathways, protein activity, conformational changes, and numerous information using the STRING database and KEGG pathway database. Ten inhibitor compounds that expressed highest activity were selected for further analysis were: (20R,22R)-5beta,6beta-Epoxy-4beta,12beta,20-trihydroxy-1-oxowith-2-en-24-enolide, irinotecan, flavopyridol, teniposide, exatecan, daphnoretin, indirubin, topitecan, wentilactone, and evidiamine. The native ligand Lapatinib was used as positive control in this analysis. The analysis was accomplished by molecular docking using Vina 4 in the PyRx software. Interactions between the ligands and residues were investigated using LIGPLOT+ 2.2. The In-silico analysis of the ten candidate compounds revealed that (20R, 22R)-5beta, 6beta-Epoxy-4beta, 12beta, 20-trihydroxy-1-oxowith-2-en-24-enolide expressed the lowest binding energy value, which is -10.4 kcal/mol, indicated the closest binding energy value to Lapatinib as the control. Based on the interaction of amino acids, (20R,22R)-5beta, 6beta-Epoxy-4beta, 12beta, 20-trihydroxy-1-oxowith-2-en-24-enolide has excellent potential to be utilized as next inhibitor com- pound candidates for EGFR protein, because it binds to the Lys745 residue. It mirrors the positive control and has a binding energy on the range of the specified acceptable parameters.

1. Introduction

Epidermal Growth Factor Receptor (EGFR) protein belongs to the ErbB receptor family and a cell surface receptor for members of the epidermal growth factor family of extracellular ligands. EGFR has a key role in signal transduction processes by regulating major cellular functions, such as cell proliferation and apoptosis. The EGFR protein is a tyrosine kinase receptor with a size of 170 kDa located at the surface of the cell and activated by binding the specific ligands, including Epidermal Growth Factor (EGF). EGFR has four functional domains, namely: extracellular ligand-binding domain, transmembrane domain, intracellular tyrosine kinase domain, and C-terminal regulatory do- main.

The C-terminal regulatory domain has several phosphorylated tyrosine kinase domains that specializes in ligand binding. After binding to the ligand, dynamic conformational changes will occur in the extracellular and intracellular domains of the kinase receptor eventuating to transphosphorylation of tyrosine residues in the C-terminal regular tory domain.

This will provide a docking location for downstream molecules that leads to the activation of multiple pathways, such as mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase/AKT, signal transduction and the activation of the STAT3 & STAT5 transcription pathways. Activation of

these pathways will eventuate in the blocking process of apoptosis, proliferation, invasion, and metastasis is all important for the cancer phenotype [1]. EGFR is a transmembrane receptor protein tyrosine kinase that is expressed on the epithelial, mesenchymal, and neurogenic tissues. Overexpression of EGFR is implicated in the pathogenesis of many tissue malignancies, including non-small cell lung cancer.

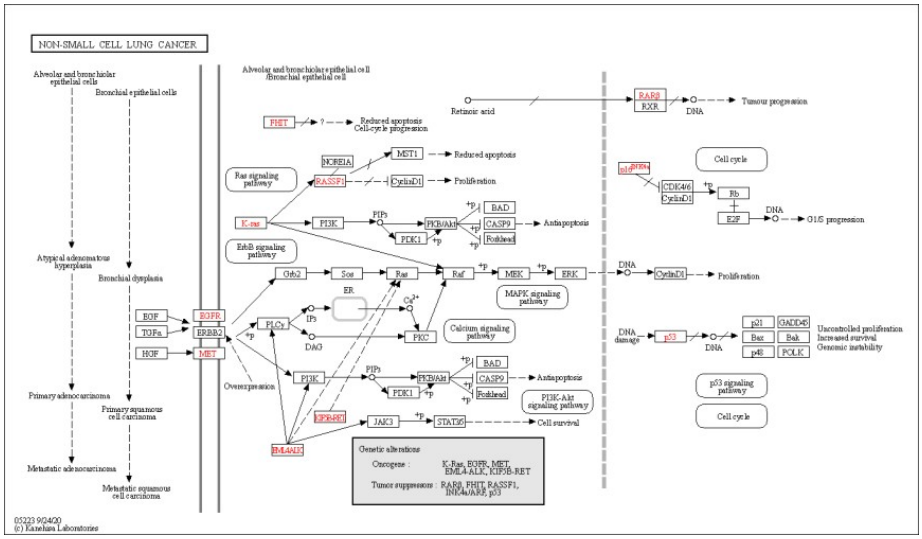


Figure 1. Non-small Lung Cancer Pathway from KEGG

When a specific ligand such as EGF and TGFα binds to EGFR, the EGFR which should functions normally will undergo a conformational change and phosphorylation occurs in the intracellular domain which eventuate the downstream signal transduction from numerous pathways such as Raf1-extracellular signal-regulated kinase, PI3K/AKT, signal transduction, and the activation of the STAT transcription factor. Depending on the pathway (Figure 1), result of the signal could be an uncontrollable cell proliferation as the result of disrupted of apoptotic mechanism [2]. Lung cancer is a complex type of disease and is one of the main causes of cancer mortality in the world [3].

The cancer survival rate can be elevated, depends on the stage that detected at initial medical diagnosis. Therefore, the screening process aims to detect early-stage lung cancer [4]. There are several acknowledged compounds that naturally available and have potential activities as the binding inhibitor between EGF and EGFR protein. Thanks to the discovery of the anticancer properties in these compounds, the therapeutic drugs for cancer derived from these compounds had been successfully manufactured.

For example, there are ethyl acetate and ethanol extracts which contains friedelan-3-one, stigmast-5-en-3-ol, palmitic acid, kaur-16-en-18-oic acid, and benzoic acid that discovered from Thai herbal plants such as Bridelia ovata, Croton oblongifolius, and Erythrophleum succirubrum. This signifies the urgent necessity to perform the screening process of natural compounds that importantly involve in the treatment or prevention of lung cancer by inhibiting EGF binding with EGFR protein to instigate apoptosis [5]. Therefore, this research aimed to identify the compounds that have potential as candidates of EGFR protein inhibitor.

2. Methodology

2.1 Preparation of EGFR Protein

An assessment was accomplished based on several parameters regarding EGFR protein activities such as pathways, protein activity, conformational changes, and other supplementary parameters. The assessment supported by the STRING database (https://stringdb.org/), the KEGG pathway database (https://www.genome.jp/ kegg/path- way.html), and other scientific research reports. The study revealed that the EGFR protein with the code PDB: 1XKK binds to the ligand of the Lapatinib compound as a compound that has been used for the lung cancer treatment [6]. The protein was further modified using ChimeraX to remove water molecules, free compounds, Lapatinib ligands, and provide polar hydrogen bonds to these proteins.

2.2 Literature Study of EGFR Protein Inhibitor Candidates and Compounds Preparation

Studies have also been accomplished. The potential compounds that capable to inhibit the activity of EGFR protein by replacing the attachment of specific ligands that trigger lung cancer has been investigated through various journals and traced into PubChem (https://pubchem.ncbi.nlm.nih.gov/) to obtain their structure. Forty-five inhibitor candidate compounds were selected for further analysis and one positive control were used to redock. The positive control in this analysis was Lapatinib, it was the native ligand and conformed inhibitor compound in the activity of EGFR protein [6]. The 3D structure of the compounds was stored in .SDF form and then was submitted into the PyRx program to be measured by application's Openable to determine its minimum energy and minimize the structure of the ligand until it was ready for dock ing.

2.3 Candidate Compound Screening for EGFR Protein Inhibitor (Molecular Docking)

Screening was accomplished by molecular docking using Vina 4 within the PyRx program version 0.8 [7]. Docking location for center position X: 18.3392, Y: 35.8643, Z: 38.1427 and grid box dimensions X: 16.4623 Å, Y: 13.3738 Å, Z: 20.3377 Å. The analysis is considered success if the binding energy results' value id exceed or draw close to the binding energy results of the Lapatinib as control compound.

2.4.Visualization and Amino Acids Interaction Analysis of EGFR Proteins with Candi- date Compounds

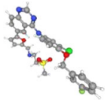
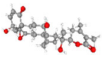
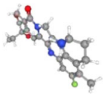
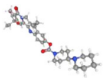
The EGFR protein was visualized with ligands of secondary metabolite com- pounds that passed the predetermined parameters using the Chimera X [8], to inspect the position of the hydrogen bond between the protein and the candidate compound. The visualized compounds were stored in .PDB form. Then each compound was investigated using LIGPLOT+ software version 2.2 to examine the interaction between protein amino acids and ligands [9].

3. Result and Discussion

The reason of the using Lapatinib as the control as candidate comparison is the ability to elevate cytotoxicity against lung cancer cells. Lapatinib exerts its activity intra- cellularly by competing with ATP for the ATP-binding domain in the cytoplasmic tail of the tyrosine kinase receptor. Many natural secondary metabolite compounds derived from various plants have anticancer properties [10]. The analysis was accomplished based on the binding energy between the EGFR protein and the 10 compounds to determine excellent candidates for attachment inhibitor of specific ligands to EGFR pro teins as lung cancer origin.

The molecular docking analysis was run using PyRx program by comparing the binding energy score of 10 compounds against the binding energy score of the Lapatinib control compound, which is -11.0 and setting a minimum score parameter of -10.0 to ensure an effect is produced. The results of the candidate compounds obtained are as follows:

Table 1. Screening Results of Secondary Metabolite Candidates to Control (Lapatinib) based on the Binding Energy Score

No.	Name of Compound	PubChemID	Binding Energy (kcal/mol)	Model 3D Ligands
1	Lapatinib (Positive control)	208908	-11.0	
2	(20R,22R)-5beta,6beta-Epoxy- 4beta,12beta,20-tri-hydroxy-1-oxowith-2-en-24-enolide	46872824	-10.4	
3	Exatecan	151115	-10.2	
4	Irinotecan	60838	-10.1	

Based on the results of the binding energy analysis between candidate compounds and EGFR protein using PyRx, six candidate compounds were considered to be potential inhibitors. They obtained the highest binding energy value among the 10 compounds; the values were close to the binding energy value of Lapatinib. Those six candidate compounds are: a) (20R,22R)-5beta,6beta-Epoxy-4beta,12beta,20-trihydroxy-1- oxowith-2-en-24-enolide with ΔG = -10.4 kcal/mol, b) Exatecan with ΔG = -10.2 kcal/mol, c) Irinotecan with ΔG = -10.1 kcal/mol, d) Flavopyridol with ΔG = -10.1 kcal/mol, and e) Teniposide with ΔG = -10.1 kcal/mol.

Meanwhile, the binding energy value between Lapatinib and EGFR protein is -11.0 kcal/mol. The binding energy demonstrates the affinity/bond between the candidate compound (ligand) and the EGFR protein. When the smaller/negative binding energy is obtained, the stronger/stable the bonds are formed [11]. Furthermore, the results of the docking were visualized between the six best candidate compounds and Lapatinib with EGFR protein using Chimera X.

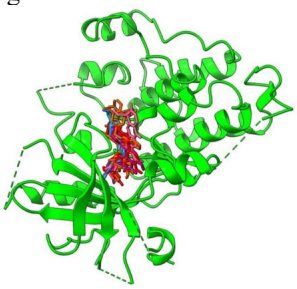


Figure 2. Protein - Ligand Visualization Results based on Docking Results between EGER (PDB: 1XKK)

All the ligands above bind at the same location and a little is seen in some compounds that have hydrogen bonds with proteins while some such as teniposide and irinotecan do not have bonds with proteins. The interaction between amino acid of EGFR protein with ligands of the five candidate compounds and Lapatinib was analyzed using the LIGPLOT+ program.

Table 2. Results of the Interaction of EGFR Protein Amino Acids with Ligands

Name of complex	*Hydrogen Bond*	*Hydrophobic Bond*
EGFR – Lapitanib	Thr790: 3.04 Å Lys745: 3.02 Å	Met1002, Leu792, Gly796, Met793, Ala743, Leu844, Gln791, Met766, Thr854, Cys775, Phe856, Asp855, Leu858, Asn842, Gly721, Val726, Ser720, Arg841, Leu718, Gly719
EGFR – (20R,22R)-5beta,6beta- Epoxy-4beta,12beta,20- trihydroxy-1- oxowith-2-en-24- enolide	Asp855: 2.88 Å Lys745: 3.18 Å	Leu718, Ile744, Leu844, Thr790, Ala743, Leu777, Ile789, Arg841, Thr854, Leu788, Val726, Gly719
EGFR – Irinotecan	-	Leu718, Gly796, Leu844, Thr790, Ala743, Leu777, Cys775, Met766, Arg776, Phe856, Thr854, Asp855, Lys745, Val726, Gly719
EGFR - Flavopiridol	Met793: 2.93 Å	Cys797, Leu718, Met1002, Gly796, Leu792, Ala743, Leu844, Lys745, Thr790, Asp855, Gly719, Thr854, Gly721, Val726, Ser720
EGFR - Teniposide	-	Arg841, Cys797, Leu718, Met1002, Gly796, Gly719, Ala743, Thr854, Leu844, Leu777, Thr790, Arg776, Cys775, Asp855, Met766, Lys745, Val726, Ser720, Gly721, Leu799
EGFR – Exatecan	Asp855: 3.30 Å	Leu799, Asp800, Arg841, Met793, Leu844, Leu792, Ala743, Thr854, Val726, Gly721,

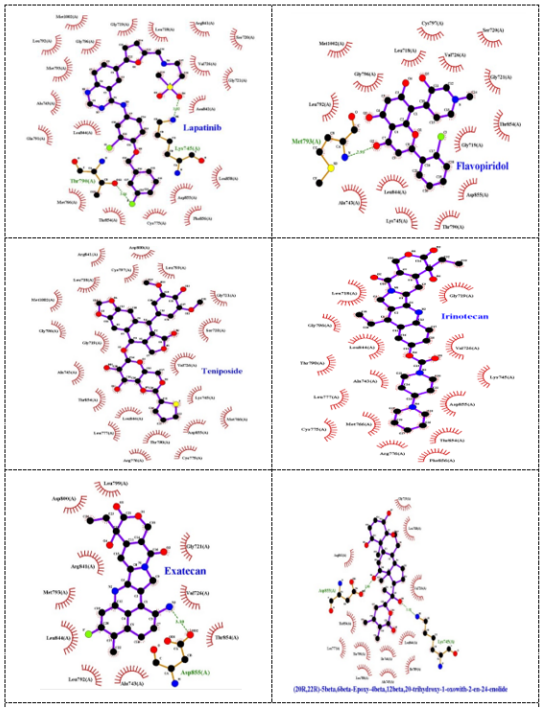


Figure 3. Interaction of amino acid residues within EGFR protein with potential ligands

The observation of hydrogen bonds and amino acids that interact between EGFR proteins and ligands of candidate compounds can be done using the LIGPLOT+ program. The interaction analysis between EGFR proteins with these ligands revealed that the Lapatinib formed two hydrogen bonds with Thr790 and Lys745 from the EGFR ATP binding pocket. The amino acids in EGFR that interact with Lapatinib are Met1002, Leu792, Gly796, Met793, Ala743, Leu844, Gln791, Met766, Thr854, Cys775, Phe856, Asp855, Leu858, Asn842, Gly721, Val726, Ser720, Arg841, Leu718, and Gly719.

Irinotecan and Teniposide compounds were unable to form hydrogen bonds with EGFR. It is probably due to a different active site that influenced the different inhibition path from the binding pocket. In case of the flavopyridol candidate compound, it has one hydrogen bond with Met793 from EGFR at a distance of 2.93 Å. Exatecan compounds has one hydrogen bond with Asp855 from the EGFR at a distance of 3.30 Å.

Then the (20R,22R)-5beta,6beta-Epoxy- 4beta,12beta,20-trihydroxy-1-oxowith-2-en-24-enolide compound formed two hydrogen bonds with Asp855 and Lys745. Lys745 were the same amino acid that Lapatinib, as the positive control, have interaction with. Therefore, with the interaction with Lys745, the (20R,22R)- 5beta,6beta-Epoxy-4beta,12beta,20-trihydroxy-1-oxowith-2-en-24-enolide compound have a potential as an EGFR protein inhibitor that could be further developed. These results indicate that (20R,22R)-5beta,6beta-Epoxy-4beta,12beta,20-trihydroxy-1-oxowith-2-en-24-enolide compounds are the most excellent candidate as the EGFR protein inhibitor, which has similar mechanism pathway with Lapatinib.

4. Conclusion

Based on literature review and in silico analysis, this research obtained the top ten candidate compounds based on the lowest binding affinity and sorted it into the five most excellent candidate compounds. The binding energy analysis discovered that (20R,22R)-5beta,6beta- Epoxy-4beta,12beta,20-trihydroxy-1-oxowith-2-en-24-enolide compounds have the lowest binding energy value, which is -10.4 kcal/mol, which has the closest value to the binding energy of Lapatinib (positive control). The interaction of amino acids confirmed that, (20R,22R)- 5beta,6beta-Epoxy-4beta,12beta,20-trihydroxy-1-oxowith-2-en-24-enolide is the most promising compound to be further developed as EGFR protein inhibitor opposing lung cancer development, as it binds to the Lys745 residue, copying the Lapatinib mechanism resulting binding energy within the specified acceptable parameters.

Acknowledgment

Authors thank to Faculty of Biotechnology, University of Surabaya for supporting this research.

References

[1] Nwizu, Tobenna and Salgia R 2012 Molecular Biology of Lung Cancer and Therapeutic Implications Emerging Cancer Therapeutics 3(1):1–10.
[2] Bethune G, Bethune D, Ridgway N and Xu Z 2010 Epidermal growth factor receptor (EGFR) in lung cancer: an overview and update Journal of Thoracic Disease 2(1):48–51.
[3] Parris B, O'Farrell H, Fong K and Yang I 2019 Chronic obstructive pulmonary disease (COPD) and lung cancer: common pathways for pathogenesis. Journal Of Thoracic Disease, 1(1), S2155-S2172.
[4] Wu G X and Raz D 2016 Lung Cancer Screening Cancer Treatment and Research 170:1-23.
[5] Poofery J, Khaw-on P, Subhawa S, Sripanidkulchai B, Tantraworasin A, Saeteng S, Siwachat S, Lertprasertsuke N and Banjerdpongchai R 2020 Potential Thai Herbal Extracts on Lung Cancer Treatment by Inducing Apoptosis and Synergizing Chemotheraphy Molecules 25(231):1–30.
[6] Elrayess R, Aziz Y M A, Elgawish M S, Elewa M, Yassen A S A, Elhady S S, Elshihawy H A and Said M M 2020 Discovery of Potent Dual EGFR/HER2 Inhibitors Based on Thiophene Scaffold Targeting H1299 Lung Cancer Cell Line Pharmaceuticals,14(1), 9.
[7] Dallakyan S and Olson A J 2015 Small-Molecule Library Screening by Docking with PyRx.

Methods in molecular bology 1263: 243-250.
[8] Pettersen E F, Goddard T D, Huang C C, Meng E C, Couch G S, Croll T I, Morris J H and Ferrin T E 2021 UCSF ChimeraX: Structure visualization for research ers, educators, and developers Protein Science, 30(1): 70–82.
[9] Laskowski R A and Swindells M B 2011 LigPlot+: multiple ligand-protein interaction diagrams for drug discovery J. Chem. Inf. Model., 51:2778-2786.
[10] Nirmala M J, Samundeeswari A and Sankar P D 2011 Natural plant resources in anti-cancer therapy-A review. Research in Plant Biology 1(3):1-14.
[11] Laksmiani N L, Paramita N L P V and Wirasuta I M A G 2016 In vitro and in silico anti-oxidant activity of purified fractions from purple sweet potato ethanolic extract. International Journal of Pharmacy and Pharmaceutical Sciences 8(8): 177-181.