EVALUATING THE POTENTIAL OF SHOC2, CDK3, AND EGFR AS DRUG TARGET IN LUNG CANCER THROUGH *IN-SILICO* STUDIES

Dr. S. Parijatham Kanchana
Assistant Professor
Department of Biochemistry,
St. Francis College for women, Hyderabad

Ms. Prasannan Pooja
Department of Biochemistry, St. Franics College for Women, Hyderabad

Dr. Rajashekar Chikati
Department of Biochemistry, Yogi Vemana University

https://doi.org/10.34293/ctbtls.2025.ch010

Abstract

Cancer is listed as one of the main public health concerns. Statistics currently claim that new cancer cases number in India will go up from approximately 1.46 million in 2022 to 1.57 million by 2025. Lung cancer is among the most severe causes of death due to its aggressive nature, late detection, and resistance to standard treatments. Targeting the key molecular pathways that promote the growth of cancer is required for the generation of successful curative medicines. This study focuses on the drug ability of proteins SHOC2 leucine-rich repeat scaffold protein, cyclin-dependent kinase3 (CDK3), and epidermal growth factor receptor (EGFR) for lung cancer through molecular docking studies. Using 'PyMOL and HDOCK software', molecular docking was performed to assess interactions between selected anticancer drugs Osimertinib, Gefitinib, Afatinib, Sorafenib, and Ceritinib. These drugs were selected due to their proven inhibitory effects on cancer-related signalling pathways. The docking results demonstrated varying binding affinities for the three Protein selected, which has strongest binding (-176.45) with Osimertinib, while Gefitinib, Afatinib, Sorafenib, and Ceritinib have less binding affinity (-116.37 to – 165.59). Docking data revealed different binding strengths for the selected drugs, which shows the strong binding interaction with certain target proteins, thus demonstrating a prospect for further development as targeted agents. These findings suggest that SHOC2, CDK3, and EGFR may be viable drug targets for lung cancer treatment. This study provides a computational framework for further in vitro and in vivo validation for drug discovery.

Keywords*: Lung cancer, SHOC2, CDK3, EGFR, Anti-cancerous drug.*

Introduction

Lung cancer remains one of the most prevalent and deadliest malignancies worldwide, with a five-year survival rate of approximately 18%, highlighting the urgent need for advanced research and novel therapeutic strategies (Siegel *et al*., 2023). The main reasons of the high rate of deaths are late-stage diagnosis, resistance to traditional therapy, and the numerous biological processes behind tumor growth (Travis *et al*., 2011). Using the in silico approaches to examine important proteins linked to the disease, our research aims to clarify the molecular complexities of lung cancer (Latimer & Mott, 2015). The unrestrained expansion of tumor-like cells in lung tissues is the root cause of lung cancer, which results in

tumors that obstruct regular breathing (Minna *et al.,* 2002). The two primary kinds of lung cancer are "Small Cellular Lung Cancer" (SCLC) and "Non-Small Cellular Lung Cancer" (NSCLC). NSCLC, which comprises subtypes such carcinoma of squamous and giant cell carcinoma, makes up about 85% of cases (Ettinger et at., 2010). The anatomy of lung is shown in figure 1.1

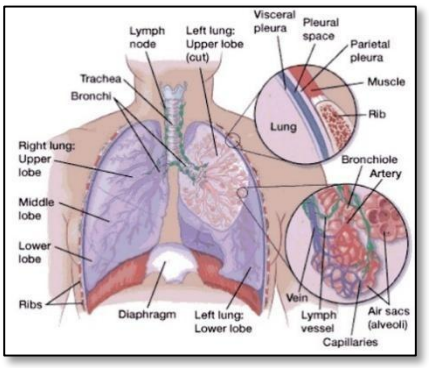


Figure 1 Overview of the Physiological Anatomy of the Lung.

1.1 Pathogenesis of Lung Cancer

Lung cancer arises from environmental exposures, genetic mutation, and molecular changes that lead in uncontrolled formation of cell growth, tumor development, and metastasis (Travis, 2011). Exposure to carcinogens from industrial chemicals, tobacco smoking, and air pollution harms DNA, which produces mutations that affect vital regulatory mechanisms (Latimer & Mott, 2015). Many genetic changes are important variables in the pathophysiology of lung cancer. Among these, tumor-causing mutations in the genes for EGFR, ALK, and KRAS are important for inducing unchecked cell division (Ettinger *et al.,* 2010). About 25–30% of lung adenocarcinomas have KRAS mutations, which cause the MAPK and PI3K pathways to remain activated over time, promoting the growth of the tumor (Sarkar, 2007).

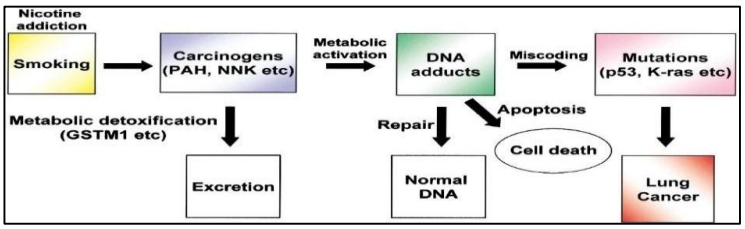


Figure 1.1 Pathogenesis of Lung Cancer Pathway

1.2 Diagnosis and Treatment

The method of diagnosis of suspected lung cancer is symptoms of abnormal pain, imaging methods, biopsy, and laboratory testing are used with abnormal chest radiograph findings or symptoms caused by either local or systemic effects of the tumor. A tissue biopsy, obtained through bronchoscopy, needle aspiration, or surgical methods, is required for a definitive diagnosis. Molecular testing for EGFR, ALK, and PD-L1 helps guide targeted

therapy decisions (Travis *et al.,* 2021). Locally advanced cases may require a combination of surgery, radiation therapy, and chemotherapy. Early detection significantly improves oncogenesis (Planchard *et al.,* 2023).

Proteins play a crucial role in cancer by regulating cell growth, division, and survival (weber, 2007).

Moreover, proteins involved in apoptosis, DNA repair, and immune regulation impact tumor development and response to therapy due to their inactivation or abnormal functioning making protein as a key-target for cancer Diagnosis and Treatment (Hanhan & Weinberg,2011)

A leucin repeat rich scaffolding protein termed SHOC2 integrates the RAS/RAF/MEK/ERK pathway's activation, which is essential for cells to survive, distinguish, and proliferating (Young, 2013). It is required for proper signal transduction and works by producing a SHOC2- MRAS-PP1C complex promoting RAF dephosphorylation, which in turn stimulates ERK pathway activation (Wang *et al.,* 2020).

Amongst the CDKs that are crucial to the advancement of the cell cycles is "cyclin-dependent kinase 3" (CDK3). The G1/S transition is related to CDK3, responsible for the initiation of replicating DNA and the advancement of the cell cycle. By modulating the MAPK/ERK and PI3K/AKT signal pathways, CDK3 has been proven to boost the development travel, and invasion of tumor cells (Chen *et al*., 2018).

The epidermal tyrosine kinase receptor termed the Epidermal Growth Factor Receptor (EGFR) comes to the ErbB receptor family. By regulating different downstream pathways, including RAS/RAF/MEK/ERK, PI3K/AKT, and JAK/STAT, EGFR helps the basic mechanisms of cellular growth, transformation, and lifespan (Yarden & Pines, 2012).

2. METHODOLOGY

2.1 Homology Modelling

Biological sources like NCBI, BLAST, PDB, and Clustal omega were all utilized for the selection of target and template proteins. Proteins with fewer than 600 amino acids were selected that had gone through review, and those with relevant research were within those chosen criteria.

2.2 Model Generation and Validation

Model generation and validation were performed using several tools: SWISS-MODEL, an automated platform for constructing comparative models of 3D protein structures; SAVES v6.0, a toolkit featuring "ERRAT, VERIFY3D, PROVE, PROCHECK, WHAT CHECK" for evaluating various stereo-chemical parameters of protein structures; and PROSA, is used to detect interactive designed for errors in 3D protein structures.

2.3 Active Site Prediction

CASTp, a website that offers tools for identifying, and evaluating the geometrical and topological characteristics of protein structures, were applied to predict the active sites of the produced models.

2.4 Selection of Ligands

For molecular docking, the ligands were chosen from PubChem, it is a public database providing data on chemical compounds, their properties, and biological interactions of molecules, they were developed as part of the US "National Institutes of Health" (NIH) Osimertinib, Gefitinib, Afatinib, Sorafenib and Ceritinib were among the ligands that were chosen.

2.5 Docking of Ligand and Macromolecule

Docking was carried out using HDOCK which is a web-based platform that produces results, on binding scores between desired biological functions. while PyMOL was used to visualize the interactions between the ligands and the protein. They help view them in 3D.

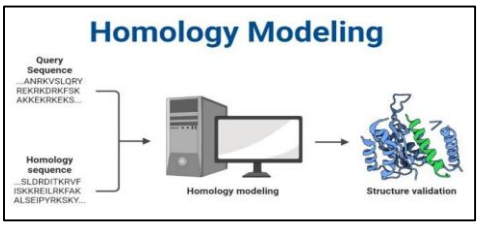


Figure 2.1: Methodology of In-silico studies

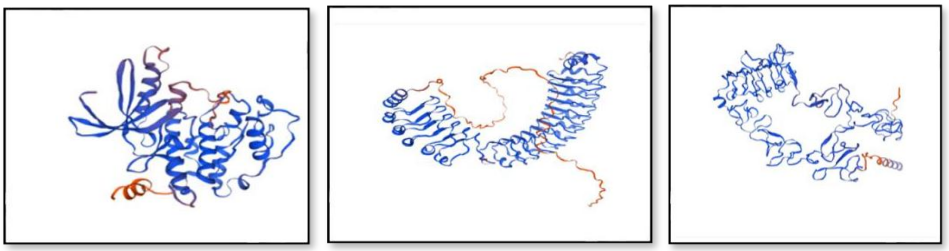
3. Results & Discussion

3.1 Homology Modelling

Selection of target protein: The accession codes **"KAI4077457.1", "NP_001249.1"**, and **"AAI18666.1"** were obtained from NCBI to retrieve the FASTA sequences of "SHOC2 leucine-rich repeat scaffold protein", "Cyclin-dependent kinase 3" and "Epidermal growth factor receptor". Protein accession codes: A5Si,11fn, A0V3 were taken from RCSB PDB. The protein with the fewest unmodeled regions was selected these were as the template for Shoc2 Cdk3 and Egfr was selected as the template for future processes and also having 85%-100% query coverage. A sequence alignment was performed using Clustal Omega by pasting the FASTA sequences of the target and template proteins (Sievers & Higgins, 2018).

3.2 Model Generation

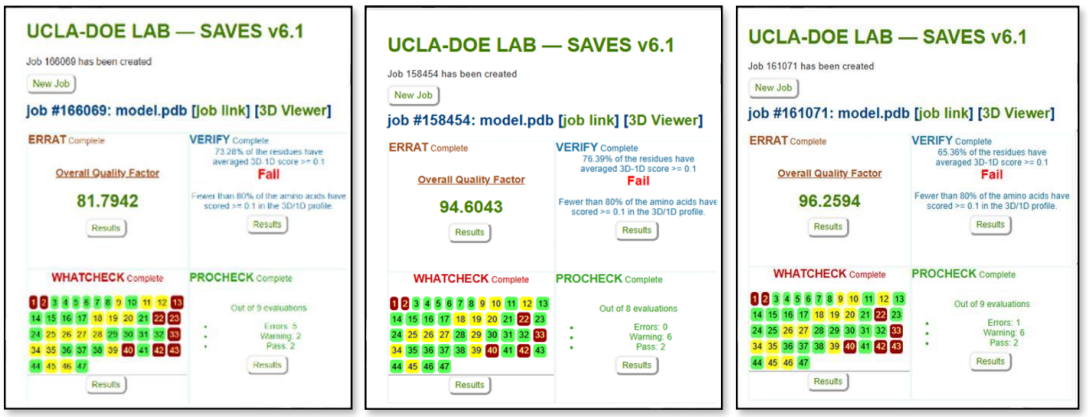
The alignment for model generation were created by modifying the code to include data for the selected target and template proteins. Using Swiss model software, which is a web-based tool used to build a high-quality 3D model in PDB format was generated (Waterhouse *et al.*, 2018). The generated models were assessed, and the one with the least objective function was selected.



(a) (b) (c) Figure 3.1(a),(b),(c): A generated model for using the Swiss model

3.3 Model Validation

Validation of model were done using Saves 6.0 and the following results are obtained in Verify 3D (Figure 3.2). The Ramachandran plots having 81.7%, 94.6%& 96.2% residues are in the favoured region (Figure 3.3) were also downloaded (Lüthy *et al.*, 1993).



(a) (b) (c) Figure 3.2 (a),(b),(c): Results of model validation from SAVES v6.0

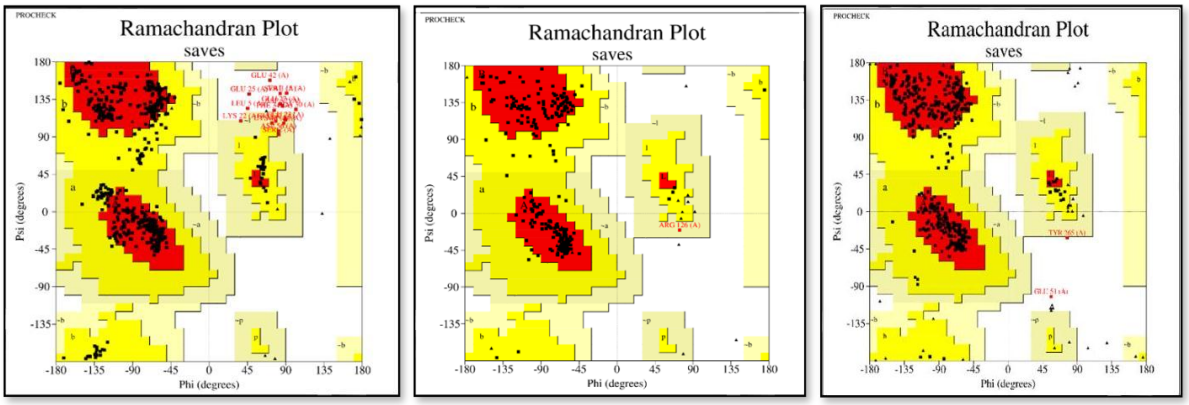
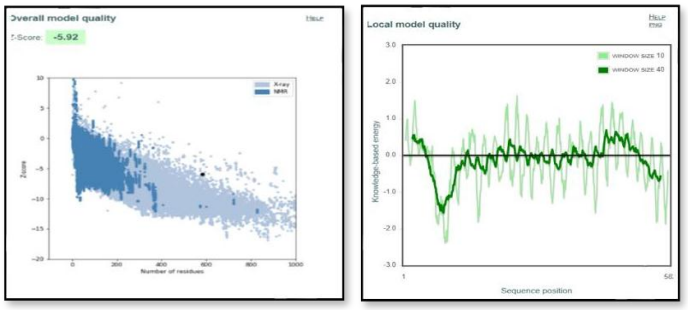
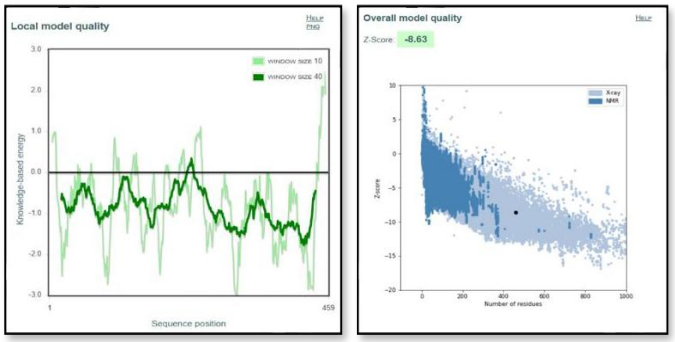


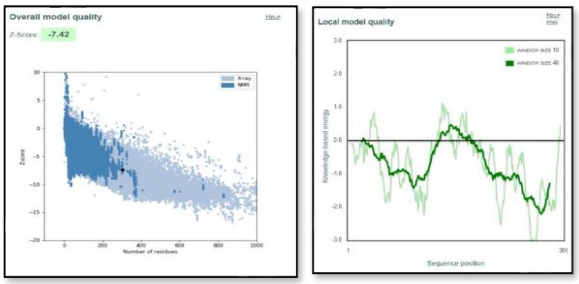
Figure 3.3 (a), (b), (c): Ramachandran Plots obtained from Procheck Verify 3D is a used for the quality accession of the protein models. The model is only considered of good quality if it's over 80% for amino acids achieve a score are of ≥0.2 in the 3D/1D profile (Lüthy *et al.*, 1992). The model failed the Verify 3D assessment (Figure 3.2). Model that are residue of 90% over are the highly favoured region, in the **Ramachandran plot** they would be more valid for further studies. Further model validation was performed using ProSA web which gave the graphs below. (Figure 3.4) shows the graphs of model overall quality specifying the Z score and model local quality (Wiederstein & Sippl, 2007).



(a)



(b)

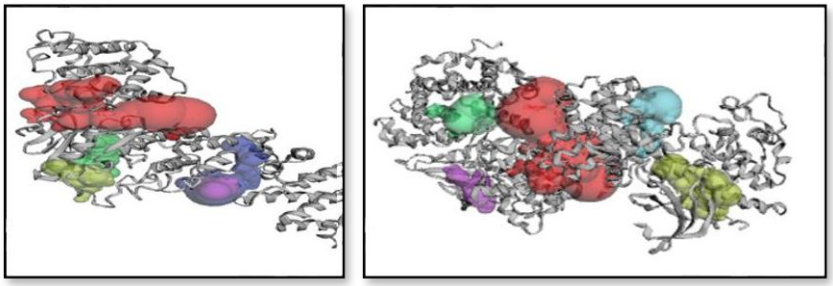


(c)

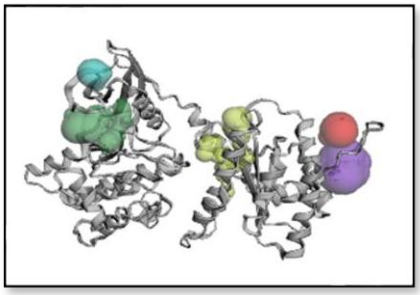
Figure 3.4 (a),(b),(c): Graphs obtained from ProSA showing the overall mode quality and local model quality The **ProSA** shows the validating of "Overall Model Quality" graphs display z-scores, indicating overall model quality, with "X-ray crystallography" (light blue) and "NMR spectroscopy" (dark blue) the z-scores plotted against protein length. The z scores for Shoc2, Cdk3, and Egfr are -5.92, -7.42, and -8.63, respectively. "Local Model Quality" energy plots show residue energies across the amino acid sequence, with positive values suggesting problematic regions. Shoc2, Cdk3 and Egfr are having only negative values.

3.4 Active-site Determination

CASTp fold is used to visualize the active site in protein pockets. Structure of protein are in PDB format with a probe radius (default 1.4 Å) are inputs to CASTp, which predicts the active sites based on these features**.**



(a) (b)



(c)

Figure: 3.5 (a), (b), (c): Active Pockets of Shoc2, Ccdk3 and Egfr Proteins

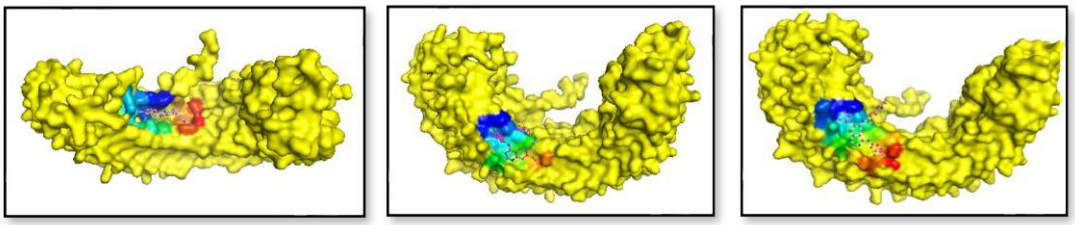
3.5 Selection of Ligands

Osimertinib, Gefitinib, Afatinib, Sorafenib and Ceritinib are drugs used in the treatment of lung cancer. The drugs were downloaded in the smile format and further converted to the PDB format using Open Babel. PDB format were used perform docking study and the interactions between each drug (acting as a ligand) and the proteins.

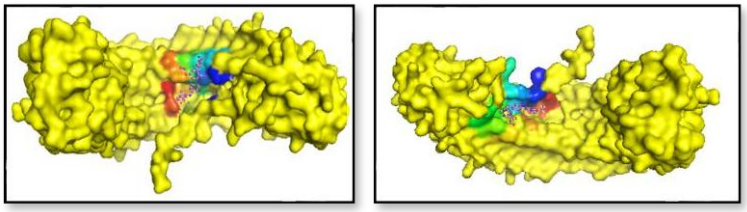
3.6 Docking of Ligand and Macromolecule

Docking ligand and macromolecule (protein) were performed using HDOCK. The results in the form of a table containing the ligand's bond affinity and structure. Docking score were saved in the PDB format. Visualizing of the interactions between ligands and protein were done by using PyMOL in the form of 3D structures as given below.

Interaction of Shoc2 with Osimertinib, Gefitinib, Afatinib, Sorafenib and Ceritinib

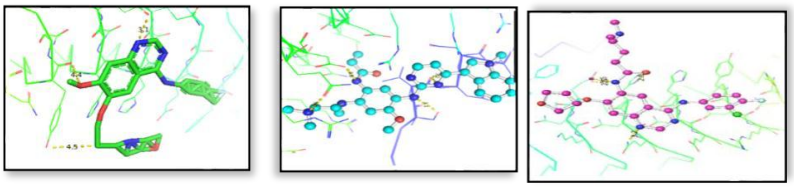


(a) (b) (c)

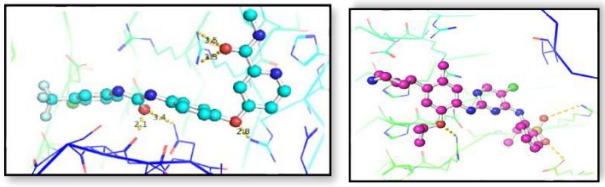


(d) (e)

Figure 3.6 (a), (b), (c), (d), (e) Docking result Of SHOC2 in PyMOL, Protein is in Surface form



(a) (b) (c)



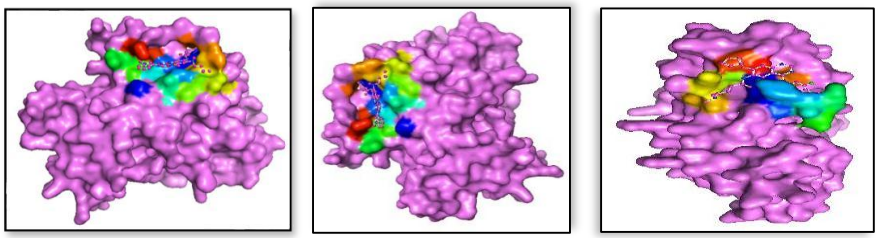
(d) (e)

Figure 3.6 (a), (b), (c), (d), (e): Shows the Docking of shoc2 with Osimertinib, Gefitinib, Afatinib, Sorafenib and Ceritinib by PyMOL

Table 1: Shows the Docking Score of Selected Ligands with SHOC2

SHOC2	*OSIMERTINIB*	*GEFITNIB*	*AFATINIB*	*SORAFENIB*	*CERITINIB*
DOCKING SCORE	**-176.45**	**-157.40**	**-149.66**	**165.59**	**-151.46**

The results from the (table 1) docking score indicate that Osimertinib showing (-176.45) "highest binding affinity" with SHOC2 followed by (-165.59 to 151.46) for Sorafenib, Gefitinib, Ceritinib, however Afatinib (-149.66) indicate a weaker binding interaction among rest.



(a) (b) (c)

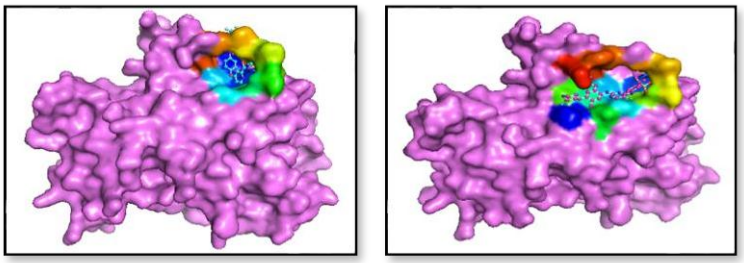
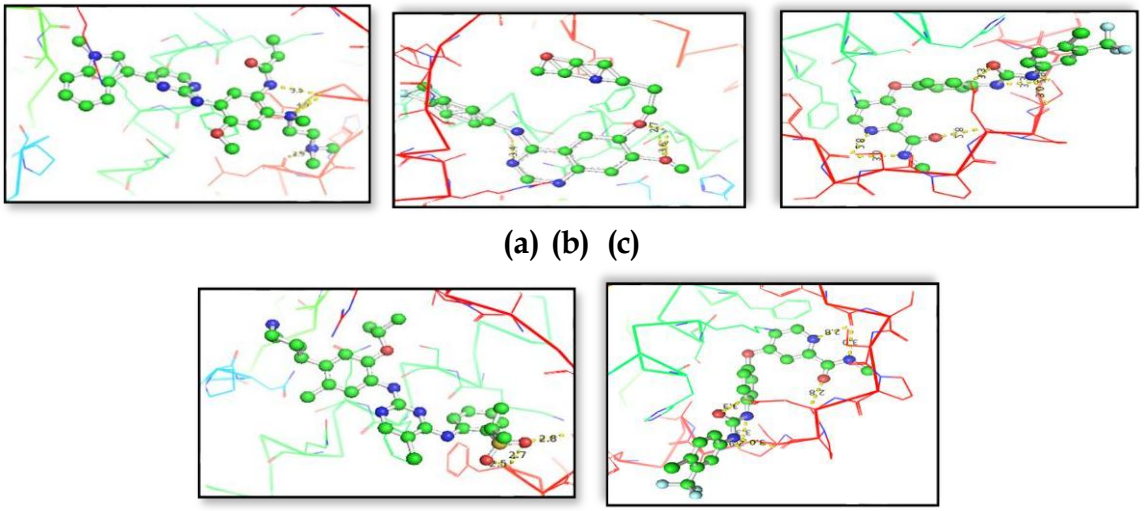


Figure 3.6 (a) (b) (c) (d) (e) Docking Result of CDK3 in PyMOL, Protein is in surface form



(a) (b) (c)

(d) (e)

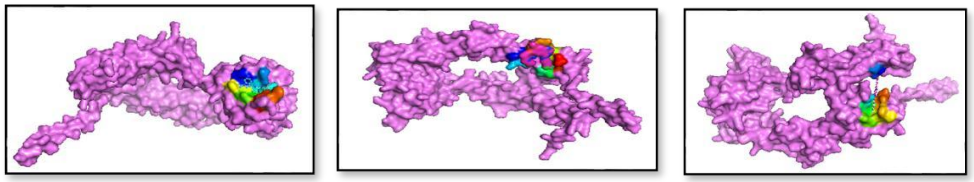
Figure 3.7 (a), (b), (c), (d), (e): Shows the Docking of Cdk3 with Osimertinib, Gefitinib, Afatinib, Sorafenib and Ceritinib by PyMOL

Table 2: Shows the docking score of selected ligands with CDK3

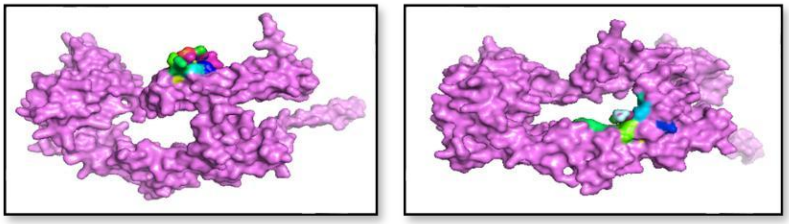
CDK3	*OSIMERTINIB*	*GEFITNIB*	*AFATINIB*	*SORAFENIB*	*CERITINIB*
DOCKING SCORE	**-144.40**	**-116.37**	**-128.33**	**-146.47**	**-135.17**

The results from the (table 2) docking score indicate that Osimertinib showing (-144.40) highest binding affinity with CDK3 followed by (-146.47 to -128.33) for Sorafenib, Ceritinib, Afatinib, however Gefitinib (-116.37) indicate a weaker binding interaction among rest.

Interaction of Egfr with Osimertinib, Gefitinib, Afatinib, Sorafenib and Ceritinib

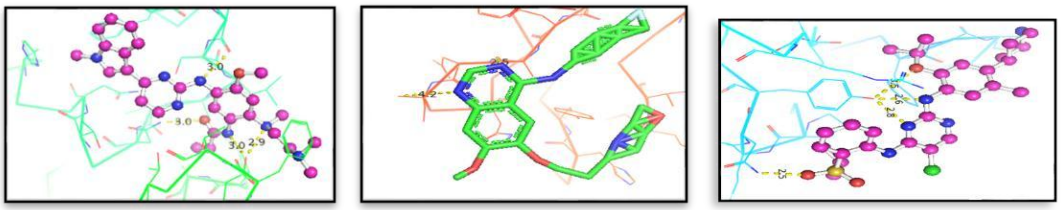


(a) (b) (c)

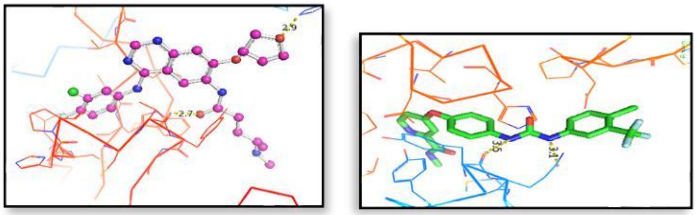


(d) (e)

Figure 3.6 (a) (b) (c) (d) (e) Docking Result of EGFR in Pymol, Protein is in surface form



(a) (b) (c)



(d) (e)

Figure 3.8 (a), (b), (c), (d), (e): Shows the Docking of Egfr with Osimertinib, Gefitinib, Afatinib, Sorafenib and Ceritinib by PyMOL

Table 3: Shows the Docking Score of Ligands Selected with EGFR

EGFR	*OSIMERTINIB*	*GEFITNIB*	*AFATINIB*	*SORAFENIB*	*CERITINIB*
DOCKING SCORE	***-168.69***	***-138.02***	***-131.20***	***-158.31***	***-162.48***

The results for docking score from (table 3) indicate that Osimertinib showing (-168.69) "highest binding affinity" with EGFR followed by (-162.48 to 138.02) Ceritinib, Sorafenib, Gefitinib, however Afatinib (-131.20) indicate a weaker binding interaction among rest.

4. Conclusion

Lung cancer remains one of the deadliest diseases, having limited options for treatment and a high mortality rate. This study explored how certain cancer-fighting drugs as Osimertinib, Gefitinib, Afatinib, Sorafenib, and Ceritinib interact among key proteins involved in lung cancer (SHOC2, CDK3, and EGFR) through molecular docking. These results are viewed that Osimertinib had the strongest binding affinity with all three proteins, suggesting it may be the most effective among the tested drugs.

Gefitinib, Sorafenib, and Ceritinib also showed promising interactions, while Afatinib

had comparatively weaker binding. This study offers important findings for further experimental research is needed to confirm these results and understand their real-world impact. By identifying promising drug-target interactions, this research contributes to the ongoing efforts to develop more effective and personalized treatments for cancer in lungs, bringing hope for better outcomes in the future

References:

1. Adams, S. J., Stone, E., Baldwin, D. R., Vliegenthart, R., Lee, P., & Fintelmann, F. J. (2023). Lung cancer screening. The Lancet, 401(10374), 390-408.
2. Ettinger, D. S., Akerley, W., Bepler, G., Blum, M. G., Chang, A., Cheney, R. T., ... & Yang, S. C. (2010). Non–small cell lung cancer. Journal of the national comprehensive cancer network, 8(7), 740-801.
3. Fruehauf, J. (2005). EGFR function and detection in cancer therapy. Journal of Experimental Therapeutics & Oncology, 5(3).
4. Hammerschmidt, S., & Wirtz, H. (2009). Lung cancer: current diagnosis and treatment. Deutsches Ärzteblatt International, 106(49), 809.
5. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. cell, 144(5), 646-674.
6. Hauseman, Z. J., Fodor, M., Dhembi, A., Viscomi, J., Egli, D., Bleu, M., ... & King, D. A. (2022). Structure of the MRAS–SHOC2–PP1C phosphatase complex. Nature, 609(7926), 416-423.
7. Kaplan, F. M., Kugel, C. H., Dadpey, N., Shao, Y., Abel, E. V., & Aplin, A. E. (2012). SHOC2 and CRAF mediate ERK1/2 reactivation in mutant NRAS-mediated resistance to RAF inhibitor. Journal of Biological Chemistry, 287(50), 41797-41807.
8. Latimer, K. M., & Mott, T. F. (2015). Lung cancer: diagnosis, treatment principles, and screening. American family physician, 91(4), 250-256.
9. Liau, N. P., Johnson, M. C., Izadi, S., Gerosa, L., Hammel, M., Bruning, J. M., ... & Sudhamsu, J. (2022). Structural basis for SHOC2 modulation of RAS signalling. Nature, 609(7926), 400-407.
10. Minna, J. D., Roth, J. A., & Gazdar, A. F. (2002). Focus on lung cancer. Cancer cell, 1(1), 49-52.
11. Nicholson, A. G., Tsao, M. S., Beasley, M. B., Borczuk, A. C., Brambilla, E., Cooper, W. A., ... & Travis, W. D. (2022). The 2021 WHO classification of lung tumors: impact of advances since 2015. Journal of Thoracic Oncology, 17(3), 362-387.
12. Nicholson, R. I., Gee, J. M. W., & Harper, M. E. (2001). EGFR and cancer prognosis. European journal of cancer, 37, 9-15.
13. Oser, M. G., Niederst, M. J., Sequist, L. V., & Engelman, J. A. (2015). Transformation from non-small-cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin. The Lancet Oncology, 16(4), e165-e172.

14. Planchard, D., Popat, S. T., Kerr, K., Novello, S., Smit, E. F., Faivre-Finn, C., ... & ESMO Guidelines Committee. (2018). Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology, 29, iv192 iv237.
15. Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. CA: a cancer journal for clinicians, 73(1), 17-48.
16. Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. Protein Science, 27(1), 135-145.
17. Teo, T., Kasirzadeh, S., Albrecht, H., Sykes, M. J., Yang, Y., & Wang, S. (2022). An overview of CDK3 in cancer: clinical significance and pharmacological implications. Pharmacological Research, 180, 106249.
18. Travis, W. D. (2011). Pathology of lung cancer. Clinics in chest medicine, 32(4), 669-692.
19. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Research, 46(W1), W296–W303
20. Yarden, Y., & Pines, G. (2012). The ERBB network: at last, cancer therapy meets systems biology. Nature Reviews Cancer, 12(8), 553-563.