IOSR Journal of Pharmacy and Biological Sciences (IOSR-JPBS)
e-ISSN:2278-3008, p-ISSN:2319-7676. Volume 21, Issue 4 Ser. 2 (July. – August. 2026), PP 28-37
www.Iosrjournals.org

NGS-Based Identification and Bio python-Driven Structural Mapping of EGFR Variants in Cancer

Ipsita Dey[1] Vaishnavi Pimparkar[1] Uma Kumari[*1]
[1]*Project Trainee at Bioinformatics Project and Research Institute, Noida - 201301, India*
[1]*Project Trainee at Bioinformatics Project and Research Institute, Noida - 201301, India*
[1*]*Senior Bioinformatics Scientist, Bioinformatics Project and Research Institute, Noida - 201301, India*
Corresponding Author: Uma Kumari(uma27910@gmail.com)

Abstract
EGFR, a receptor tyrosine kinase often mutated in cancer, has variants affecting its structure, signalling, drug interactions, and treatment response in non-small cell lung cancer. This study introduces a combined NGS and Bio python-based computational method for identifying and mapping EGFR variants using PDB ID: 8D76. NGS data helps detect clinically significant variants, which are then analysed and mapped with combining Bio python. Sequence similarity and conservation are evaluated using, while InterProScan identifies functional domains and conserved regions. Structural integrity and residue flexibility are assessed through ERRAT and B-factor analysis showing stable and flexible residue. Variants structural impacts are examined via RMSD analysis. Protein interaction networks and signaling pathways are studied using STRING and pathway analysis to understand cancer progression. Molecular docking explores interactions between drugs and EGFR, with ADMET analysis evaluating their pharmacokinetics and toxicity. In toxicity prediction classified sotorasib as class 5, while other ligand was placed in class 4. This comprehensive approach integrates genomic, sequence, structural, pathway, and pharmacological data to identify potentially critical EGFR variants as a central hub gene. It aims to improve understanding of EGFR alterations and support personalized cancer treatments and drug development.

Keywords: *EGFR variants, Next-generation sequencing (NGS), Bio-python, Structural bioinformatics, Molecular docking, Targeted therapy, ADMET analysis, Tyrosine kinase.*

Date of Submission: 11-08-2026 Date of Acceptance: 22-08-2026

1.Introduction

Lung cancer is one of the major causes of cancer-related mortality worldwide, with non-small-cell lung cancer (NSCLC) accounting for approximately 85% of cases. Molecular alterations in oncogenes such as EGFR and KRAS play important roles in the development of NSCLC and in therapeutic response. EGFR, also known as HER1 or ErbB1, is a 170-kDa transmembrane receptor tyrosine kinase belonging to the ErbB family. **[1,2]** Activation of EGFR by epidermal growth factor initiates phosphorylation and downstream signalling cascades that regulate cell proliferation, survival, migration, and other cellular processes. Dysregulation of EGFR signalling can therefore contribute to tumor progression and metastasis.**[3]**

EGFR-targeted tyrosine kinase inhibitors (TKIs), including gefitinib, erlotinib, and osimertinib, have significantly improved treatment options for patients with EGFR-mutated NSCLC. The most clinically important EGFR alterations occur within exons 18–21 of the tyrosine kinase domain. **[4,5]** Exon 19 deletions (Del19) and L858R are classical activating mutations that generally increase sensitivity to EGFR-TKIs. In contrast, T790M is a gatekeeper mutation that increases ATP affinity and produces resistance to earlier-generation TKIs. **[6]** The third-generation inhibitor osimertinib can target T790M; however, acquisition of the C797S mutation can disrupt the covalent interaction required for osimertinib activity, thereby contributing to resistance. The spatial relationship between T790M and C797S can further influence therapeutic response. **[7,8]**

Next-generation sequencing (NGS) provides a cost-effective approach for simultaneously identifying large numbers of genomic variants, enabling researchers to prioritize clinically and biologically important alterations. **[9,10]** Computational analysis can subsequently connect these genomic changes with protein-level consequences. Bio python facilitates sequence processing, mutation mapping, and comparison of wild-type and mutant EGFR proteins, while structural resources and visualization tools such as PDB and PyMOL support three-dimensional protein analysis. **[11,12]**

Furthermore, computer-aided drug design (CADD) enables structure-based and ligand-based investigation of potential therapeutics. Molecular docking can help evaluate interactions between EGFR and candidate compounds, while ADMET analysis provides information regarding absorption, distribution, metabolism, excretion, and toxicity. Integrating NGS, Bio python, structural analysis, CADD, and ADMET

therefore provides a comprehensive computational strategy for understanding EGFR variants, drug resistance, and precision therapeutic opportunities in NSCLC. **[13,14]**

2.Materials and Methodology

Structural basis: PDB ID 8D76 shows an EGFR protein–ligand crystal structure studied for targeted therapy in lung cancer. Key EGFR mutations like Exon 19 deletion, L858R, T790M, and C797S play significant roles in non-small-cell lung cancer (NSCLC).**[15]** NGS and sequence analysis involve using NGS-based variant detection in combination with MMDB, FASTA data, sequence similarity search, Bio python, and Jalview. BLAST verifies sequence identity and similarity, and Jalview facilitates multiple sequence alignment, evolutionary studies, and detection of conserved regions. **[16]** Functional annotation is performed using InterProScan to identify protein domains, conserved sites, repeats, and Gene Ontology (GO) annotations. Pathway analysis is employed to interpret EGFR-associated signalling pathways and understand the biological effects of mutations.**[17]** Structural analysis employs Bio python and PyMOL for sequence processing, mutation mapping, and 3D visualization. SAVES/ERRAT assesses structural quality, while B-factor analysis reveals residue flexibility. RMSD measures differences between wild-type and mutant EGFR structures. **[18]** Interaction and therapeutic analysis: protein enrichment analysis assesses EGFR protein interaction networks, while CB-Dock2 predicts protein–ligand binding interactions and helps identify potential therapeutic compounds**. [19]** Drug safety assessment using ADMET and ProTox 3.0 analyses examines pharmacokinetic features and toxicity endpoints, aiding in the prioritization of potentially safer candidates for EGFR targeting. Overall,this comprehensive workflow combining NGS, Bio python, structural analysis, pathway mapping, docking, and ADMET predicts how EGFR genetic variants influence structure and potential therapies. **[20,21]**

3.Result

3.1 Retrieval of sequence and Conversion

The structure of 8D76 was retrieved from the Protein Data Bank. Its amino acid sequence was extracted in FASTA format for analysis. FASTA is a standard format for representing protein sequences, facilitating computational tasks like sequence alignment. The FASTA sequence of 8D76 is shown below.

PDB: 8D76_A

MGSSHHHHHHSSGENLYFQGSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAI
KELREATSPKANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLIMQLMPFGCLLDYVREHKDNIGSQY
LLNWCVQIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGRAKLLGAEEKEYHAEGGKVPIK
WMALESILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMRK
CWMIDADSRPKFRELIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYL
IPQQG

3.2 Sequence similarity search and Structure Validation analysis

This analysis indicates a query coverage of 96% to 100%. The 8D76 query protein closely resembles the known human EGFR structure, with a highly significant alignment. It demonstrates high sequence similarity, and analysis suggests that the 8D76 protein functions as an EGFR receptor tyrosine kinase with conserved catalytic domains. The ERRAT analysis of the protein structure (PDB ID: 8D76, Chain A) showed an overall quality factor of 99.642%, which is well above the acceptable threshold of 95%.

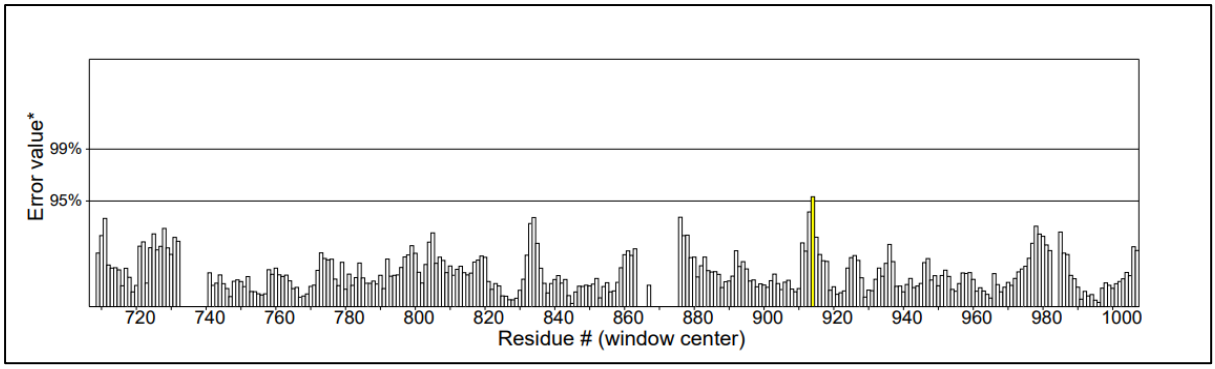


Figure 1: ERRAT Structure Validation Analysis

3.3 Pathway Analysis of EGFR Variants in Cancer

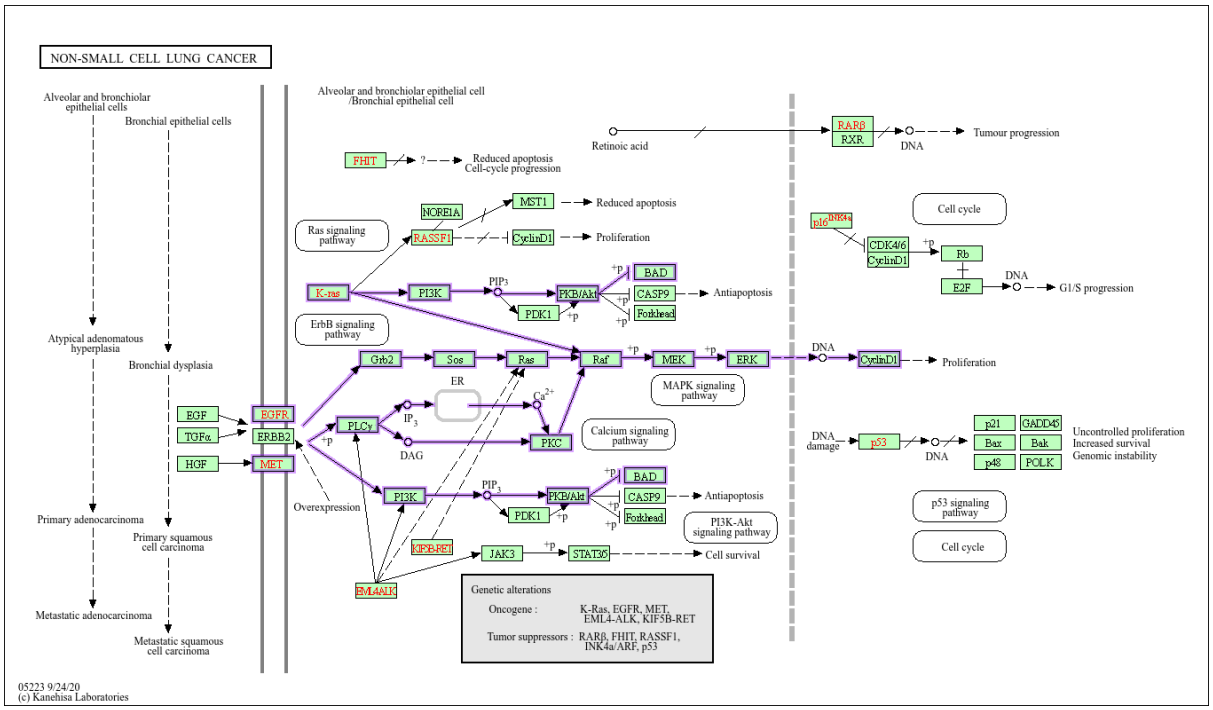


Figure 2: KEGG pathway analysis of EGFR variants in Cancer (Mutation activation is highlighted)

Non-small cell lung cancer (NSCLC) develops due to abnormal activation of receptor tyrosine kinases such as EGFR, MET, and ERBB2, which activate downstream signaling pathways. A major pathway, RAS–RAF–MEK–ERK (MAPK), is triggered through EGFR-KRAS-RAF-MEK-ERK, promoting cell proliferation, differentiation, and tumour growth; mutations in KRAS and BRAF lead to continuous activation of this cascade. At the same time, mutations in cell cycle regulators (TP53, RB1, CDK4/6, Cyclin D1) disable normal checkpoints, resulting in uncontrolled division. Reduced apoptosis, caused by increased BCL-2 activity or TP53 mutations, also supports tumor survival by decreasing caspase-mediated cell death via the BAX/BCL-2 pathway. These disruptions focus on a common set of mutated genes—EGFR, KRAS, ALK, ROS1, MET, BRAF, PIK3CA, TP53, and RB1—that are routinely examined using next-generation sequencing (NGS). Clinically, these mutations serve as biomarkers to guide targeted therapies (such as inhibitors for EGFR, ALK, ROS1, BRAF, and MET), help interpret NGS data, and facilitate drug discovery through molecular docking and pathway analysis.

3.4 Functional Domains Analysis

InterProScan analysis identified the query protein as belonging to the Receptor Tyrosine Kinase (RTK) family, a group of signalling proteins that act as upstream regulators of pathways such as MAPK/ERK and PI3K/Akt. Conserved residues linked to ATP binding, catalytic activity, and Mg^{2+} coordination were also identified, supporting the presence of functional kinase machinery. The active site was mapped between residues 159–171, and the ATP-binding region between residues 44–71, both essential for phosphorylation-based signalling. Two receptor tyrosine kinase domains spanning residues 22–118 and 119–311 further confirmed the protein's role in ATP-dependent signal transduction. Overall, these findings confirm that the protein is a member of the RTK superfamily, with conserved catalytic and ATP-binding regions, and functions as an ATP-dependent kinase involved in key cellular signalling and regulatory processes.

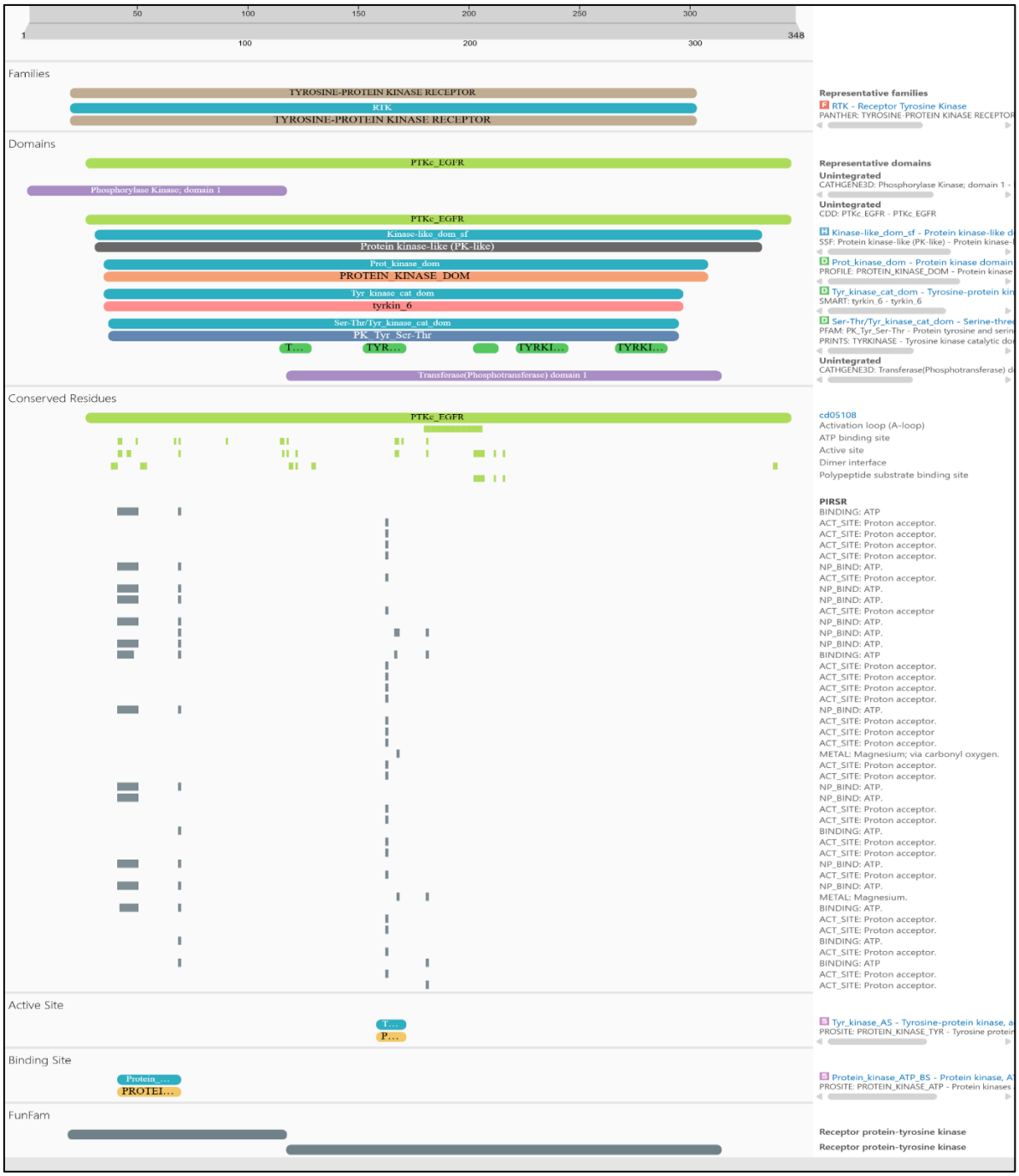


Figure 3: functional domain analysis of sample 8D76

3.5 RMSD Analysis

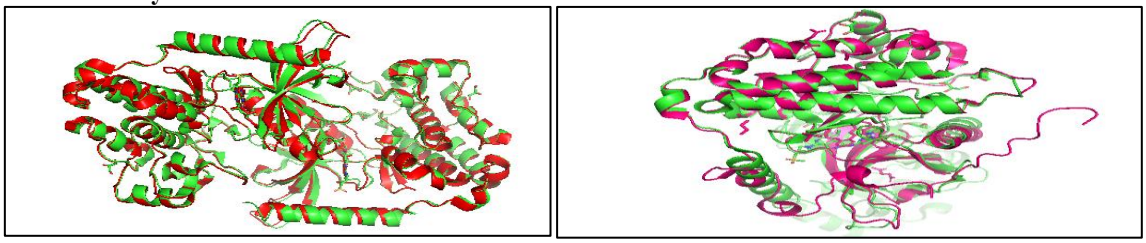


Figure 4: RMSD Calculation analysis with score 0.325 A (where Green 8D76, Red 8D73). Figure 5: RMSD Calculation analysis with score 0.458 A (where Green 8D76, Pink 9QXN)

Structural alignment of PDB 8D76 with PDB 8D73 was conducted using PyMOL. Out of 5076 atoms, 3889 were successfully aligned, resulting in an RMSD of 0.325, indicating excellent alignment typical of highly similar or identical structures. Similarly, alignment of PDB 8D76 with PDB 9QXN was performed with PyMOL. From 2515 atoms, 1730 aligned successfully, with an RMSD of 0.458, also reflecting excellent alignment often seen in highly similar or identical structures, such as different models from the same crystallographic data or high-quality homologs.

3.6 B-Factor Analysis

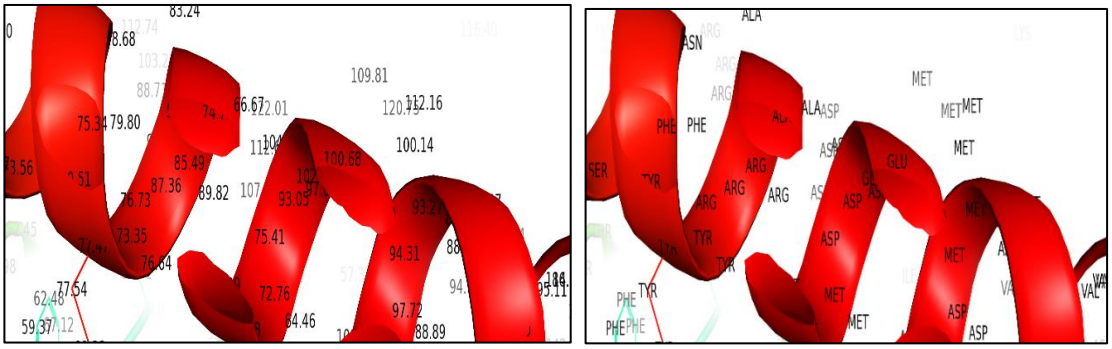


Figure 6: B-factor Analysis showing residue and B-factor positions in protein secondary structure (alpha-helix) from PDB ID 8D76

B-factor analysis shows that stable residues like ALA, TYR, PHE, and ARG have low B-values (64-80 Å), indicating stability. In contrast, flexible residues such as MET, ASP, and GLU exhibit high B-values, suggesting greater flexibility.

3.7 Protein-protein interaction analysis

The STRING network identified EGFR as the central hub with the highest number of interactions, highlighting its key role in growth factor signaling. EGF was the main ligand activating EGFR, while ERBB2 and ERBB3 were important downstream partners that facilitate receptor heterodimerization and amplify signaling. Additional interactors such as PIK3CA, GAB1, HBEGF, EREG, CBL, CDH1, and DCN form a complex network that regulates cell growth, proliferation, migration, and survival. These results indicate that EGFR and its associated proteins—particularly ERBB2, ERBB3, PIK3CA, GAB1, and EGF—are crucial targets and potential biomarkers for diseases such as cancer. The network also underscores the importance of targeting multiple pathway nodes to enhance therapy effectiveness and reduce drug resistance.

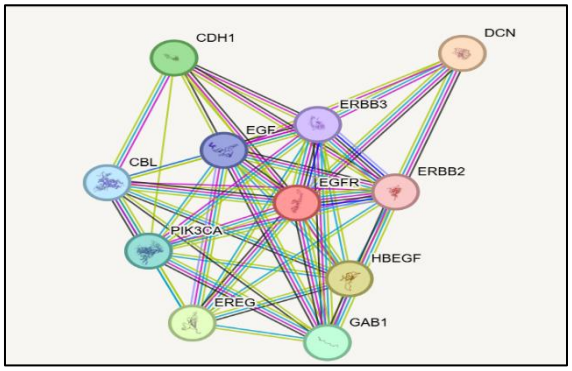


Figure 7: Protein-protein interaction network generated using the STRING database

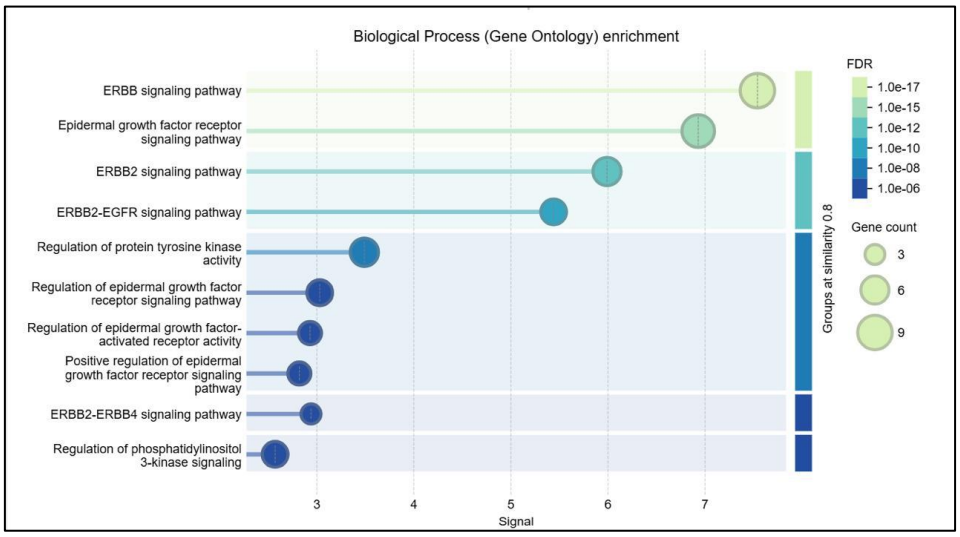


Figure 8: Bubble plot in protein enrichment analysis using STRING

3.8 Bio python

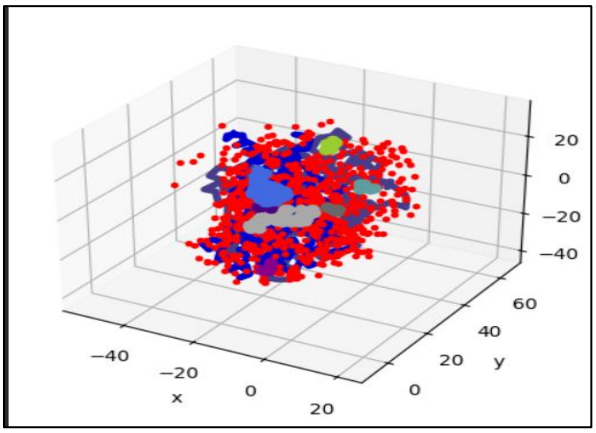
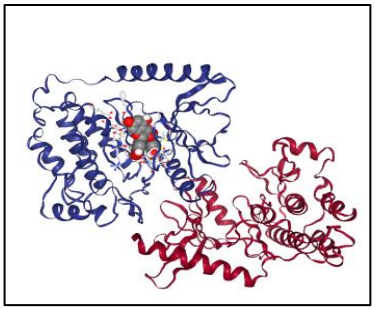


Figure 9. Three-dimensional scatter plot showing the spatial distribution and clustering of EGFR structural coordinates, highlighting distinct residue groups and potential mutation-associated regions within the protein model

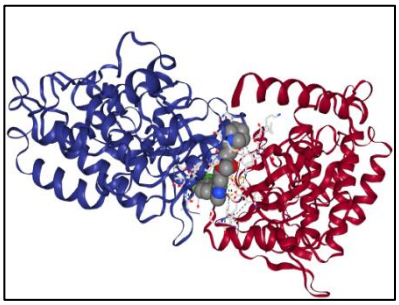
The 3D scatter plot shows the spatial distribution of mapped EGFR structural coordinates or mutation-related residues within the protein model. Red and blue points denote different residue groups or structural states, while highlighted areas indicate possible mutation hotspots. The clustering pattern implies localized structural variation that could affect protein stability, flexibility, and ligand-binding interactions.

3.9 Molecular Docking

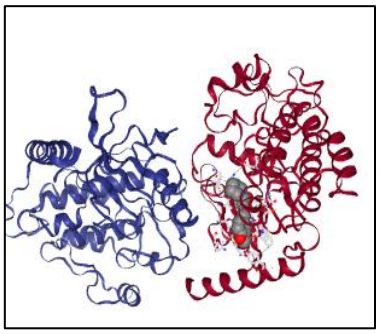
The 3D structures of all chemicals and the target protein were obtained from PubChem. In molecular docking, the red ribbon indicates the EGFR Kinase (a single chain of the receptor), while the blue ribbon shows the associated protein chain within the domain structural complex. The ligand is encircled by several amino acid residues, indicating hydrogen bonds and hydrophobic interactions that help stabilize the protein–ligand complex. Docking results reveal that binding pockets C1, C2, and C3 are the most favorable sites for ligands, with the highest docking scores, suggesting the formation of stable protein–ligand complexes with strong binding affinity.



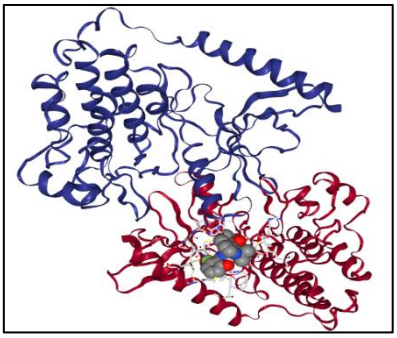
Epigallocatechin Gallate



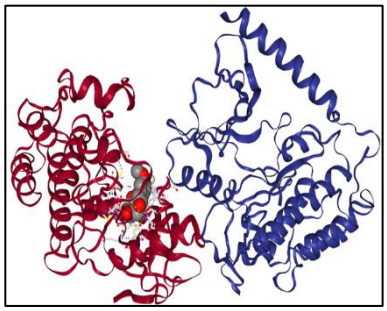
Carboplatin



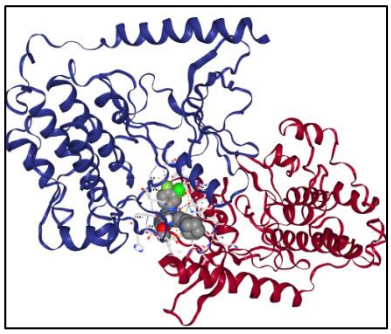
Piperine



Sotorasib

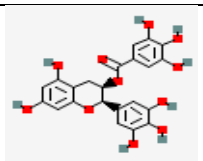
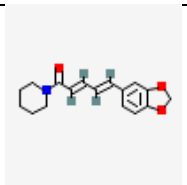


Curcumin



Dacomitinib

Figure 10 :- 3D structure of Molecular docking of ligands with EGFR protein(8D76) by using CB-dock

PubChem ID	Drug	CurPocket ID	Vina score (kcal/mol)	Cavity volume (Å³)	Center (x,y,z)	Remarks
65064	 Epigallocatechin Gallate	C2	-10.0	1197	-33,45,-10	Excellent binding
638024	 Piperine	C1	-8.4	6695	-14,27,-11	Strong binding

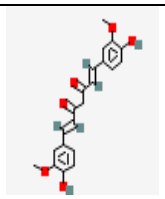
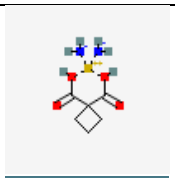
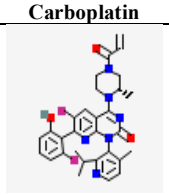
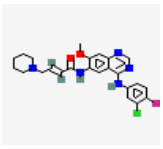
969516	 **Curcumin**	**C3**	**-9.1**	**1186**	**-12,13,0**	**Strong binding**
426756	 **Carboplatin**	**C1**	**-10.2**	**6695**	**-14,27,-11**	**Excellent binding**
137278711	 **Sotorasib**	**C3**	**-10.4**	**1186**	**-12,13,0**	**Excellent binding**
11511120	 **Dacomitinib**	**C1**	**-9.6**	**6695**	**-14,27,-11**	**Strong binding**

Table 1: Summarises the molecular docking results of Ligands against different predicted binding pockets of the target protein.

3.10 ADMET Analysis

After molecular docking, the drugs with binding compatibility and favourable ADMET properties with the EGFR binding site, Epigallocatechin Gallate, Curcumin, Piperine, Carboplatin, Sotorasib, and Dacomitinib, were selected based on their anticancer potential report and safety profiles. ProTox 3.0 classifies Epigallocatechin Gallate, Curcumin, Piperine, Carboplatin, and Dacomitinib as Class 4 (Low to moderate toxicity) and Sotorasib as Class 5 (practically non-toxic). The analysed lethal dose 50 (LD50) of Epigallocatechin Gallate, Curcumin, Piperine, Carboplatin, Sotorasib, and Dacomitinib are 1000 mg/kg, 2000 mg/kg, 330mg/kg and 343mg/kg, 1000 mg/kg,2500 mg/kg, respectively. ProTox 3.0 indicates that the drugs generally show toxicity across most categories (e.g., hepatotoxicity and carcinogenicity in most cases), with favourable metabolic profiles, as shown in Table 2. The Toxicity comparison demonstrates that the input drugs have a higher molecular weight than the dataset average. The predicted toxic dose falls within the normal distribution of the ProTox-3.0 dataset, indicated that the model's toxicity prediction is reliable.

Parameters	Epigallocatechin Gallate	Piperine	Curcumin	Carboplatin	Sotorasib	Dacomitinib
Hepatotoxicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Neurotoxicity	Inactive	Active	Inactive	Inactive	Active	Active
Nephrotoxicity	Active	Inactive	Active	Inactive	Inactive	Inactive
Respiratory toxicity	Active	Active	Inactive	Active	Active	Active
Cardiotoxicity	Inactive	Inactive	Active	Inactive	Inactive	Inactive
Carcinogenicity	Inactive	Active	Inactive	Inactive	Inactive	Inactive
Immunotoxicity	Inactive	Active	Active	Inactive	Active	Active
Mutagenicity	Inactive	Inactive	Inactive	Inactive	Active	Active
Cytotoxicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Clinical toxicity	Active	Active	Active	Inactive	Active	Active
Aryl hydrocarbon Receptor (AhR)	Inactive	Active	Inactive	Inactive	Inactive	Inactive
Androgen Receptor (AR)	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Estrogen Receptor Alpha (ER)	Inactive	Active	Active	Inactive	Inactive	Inactive
PPAR-Gamma	Inactive	Inactive	Active	Inactive	Inactive	Inactive
Heat shock factor response element	Inactive	Inactive	Active	Inactive	Inactive	Inactive
Phosphoprotein p53	Inactive	Inactive	Active	Inactive	Inactive	Inactive

ATAD5	Inactive	Active	Inactive	Inactive	Inactive	Inactive
Thyroid hormone receptor alpha (THRα)	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
GABA receptor	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Acetylcholinesterase (AChE)	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Cytochrome CYP1A2	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Cytochrome CYP2C9	Inactive	Inactive	Active	Inactive	Inactive	Active
Cytochrome CYP2D6	Inactive	Active	Inactive	Inactive	Inactive	Active
Cytochrome CYP3A4	Inactive	Inactive	Active	Inactive	Inactive	Inactive

Table 2: ProTox-3.0 in silico toxicity prediction of Ligands, natural phytochemicals investigated as potential therapeutic candidates for NSCLC.

I. Conclusion

In conclusion, this study highlights how combining various computational and bioinformatics tools improves EGFR gene mapping and related analyses. NGS provided the essential raw data, while Bio python facilitated efficient sequence retrieval, parsing, and manipulation. RMDs enabled robust statistical modeling and ensured reproducibility. BLAST was crucial for sequence alignment and homology search, and InterProScan offered functional annotation of protein domains. Network analysis via STRING uncovered protein–protein interactions, and KEGG pathway mapping situated EGFR within larger biological systems. Structural insights were gained through B-factor evaluation, revealing protein flexibility. Molecular docking predicted ligand–protein binding affinities, and ADMET analysis evaluated pharmacokinetics and toxicity, supporting drug discovery relevance. Together, these tools formed an integrated pipeline linking raw genomic data to functional, structural, and therapeutic insights. Their combined use demonstrates the power of computational biology in precision medicine, enabling accurate mapping, functional annotation, and drug validation. This integrated approach advances EGFR research and provides a solid foundation for future targeted therapy and personalised healthcare studies.

Reference

[1]. Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians, 73*, 17–48.

[2]. Alsaed, B., Bobik, N., Laitinen, H., et al. (2025). Shaping the battlefield: EGFR and KRAS tumor mutations' role on the immune microenvironment and immunotherapy responses in lung cancer. *Cancer Metastasis Reviews, 44*, 56. https://doi.org/10.1007/s10555-025-10272-4

[3]. Bhoomika, A., Kumari, U., & Bhavana, K. S. (2023). In silico computer-aided drug designing of the mutated EGFR kinase domain of NSCLC (non-small cell lung cancer). *Journal of Emerging Technologies and Innovative Research, 10*(12), B226–B235.

[4]. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: Protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research, 51*(D1), D638–D646. https://doi.org/10.1093/nar/gkac1000

[5]. Banerjee, P., Kemmler, E., Dunkel, M., & Preissner, R. (2024). ProTox 3.0: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Research, 52*(W1), W513–W520. https://doi.org/10.1093/nar/gkae303

[6]. Johri, V., Tyagi, E., & Kumari, U. (2024). Next generation sequencing analysis of NSP5, NSP6 and NSP7 SARS-CoV-2 nonstructural proteins with BioPython. *South Eastern European Journal of Public Health*, 1653–1666. https://doi.org/10.70135/seejph.vi.2936

[7]. Kumari, U., & Johri, V. (2025). Integrating NGS and CADD for the identification of medicinal plant compounds targeting EGFR mutants in non-small cell lung cancer. *Journal of Emerging Technologies and Innovative Research, 12*(8).

[8]. Gurung, R., & Kumari, U. (2026). NGS and structure-based drug designing of natural compound against drug resistant EGFR in non-small cell lung cancer. *IOSR Journal of Biotechnology and Biochemistry, 12*(2), 36–46.

[9]. Kumari, U., & Agrawal, N. (2023). NGS and mutational profile analysis of non-small-cell lung carcinoma (NSCLC). *International Journal for Research in Applied Science and Engineering Technology, 11*, 3090–3094. https://doi.org/10.22214/ijraset.2023.50880

[10]. Kumari, U., & Gupta, S. (2023). NGS and sequence analysis with Biopython for prospective brain cancer therapeutic studies. *International Journal for Research in Applied Science and Engineering Technology, 11*. https://doi.org/10.22214/ijraset.2023.50885

[11]. Kumari, U., Johri, V., Dhopate, S., & Jha, T. (2024). Structure-based drug design for the prediction of epitopes for targeting malignant brain tumours. *Journal of Emerging Technologies and Innovative Research, 11*(7).

[12]. Bandbe, T., Saikia, J., & Kumari, U. (2025). NGS analysis of human papillomavirus type 18 E2 DNA-binding domain bound to its DNA target with Biopython. *South Eastern European Journal of Public Health, 26*(S2), 3781–3792. https://doi.org/10.70135/seejph.vi.5856

[13]. Kumari, U., & Bajaj, T. (2025). NGS data analysis and active site identification of alpha-1-acid glycoprotein bound to potent anti-tumor compound UCN-01 in malignant brain tumor. *International Journal of Emerging Technologies and Innovative Research, 12*(3), G107–G116. https://doi.org/10.1729/Journal.44259

[14]. Kumari, U., & Renu. (2025). Structure analysis and molecular docking of mesothelin-207 fragment in human cancer. *Journal of Emerging Technologies and Innovative Research, 12*(3), G98–G106. https://doi.org/10.1729/Journal.44261

[15]. Kumari, U., Verma, G., & AnbuMegala, M. (2025). In silico drug design in human apoptosis-inducing factor (AIF) in lung cancer. *International Journal of Emerging Technologies and Innovative Research, 12*(2), D139–D148. https://doi.org/10.1729/Journal.43606

[16]. Kumari, U., & Mehrotra, K. (2025). NGS and proteomic gene expression analysis in studies of NUDT5 silencing hormone signaling in breast cancer. *International Journal of Emerging Technologies and Innovative Research, 12*(2), C630–C638. https://doi.org/10.1729/Journal.43563

[17]. Kumari, U., & Krishnakumar, P. (2024). Analysis of gastro-intestinal stromal tumor by computer aided drug designing and bioinformatics tools. *International Journal of Emerging Technologies and Innovative Research, 11*(2), C56–C65.

[18]. Kumari, U., & Arya, P. (2024). Molecular docking approach for human oncology research with chemoinformatic. *International Journal of Emerging Technologies and Innovative Research, 11*(2), A658–A666.

[19]. Kumari, U., Bhoomika, A., & Bhavana, K. S. (2023). In silico computer-aided drug designing of the mutated EGFR kinase domain of NSCLC (non-small cell lung cancer). *International Journal of Emerging Technologies and Innovative Research, 10*(12), B226–B235. https://doi.org/10.1729/Journal.37000

[20]. Kumari, U., Srivastava, N., & Tripathi, K. (2024). Validation of NGS and multi-omics approach for active site identification CDK2-associated protein 1; oral cancer. *International Journal for Multidisciplinary Research, 6*(2). https://doi.org/10.36948/ijfmr.2024.v06i02.15979

[21]. Kumari, U. (2024). Computational insights to the drug discovery on MET tyrosine kinase inhibition in tumor efficacy. *International Journal of Emerging Technologies and Innovative Research, 11*(5), J408–J422.