IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)
ISSN: 2455-264X, Volume 12, Issue 3 Series 1 (May. – June. 2026), PP 16-24
www.iosrjournals.org

AI Based Multi-Omics Analysis And Structure-Based Drug Discovery Pipeline Targeting Kras4a Mutations For Precision Oncology

Sonali, Uma Kumari
Project Trainee At Bioinformatics Project And Research Institute, Noida - 201301, India
Senior Bioinformatics Scientist, Bioinformatics Project And Research Institute, Noida - 201301, India

Abstract:
Overall analysis presents an integrated computational framework for precision oncology targeting KRAS4A mutations in Non-Small Cell Lung Cancer (NSCLC). The workflow combines Next-Generation Sequencing (NGS), Computer-Aided Drug Design (CADD), Artificial Intelligence (AI), and bioinformatics tools to systematically identify and evaluate potential therapeutic targets. Structural analysis of KRAS4A (PDB ID: 9GGY) revealed a stable conformation with a well-defined binding pocket between switch I and switch II regions. Sequence similarity and evolutionary relationships were assessed using sequence similarity and protein embedding models, including ProtBERT and ESM, highlighting conserved domains and mutation-driven variability. Machine learning approaches, including feature selection, clustering, and classification algorithms, enabled the identification of key biomarkers and cancer stem cell (CSC) populations associated with tumour progression and therapy resistance. Molecular docking studies demonstrated strong binding affinities of targeted inhibitors such as adagrasib, sotorasib, and afatinib within the KRAS4A active site. Toxicity profiling using ProTox-III supported the safety assessment of candidate compounds. Structural validation using Ramachandran analysis and Root mean square confirmed the reliability of the protein model. Overall, the proposed pipeline provides a reproducible and scalable framework integrating multi-omics data and AI-driven analysis for efficient drug discovery and personalised cancer therapy.

Keywords: *KRAS4A; NSCLC; NGS; CADD; Artificial Intelligence; ML; Molecular Docking; Precision Oncology; Biopython*

Date of Submission: 26-04-2026 Date of Acceptance: 06-05-2026

I. Introduction

Lung cancer is one of the major factors of cancer-related fatality rate globally, leading to 1.8 million deaths per year and representing a serious global health burden. Non-small cell lung cancer (NSCLC) accounts for 85% of all lung cancer cases and is marked by advanced molecular heterogeneity caused due to genetic variations. While targeted therapies and immunotherapy have significantly contributed to advancements, treatment remains limited because of late diagnosis, tumour progression and uncontrolled cell growth. KRAS is one of the most mutated genes in NSCLC, specifically in lung adenocarcinoma [1,2,3].

KRAS is one of the most mutated genes in NSCLC, specifically in lung adenocarcinoma. KRAS encodes a small GTPase that acts as a molecular switch, cycling between inactive GDP-bound and active GTP-bound states to manage downstream signalling pathways such as MAPK/ERK and PI 3K/AKT [4]. Mutations in the KRAS gene lead to the impairment of GTPase activity, which ultimately causes an active downstream signalling pathway that further causes cell proliferation and tumour development. Besides, KRAS-mutant tumours show major biological diversity that leads to poor diagnosis and tolerance to traditional therapies [4]. Advanced NSCLC with alterations that cause cells to become cancerous, such as EFGR mutations, are treated using targeted therapies, including EFGR tyrosine kinase inhibitors (EFGR-TKIs). KRAS is a GTPase that controls cell signalling through GTP hydrolysis. Most KRAS mutations occur at codons 12, 13 and 61, with G12 being the most frequent. The MAPK pathway promotes cell growth and tumour formation. RAS also activates PI3, which leads to AKT activation of the mTOR pathway. This pathway supports cancer growth and its survival. Yet, drugs that directly target KRAS show limited effectiveness and safety. Hence, chemotherapy and immunotherapy are broadly used as first-line treatment for NSCLC with KRAS mutations [5]. KRAS has two isoforms, KRAS4A and KRAS4B, which differ in exon 4, at their C-terminal regions and membrane binding regions. While KRAS4B has been studied to a greater extent, studies suggest that KRAS4A plays an analytical role in tumour biology, including metabolic regulation, oxidative stress adaptation, and cancer stem cell maintenance [6]. Recent studies

indicate that KRAS4A activity relates to patient prognosis and plays an important role in treatment response, especially in immunotherapy [7].

The protein structure with the PDB ID 9GGY represents human KRAS4A in its GDP-bound state with a small molecular inhibitor. At high resolution (~ 1.27A), this structure shows a drug-like pocket located between switch I and switch II regions, which plays an important role in effector binding and cell signalling [8]. NGS has emerged as a crucial approach in cancer research for evaluating the genetic complexity of tumours by enabling detailed identification of genetic variations at high resolution. NGS allows detailed analysis of cancer genomes, identifies mutations, structural variations, and genomic changes involved in cancer development. This information is essential for precision medicine as it allows treatments to be performed according to the needs of the patient's genetic profile [9].

CADD is a modern drug discovery tool that integrates computational chemistry, structural biology and bioinformatics. CADD enables us to obtain three-dimensional structures of cancer-related proteins and enzymes from publicly available databases. These 3D models can be used for in-silico drug discovery. CADD, along with molecular docking, facilitates the identification of therapeutic compounds with improved efficiency and efficacy. CADD is categorised into structure-based drug design (SBDD) and ligand-based drug design (LBDD). Structure-based techniques focus on high-resolution 3D structures that fit into the target binding site. In the KRAS4A crystal structure, a druggable binding pocket is located between the switch I and switch II regions, which are important for signal transduction and signalling [10,11]. Artificial intelligence and machine learning techniques are also used as powerful tools for data analysis. An AI model can identify hidden patterns within complex datasets, enabling improved prediction of drug-target interactions, biomarker discovery, and therapeutic response [12]. Bio Python is a widely used open-source bioinformatics toolkit. It is used for sequence analysis, structural parsing, and data integration. It assists in handling NGS data, aligning sequences, identifying mutations, and mapping them on to protein structures such as 9GGY [13]. While targeted therapies such as adagrasib and sotorasib have improved, issues remain in pharmacokinetics, ligand binding for the KRAS4A switch vs PDB ID: 9GGY, and toxicity problems exacerbated by tumour heterogeneity/cancer stem cell (CSC) populations identified by multi-omics. Current drug discovery pipelines usually silo Next-Generation Sequencing (NGS), Computer-Aided Drug Design (CADD) and AI with little or no integration for reproducibly discovering biomarkers for personalised oncology. This paper proposes a unified computational framework that incorporates NGS, CADD, AI models (e.g., ProtBERT, ESM), molecular docking and ProTox-III toxicity profiles to allow for the accurate prediction, optimisation, and validation of therapeutic ligands and to afford drug developers an effective methodology to conduct drug discovery in KRAS4A-mutated NSCLC.

II. Materials And Methods

Structural data retrieval

The 3D structure of human KRAS4A was retrieved from the Protein Data Bank (PDB) using PDB ID: **9GGY | pdb_00009ggy,** representing KRAS in its GDP-bound state with a small-molecule inhibitor [14].

Protein Visualisation:

3D structural visualisation using RasMol (v2.7.5) was done to interpret molecular structure in different representations, including ribbon, surface and stick models. These techniques allowed the identification of secondary structural elements and ligand-binding pockets [29,30] Wireframe and backbone command used to identify N-terminal and C-terminal regions. Space fill is used for the van der Waals surface representation. Structural preparation is done by removing water molecules and non-essential heteroatoms, followed by hydrogen addition [15].

Computational environment and tool deployment:

Google Colab- cloud computing environment used for 9GGY structural data. Bio Python (v1.86) was used for PDB parsing and sequence alignment. Matplotlib is used for data visualisation, including validation metrics and Cancer Stem Cell (CSC) identification plot. Ramachandran Plot: used to check backbone geometry. Ramachandran analysis is used to classify residues into General, Glycine, Proline, etc. [16,17,18].

Advanced data processing and Multi-omics Analysis

The Pandas library was used as the primary tool for data manipulation. Heatmaps were generated to visualise the correlation matrix between the KRAS4A site and ligand binding affinity. Protein Language Models and Embedding Generation: ESM Protein language model evaluates the co-evolutionary pattern of the KRAS family. Phylogeny-based LLM Embedding used to differentiate the KRAS4A variant from 4B, even with 92.02 % similarity. Principal Component Analysis (PCA) is used for initial dimensionality reduction. UMAP (Uniform Manifold Approximation Projection) is utilised to cluster embeddings in 2D and 3D formats. Latent Space Density Estimation locates areas of high instability. K-means clustering differentiates data and helps to identify stable vs

unstable binding pockets. Clustering Analysis was done to train the AI model so that it can differentiate between Bulk tumour and Cancer Stem Cell. Statistical feature selection, Analysis of Variance (ANOVA), removes features that don't contribute to the variance in drug efficacy.

Sequence alignment visualisation

Protein-protein BLAST **(**BLASTp) to measure homology and evolutionary context. Protein sequence analysis of KRAS4A was done using BLASTp to determine sequence identity. [19] The amino acid sequence was retrieved in FASTA format and was used as a query. It measures percent identity and alignment score. Sequences with high identity and low E-values were considered to be good [20].

3D visualization and pocket analysis

For advanced structural analysis, PyMOL served as the primary platform. It was used to visualise ligand-binding pockets, assess parameters such as bond angles, and generate a high-resolution quality image. The protein structure is prepared by removing water molecules and other heteroatoms that interfere during analysis. Polar hydrogen atoms to stabilise the structure. For structure visualisation, a stick model is used to clearly see atomic bonding, a surface model to identify the binding pocket, space fill model to see Vander Waal interactions [21,22,23].The root mean square deviation value is calculated to identify structural differences of aligned structures. A low RMSD value (~0.112 A) indicated high conformational stability [24,28].

Molecular docking

Molecular docking was performed using Auto Dock Vina to evaluate ligand binding with the KRAS4A protein (PDB ID: 9GGY), following standard protein and ligand preparation protocols. A grid box was defined around the active site (Switch I and II regions), and docking simulations were conducted to predict optimal binding conformations based on minimum binding energy [25,26].

Toxicity analysis

In silico toxicity prediction carried out using Protox Toxicity Prediction System. Some evaluated parameters: Toxicity class, Organ toxicity, Carcinogenicity, and Mutagenicity**.** Toxicity classification: Class I: Fatal, Class II: Toxic, etc. Model reliability: Low similarity values indicated reduced confidence [27].

III. Results And Discussion

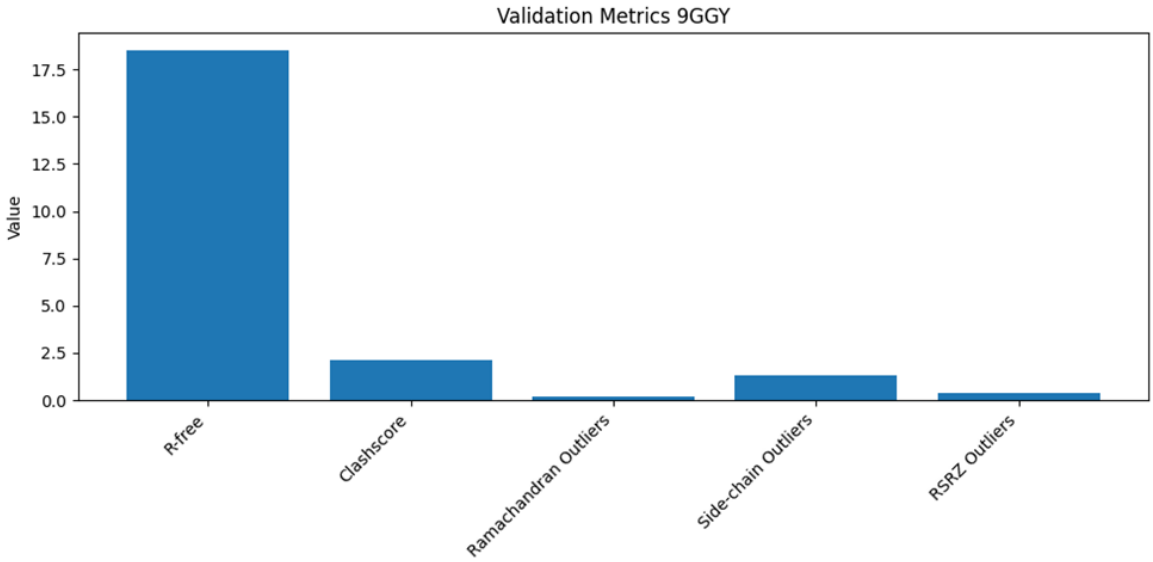


Figure 1: Validation metrics bar plot (represents the structural quality of KRAS PDB: 9GGY)

R free- 18.5% indicate high crystallographic reliability (Figure 1). Clash score- 2.1 reflects minimal steric conflict. Ramachandran outliner- 0.2% excellent backbone geometry. Sidechain outliner 1.3% - within acceptable limit.

PDB 9GGY is structurally validated and biologically meaningful for cancer-related computational studies.

Cancer Stem cell Bio programming

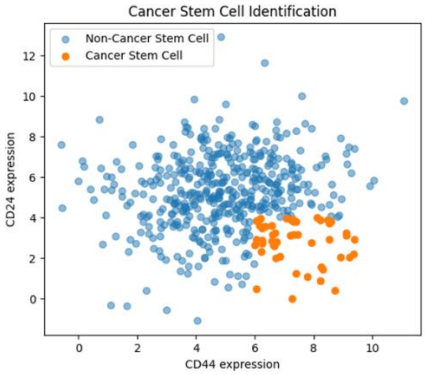


Figure 2: CD44 vs CD24 scatter plot.

The separation of clusters indicates heterogeneity within the tumour microenvironment 9Figure 2). CSC identification highlights therapy-resistant subpopulation driving tumour aggressiveness [29].

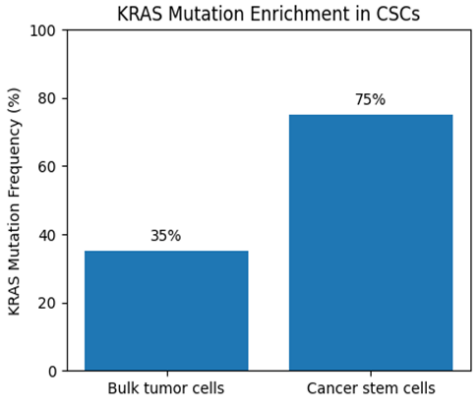


Figure 3: KRAS mutation frequency bar plot - Bulk tumour cells (35%), Cancer stem cell (75%).

KRAS mutations are preferentially enriched in CSC, explaining resistance, metastasis and recurrence (Figure 3). KRAS-driven CSCs represent a high-value therapeutic target.

Similarity Heatmap Visualisation and Prot BERT: Protein Bidirectional Encoder Representation from Transformer

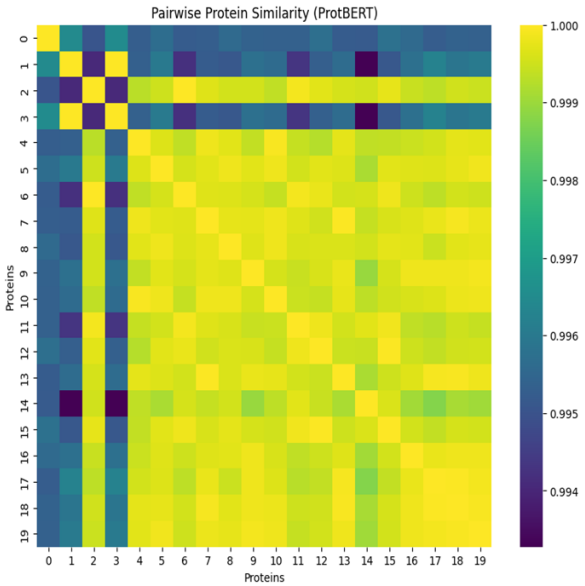


Figure 4: Representation of cosine similarity heatmap (ProtBERT embeddings) of multiple proteins.

High overall similarity (Yellow region ~0.998-1.000). Most of the matrix in the above figure is bright yellow, indicating very high similarity between proteins. (Similar sequence, conserved domain) (Figure 4). Diagonal line (self-similarity = 1.0) The bright diagonal from top left to bottom right represents each protein compared with itself. Low similarity block (purple/blue region ~0.994-0.996), these protein matrices indicate structurally and functionally different, possible outliers or divergent sequences, belong to different families of protein. Large yellow blocks suggest a cluster of highly similar proteins (ProtBERT). Small differences (e.g., 0.995 vs 0.999) are minor mutations and domain variation. The heatmap demonstrates that Prot BERT embedding successfully captures proteins with a similar pattern.

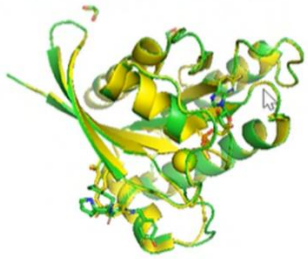


Figure 5: RMSD Calculation 9ggy(green) 9G0Y (yellow)

Root mean square deviation measures the average distance between corresponding atoms 9ggy green and 9g0y yellow. In this analysis, observe that 0.112 is an extremely low value represent very high structural similarity, with no major conformational differences (Figure 5). In PyMOL, alignment gives RMSD = 0.112 Å overall protein backbone, nearly identical, and the binding pocket geometry is conserved. 9ggy vs 9g0y, that structure has structural conserved have structural similarity superimposition indicating an exceptionally high degree of structural similarity.

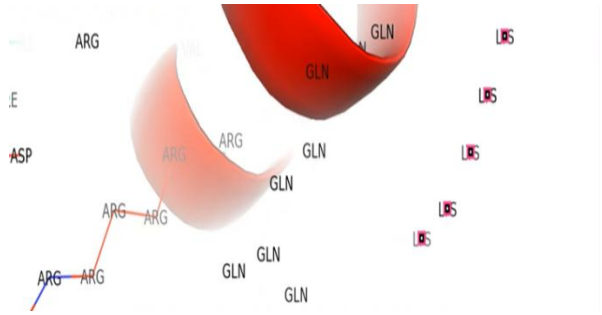


Figure 6: B factor analysis representing b factor value lys (67.40) residue position 169 Red color in Pymol (spectrum b) – high atomic mobility strctural

PDB 9GGY correspond to a KRAS G12C mutant structure, commonly associated with non-small cell lung cancer (NSCLC) Lysine residue in the loop region has B=67.40 (high flexibility loop that's the region it participate ligand binding, effector recognition, and experimental factor(Local disorder). B-factor of 67.40A in LYS (loop region) of 9GGY indicates highly flexible and structurally dynamic my contribute to conformational changes involved in KRAS signalling or drug binding.

Afatinib PubChem CID 10184653

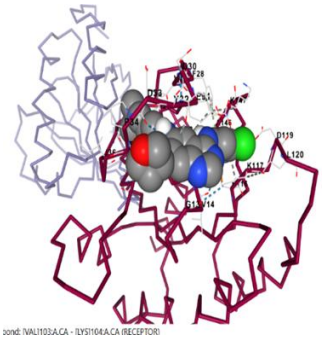


Figure 7: Molecular Docking (protein 871 and ligand Afatinib with score -8.3)

The proteomic structure demonstrates that Afatinib bind irreversibly within the pocket C2 in KRAS4a (Figure 7).

Figure 8: Molecular Docking table with score -8.3 kcal/mol

PDB ID	Ligand	Autodock vina score	Center	Remarks
9GGY	Afatinib	-8.3	C2 pocket	Strong
9GGY	Adagrasib	-9.1	C2 pocket	Very strong
9GGY	Sotorasib	-8.1	C2 pocket	Strong

Among the above binding sites best binding pocket C2 with vina score -8.3 kcal/mol, indicate the strongest predicted binding affinity/ cavity volume (859A) and is apropriate for Aftanib size C1, which has a larger volume (1241A) but weaker binding (-6.4), C5, C3, and C4 show moderate to weaker binding (Figure 8).

Adagrasib

Adagrasib is a valuable drug for the targeted therapy for KRAS G12C-mutated advanced NSCLC - especially after initial treatments have failed - because it directly blocks a key driver mutation in these cancer cells, leading to tumour shrinkage and delayed progression in many patients. Only cancers with the KRAS G12C mutation respond to this drug, which occurs in a subset of NSCLC cases (roughly around 10-14%). In December 2022, the U.S. Food and Drug Administration (FDA) granted accelerated approval for Adagrasib (Krazati) to treat adult patients with locally advanced or metastatic NSCLC that has the KRAS G12C mutation.

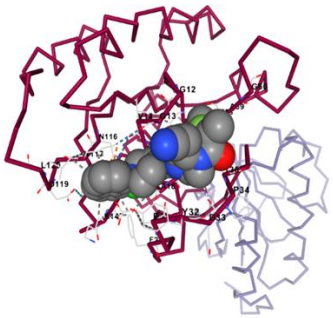


Figure 9: Molecular docking (Protein 8T71 with ligand Adagrasib score -9.1 c2)

In the backbone 3D structures, the grey space filling model represents the Adagrasib ligand position near the G12 regions mutation. Docking results confirm that Adagrasib exhibits strong affinity toward the KRAS G12C structure (8GGY), with proper orientation within the Switch II pocket. The presence of Cys12 enables covalent inhibition, validating its therapeutic potential in KRAS G12C-mutant NSCLC.

Sotorasib drug:

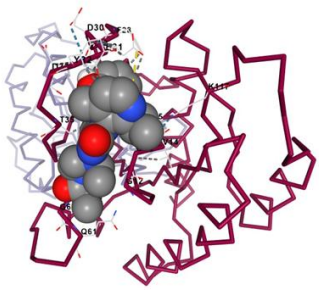


Figure 10: Molecular Docking (Protein-ligand (Sotorasib) interaction score C2(-8.1)

Residue chain:

Chain A: GLY12 GLY13 VAL14 GLY15 LYS16 SER17 ALA18 PHE28 VAL29 ASP30 GLU31 TYR32 ASP33 PRO34 THR35 ALA59 GLY60 GLN61 GLU62 ASN85 ASN86 ASN116 LYS117 ASP119 LEU120 SER122 SER145 ALA146 LYS147

These residues belong to the Switch I region (30-38), P-loop region (10–17), and Switch II region (59–67). The interactions include Hydrogen bonds (dashed blue lines), Hydrophobic interactions, Van der Waals contacts. This indicates Sotorasib is positioned near the Switch II pocket (SII-P)- the known druggable pocket in KRAS (Figure 10).

From the Molecular docking table: Best binding pocket: C2, Vina score = -8.1 kcal/mol- Indicates good predicted binding affinity

Generally, values below -7.0 kcal/mol suggest a stable interaction. Pocket volume (859 Å^3) is appropriate for Sotorasib accommodation. Docking grid size (24×24×24 Å) adequately covers the Switch II region.

Toxicity assessment of Sotorasib

Protox was used for assessing the toxicity of the proposed compounds for their screening, adsorption, permeability, reducing cost and time.

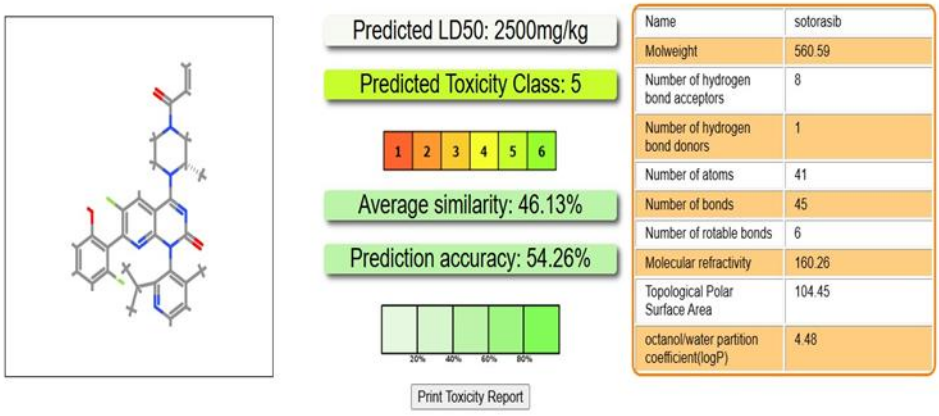


Figure 11: Toxicity Assessment Analysis

Predicted LD50 value of 2500 mg/kg places the compound in toxicity class 5, indicating non-toxicity (Figure 11). The average similarity of 46.13% and prediction accuracy 54.26 % are relatively low, indicating the model has limited confidence due to weak similarity with the compounds in its training dataset. Harmful only at relatively high doses, this suggests the compound is safe within therapeutic limits. Model reliability parameters: Average similarity 46.13%, moderate similarity of the compound. Prediction accuracy: 54.26- Moderate confidence level (accuracy is acceptable). Overall, the compound shows favourable drug-like properties (Figure 11). Further experimental validation (in vitro and in vivo toxicity studies) is necessary before confirming safety.

Graph:

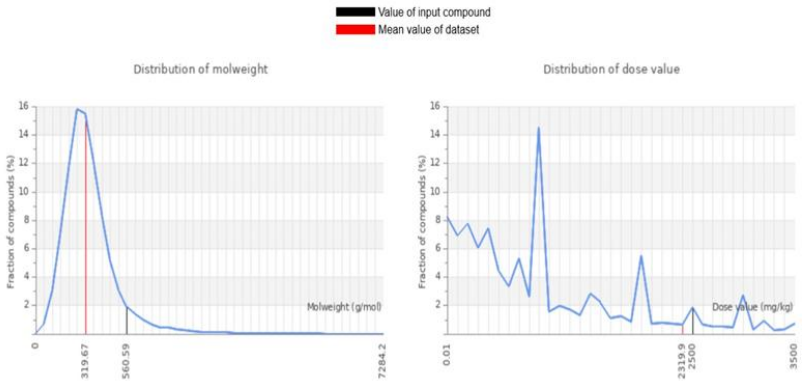
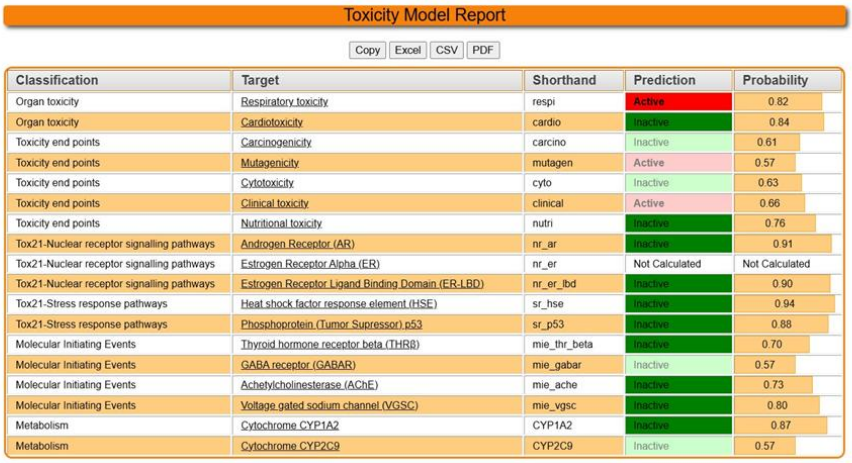


Figure 12: Representation of the distribution of molecular weight and distribution and molecular dose value.

The black line represents the value of the input compound, and the red line shows the mean value of the dataset (this represents the average value of all compounds in the dataset). The drug is acceptable due to the dose value within the normal range pharmacologically acceptable. The plot shows the distribution for compounds concentrated in the low molecular weight region (200-400 g/mol), with the mean molecular weight (319.67 g/mol) indicated by the red line (Figure 12).

Report:



Toxicity Model Report

Copy Excel CSV PDF

Classification	Target	Shorthand	Prediction	Probability
Organ toxicity	Respiratory toxicity	respi	Active	0.82
Organ toxicity	Cardiotoxicity	cardio	Inactive	0.84
Toxicity end points	Carcinogenicity	carcino	Inactive	0.61
Toxicity end points	Mutagenicity	mutagen	Active	0.57
Toxicity end points	Cytotoxicity	cyto	Inactive	0.63
Toxicity end points	Clinical toxicity	clinical	Active	0.66
Toxicity end points	Nutritional toxicity	nutri	Inactive	0.76
Tox21-Nuclear receptor signalling pathways	Androgen Receptor (AR)	nr_ar	Inactive	0.91
Tox21-Nuclear receptor signalling pathways	Estrogen Receptor Alpha (ER)	nr_er	Not Calculated	Not Calculated
Tox21-Nuclear receptor signalling pathways	Estrogen Receptor Ligand Binding Domain (ER-LBD)	nr_er_lbd	Inactive	0.90
Tox21-Stress response pathways	Heat shock factor response element (HSE)	sr_hse	Inactive	0.94
Tox21-Stress response pathways	Phosphoprotein (Tumor Supressor) p53	sr_p53	Inactive	0.88
Molecular Initiating Events	Thyroid hormone receptor beta (THRß)	mie_thr_beta	Inactive	0.70
Molecular Initiating Events	GABA receptor (GABAR)	mie_gabar	Inactive	0.57
Molecular Initiating Events	Achetylcholinesterase (AChE)	mie_ache	Inactive	0.73
Molecular Initiating Events	Voltage gated sodium channel (VGSC)	mie_vgsc	Inactive	0.80
Metabolism	Cytochrome CYP1A2	CYP1A2	Inactive	0.87
Metabolism	Cytochrome CYP2C9	CYP2C9	Inactive	0.57

Figure 13: Toxicity Model Report

The compound shows active respiratory toxicity (0.82), indicating adverse effects on the respiratory system. Carcinogenicity is predicted to be inactive (0.61), indicating a low cancer-causing potential, while mutagenicity is active (0.57), suggesting a risk of genetic mutations. Cytotoxicity is inactive (0.63), but clinical toxicity is active (0.66), indicating possible adverse effects in physiological conditions. Nuclear Receptor Signalling Pathways (Tox21) inactive interaction with Androgen Receptor (AR) and Estrogen Receptor Ligand Binding Domain (ER-LBD) with high confidence (0.91 and 0.90, respectively). Stress Response Pathways (HSE:0.94), (p53 0.88) unlikely to induce cellular stress. Thyroid hormone receptor beta (THRβ) (0.70), GABA receptor (0.57), Acetylcholinesterase (AChE) (0.73), Voltage-gated sodium channels (VGSC) (0.80) suggest a low probability of neurotoxicity or endocrine disruption at the receptor level. CYP1A2 inhibition: inactive (0.87), CYP2C9 inhibition: inactive (0.57) indicates a low likelihood of metabolic interference or drug–drug interactions mediated by these cytochrome P450 enzymes.

It is important for industrial research in Compound screening and selection, toxicity and dose assessment, and distribution analysis, allowing the industries to quickly evaluate whether a compound fits within the acceptable chemicals and pharmacological space.

IV. Conclusion

Comprehensive computational pipeline for targeting KRAS4A mutations in NSCLC through the integration of structural biology, multi-omics analysis, and artificial intelligence. The KRAS4A structure (9GGY) demonstrated high structural integrity and suitability for drug design studies. Sequence and embedding analyses revealed both conservation and functional divergence among KRAS isoforms, supporting the biological relevance of mutation-specific targeting. Machine learning techniques effectively identified critical features and distinguished cancer stem cell populations, emphasising their role in tumour recurrence and therapeutic resistance. Molecular docking results confirmed that KRAS G12C-specific inhibitors, particularly Adagrasib and Sotorasib, exhibit strong binding affinity within the switch II pocket, validating their role as targeted therapies. However, their mutation-specific activity highlights the importance of precision medicine approaches. Toxicity prediction further strengthened the evaluation of drug candidates.

Overall, this study provides a scalable and reproducible framework that can be applied to other oncogenic targets, bridging computational biology and clinical oncology for improved therapeutic development.

References:

[1]. Corral De La Fuente E, Olmedo Garcia ME, Gomez Rueda A, Lage Y, Garrido P. Targeting KRAS In Non-Small Cell Lung Cancer. Front Oncol. 2022 Jan 10;11:792635. Doi: 10.3389/Fonc.2021.792635. PMID: 35083149; PMCID: PMC8784727.Mechanisms And Current Advances In Treating KRAS-Mutated Lung Cancer

[2]. Zhao, R., Shu, Y., Xu, W. Et Al. The Efficacy Of Immunotherapy In Non-Small Cell Lung Cancer With KRAS Mutation: A Systematic Review And Meta-Analysis. Cancer Cell Int 24, 361 (2024). Https://Doi.Org/10.1186/S12935-024-03498-9

[3]. Ritu, K., Kumar, P., Singh, A. Et Al. Untangling The KRAS Mutated Lung Cancer Subsets And Its Therapeutic Implications. Mol Biomed 2, 40 (2021). Https://Doi.Org/10.1186/S43556-021-00061-0

[4]. Ghosh S, Bhuniya T, Dey A, Koley M, Roy P, Bera A, Gol D, Chowdhury A, Chowdhury R, Sen S. An Updated Review On KRAS Mutation In Lung Cancer (NSCLC) And Its Effects On Human Health. Appl Biochem Biotechnol. 2024 Jul;196(7):4661-4678. Doi: 10.1007/S12010-023-04748-8. Epub 2023 Oct 28. PMID: 37897621.

[5]. He, Q. Et Al. (2025) 'First-Line Treatments For KRAS-Mutant Non-Small Cell Lung Cancer: Current State And Future Perspectives', Cancer Biology & Therapy, 26(1). Doi: 10.1080/15384047.2024.2441499.

[6]. Ma, Q., Zhang, W., Wu, K. Et Al. The Roles Of KRAS In Cancer Metabolism, Tumor Microenvironment And Clinical Therapy. Mol Cancer 24, 14 (2025). Https://Doi.Org/10.1186/S12943-024-02218-1

[7]. Chiari R, Palladino S, Emili R, De Lisa M, Sarti D, Catalano V, Magnani M, Graziano F, Ruzzo A. KRAS4A And KRAS4B In Liquid Biopsy Of Metastatic Lung Adenocarcinoma Patients Treated With Pembrolizumab Or Chemotherapy Plus Pembrolizumab. Sci Rep. 2023 Nov 29;13(1):21036. Doi: 10.1038/S41598-023-48304-0. PMID: 38030703; PMCID: PMC10687227.

[8]. "9GGY: KRAS4A Structure." RCSB Protein Data Bank, Https://Www.Rcsb.Org/Structure/9GGY.

[9]. Goodwin, S., Mcpherson, J. & Mccombie, W. Coming Of Age: Ten Years Of Next-Generation Sequencing Technologies. Nat Rev Genet 17, 333–351 (2016). Https://Doi.Org/10.1038/Nrg.2016.49

[10]. Panwar U, Murali A, Khan MA, Selvaraj C, Singh SK. Virtual Screening Process: A Guide In Modern Drug Designing. Methods Mol Biol. 2024;2714:21-31. Doi: 10.1007/978-1-0716-3441-7_2. PMID: 37676591.

[11]. Opeyemi Iwaloye, Paul Olamide Ottu, Femi Olawale, Olorunfemi Oyewole Babalola, Olusola Olalekan Elekofehinti, Babatomiwa Kikiowo, Abayomi Emmanuel Adegboyega, Henry Nnaemeka Ogbonna, Covenant Femi Adeboboye, Ibukun Mary Folorunso, Aderonke Elizabeth Fakayode, Moses Orimoloye Akinjiyan, Sunday Amos Onikanni, Sergey Shityakov, Computer-Aided Drug Design In Anti-Cancer Drug Discovery: What Have We Learnt And What Is The Way Forward? Informatics In Medicine Unlocked,Volume 41, 2023, 101332, ISSN 2352-9148, Https://Doi.Org/10.1016/J.Imu.2023.101332.

[12]. Vamathevan, J., Clark, D., Czodrowski, P. Et Al. Applications Of Machine Learning In Drug Discovery And Development. Nat Rev Drug Discov 18, 463–477 (2019). Https://Doi.Org/10.1038/S41573-019-0024-5

[13]. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, De Hoon MJ. Biopython: Freely Available Python Tools For Computational Molecular Biology And Bioinformatics. Bioinformatics. 2009 Jun 1;25(11):1422-3. Doi: 10.1093/Bioinformatics/Btp163. Epub 2009 Mar 20. PMID: 19304878; PMCID: PMC2682512.

[14]. Schuettelkopf, A.W., Parry, C.W., Pellicano, F., Beyer, K.S., Et Al. (2025). Human Kras4a (GDP) In Complex With Compound 29. RCSB Protein Data Bank. DOI: Https://Doi.Org/10.2210/Pdb9ggy/Pdb

[15]. Sayle, R. A., & Milner-White, E. J. (1995). RASMOL: Biomolecular Graphics For All. Trends In Biochemical Sciences, 20(9), 374-376. Https://Doi.Org/10.1016/S0968-0004(00)89080-5

[16]. Kumari U, Adya AP, Satheesan S. Drug Discovery And Biopython Analysis Of MBP–MCL1 In Myeloid Cell Leukemia. J Emerg Technol Innov Res. 2025;12(1):F178–F187.

[17]. Kumari U, Gupta S. NGS And Sequence Analysis With Biopython For Prospective Brain Cancer Therapeutic Studies. Int J Res Appl Sci Eng Technol. 2023. Https://Doi.Org/10.22214/Ijraset.2023.50885

[18]. Kumari U, Kaur G, Et Al. Biopython/Network Of Protein Identification And NGS Analysis Of Glioma Cancer ATP Competitive Type III C-MET Inhibitor. J Emerg Technol Innov Res. 2024;11(2):41–51. Http://Doi.One/10.1729/Journal.40229

[19]. Madden T. The BLAST Sequence Analysis Tool. The NCBI Handbook. 2013 Mar 15;2(5):425-36.

[20]. Kumari U, Mehrotra K. NGS And Proteomic Gene Expression Analysis Of NUDT5 In Breast Cancer. J Emerg Technol Innov Res. 2025;12(2):C630–C638. Http://Doi.One/10.1729/Journal.43563

[21]. Kumari U, Mehrotra K. NGS And Proteomic Gene Expression Analysis Of NUDT5 In Breast Cancer. J Emerg Technol Innov Res. 2025;12(2):C630–C638. Http://Doi.One/10.1729/Journal.43563

[22]. Kumari U, Agrawal N. NGS And Mutational Profile Analysis Of Non-Small-Cell Lung Carcinoma (NSCLC). Int J Res Appl Sci Eng Technol. 2023;11:3090–3094. Https://Doi.Org/10.22214/Ijraset.2023.50880

[23]. Kumari U, Bajaj T. NGS Data Analysis And Active Site Identification Of Alpha-1-Acid Glycoprotein Bound To Anti-Tumor Compound UCN-01. J Emerg Technol Innov Res. 2025;12(3):G107–G116. Http://Doi.One/10.1729/Journal.44259

[24]. Pearson, W.R. (2020). Sequence Similarity Searching Evolution. Curr Opin Struct Biol. DOI: 10.1016/J.Sbi.2020.05.002

[25]. Kumari U, Renu, Et Al. Structure Analysis And Molecular Docking Of Mesothelin-207 Fragment In Human Cancer. J Emerg Technol Innov Res. 2025;12(3):G98–G106. Http://Doi.One/10.1729/Journal.44261

[26]. Chaudhary S, Kumari U. NGS, Molecular Docking And Network Pharmacology Reveal Potent Inhibitor For Lung Cancer Treatment. J Emerg Technol Innov Res. 2024;11(9):F116–F126. Https://Doi.One/10.1729/Journal.41696

[27]. Kumari U, Verma G, Anbumegala M. In Silico Drug Design Targeting Apoptosis Inducing Factor (AIF) In Lung Cancer. J Emerg Technol Innov Res. 2025;12(2):D139–D148. Http://Doi.One/10.1729/Journal.43606

[28]. Kumari U, Johri V, Dhopate S, Jha T. Structure-Based Drug Designing For Prediction Of Epitopes Targeting Malignant Brain Tumor. J Emerg Technol Innov Res. 2024;11(7).

[29]. Kumari U, Belokar P, Esompalli A, Deshpande S, Kumkar M, Tripathi A. Next-Generation Sequencing To Investigate P53 Mutant Y234C For Targeted Cancer Therapy. IOSR J Pharm Biol Sci. 2025;20(4):63–72.

[30]. Kumari U, Raj K, Maheshwari S, Et Al. Drug Discovery Against FGFR2 Mutation In Lung Cancer Using NGS And Medicinal Plants. J Emerg Technol Innov Res. 2025;12(5):H569–H583. Https://Doi.Org/10.56975/Jetir.V12i5.562943