Mugla Journal of Science and Technology

STRUCTURE-BASED OPTIMIZATION OF ADAGRASIB (MRTX849) ANALOGUES: ADVANCED COMPUTATIONAL FRAMEWORK FOR KRAS G12D INHIBITOR DESIGN

Ayşegül VAROL*, Department of Pharmaceutical Biology, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University-Mainz, 55128, Mainz, Germany, avarol@uni-mainz.de
(https://orcid.org/0009-0001-3631-6811)

Received: 19.11.2025, Accepted: 22.12.2025
*Corresponding author
Research Article
DOI: 10.22531/muglajsci.1826880

Abstract

The KRAS G12D mutation poses a major therapeutic challenge, particularly in pancreatic and colorectal cancers where current treatments are limited. While covalent inhibitors for KRAS G12C have reached clinical success, developing effective G12D-targeted agents remains difficult due to its unique structural and biochemical features. This study introduces a computational framework for structure-based optimization of adagrasib analogues targeting KRAS G12D. Using an Advanced Molecular Design Platform, fifty derivatives were designed by modifying positions 17–25 of the tetracyclic scaffold with medicinal chemistry-guided R-group substitutions. Molecular docking against KRAS G12D (PDB: 7RPZ) identified several high-affinity candidates (−6.9 to −9.6 kcal/mol) outperforming adagrasib (−7.7 kcal/mol). Structure–activity analysis revealed isopropyl substitution at position 17 as optimal, with Deriv-34 achieving the strongest binding (−9.6 kcal/mol) via key interactions with ARG-68, GLU-62, and TYR-96. Principal component analysis highlighted hydroxylated derivatives with superior drug-likeness (QED = 0.384) and synthetic feasibility. Comprehensive ADME profiling guided lead prioritization, defining a rational pipeline for KRAS G12D inhibitor design. This integrated computational approach provides a promising foundation for experimental validation and advances targeted therapy development against KRAS-driven cancers.

Keywords: KRAS G12D Mutation, Adagrasib analogues, Molecular docking, Drug-likeness, Targeted cancer therapy

ADAGRASIB (MRTX849) ANALOGLARININ YAPI TEMELLI OPTİMİZASYONU: KRAS G12D İNHİBİTÖR TASARIMI İÇİN GELİŞMİŞ HESAPLAMALI ÇERÇEVE

Özet

KRAS G12D mutasyonu, özellikle mevcut tedavi seçeneklerinin sınırlı olduğu pankreas ve kolorektal kanserlerde önemli bir terapötik zorluk oluşturmaktadır. KRAS G12C mutasyonu için kovalent inhibitörler klinik başarıya ulaşmış olsa da, G12D'yi hedefleyen etkili ajanların geliştirilmesi yapısal ve biyokimyasal farklılıklar nedeniyle güçtür. Bu çalışma, KRAS G12D'yi hedefleyen adagrasib analoglarının yapı temelli optimizasyonu için bir hesaplamalı çerçeve sunmaktadır. Gelişmiş Moleküler Tasarım Platformu kullanılarak, 17–25 pozisyonlarındaki R-gruplarının ilaç kimyası odaklı değiştirilmesiyle elli türev tasarlanmıştır. KRAS G12D (PDB: 7RPZ) üzerinde yapılan moleküler kenetlenme analizleri, adagrasib'e (−7.7 kcal/mol) kıyasla daha güçlü bağlanan adaylar (−6.9 ila −9.6 kcal/mol) belirlemiştir. Yapı-aktivite analizinde 17. pozisyondaki izopropil grubunun en uygun olduğu ve Deriv-34'ün ARG-68, GLU-62 ve TYR-96 ile etkileşimler üzerinden en yüksek bağlanma gücünü (−9.6 kcal/mol) gösterdiği saptanmıştır. Temel bileşen analizi, hidroksillenmiş türevlerin daha yüksek ilaç benzerliği (QED = 0.384) ve sentez kolaylığına sahip olduğunu ortaya koymuştur. Kapsamlı ADME değerlendirmesiyle adayların önceliklendirilmesi, KRAS G12D inhibitör tasarımı için rasyonel bir yol haritası sunmaktadır.

Anahtar Kelimeler: KRAS G12D mutasyonu, Adagrasib analogları, Moleküler kenetlenme, İlaç-benzerliği, Hedefe yönelik kanser tedavisi

Cite

Varol, A., (2025). "Structure-Based Optimization of Adagrasib (MRTX849) Analogues: Advanced Computational Framework for KRAS G12D Inhibitor Design", Mugla Journal of Science and Technology, 11(2), 135-144.

1. Introduction

KRAS mutations are among the most prevalent oncogenic drivers in human malignancies, occurring in approximately 30% of all cancers with particularly high frequencies in pancreatic (90%), colorectal (45%), and lung adenocarcinomas (35%) [1, 2]. Among these, the G12D mutation represents a significant clinical

challenge, accounting for 41% of KRAS-mutant pancreatic ductal adenocarcinomas and 28% of KRAS-mutant colorectal cancers [3, 4]. This point mutation substitutes aspartic acid for glycine at position 12, resulting in constitutive activation of KRAS and persistent downstream signaling through MAPK and PI3K pathways, thereby promoting uncontrolled cellular proliferation, survival, and metabolic reprogramming [5, 6]. For decades, KRAS was considered "undruggable" due to its high binding affinity for GTP/GDP, lack of deep hydrophobic pockets, and smooth protein surface that offers limited opportunities for small molecule binding [7, 8]. However, groundbreaking research over the past decade has transformed this paradigm, culminating in the recent clinical approval of sotorasib (AMG 510) and adagrasib (MRTX849) for mutant KRAS non-small cell lung cancer [9-11]. These irreversible inhibitors form a covalent bond with the mutant cysteine residue and lock KRAS in its inactive GDP-bound state, providing proof-of-concept for direct KRAS targeting [12, 13]. Despite these advances, developing effective inhibitors against other KRAS mutations, particularly G12D, presents distinct challenges. Unlike G12C, the aspartic acid substitution lacks a nucleophilic thiol group required for covalent bond formation, necessitating alternative targeting strategies [14]. Recent crystallographic studies have revealed that adagrasib derivatives can be modified to interact with the G12D mutant through specific non-covalent interactions, exploiting subtle conformational differences in Switch-II pocket geometry [15, 16]. Notably, MRTX1133, a KRAS G12D-selective inhibitor derived from the adagrasib scaffold, has shown promising preclinical activity by leveraging increased polarity and specific hydrogen bonding networks [17-19]. Computational approaches have become increasingly valuable in guiding KRAS inhibitor optimization by identifying favorable substitution patterns and predicting binding modes [20-22]. Structure-based drug design, molecular docking, and multiparameter optimization techniques have facilitated rapid exploration of chemical space while balancing potency with drug-like properties [23-25]. Moreover, recent advances in machine learning approaches have enhanced our ability to predict protein-ligand interactions and guide lead optimization strategies for challenging oncogenic targets [26-28].

Building upon these foundations, we report herein a systematic computational framework for structure-based optimization of adagrasib analogues targeting the KRAS G12D mutant. Our approach integrates molecular docking, physicochemical property analysis, and structure-activity relationship studies to identify promising chemical modifications that enhance binding affinity while maintaining favorable drug-like characteristics. By strategically exploring substitution patterns at positions 17-25 of the tetracyclic scaffold, we have identified several derivatives with significantly improved predicted binding affinities and refined molecular interactions with key residues in the Switch-II pocket. These findings provide valuable insights for the rational design of potent and selective KRAS G12D inhibitors and establish a comprehensive computational workflow applicable to other challenging oncogenic targets.

2. Materials and Methods

2.1. Computational Environment and Software Framework

All computational analyses were conducted using a custom-developed Advanced Molecular Design Platform implemented in Python 3.8. The cheminformatics foundation relied on RDKit (2023.09.5) [29], an open-source toolkit for computational chemistry and drug discovery, supplemented by NumPy (1.24.0), pandas (2.0.3), and scikit-learn (1.3.0) for data processing and statistical analyses. Visualization capabilities were provided by Matplotlib (3.7.2), Seaborn (0.12.2), and py3Dmol for interactive 3D molecular representation. This integrated computational environment facilitated seamless molecular manipulation, property calculation, and statistical analysis throughout the investigation.

2.2. Molecular Standardization Protocol

Input molecules underwent a rigorous standardization protocol to ensure consistent chemical representation. The molecular preprocessing workflow began with SMILES parsing using explicit hydrogen preservation, followed by application of the MolVS Standardizer for tautomer handling, charge neutralization, and stereochemistry validation. Three-dimensional conformer generation employed the ETKDG v3 algorithm [30] with a fixed random seed (42) to ensure reproducibility. For molecules where initial 3D embedding proved challenging, the system automatically implemented alternative embedding strategies with increased sampling attempts. Conformers were subsequently optimized using the Merck Molecular Force Field (MMFF94) with a maximum of 500 iterations, followed by hydrogen removal, stereochemistry refinement, and molecular sanitization to ensure chemically valid structures throughout the analysis pipeline.

2.3. R-Group Library Design and Selection Strategy

A comprehensive library of 31 R-groups representing medicinally relevant chemical moieties was established for molecular diversification. The library encompassed four primary chemical classes: (1) halogens and small functional groups (F, Cl, CF_3) assigned high probability weights (0.7-0.9); (2) alkyl substituents (methyl, ethyl, cycloalkyl) with medium weights (0.5-0.7); (3) polar functional groups (hydroxyl, amino, carboxyl) with variable weights (0.3-0.6); and (4) aromatic and heterocyclic systems with lower weights (0.2-0.4). The R-group selection algorithm employed probability-weighted random sampling, systematically favoring synthetically tractable modifications while maintaining

structural diversity. This approach effectively mimicked medicinal chemistry decision-making processes in lead optimization campaigns, balancing synthetic accessibility with chemical space exploration.

2.4. Modification Site Identification and Selection Criteria

Modification sites on parent molecules were algorithmically identified using structure-based reactivity criteria. The site identification algorithm evaluated each non-hydrogen atom against predefined reactivity rules that prioritized: (1) aromatic carbon atoms conducive to electrophilic aromatic substitution, and (2) heteroatoms (N, O, S) with available hydrogen atoms suitable for nucleophilic substitution reactions. These criteria were established to align with conventional synthetic chemistry approaches, focusing on positions with demonstrated reactivity in medicinal chemistry transformations. The modification site selection for individual derivatives employed random sampling from the qualified pool of sites, ensuring diverse exploration of substitution patterns across the parent scaffold while maintaining synthetic plausibility.

2.5. Derivative Generation and Molecular Editing Methodology

The derivative generation process implemented a systematic transformation of parent molecules into structurally diverse analogs through well-defined molecular editing operations. For each derivative, the workflow executed the following sequence: (1) random selection of a modification site from pre-identified candidates; (2) probabilistically weighted selection of an R-group from the library; (3) conversion of the parent molecule to an editable representation with explicit hydrogens; (4) identification and removal of a hydrogen atom at the selected position; (5) combination of the edited parent molecule with the selected R-group; (6) formation of a new covalent bond between the modification site and R-group; and (7) structure processing including sanitization, hydrogen removal, and canonical SMILES generation. A duplicate detection system based on canonical SMILES representations prevented redundancy, focusing computational resources on unique chemical entities. Generation continued until either the target number of derivatives was produced or the maximum attempt threshold was reached, with comprehensive error handling implemented throughout the process.

2.6. Molecular Property Calculation and ADMET Prediction

Comprehensive molecular characterization was achieved through calculation of physicochemical and drug-likeness properties for each generated compound. Standard descriptors included molecular weight, calculated octanol-water partition coefficient (LogP), hydrogen bond donors/acceptors (HBD/HBA), topological polar surface area (TPSA), rotatable bond count, and aromatic ring count, all computed via RDKit's descriptor module. Advanced metrics included quantitative estimate of drug-likeness (QED) [31], a custom synthetic accessibility score incorporating ring count and structural complexity factors, and a natural product likeness score based on heteroatom distribution patterns. ADMET properties were estimated using rule-based models that incorporated molecular descriptors to predict blood-brain barrier permeability and cytochrome P450 2D6 inhibition potential. These predictive models applied established medicinal chemistry heuristics including Lipinski's Rule of Five parameters, providing preliminary insights into pharmacokinetic behavior.

2.7. Chemical Space Analysis and Statistical Methods

Chemical space exploration was visualized using Principal Component Analysis (PCA) of the calculated molecular properties. The analytical procedure first standardized property data using z-score normalization (mean=0, standard deviation=1) via scikit-learn's StandardScaler, followed by dimensionality reduction using PCA to identify orthogonal vectors capturing maximal variance in the dataset. The first two principal components were extracted and used to generate two-dimensional projections of the chemical space, facilitating visualization of relationships between derivatives and parent compounds in terms of their physicochemical profiles. This approach enabled quantitative assessment of structural diversity and property distributions across the generated molecular library, identifying clusters and outliers within the multidimensional property space.

2.8. Drug-likeness Assessment and Lead Prioritization

Generated compounds were systematically evaluated for drug-likeness using Lipinski's Rule of Five [32], with violations tallied based on established criteria: molecular weight >500 Da, calculated LogP >5, hydrogen bond donors >5, and hydrogen bond acceptors >10. The assessment was implemented as a series of conditional statements evaluating each property against its corresponding threshold. Compounds were subsequently categorized based on violation counts, facilitating the prioritization of molecules with favorable drug-like characteristics. This systematic evaluation of pharmaceutical relevance complemented the structural diversity and synthetic accessibility assessments, providing a comprehensive framework for lead candidate identification and optimization.

2.9. Molecular Docking Studies

The X-ray crystallographic structure of the KRAS G12D mutant in complex with the inhibitor MRTX-1133 (PDB ID: 7RPZ) was retrieved from the RCSB Protein Data Bank. Prior to docking, the protein structure was refined by removing water molecules and co-crystallized ligands not involved in the binding interface, followed by the addition of polar hydrogen atoms and charges

using AutoDock Tools v1.5.6 [33]. Ligands of interest were energy-minimized and similarly prepared for docking. The docking grid was centered around the MRTX-1133 binding site, with dimensions set to 25 × 25 × 25 Å to accommodate ligand flexibility and fully enclose the active pocket. Docking simulations were carried out using AutoDock Vina, which employs a scoring function optimized through stochastic global optimization algorithms, including Genetic Algorithm with Local Search (GA-LS) [34, 35]. For each ligand, twenty binding conformations were generated, and the top-ranked poses based on binding affinity were selected for further analysis. The molecular interactions between the KRAS G12D protein and ligands were visualized using PyMOL and BIOVIA Discovery Studio Visualizer [36].

3. Results and Discussion

3.1. Design and Generation of Novel Derivatives

The study employed a rational design approach to develop a library of fifty derivatives based on a tetracyclic parent scaffold (MW = 604.13 Da, cLogP = 4.73). Modifications were strategically introduced at positions 17–25 to explore structure-activity relationships while maintaining core scaffold integrity (Figure 1). Halogen substitutions (F, Cl, Br, I) were incorporated to investigate hydrophobic interactions and electronic effects, while alkyl chains of varying lengths (methyl to butyl) probed steric tolerance limits. The design further incorporated electronic modulators, including strong electron-withdrawing groups (CF_3, CN, NO_2) and electron-donating moieties (OH, NH_2, OMe), to systematically alter electronic properties. Heterocyclic substitutions (morpholine, piperidine, furan) were introduced to assess the impact of polarity and conformational constraints on molecular properties. This comprehensive diversification strategy yielded ligands with significant variations in physicochemical properties, as detailed in Table 1, which presents representative examples spanning the full range of molecular diversity. The hydrogen bond acceptor count remained within Lipinski's recommended range (8–10) across all derivatives, while only three compounds (Deriv-10, -24, -46) were designed with hydrogen bond donors to strategically balance permeability and polarity.

3.2. Physicochemical Property Analysis

Detailed analysis of the derivative library revealed distinct trends in physicochemical properties critical for drug development. All compounds exhibited molecular weights exceeding 500 Da, reflecting the inherent complexity of the tetracyclic scaffold. Calculated LogP values ranged from 4.32 to 6.40, indicating moderate to high lipophilicity across the series. Polar surface area measurements showed considerable variation (88.83–131.97 $Å^2$), with electron-donating substituents such as amino and hydroxyl groups significantly enhancing molecular polarity. For instance, Deriv-10 (NH_2-substituted) achieved a TPSA of 114.85 $Å^2$, representing a 29% increase over the parent compound. Halogenated derivatives followed predictable lipophilicity trends, with fluorine substitutions yielding the lowest cLogP values (4.87) and bromine substitutions the highest (5.50). Interestingly, the iodinated Deriv-7 (MW = 730.03 Da) exhibited a slightly lower cLogP (5.34) than its brominated counterpart, suggesting that steric bulk may partially offset the lipophilicity contribution of the heavy halogen. These systematic variations in molecular properties provided a robust foundation for establishing structure-property relationships.

3.3. Drug-Likeness and Lead Prioritization

Comprehensive assessment of the molecular property distribution profiles (Figure 2) provided critical insights into the structure-property relationships governing our derivative library. Multivariate analysis of physicochemical parameters revealed that the majority of compounds (54%) violated two of Lipinski's Rule of Five criteria, predominantly due to elevated molecular weight (MW >500 Da) and lipophilicity (cLogP >5.0). This group included derivatives with bulky aromatic substitutions, such as the phenyl-substituted Deriv-37 (cLogP = 6.40). A significant subset (46%) violated only the molecular weight criterion, encompassing various polar analogs including the cyano-substituted Deriv-4 (MW = 629.14 Da, cLogP = 4.60). Notably, hydroxylated derivatives (Deriv-24 and -46) emerged as exceptional cases, fully complying with all Lipinski parameters while achieving the highest Quantitative Estimate of Drug-likeness (QED) scores (0.384) in the series. Synthetic accessibility scores remained favorable throughout the library (0.55–0.75), with the top-performing hydroxylated derivatives maintaining excellent synthetic tractability (SA = 0.60). This comprehensive assessment enabled clear prioritization of lead compounds based on balanced drug-like properties and synthetic feasibility.

3.4. Structure-Property Relationships and Pharmacokinetic Insights

Principal component analysis (Figure 3) revealed distinctive clustering patterns, with PC1 and PC2 accounting for 35.4% and 24.5% of variance, respectively. The parent compound (red star) occupies a unique position in chemical space. In-depth analysis of structure-property relationships revealed several key trends with implications for further optimization. Polar substituents, particularly hydroxyl and amino groups, demonstrated superior property modulation, increasing polar surface area by 20–30% while reducing cLogP by 0.4–1.0 units compared to alkyl analogs. Positional effects were relatively modest, with identical substituents at different positions (e.g., 17 vs. 25) showing minimal property variations (ΔcLogP <0.3). Bulky substitutions, including OCF_3 and phenyl groups, consistently impaired drug-likeness by simultaneously increasing molecular weight and lipophilicity. Pharmacokinetic predictions indicated all derivatives

would likely be excluded from the CNS (BBB-impermeable) while showing minimal CYP2D6 inhibition potential, suggesting favorable safety profiles for peripheral targets. Carboxymethyl ester derivatives (Deriv-1, -19) presented an interesting case, increasing hydrogen bond acceptor count without substantial lipophilicity penalties (ΔcLogP <0.2 versus parent). Cyano-substituted analogs (Deriv-4, -6) demonstrated particular promise, combining moderate polarity (TPSA ≈112 Å^2) with potential metabolic stability advantages.

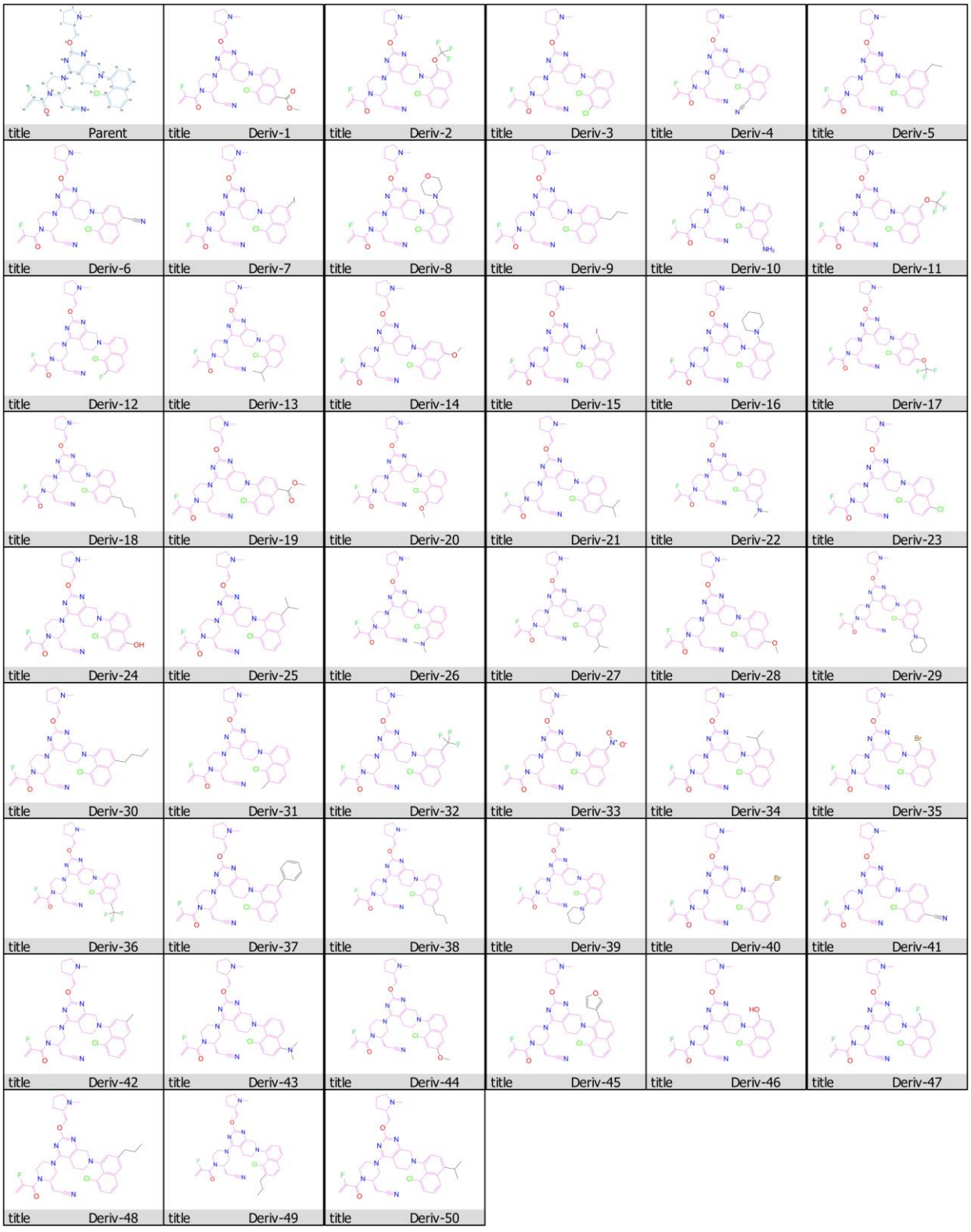

Figure 1. Comprehensive structural library of designed derivatives. Parent scaffold (top left) and 50 derivatives showing systematic R-group variations at positions 17-25.

Table 1. Comprehensive characterization of representative derivatives

Ligand	R-Group	MW (Da)	LogP	HBD	HBA	TPSA (Å²)	RotB	QED	SAScore	Violations	Notable Features
Parent	-	604.13	4.73	0	8	88.83	7	0.358	0.60	1	Baseline scaffold
Deriv-1	COOMe	662.17	4.52	0	10	115.13	8	0.252	0.70	1	High polarity
Deriv-4	CN	629.14	4.60	0	9	112.62	7	0.346	0.60	1	Balanced profile
Deriv-7	I	730.03	5.34	0	8	88.83	7	0.238	0.60	2	Heavy halogen
Deriv-10	NH_2	619.15	4.32	1	9	114.85	7	0.305	0.60	0	Optimal Ro5 compliance
Deriv-18	n-Bu	660.24	6.08	0	8	88.83	10	0.242	0.75	2	High flexibility
Deriv-24	OH	620.13	4.44	1	9	109.06	7	0.384	0.60	0	Best QED
Deriv-28	OMe	634.16	4.74	0	9	98.06	8	0.326	0.65	1	Methoxy variant
Deriv-37	Ph	680.23	6.40	0	8	88.83	8	0.198	0.55	2	Aromatic bulk
Deriv-46	OH	620.13	4.44	1	9	109.06	7	0.384	0.60	0	Positional isomer

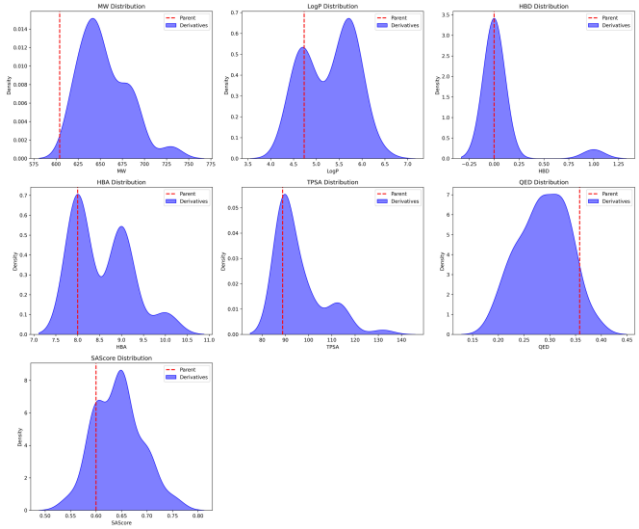

Figure 2. Distribution of physicochemical properties (MW, LogP, HBD, HBA, TPSA, QED, and SAScore) for a parent compound (red dashed line) and its derivatives (blue filled distribution).

3.5. Molecular Docking Analysis

Molecular docking studies against the KRAS G12D mutant (PDB ID: 7RPZ) demonstrated significant variations in binding affinity across the derivative series, with ΔG values ranging from -6.9 to -9.6 kcal/mol (Table 2). The parent compound exhibited moderate binding (-7.7 kcal/mol) through conventional hydrogen bonds, carbon hydrogen bonds, pi-anion, pi-alkyl, alkyl and halogen (fluorine) interactions (Figure 4 and Table 3). Notably, several derivatives showed substantially enhanced binding, with Deriv-34 (17-isopropyl substitution) emerging as the most potent inhibitor (-9.6 kcal/mol), forming critical interactions. Similarly, Deriv-4 (23-cyano) and Deriv-5 (8-ethyl) demonstrated strong binding affinities (-9.5 and -9.3 kcal/mol, respectively). Structure-activity relationship analysis revealed that isopropyl substitution at position 17 provided optimal binding, outperforming bulkier groups, while position 23 favored compact polar groups over halogens, and position 18 showed preference for small alkyl chains. The docking reliability was confirmed by consistent reproduction of key interactions across multiple runs and progressive improvement in binding energy with strategic modifications. These results demonstrate that targeted structural optimization, particularly at positions 17, 18, and 23, can significantly enhance KRAS G12D inhibition through optimized hydrophobic contacts and additional specific interactions, with Deriv-34 representing the most promising candidate for further development (Figure 4, Table 2 and Table 3).

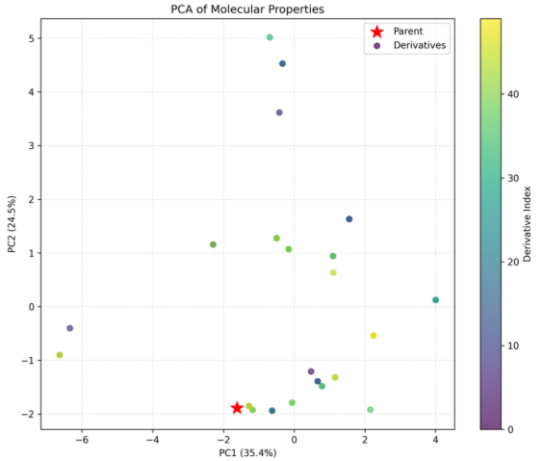

Figure 3. Principal Component Analysis of molecular properties. Score plot showing PC1 (35.4% variance, dominated by MW and cLogP) versus PC2 (24.5% variance, driven by HBA and TPSA).

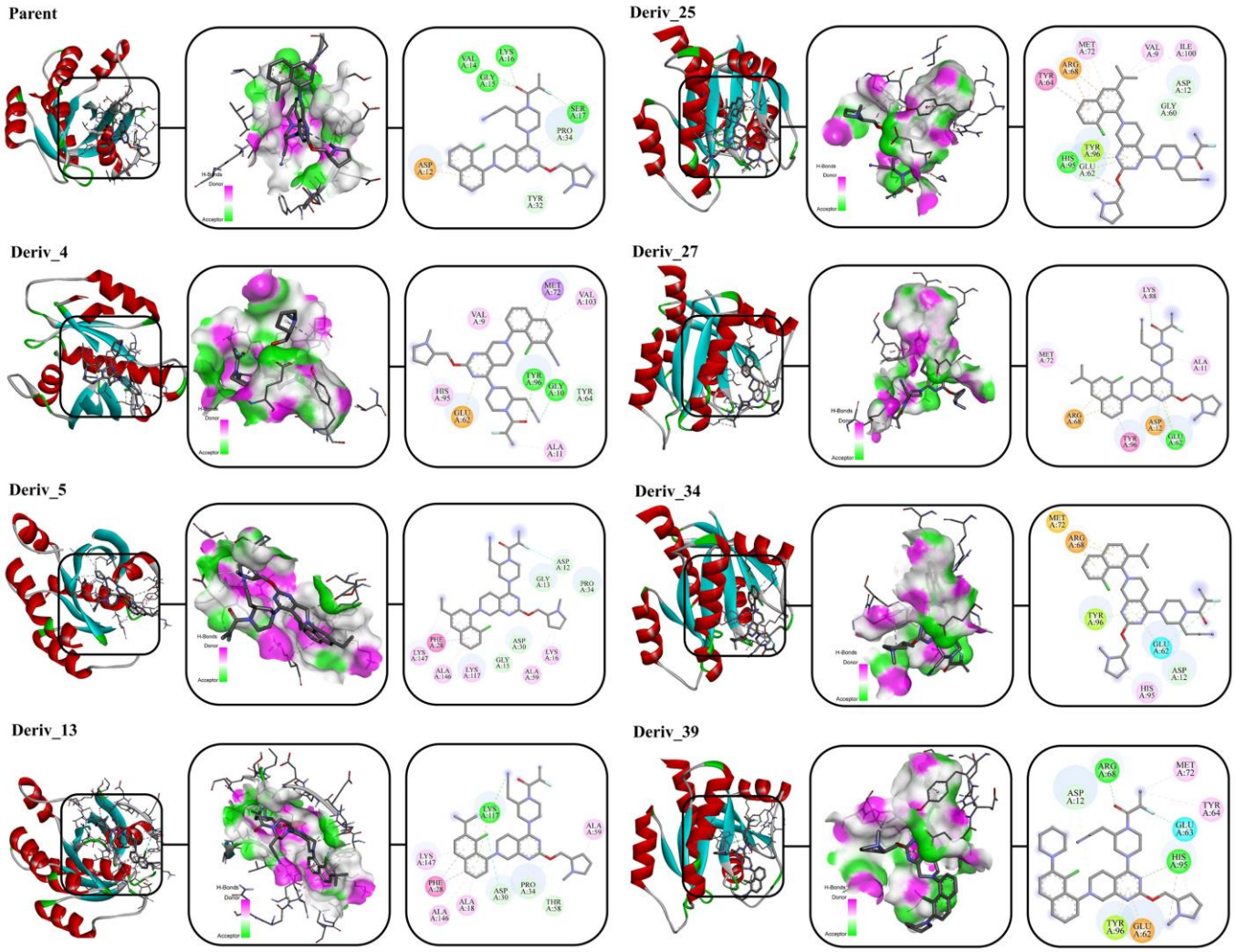

Figure 4. Molecular docking results of parent compound and derivatives against KRAS G12D mutant (PDB ID: 7RPZ).

Table 2. Calculated binding affinities (ΔG, kcal/mol) of parent compound and derivatives against KRAS G12D mutant (PDB ID: 7RPZ).

Ligand	ΔG (kcal/mol)	Ligand	ΔG (kcal/mol)	Ligand	ΔG (kcal/mol)	Ligand	ΔG (kcal/mol)
Parent	-7.7	*Deriv-13*	**-9.2**	*Deriv-26*	-7.6	***Deriv-39***	**-9.1**
Deriv-1	-8.7	*Deriv-14*	-8.3	***Deriv-27***	**-9.0**	*Deriv-40*	-6.9
Deriv-2	-7.6	*Deriv-15*	-8.0	*Deriv-28*	-7.6	*Deriv-41*	-8.1
Deriv-3	-8.2	*Deriv-16*	-7.6	*Deriv-29*	-8.7	*Deriv-42*	-8.2
Deriv-4	**-9.5**	*Deriv-17*	-8.0	*Deriv-30*	-6.9	*Deriv-43*	-8.0
Deriv-5	**-9.3**	*Deriv-18*	-8.1	*Deriv-31*	-7.8	*Deriv-44*	-8.3
Deriv-6	-7.2	*Deriv-19*	-8.0	*Deriv-32*	-7.7	*Deriv-45*	-7.7
Deriv-7	-8.6	*Deriv-20*	-8.2	*Deriv-33*	-8.1	*Deriv-46*	-7.8
Deriv-8	-8.6	*Deriv-21*	-7.9	***Deriv-34***	**-9.6**	*Deriv-47*	-8.4
Deriv-9	-7.9	*Deriv-22*	-7.1	*Deriv-35*	-8.5	*Deriv-48*	-8.2
Deriv-10	-8.4	*Deriv-23*	-8.4	*Deriv-36*	-8.8	*Deriv-49*	-8.7
Deriv-11	-8.8	*Deriv-24*	-7.8	*Deriv-37*	-7.5	*Deriv-50*	-8.3
Deriv-12	-7.5	***Deriv-25***	**-9.2**	*Deriv-38*	-8.2		

Table 3. Key molecular interactions of high-affinity derivatives with KRAS G12D active site residues.

Ligands	ΔG (kcal/mol)	R-Group	Interactions
Parent	**-7.7**	**N/A**	Conventional Hydrogen Bond (VAL A:14, GLY A:15, LYS A:16 and SER A:17); Carbon Hydrogen Bond (PRO A:34 and TYR A:32); Pi-Anion (ASP A:12); Alkyl (PRO A:34); Pi-Alkyl (PRO A:34); Halogen(Fluorine) (SER A:17)
Deriv-4	**-9.5**	**23-CN**	Conventional Hydrogen Bond (GLY A:10 and TYR A:96); Carbon Hydrogen Bond (TYR A:64); Pi-Anion (GLU A:62); Alkyl (ALA A:11); Pi-Alkyl (VAL A:9, MET A:72, HIS A:95, TYR A:96 and VAL A:103); Pi-Sigma (MET A:72)
Deriv-5	**-9.3**	**18-Et**	Carbon Hydrogen Bond (ASP A:12, GLY A:13, GLY A:15,ASP A:30 and PRO A:34); Pi-Pi T-shaped (PHE A:28); Alkyl (LYS A:16, ALA A:59 and LYS A:147); Pi-Alkyl (LYS A:117, ALA A:146 and LYS A:147); Halogen(Fluorine) (ASP A:12); Halogen(Chlorine) (GLY A:15)
Deriv-13	**-9.2**	**23-i-Pr**	Conventional Hydrogen Bond (LYS A:117); Carbon Hydrogen Bond (ASP A:30, PRO A:34 and THR A:58); Pi-Cation (LYS A:117); Pi-Pi T-shaped (PHE A:28); Alkyl (PRO A:34 and ALA A:59); Pi-Alkyl (ALA A:18, LYS A:117, ALA A:146 and LYS A:147); Pi-Sigma (LYS A:117); Halogen(Chlorine) (ASP A:30)
Deriv-25	**-9.2**	**18-i-Pr**	Conventional Hydrogen Bond (HIS A:95); Carbon Hydrogen Bond (ASP A:12, GLY A:60 and GLU A:62); Pi-Cation (ARG A:68); Pi-Anion (GLU A:62); Pi-Pi Stacked (TYR A:64); Alkyl (VAL A:9 and ILE A:100); Pi-Sigma (HIS A:95); Pi-Alkyl (ARG A:68, MET A:72 and TYR A:96); Pi-Lone Pair (TYR A:96)
Deriv-27	**-9.0**	**24-i-Pr**	Conventional Hydrogen Bond (GLU A:62 and LYS A:88); Pi-Donor Hydrogen Bond (ARG A:68); Pi-Cation (ARG A:68); Pi-Anion (ASP A:12 and GLU A:62); Pi-Pi T-shaped (TYR A:96); Alkyl (ALA A:11, MET A:72 and LYS A:88)
Deriv-34	**-9.6**	**17-i-Pr**	Carbon Hydrogen Bond (ASP A:12); Pi-Donor Hydrogen Bond (ARG A:68); Pi-Cation (ARG A:68); Pi-Anion (GLU A:62);Pi-Alkyl (ARG A:68, HIS A:95 and TYR A:96); Halogen(Fluorine) (GLU A:62); Pi-Sulfur (MET A:72); Pi-Lone Pair (TYR A:96)
Deriv-39	**-9.1**	**23-Piper**	Conventional Hydrogen Bond (ARG A:68 and HIS A:95);Carbon Hydrogen Bond (ASP A:12); Pi-Anion (GLU A:62); Alkyl (MET A:72); Pi-Sigma (HIS A:95); Pi-Alkyl (TYR A:64, HIS A:95 and TYR A:96); Halogen(Fluorine) (GLU A:63); Pi-Lone Pair (TYR A:96)

4. Conclusion

Herein we present a computational framework for the structure-based optimization of adagrasib analogues targeting the KRAS G12D mutant. Through this integrated molecular design approach, several promising lead compounds with significantly enhanced predicted binding affinities were identified relative to the parent scaffold. Comprehensive analysis of fifty derivatives elucidated structure-activity relationships that may inform subsequent medicinal chemistry efforts. Notably, isopropyl substitutions at position 17 (exemplified by Deriv-34) exhibited superior binding profiles via optimal engagement with key residues in the Switch-II pocket. These specific interactions contributed to a binding affinity enhancement of approximately 2 kcal/mol, potentially corresponding to a 25-fold increase in inhibitory potency. Physicochemical characterization revealed that hydroxylated derivatives (Deriv-24 and Deriv-46) demonstrated favorable Lipinski compliance concomitant with optimal Quantitative Estimate of Drug-likeness scores. Principal component analysis of molecular descriptors provided valuable insights into property space exploration, establishing a foundation for property-guided optimization strategies. Several methodological limitations warrant acknowledgment, including the necessity for experimental validation of computational predictions and the absence of protein flexibility considerations within the framework. The most promising candidates identified herein merit synthesis and experimental evaluation through biochemical binding assays followed by cellular viability assessments in relevant cancer models.

This investigation contributes to the expanding repertoire of computational methodologies for targeting historically intractable oncoproteins and represents a significant advancement toward the development of efficacious therapeutics for KRAS-driven malignancies.

5. References

[1] Bento, A.P., Hersey, A., Félix, E., Landrum, G., Gaulton, A., Atkinson, F., Bellis, L.J., De Veij, M., Leach, A.R., 2020. An open source chemical structure curation pipeline using RDKit. Journal of Cheminformatics 12, 1-16.

[2] Bickerton, G.R., Paolini, G.V., Besnard, J., Muresan, S., Hopkins, A.L., 2012. Quantifying the chemical beauty of drugs. Nature chemistry 4, 90-98.

[3] Buscail, L., Bournet, B., Cordelier, P., 2020. Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. Nature reviews Gastroenterology & hepatology 17, 153-168.

[4] Canon, J., Rex, K., Saiki, A.Y., Mohr, C., Cooke, K., Bagal, D., Gaida, K., Holt, T., Knutson, C.G., Koppada, N., 2019. The clinical KRAS (G12C) inhibitor AMG 510 drives anti-tumour immunity. Nature 575, 217-223.

[5] Cheng, H., Li, P., Chen, P., Irimia, A., Bae, J.H., Brooun, A., Fagan, P., Lam, R., Lin, B., Zhang, J., 2023. Structure-based design and synthesis of potent and selective KRAS G12D inhibitors. ACS Medicinal Chemistry Letters 14, 1351-1357.

[6] Christensen, J.G., Olson, P., Briere, T., Wiel, C., Bergo, M.O., 2020. Targeting Krasg12c‐mutant cancer with a mutation‐specific inhibitor. Journal of internal medicine 288, 183-191.

[7] Cox, A.D., Fesik, S.W., Kimmelman, A.C., Luo, J., Der, C.J., 2014. Drugging the undruggable RAS: Mission possible? Nature reviews Drug discovery 13, 828-851.

[8] Eberhardt, J., Santos-Martins, D., Tillack, A.F., Forli, S., 2021. AutoDock Vina 1.2. 0: New docking methods, expanded force field, and python bindings. Journal of Chemical Information and Modeling 61, 3891-3898.

[9] Fell, J.B., Fischer, J.P., Baer, B.R., Blake, J.F., Bouhana, K., Briere, D.M., Brown, K.D., Burgess, L.E., Burns, A.C., Burkard, M.R., 2020. Identification of the clinical development candidate MRTX849, a covalent KRASG12C inhibitor for the treatment of cancer. Journal of medicinal chemistry 63, 6679-6693.

[10] Hallin, J., Engstrom, L.D., Hargis, L., Calinisan, A., Aranda, R., Briere, D.M., Sudhakar, N., Bowcut, V., Baer, B.R., Ballard, J.A., 2020. The KRASG12C inhibitor MRTX849 provides insight toward therapeutic susceptibility of KRAS-mutant cancers in mouse models and patients. Cancer discovery 10, 54-71.

[11] Hobbs, G.A., Der, C.J., Rossman, K.L., 2016. RAS isoforms and mutations in cancer at a glance. Journal of cell science 129, 1287-1292.

[12] Isert, C., Atz, K., Schneider, G., 2023. Structure-based drug design with geometric deep learning. Current Opinion in Structural Biology 79, 102548.

[13] Issahaku, A.R., Mukelabai, N., Agoni, C., Rudrapal, M., Aldosari, S.M., Almalki, S.G., Khan, J., 2022. Characterization of the binding of MRTX1133 as an avenue for the discovery of potential KRASG12D inhibitors for cancer therapy. Scientific Reports 12, 17796.

[14] Janes, M.R., Zhang, J., Li, L.-S., Hansen, R., Peters, U., Guo, X., Chen, Y., Babbar, A., Firdaus, S.J., Darjania, L., 2018. Targeting KRAS mutant cancers with a covalent G12C-specific inhibitor. Cell 172, 578-589. e517.

[15] Jones, R.P., Sutton, P.A., Evans, J.P., Clifford, R., McAvoy, A., Lewis, J., Rousseau, A., Mountford, R., McWhirter, D., Malik, H.Z., 2017. Specific mutations in KRAS codon 12 are associated with worse overall survival in patients with advanced and recurrent colorectal cancer. British journal of cancer 116, 923-929.

[16] Kessler, D., Gmachl, M., Mantoulidis, A., Martin, L.J., Zoephel, A., Mayer, M., Gollner, A., Covini, D., Fischer, S., Gerstberger, T., 2019. Drugging an undruggable pocket on KRAS. Proceedings of the National Academy of Sciences 116, 15823-15829.

[17] Lanman, B.A., Allen, J.R., Allen, J.G., Amegadzie, A.K., Ashton, K.S., Booker, S.K., Chen, J.J., Chen, N., Frohn, M.J., Goodman, G., 2019. Discovery of a Covalent Inhibitor of KRASG12C (AMG 510) for the Treatment of Solid Tumors. ACS Publications.

[18] Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews 23, 3-25.

[19] Lito, P., Solomon, M., Li, L.-S., Hansen, R., Rosen, N., 2016. Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. Science 351, 604-608.

[20] Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M.-D., Ritchie, D.W., 2010. HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic acids research 38, W445-W449.

[21] McFall, T., Diedrich, J.K., Mengistu, M., Littlechild, S.L., Paskvan, K.V., Sisk-Hackworth, L., Moresco, J.J., Shaw, A.S., Stites, E.C., 2019. A systems mechanism for KRAS mutant allele–specific responses to targeted therapy. Science signaling 12, eaaw8288.

[22] Mehmood, A., Kaushik, A.C., Wang, Q., Li, C.-D., Wei, D.-Q., 2021. Bringing structural implications and deep learning-based drug identification for KRAS mutants. Journal of Chemical Information and Modeling 61, 571-586.

[23] Molina-Arcas, M., Hancock, D.C., Sheridan, C., Kumar, M.S., Downward, J., 2013. Coordinate direct input of both KRAS and IGF1 receptor to activation of PI3 kinase in KRAS-mutant lung cancer. Cancer discovery 3, 548-563.

[24] Moore, A.R., Rosenberg, S.C., McCormick, F., Malek, S., 2020. RAS-targeted therapies: is the undruggable drugged? Nature reviews Drug discovery 19, 533-552.

[25] Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry 30, 2785-2791.

[26] Mukherjee, S., Mitra, I., Saha, R., Dodda, S.R., Linert, W., Moi, S.C., 2015. In vitro model reaction of sulfur containing bio-relevant ligands with Pt (ii)

complex: Kinetics, mechanism, bioactivity and computational studies. RSC Advances 5, 76987-76999.
[27] Ostrem, J.M., Shokat, K.M., 2016. Direct small-molecule inhibitors of KRAS: from structural insights to mechanism-based design. Nature reviews Drug discovery 15, 771-785.
[28] Pant, S., Verma, S., Pathak, R.K., Singh, D.B., 2022. Structure-based drug designing, Bioinformatics. Elsevier, pp. 219-231.
[29] Pantsar, T., 2020. The current understanding of KRAS protein structure and dynamics. Computational and structural biotechnology journal 18, 189-198.
[30] Poorebrahim, M., Abazari, M.F., Moradi, L., Shahbazi, B., Mahmoudi, R., Kalhor, H., Askari, H., Teimoori-Toolabi, L., 2022. Multi-targeting of K-Ras domains and mutations by peptide and small molecule inhibitors. PLoS Computational Biology 18, e1009962.
[31] Prior, I.A., Hood, F.E., Hartley, J.L., 2020. The frequency of Ras mutations in cancer. Cancer research 80, 2969-2974.
[32] Riniker, S., Landrum, G.A., 2015. Better informed distance geometry: using what we know to improve conformation generation. Journal of Chemical Information and Modeling 55, 2562-2574.
[33] Waters, A.M., Der, C.J., 2018. KRAS: the critical driver and therapeutic target for pancreatic cancer. Cold Spring Harbor perspectives in medicine 8, a031435.
[34] Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., Baker, D., 2020. Improved protein structure prediction using predicted interresidue orientations. Proceedings of the National Academy of Sciences 117, 1496-1503.
[35] Yuan, T.L., Amzallag, A., Bagni, R., Yi, M., Afghani, S., Burgan, W., Fer, N., Strathern, L.A., Powell, K., Smith, B., 2018. Differential effector engagement by oncogenic KRAS. Cell reports 22, 1889-1902.
[36] Zeng, M., Lu, J., Li, L., Feru, F., Quan, C., Gero, T.W., Ficarro, S.B., Xiong, Y., Ambrogio, C., Paranal, R.M., 2017. Potent and selective covalent quinazoline inhibitors of KRAS G12C. Cell chemical biology 24, 1005-1016. e1003.