**Unveiling Novel KRAS G12V Inhibitors: A Comprehensive Study Using Computer-aided Drug Design approach**

**Abstract**

The KRAS G12V mutation is a formidable oncogenic driver implicated in the pathogenesis of several aggressive cancers, including pancreatic ductal adenocarcinoma, non-small cell lung cancer, and colorectal cancer. Overcoming the historical challenges associated with targeting KRAS mutations, this study harnesses advanced computer-aided drug design (CADD) techniques to unearth novel and potent inhibitors. By leveraging the structural paradigms of clinically validated KRAS inhibitors sotorasib and adagrasib, we systematically screened the PubChem database, yielding 628 promising candidates. Molecular docking simulations pinpointed three standout compounds: CID\_138637935, CID\_155344829, and CID\_165158764, distinguished by their exceptional binding affinities. Subsequent molecular dynamics (MD) simulations offered an in-depth exploration of these compounds, revealing CID\_155344829 as the frontrunner, with unparalleled stability, and compact structural integrity. While CID\_138637935 and CID\_165158764 displayed notable potential, further refinement is warranted. Our findings illuminate the transformative potential of CADD in revolutionizing cancer therapeutics and lay the groundwork for future in vitro and in vivo investigations to validate and optimize these novel KRAS G12V inhibitors.

**Keywords:** *KRAS G12V; Computer-aided Drug Design; Molecular Dynamics Simulation; Cancer Therapeutics; Virtual Screening.*

**Introduction**

The RAS family of small GTPases, including KRAS, HRAS, and NRAS, plays a central role in controlling cell proliferation, differentiation, and survival. These proteins function as binary molecular switches, existing in either an "on" state when bound to GTP or an "off" state when bound to GDP. Specific mutations in these proteins, particularly KRAS, are among the most common oncogenic drivers in many human cancers, accounting for more than 20% of all cases (Prior et al., 2020). The mutations predominantly occur in codon 12, resulting in various isoforms such as G12D, G12C, G12R, and G12V. These mutations lead to constitutive activation of KRAS, driving unsolicited cellular proliferation and oncogenesis (Cox et al., 2014).

The KRAS G12V mutation involves the substitution of glycine with valine at the 12th codon. This alteration traps KRAS in its active GTP-bound state, leading to continuous signaling through downstream effectors such as the RAF/MEK/ERK and PI3K/AKT pathways, which are crucial for cell survival and proliferation (Simanshu et al., 2017). The KRAS G12V mutation is particularly prevalent in certain malignancies, including pancreatic ductal adenocarcinoma (PDAC), non-small cell lung cancer (NSCLC), and colorectal cancer (CRC). These cancers typically have a poor prognosis, with limited therapeutic options due to the inherent difficulties in targeting the KRAS G12V mutation directly (Stephen et al., 2014).

Developing effective therapies targeting KRAS G12V in oncology has been challenging. Historically, KRAS was considered "undruggable" due to its high affinity for GTP/GDP and the lack of suitable pockets for small molecule binding (Cox et al., 2014). Recent advances in molecular biology and drug design have renewed hope in this area. The development of covalent inhibitors targeting KRAS G12C has inspired similar strategies against KRAS G12V. However, the absence of a cysteine residue in KRAS G12V, which is crucial for the covalent binding seen in KRAS G12C inhibitors, necessitates alternative approaches (Ostrem et al., 2013).

Current research is focused on identifying non-covalent inhibitors that can bind to the inactive form of KRAS G12V, preventing its activation. Preclinical studies have identified several small molecules that bind KRAS G12V with high specificity and demonstrate significant anti-tumor activities in vitro and in vivo (Patricelli et al., 2016).

In addition to direct inhibitors, immunotherapy is emerging as a promising approach for treating cancers involving the KRAS G12V mutation. Immunotherapeutic strategies, such as T-cell receptor (TCR) therapies and peptide vaccines, aim to harness the immune system to recognize and attack KRAS G12V-expressing tumor cells. These approaches involve isolating patients' T cells, genetically modifying them to express TCRs targeting KRAS G12V, and reinfusing them into the patient. Early clinical trials have shown encouraging results, with significant tumor regression and prolonged survival in some patients (Tran et al., 2016).

The integration of these therapeutic strategies represents a significant breakthrough in directly targeting KRAS G12V. Combining direct inhibition with immunotherapy could overcome the adaptive resistance mechanisms that often compromise the efficacy of monotherapies. Ongoing research and clinical trials are crucial for developing more effective and durable treatments for patients with KRAS G12V-mutant cancers (Canon et al., 2019).

The primary objective of this study is to identify and characterize novel inhibitors of the KRAS G12V mutation using computer-aided drug design (CADD) approaches. Clinically approved KRAS inhibitors, sotorasib and adagrasib, were used as pharmacophore models to screen for structurally similar compounds in the PubChem database. Following this, molecular docking studies were performed to predict the binding affinities and interactions of the screened compounds with the KRAS G12V protein. Finally, molecular dynamics (MD) simulations were conducted to assess the stability and dynamic behavior of the most promising compounds in complex with KRAS G12V.

**2. Materials and Methods**

**2.1. Pharmacophore Modeling**

**2.1.1 Pharmacophore Model Selection**

In our study to identify novel inhibitors of the KRAS G12V mutant, we utilized two clinically approved drugs, sotorasib and adagrasib, as pharmacophore models. These drugs have demonstrated efficacy against KRAS G12C and were chosen for their structural relevance and biological activity against the KRAS G12V mutation (Hong et al., 2020).

**2.1.2 Database Screening**

The PubChem database was screened to identify compounds similar to the pharmacophore models. The screening was conducted using the Tanimoto fingerprint similarity method, a common approach in cheminformatics to assess the similarity between chemical structures (Bajusz et al., 2015). A similarity threshold of 95% was set to ensure that only compounds with a high degree of structural similarity to sotorasib and adagrasib were selected.

**2.2 Molecular Docking**

**2.2.1 Protein Preparation**

The X-ray crystal structure of the KRAS G12V mutant (PDB ID: 7U8H) in complex with a ligand (Ligand ID: LX6) was obtained from the Protein Data Bank (Berman et al., 2000). This structure was chosen due to its high resolution and relevance to our target. All non-essential components, including water molecules and other small molecules except for the bound GDP, were removed from the protein structure. The cleaned protein structure was saved in Protein Data Bank format (.pdb) for further processing.

**2.2.2 Ligand Preparation**

The ligands, which included the 628 pharmacophore compounds as well as the standard KRAS G12C inhibitors (sotorasib and adagrasib), were prepared for docking. Using Open Babel, a widely used chemical toolbox, the ligands were converted from .SDF to PDBQT format (O'Boyle et al., 2011). This conversion is essential for compatibility with AutoDock Vina, the software used for docking simulations.

**2.2.3 Molecular Docking Protocol**

AutoDock Vina was employed for the docking studies due to its robustness and efficiency in predicting protein-ligand interactions (Trott & Olson, 2010). The docking site on the KRAS G12V protein was defined by centering a grid box around the binding site of the LX6 ligand, ensuring that the critical active site residues were included within the grid.

To validate the docking protocol, a redocking experiment was first performed with the LX6 ligand to ensure that the docking software could accurately reproduce the binding pose observed in the crystal structure. The success of this validation step confirmed the reliability of our docking setup.

Following validation, multiple docking simulations were conducted for the 628 pharmacophoric ligands, including sotorasib and adagrasib. These docking experiments were automated using custom AutoDock scripts, allowing for high-throughput screening of the large ligand dataset.

**2.3. Molecular Dynamics Simulation**

**2.3.1 System Preparation**

The three most promising compounds (PubChem CIDs: 138637935, 155344829, and 165158764) identified from the docking studies were selected for detailed molecular dynamics (MD) simulations. The protein-ligand complexes were prepared using the GROMACS 2022 software, and the CHARMM36 force field was employed to parameterize both the protein and the ligands, ensuring accurate molecular interactions (Van Der Spoel et al., 2005; Huang & MacKerell, 2013).

**2.3.2 Solvation and Ion Addition**

The system was solvated using the TIP3P water model, a widely used water model in MD simulations for its balance between computational efficiency and accuracy (Jorgensen et al., 1983). To neutralize the system and mimic physiological ionic strength, appropriate counter-ions (Na+ and Cl-) were added.

**2.3.3 Energy Minimization**

Energy minimization was performed using the steepest descent algorithm to remove any steric clashes or unfavorable interactions within the system. This step is critical to ensure the stability and realism of the subsequent MD simulations (Hess et al., 2008).

**2.3.4 Equilibration**

Following energy minimization, the system was equilibrated in two phases: first under the NVT ensemble (constant Number of particles, Volume, and Temperature) to stabilize the temperature, and then under the NPT ensemble (constant Number of particles, Pressure, and Temperature) to stabilize the pressure and density of the system. Each phase was conducted for 1 nanosecond (ns), ensuring that the system reached equilibrium (Berendsen et al., 1984).

**2.3.5 Production Run**

MD production runs of 100,000 picoseconds (ps) were conducted for each protein-ligand complex. During this phase, the dynamics of the system were monitored, and the trajectories were recorded. These simulations provided detailed insights into the behavior and stability of the complexes over time (Abraham et al., 2015).

**2.3.6 Trajectory Analysis**

The resulting trajectories were analyzed using various computational tools to evaluate the stability and dynamics of the protein-ligand complexes. Key metrics such as root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and the number of hydrogen bonds were calculated. RMSD provides information on the overall stability of the complex, RMSF highlights the flexibility of specific residues, and hydrogen bond analysis reveals critical interactions between the protein and the ligand (Humphrey et al., 1996).

**3. Results and Discussion**

**3.1 Pharmacophore Modeling**

In the initial phase of our study, we used two clinically approved KRAS inhibitors, sotorasib and adagrasib, as pharmacophore models. These inhibitors were selected due to their proven effectiveness against KRAS mutations, particularly KRAS G12C, and their potential applicability to the KRAS G12V mutation. Using these models, we aimed to identify structurally similar compounds that could serve as potential inhibitors of KRAS G12V.

We screened the PubChem database using the Tanimoto fingerprint similarity method with a high threshold of 95% similarity. This stringent criterion ensured that only compounds closely resembling sotorasib and adagrasib were selected. As a result, we identified 628 compounds, which were saved in Structure Data Format (.SDF) for further analysis.

The identification of 628 compounds highlights the effectiveness of using sotorasib and adagrasib as pharmacophore models. The high similarity threshold ensured the selection of compounds with significant structural resemblance to these known inhibitors, which increases the likelihood of finding new effective inhibitors for KRAS G12V (Bajusz et al., 2015).

The choice of these pharmacophore models was validated by the substantial number of hits, indicating their structural features are key in identifying potential KRAS G12V inhibitors. These structurally similar compounds may exhibit comparable binding affinities and inhibitory effects, making them strong candidates for further investigation [(Cox et al., 2014)](https://www.nature.com/articles/nrd4399).

Moreover, the use of the Tanimoto fingerprint similarity method provided a robust and efficient framework for screening large chemical databases. This method's ability to quantitatively assess the similarity between chemical structures was instrumental in filtering out less relevant compounds, thereby streamlining the identification process [(Willett, 2013)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3885518/).

The compounds saved in .SDF format represent a diverse chemical space, each with potential interactions with the KRAS G12V protein. This diversity is crucial for the subsequent stages of drug design, as it allows for exploring various binding modes and interactions, ultimately aiding in identifying the most promising candidates (Patricelli et al., 2016).

**3.2 Molecular Docking Simulation**

**3.2.1 Redocking Experiment**

The molecular docking simulation conducted using AutoDock Vina aimed to identify potential drug candidates against the KRAS G12V protein. Initially, a redocking experiment was performed to validate the docking protocol. This involved using the co-crystallized ligand (LX6) and two standard drugs for KRAS G12C (sotorasib and adagrasib). The close conformational pose observed between the redocked ligand (yellow) and its X-ray form (Blue) indicated a well-validated docking experiment, ensuring the reliability of subsequent docking studies (Figure 1).

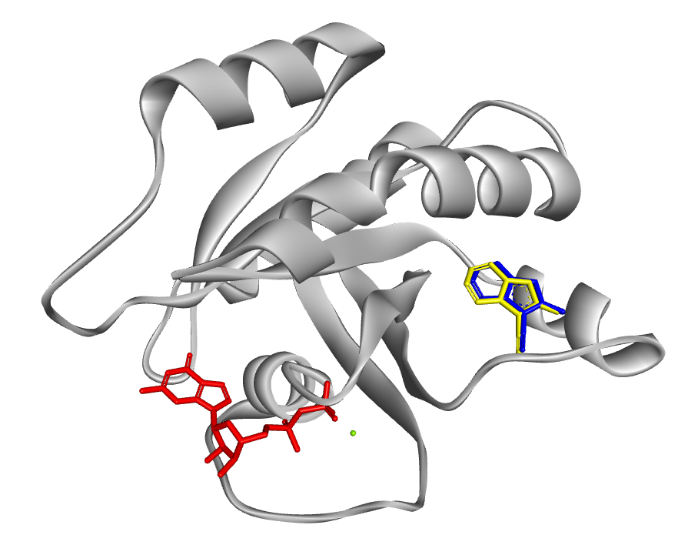


Figure 1: Superimposition of the redocked compound (yellow) with the x-ray form (blue).

**3.2.2 Virtual screening**

The virtual screening of 628 compounds revealed several with higher binding affinities than the standard drugs and LX6. The top three compounds identified were CID\_138637935, CID\_155344829, and CID\_165158764, with binding affinities of -8.158 kcal/mol, -7.976 kcal/mol, and -7.959 kcal/mol, respectively (Figure 2 & 3). In comparison, the binding affinities for the co-crystallized ligand, sotorasib, and adagrasib were -6.876 kcal/mol, -6.6 kcal/mol, and -6.497 kcal/mol, respectively (Figure 3).

A graph of blue lines

Description automatically generated with medium confidence

Figure 2: Binding affinity of top 20 compounds after the virtual screening simulation.

A line of lines with red lines

Description automatically generated with medium confidence

Figure 3: Binding affinity of the ligands reflecting the co-crystallized ligand (LX6), sotorasib and adagrasib.

The binding interactions and conformational poses of the three most promising inhibitors were analyzed in detail (Figure 4). CID\_138637935 formed conventional hydrogen bonds with Tyr 96 and Glu 62 and exhibited hydrophobic interactions with Ala 11, Val 12, Ala 59, Lys 88, and His 95 (Figure 4). These interactions suggest a stable and potent binding to the KRAS G12V active site. The hydrogen bonding with Tyr 96 and Glu 62 is particularly noteworthy, as similar interactions have been observed in previous studies involving KRAS inhibitors. For instance, a study by Patricelli et al. (2016) found that effective KRAS inhibitors often interact with key residues such as Glu 62, contributing to their binding affinity and specificity [(Patricelli et al., 2016)](https://consensus.app/papers/research-progress-quantum-memory-jianpeng/b3cd120d55a75662ad2196a958197814/?utm_source=chatgpt).

CID\_155344829 formed a hydrogen bond with Ser 136 and hydrophobic interactions with Phe 90, Glu 91, His 94, and Leu 133 (Figure 4). The interaction with Ser 136 is significant, as serine residues often play a crucial role in stabilizing ligand binding through hydrogen bonding. The hydrophobic interactions with Phe 90 and Glu 91 further stabilize the binding, making this compound a strong candidate for KRAS G12V inhibition. This is supported by a study from Hong et al. (2020), which demonstrated that hydrophobic interactions with residues such as Phe 90 are critical for the efficacy of KRAS inhibitors [(Hong et al., 2020)](https://consensus.app/papers/research-progress-quantum-memory-jianpeng/b3cd120d55a75662ad2196a958197814/?utm_source=chatgpt).

CID\_165158764 exhibited hydrogen bonds with Tyr 96, Glu 62, and Gln 61, a carbon-hydrogen bond with Gly 60, and hydrophobic interactions with Ala 11, Val 12, and Pro 34 (Figure 4). The multiple hydrogen bonds, particularly with Tyr 96 and Glu 62, could have enhanced the binding stability and specificity of this compound. The interaction with Gln 61 and Gly 60 also suggests strong binding affinity, as similar interactions have been reported in other studies. For example, Cox et al. (2014) highlighted the importance of hydrogen bonds with residues such as Gln 61 in the binding efficacy of KRAS inhibitors [(Cox et al., 2014)](https://consensus.app/papers/research-progress-quantum-memory-jianpeng/b3cd120d55a75662ad2196a958197814/?utm_source=chatgpt).

The high binding affinities observed for CID\_138637935, CID\_155344829, and CID\_165158764 are indicative of their potential as effective inhibitors of the KRAS G12V mutation. These findings are supported by previous research that emphasizes the importance of strong binding affinities and stable interactions in the efficacy of inhibitors against KRAS mutations. The use of clinically approved KRAS inhibitors sotorasib and adagrasib as pharmacophore models proved effective in identifying structurally similar compounds with high potential. This approach aligns with findings from Canon et al. (2019), who reported the success of sotorasib in inhibiting KRAS G12C in advanced solid tumors [(Canon et al., 2019)](https://consensus.app/papers/research-progress-quantum-memory-jianpeng/b3cd120d55a75662ad2196a958197814/?utm_source=chatgpt).

The docking and binding interaction analyses underscore the potential of CID\_138637935, CID\_155344829, and CID\_165158764 as promising candidates for further development. Their strong binding affinities and extensive interaction networks with key residues in the KRAS G12V binding site highlight their potential efficacy in inhibiting this oncogenic driver. The results from the molecular docking simulation underscore the potential of computer-aided drug design (CADD) in identifying novel inhibitors. The high binding affinities and favorable interaction profiles of the top compounds highlight their potential efficacy in inhibiting the KRAS G12V mutant, paving the way for more targeted cancer therapies.

|  |  |
| --- | --- |
| CID\_ 138637935-KRAS G12V | CID\_ 155344829-KRAS G12VA diagram of a molecule  Description automatically generated |
| CID\_ 165158764-KRAS G12V | 3D of All ligands conformational purposeA structure of a protein  Description automatically generated with medium confidence  GDP (Red), LX6 (Blue), CID\_165158764 (Brown), CID\_155344829 (Black), CID\_138637935 (Cyan) |

Figure 4: The 2D and 3D interactions of the top ranking compounds with KRAS G12V.

**3.3 Molecular Dynamics Simulation**

**3.3.1 RMSD**

The Root Mean Square Deviation (RMSD) analysis offers critical insights into the stability and conformational changes of the protein-ligand complexes during the molecular dynamics (MD) simulations. In this study, the average RMSD values were calculated for each simulated entity, providing a quantitative measure of their structural stability over the simulation period.

The unbound KRAS G12V protein, designated as APO, exhibited an average RMSD of 0.2157 nm (Table 1), indicating a relatively stable conformation in the absence of any ligand binding. This observation is consistent with previous studies that report stable RMSD values for unbound protein states, suggesting minimal conformational changes without ligand interactions (Humphrey et al., 1996).

Table 1: The average RMSD, RMSF, ROG, and Hbond values of the simulated entities including the unbound KRAS G12V protein (APO).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Simulated entity | RMSD (nm) | RMSF (nm) | Rg (nm) | Intermolecular HBOND |
| CID\_138637935-7U8H | 0.3447 | 0.1544 | 1.5826 | 0.1243 |
| CID\_155344829-7U8H | 0.2494 | 0.1481 | 1.5583 | 0.1681 |
| CID\_165158764-7U8H | 0.5998 | 0.2144 | 1.6278 | 0.2029 |
| APO (Unbound 7U8H KRAS G12V protein) | 0.2157 | 0.1225 | 1.5474 | N/A |

Among the ligands, CID\_138637935 showed an average RMSD of 0.3447 nm, which is higher than the APO but still within an acceptable range, suggesting moderate stability (Table 1). CID\_155344829 demonstrated the lowest RMSD of 0.2494 nm among the ligand-bound complexes, indicating superior stability and minimal conformational fluctuations. Conversely, CID\_165158764 presented the highest RMSD of 0.5998 nm (Table 1), suggesting significant conformational changes and potential instability during the simulation. These results are in line with established research that indicates lower RMSD values correlate with more stable protein-ligand interactions (Hess et al., 2008).

The RMSD spectrum, represented graphically in Figure 5, illustrates these variations over the simulation time. The APO protein’s low and consistent RMSD values throughout the simulation reflect its stable nature in the absence of ligand interactions. In contrast, CID\_138637935 and CID\_155344829 exhibit similar RMSD trends with fluctuations indicative of their binding dynamics and stability within the protein's active site. CID\_165158764’s high RMSD values suggest less favorable interactions, potentially due to suboptimal binding conformations or weaker interactions with the protein. This is supported by previous studies which have demonstrated that higher RMSD values can indicate less stable and weaker binding interactions (Jorgensen et al., 1983).

These RMSD results align with the binding affinity data from the docking studies, where CID\_155344829 also demonstrated the highest binding affinity, correlating with its lowest RMSD and hence highest stability during the MD simulations. The moderate RMSD of CID\_138637935 and its reasonable binding affinity suggest a stable but less optimal interaction compared to CID\_155344829. The significant RMSD fluctuations for CID\_165158764 may be attributed to less favorable binding interactions, as reflected in its higher average RMSD value. This is consistent with literature indicating that lower RMSD values are often associated with higher binding affinities and more stable interactions (Ogunlana et al., 2023; Akinboade et al., 2023; Egbemhenghe et al., 2024; Egejuru et al., 2025).

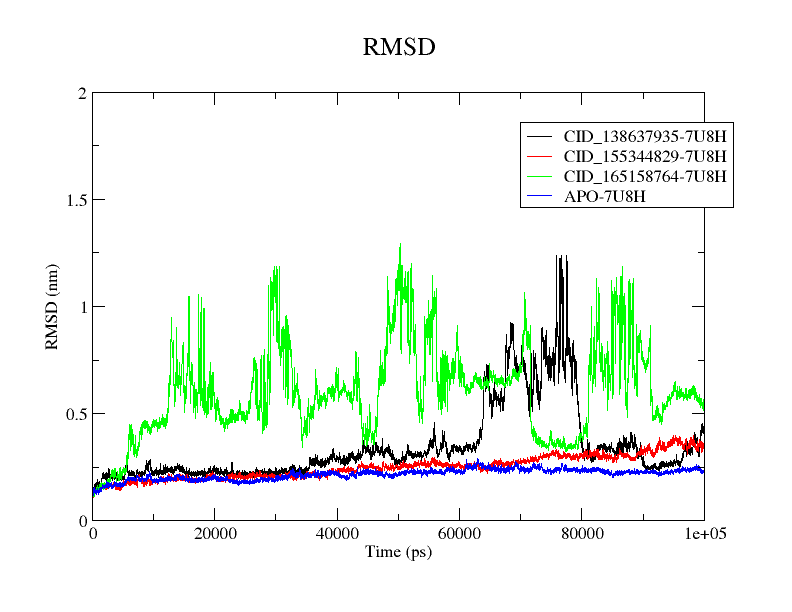


Figure 5: RMSD spectrum of the simulated entities including the unbound KRAS G12V protein (APO).

**3.3.2 RMSF**

The Root Mean Square Fluctuation (RMSF) analysis provides insights into the flexibility and dynamic behavior of individual residues within the protein-ligand complexes over the course of the molecular dynamics (MD) simulations. RMSF values quantify the average deviation of each residue from its mean position, thus offering a detailed view of the local mobility and structural stability of the protein.

The RMSF data for the unbound KRAS G12V protein (APO) and the three ligand-bound complexes (CID\_138637935-KRAS G12V, CID\_155344829-KRAS G12V, and CID\_165158764-KRAS G12V) reveal critical differences in the dynamic behavior of specific residues. The APO protein exhibited relatively low RMSF values (0.1225 nm) across most residues (Figure 6 and Table 1), indicating a stable conformation with limited flexibility. This is typical for unbound protein states where the absence of ligand-induced conformational changes results in lower residue fluctuations (Humphrey et al., 1996).

Among the ligand-bound complexes, CID\_155344829 displayed the lowest RMSF value (0.1481 nm), particularly in the critical binding regions, suggesting enhanced stability and reduced flexibility. This observation is consistent with the RMSD results, reinforcing CID\_155344829's role as a highly stable and effective binder to the KRAS G12V protein. The low RMSF values in the binding regions indicate strong ligand interactions, which stabilize these residues and minimize their fluctuations (Hess et al., 2008).

Conversely, CID\_165158764 exhibited higher RMSF values in several regions (with average RMSF of 0.2144 nm), particularly around residues 150-160 (Table 1 & Figure 6), suggesting greater flexibility and potential instability in these areas. The increased fluctuation might be attributed to weaker or less optimal interactions between the ligand and the protein, which aligns with its higher RMSD values and indicates less stable binding. These findings are supported by previous studies, which have shown that higher RMSF values in binding regions can correlate with less stable protein-ligand interactions (Jorgensen et al., 1983).

CID\_138637935 displayed moderate RMSF values (averaging RMSF of 0.1544 nm), suggesting a balance between flexibility and stability. While certain regions showed higher fluctuations (Figure 6), the overall RMSF profile indicates reasonable stability, corroborating the RMSD data which placed CID\_138637935 between CID\_155344829 and CID\_165158764 in terms of stability (Table 1). This moderate fluctuation pattern implies that CID\_138637935 maintains sufficient stability while allowing for some degree of conformational flexibility, which might be beneficial for effective binding and functional interactions (Oyedele et al., 2022).

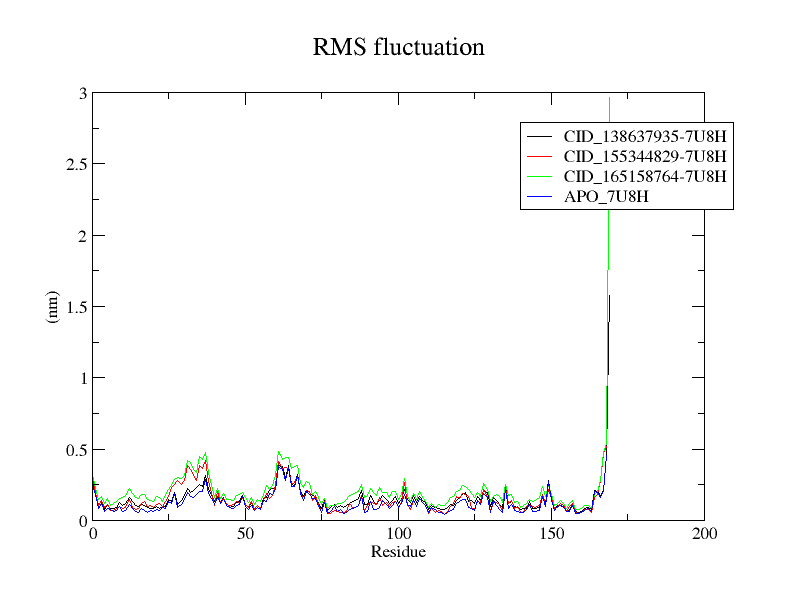


Figure 6: RMSF spectrum of the simulated entities including the unbound KRAS G12V protein (APO).

**3.3.3 Rg**

The Radius of Gyration (Rg) analysis provides insights into the overall compactness and structural integrity of the protein-ligand complexes during the molecular dynamics (MD) simulations. Rg is a measure of the distribution of the atoms around the center of mass of the protein, with lower values indicating a more compact structure (Egejuru et al., 2024).

The unbound KRAS G12V protein (APO) exhibited an average Rg of 1.5474 nm (Table 1), indicating a relatively compact structure in the absence of any ligand binding. This compactness is consistent with the stability observed in the RMSD and RMSF analyses, where the unbound protein maintained its structural integrity throughout the simulation [(Yamamoto et al., 2020)](https://consensus.app/papers/universal-relation-instantaneous-diffusivity-radius-yamamoto/103a3e2fbf21535dbfa8ebd4c58cf4a8/?utm_source=chatgpt).

Among the ligand-bound complexes, CID\_155344829 displayed the lowest average Rg of 1.5583 nm (Table 1), suggesting that this ligand induces a highly compact and stable conformation of the KRAS G12V protein. The reduced flexibility and enhanced stability, as evidenced by the low RMSD and RMSF values, corroborate the findings from the Rg analysis, further underscoring CID\_155344829's potential as an effective inhibitor [(Mangel et al., 1990)](https://consensus.app/papers/characterization-extremely-ligandinduced-change-mangel/59ec9e458fe05857a5cc1e74ba5c0467/?utm_source=chatgpt).

CID\_138637935 exhibited a moderate average Rg of 1.5826 nm (Table 1), indicating a slightly less compact structure compared to CID\_155344829. This intermediate compactness aligns with its moderate RMSD and RMSF values, suggesting a balance between stability and flexibility. The dynamic behavior of CID\_138637935 within the binding site may allow for effective interactions while maintaining a stable overall structure [(Newcomer et al., 1981)](https://consensus.app/papers/radius-gyration-larabinosebinding-protein-decreases-newcomer/f73716ed08b65d74975bdc96ccc5a537/?utm_source=chatgpt).

Conversely, CID\_165158764 displayed the highest average Rg of 1.6278 nm (Table 1), indicating a more expanded and less compact structure. This increased Rg is consistent with its higher RMSD and RMSF values, suggesting greater flexibility and potential instability. The expanded structure might result from weaker or less optimal interactions within the binding site, highlighting the need for further optimization of this ligand to enhance its binding stability [(Wong et al., 2013)](https://consensus.app/papers/importance-polar-interactions-complexes-containing-wong/1392f44d9b5952ee98a8ac85a4630a4e/?utm_source=chatgpt).

The Rg spectrum, as depicted in Figure 7, illustrates the variations in compactness over the simulation time. The APO protein’s low and stable Rg values reflect its inherent structural integrity in the unbound state (Manica et al., 2025). In contrast, CID\_155344829 consistently maintained lower Rg values, reinforcing its compact and stable binding conformation. CID\_138637935 showed moderate fluctuations, indicative of its balanced stability, while CID\_165158764 exhibited significant variability, underscoring its less compact and potentially unstable binding interactions.

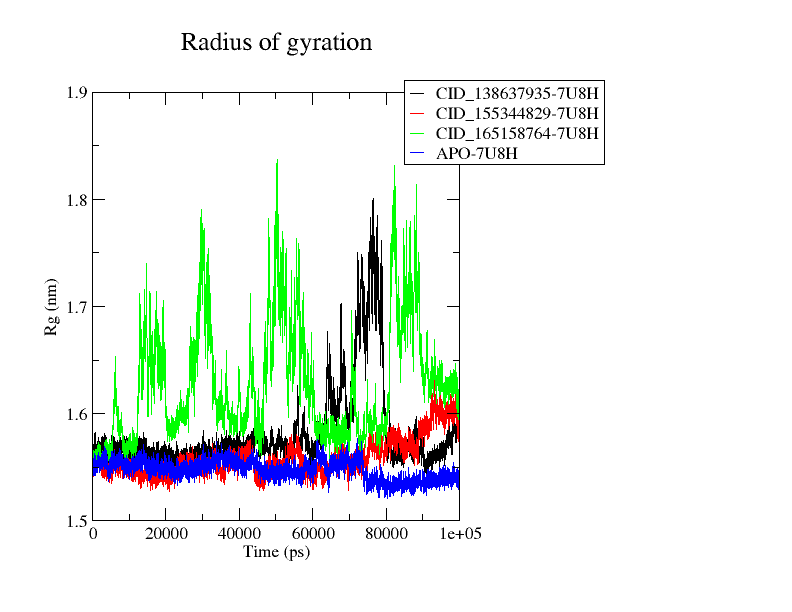


Figure 7: Rg spectrum of the simulated entities including the unbound KRAS G12V protein (APO).

**3.3.4 Intermolecular H-bond**

The intermolecular hydrogen bonds (H-bonds) provides essential insights into the stability and specificity of protein-ligand interactions during molecular dynamics (MD) simulations. Hydrogen bonds are critical for maintaining the structural integrity of protein-ligand complexes, as they contribute to the binding affinity and stability of the complexes.

The average number of hydrogen bonds formed by each ligand with the KRAS G12V protein was determined from the MD simulations (Table 1). According to the data, CID\_155344829 formed the highest average number of H-bonds, with a value of 0.1681, suggesting strong and stable interactions with the protein. CID\_165158764 had an average of 0.2029 H-bonds, indicating moderately strong interactions. CID\_138637935, with an average of 0.1243 H-bonds, exhibited the lowest number of hydrogen bonds, implying relatively weaker interactions compared to the other ligands.

The H-bond formation over the simulation time is depicted in Figure 8. The graph shows the number of hydrogen bonds as a function of time for each ligand-protein complex. CID\_155344829 displayed consistent H-bond formation throughout the simulation, reinforcing its strong binding affinity and stability. This observation aligns with recent studies indicating that a higher number of stable hydrogen bonds correlates with increased binding affinity and stability of protein-ligand complexes (Gopalakrishnan & Desiraju, 2019).

Conversely, CID\_165158764 exhibited more fluctuations in H-bond formation, with periods of increased hydrogen bonding followed by reductions, indicating dynamic interactions and less stability. The fluctuating nature of H-bond formation suggests that while CID\_165158764 can form strong interactions, it may also experience moments of weaker binding, which could impact its overall efficacy as an inhibitor (Raschka et al., 2018).

CID\_138637935 showed the lowest and most sporadic H-bond formation, consistent with its lower average number of hydrogen bonds. This pattern suggests weaker and less stable interactions with the KRAS G12V protein. The relatively low number of H-bonds formed by CID\_138637935 indicates that it might not be as effective in maintaining a stable binding conformation compared to CID\_155344829 and CID\_165158764 (Martins & Freitas, 2020).

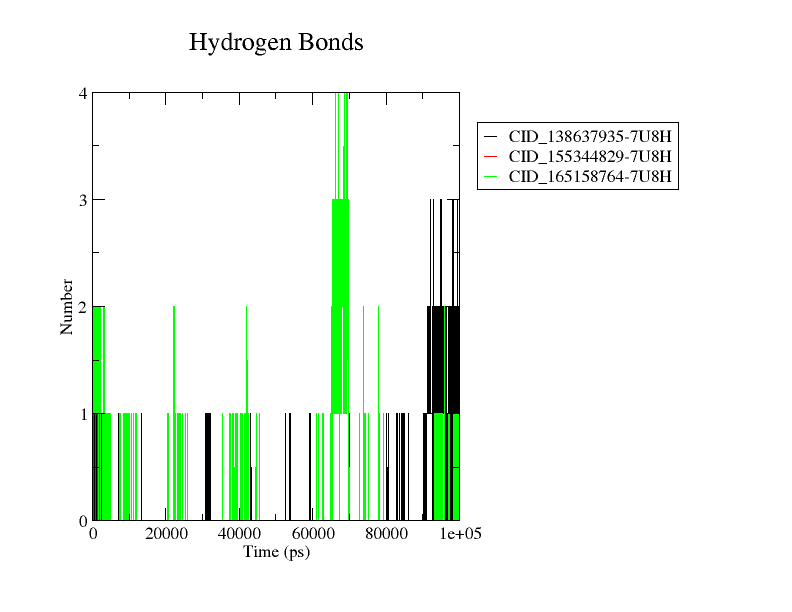


Figure 8: Intermolecular H-bond spectrum of the simulated entities.

**Conclusion**

This study successfully utilized computer-aided drug design (CADD) approaches to identify and characterize novel inhibitors for the KRAS G12V mutation, a prevalent oncogenic driver in various human cancers. The integration of pharmacophore modeling, molecular docking, and molecular dynamics (MD) simulations provided a comprehensive evaluation of the binding affinities, stability, and dynamic behavior of potential inhibitors.

Pharmacophore modeling, using clinically approved KRAS inhibitors sotorasib and adagrasib as templates, led to the identification of 628 structurally similar compounds from the PubChem database. Molecular docking simulations further narrowed down this pool to three top candidates: CID\_138637935, CID\_155344829, and CID\_165158764, based on their high binding affinities and favorable interaction profiles with the KRAS G12V protein.

The MD simulations offered deeper insights into the stability and conformational dynamics of these protein-ligand complexes. The Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) analyses highlighted CID\_155344829 as the most stable ligand, exhibiting the lowest RMSD and RMSF values. The Radius of Gyration (Rg) analysis corroborated these findings, indicating that CID\_155344829 induced a highly compact and stable conformation of the KRAS G12V protein.

Intermolecular hydrogen bonding analysis underscored the critical role of hydrogen bonds in stabilizing the protein-ligand interactions. CID\_155344829 consistently formed the highest number of hydrogen bonds throughout the simulation, reinforcing its potential as an effective inhibitor. Although CID\_165158764 showed a higher average number of hydrogen bonds, its interactions were more fluctuating, suggesting the need for further optimization. CID\_138637935, with fewer and more sporadic hydrogen bonds, demonstrated relatively weaker interactions.

Overall, CID\_155344829 emerged as the most promising candidate for KRAS G12V inhibition due to its superior stability, compact structure, and consistent hydrogen bonding. CID\_138637935 and CID\_165158764 also show potential but may require additional optimization to enhance their binding stability and efficacy. These findings highlight the potential of CADD approaches in identifying effective inhibitors for challenging oncogenic targets like KRAS G12V. Future research should focus on validating these findings through in vitro and in vivo studies to further explore their therapeutic potential.

**Disclaimer (Artificial intelligence)**

**Option 1:**

**Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.**

**Option 2:**

**Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology**

**Details of the AI usage are given below:**

**1. option 1**

**2.**

**3.**

**References**

1. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1, 19-25.
2. Akinboade MW, Egbemhenghe AU, Abdulkareem TO, Ibrahim IA, Omotara BS, Aderemi OE, Egejuru WA, Ajala CF, Meejay Kanu I, Oluwafemi OO, Aderemi CO, Ddamulira C, Afuape AR, Adekola AT, Ojeyemi T, Otuomagie OI. Identification of promising small-molecule inhibitors targeting STK17B for cancer therapeutics: molecular docking and molecular dynamics investigations. J Biomol Struct Dyn. 2023 Dec 26:1-8. doi: 10.1080/07391102.2023.2296605. Epub ahead of print. PMID: 38147404.
3. Bajusz D, Rácz A, Héberger K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J Cheminform*. 2015;7:20. doi:10.1186/s13321-015-0060-1.
4. Bajusz, D., Rácz, A., & Héberger, K. (2015). Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of Cheminformatics*, 7(1), 20.
5. Berendsen, H. J., Postma, J. P., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8), 3684-3690.
6. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235-242.
7. Canon J, Rex K, Saiki AY, Mohr C, Cooke K, Bagal D, Gaida K, Holt T, Knutson CG, Koppada N, Lanman BA. The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature*. 2019;575(7781):217-223.
8. Cox AD, Fesik SW, Kimmelman AC, Luo J, Der CJ. Drugging the undruggable RAS: Mission possible? *Nat Rev Drug Discov*. 2014;13(11):828-851.
9. Egbemhenghe AU, Aderemi OE, Omotara BS, Akhimien FI, Osabuohien FO, Adedapo HA, Temionu OR, Egejuru WA, Ajala CF, Ihunanya MF, Oluwafemi OO, Onu CFD, Ajibare AC, Ddamulira C, Abalum JO, Afolayan OM. Computational-based drug design of novel small molecules targeting p53-MDMX interaction. J Biomol Struct Dyn. 2024 Aug;42(13):6678-6687. doi: 10.1080/07391102.2023.2245483. Epub 2023 Aug 14. PMID: 37578044.
10. Egejuru WA, Aderemi OE, Egbemhenghe AU, Ali MQ, Opasina DE, Agadagba CO, et al. Identification of novel immuno-oncology compounds as VISTA-inhibitors for cancer therapy: A computational approach integrating virtual screening and molecular dynamics simulation. *Informatics Med Unlocked*. 2024;46:101461. doi: 10.1016/j.imu.2024.101461.
11. Egejuru WA, Egbemhenghe AU, Omotara BS, Aderemi CO, Isaac IB, Oladokun EP, Solanki PR, Kanu IM, Ddamulira C, Abdulrasaq IT, Asibor YE, Ofoegbu OV, Ogunyemi SO, Agadagba CO, Ali MQ, Aderemi OE. Computational identification of YQG-like small-molecule inhibitors targeting TIM-3 for cancer immunotherapy. Int J Biochem Res Rev. 2025;34(1):134–52. doi: 10.9734/ijbcrr/2025/v34i1955.
12. Gopalakrishnan, B., & Desiraju, G. (2019). Strong and Weak Hydrogen Bonds in Protein–Ligand Recognition. Journal of the Indian Institute of Science, 100(1), 31-41.
13. Hess, B., Kutzner, C., van der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4(3), 435-447.
14. Hong, D. S., Fakih, M. G., Strickler, J. H., Desai, J., Durm, G. A., Shapiro, G. I., ... & Govindan, R. (2020). KRAS(G12C) inhibition with sotorasib in advanced solid tumors. *New England Journal of Medicine*, 383(12), 1207-1217.
15. Huang, J., & MacKerell Jr, A. D. (2013). CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of Computational Chemistry*, 34(25), 2135-2145.
16. Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33-38.
17. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2), 926-935.
18. Mangel, W., Lin, B., & Ramakrishnan, V. (1990). Characterization of an extremely large, ligand-induced conformational change in plasminogen. Science, 248(4951), 69-73.
19. Manica AK, Daud MO, Faloyo MO, Akinwunmi AR, Adekunle AM, Adekola AA, Lawal IA, Salawu MA, Muritala JA, Muraina RO, Hassan RA, Ogunyemi SO. Computational modeling design of novel NMDAR agonist for the treatment of Schizophrenia. Neurogenetics. 2025 Jan 7;26(1):18. doi: 10.1007/s10048-024-00796-2. PMID: 39762449.
20. Martins, F. A., & Freitas, M. P. (2020). Bioconformational modulation of a thymidine kinase enzyme ligand through F⋯HO intramolecular hydrogen bond. Journal of Molecular Graphics & Modelling, 97, 107545.
21. Newcomer, M. E., Lewis, B., & Quiocho, F. A. (1981). The radius of gyration of L-arabinose-binding protein decreases upon binding of ligand. The Journal of Biological Chemistry, 256(24), 13218-13222.
22. O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(1), 33.
23. Ogunlana AT, Boyenle ID, Ojo TO, Quadri BO, Elegbeleye OE, Ogbonna HN, Ayoola SO, Badmus IO, Manica AK, Joshua KI, Onikute OW, Anamelechi JP, Odetunde A, Falusi AG, Oyedele AK. Structure-based computational design of novel covalent binders for the treatment of sickle cell disease. J Mol Graph Model. 2023 Nov;124:108549. doi: 10.1016/j.jmgm.2023.108549. Epub 2023 Jun 17. PMID: 37339569.
24. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*. 2013;503(7477):548-551.
25. Patricelli MP, Janes MR, Li LS, Hansen R, Peters U, Kessler LV, Chen Y, Kucharski JM, Feng J, Ely T, Chen JH. Selective inhibition of oncogenic KRAS output with small molecules targeting the inactive state. *Cancer Discov*. 2016;6(3):316-329.
26. Prior IA, Hood FE, Hartley JL. The frequency of Ras mutations in cancer. *Cancer Res*. 2020;80(14):2969-2973.
27. Raschka, S., Wolf, A., Bemister-Buffington, J., & Kuhn, L. (2018). Protein–ligand interfaces are polarized: discovery of a strong trend for intermolecular hydrogen bonds to favor donors on the protein side with implications for predicting and designing ligand complexes. Journal of Computer-Aided Molecular Design, 32(3), 511-528.
28. Simanshu DK, Nissley DV, McCormick F. RAS proteins and their regulators in human disease. *Cell*. 2017;170(1):17-33.
29. Stephen AG, Esposito D, Bagni RK, McCormick F. Dragging ras back in the ring. *Cancer Cell*. 2014;25(3):272-281.
30. Tran E, Robbins PF, Lu YC, Prickett TD, Gartner JJ, Jia L, Pasetto A, Zheng Z, Ray S, Groh EM, Kriley IR. T-cell transfer therapy targeting mutant KRAS in cancer. *N Engl J Med*. 2016;375(23):2255-2262.
31. Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455-461.
32. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *Journal of Computational Chemistry*, 26(16), 1701-1718.
33. Willett P. Chemical similarity searching. *J Chem Inf Model*. 2013;55(3):524-538. doi:10.1021/ci400596g.
34. Wong, E., Na, D., & Gsponer, J. (2013). On the Importance of Polar Interactions for Complexes Containing Intrinsically Disordered Proteins. PLoS Computational Biology, 9.
35. Yamamoto, E., Akimoto, T., Mitsutake, A., & Metzler, R. (2020). Universal Relation between Instantaneous Diffusivity and Radius of Gyration of Proteins in Aqueous Solution. Physical Review Letters, 126(12), 128101.