ACCSCIENCE PUBLISHING

Tumor Discovery

ORIGINAL RESEARCH ARTICLE

Covalent docking-based virtual screening and molecular dynamics simulations identify C02b as a potential *KRAS*(G12C) inhibitor

Safiye Merve Bostancioglu[1]* and Ahmet Acar[2]*

[1]Department of Biology, Faculty of Arts and Sciences, Marmara University, Goztepe Campus, Istanbul, Turkey

[2]Department of Biological Sciences, Middle East Technical University, Universiteler Mah., Ankara, Turkey

Abstract

Decades of efforts to target the "undruggable" *Kirsten rat sarcoma viral oncogene homolog* (*KRAS*) oncoprotein yielded promising results in 2021 with the approval of Sotorasib, a *KRAS*(G12C) inhibitor for non-small cell lung cancer patients with the *KRAS*(G12C) variant. Before Sotorasib's approval, developing *KRAS*(G12C) covalent inhibitors faced challenges, particularly their inability to preserve the *KRAS* protein in the GDP-bound state, which hindered clinical trial progression. Considering the importance of developing inhibitors targeting *KRAS*(G12C), we aimed to identify compounds analogous to Sotorasib, resulting in the discovery of a total of 174 Sotorasib scaffold-compounds. We then performed covalent docking-based virtual screening to examine the binding affinity of Sotorasib-like compounds to *KRAS*(G12C). Four compounds showed comparable binding energies according to Glide score to Sotorasib for targeting *KRAS*(G12C). Subsequently, molecular dynamics (MD) simulations were conducted for these four compounds, spanning 100 ns, 300 ns, and 500 ns durations, to identify the most stable complex with the lowest root-mean-square deviation (RMSD), similar to the *KRAS*(G12C)-Sotorasib reference complex. Additional dynamic cross-correlation matrix and PCA were performed as post-MD analyses to investigate the movements of two switches and the flexible regions of *KRAS*(G12C)-Sotorasib and -C02b complexes. As a result, among these four compounds, *KRAS*(G12C)-C02b was found as the optimal candidate. Further investigations beyond this study may provide more insight into C02b's inhibitory effect on *KRAS*(G12C), offering a deeper understanding of its potential as a therapeutic agent.

Keywords: *KRAS*(G12C); Sotorasib-like compounds; Molecular dynamics simulations

***Corresponding authors:**
Safiye Merve Bostancioglu (mervebostancioglu@marun.edu.tr); Ahmet Acar (acara@metu.edu.tr)

Citation: Acar A, Bostancioglu SM. Covalent docking-based virtual screening and molecular dynamics simulations identify C02b as a potential *KRAS*(G12C) inhibitor. *Tumor Discov*. 2025;4(1):79-98. doi: 10.36922/td.5163

Received: October 15, 2024

Revised: November 30, 2024

Accepted: December 9, 2024

Published online: December 26, 2024

Copyright: © 2024 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Cancer is a complex disease characterized by the number of key hallmarks, including sustained proliferative signaling, evasion of growth suppressors, resistance to cell death, enabling replicative immortality, inducing angiogenesis, activation of invasion and metastasis, deregulation of cellular energetics, avoidance of immune destruction, promotion of inflammation, genome instability and mutation, and alteration of

the tumor microenvironment.[1,2] These characteristics collectively contribute to the development and progression of cancer, making it a difficult and complex disease to treat. Understanding these characteristics is critical for developing cancer therapies that target specific aspects of the disease's biology, as cancer is greatly influenced by evolutionary principles, particularly through tumor evolution.[3-6] Genetic mutations and natural selection drive the development of diverse cell populations within a tumor, resulting in genetic diversity and heterogeneity.[7] This diversity enables cancer cells to adapt to various environmental challenges, such as treatments or changes in the microenvironment, resulting in the survival of drug-resistant and more aggressive cell clones.[8-11]

Kirsten rat sarcoma viral oncogene homolog (*KRAS*) signaling regulates important cellular functions through a network of intracellular signaling pathways.[12] *KRAS* functions normally as a molecular switch, cycling between an inactive (GDP-bound) and active (GTP-bound) form. When activated, GTP-bound *KRAS* drives downstream signaling pathways that regulate cell growth, survival, and proliferation, such as the MAPK/ERK and PI3K/AKT pathways.[13] *KRAS* mutations, such as the common G12C mutation, maintain *KRAS* in an active state, resulting in uncontrolled signaling and cellular responses, ultimately driving cancer development and progression by promoting cell growth, inhibiting apoptosis, enhancing metastasis, and modulating the tumor microenvironment.[14]

KRAS gene mutations, particularly the G12C variant, are prevalent in a variety of cancers, including non-small cell lung cancer (NSCLC) and colorectal cancer.[15,16] *KRAS*(G12C) inhibitors are an emerging class of targeted cancer therapy that targets a specific mutation (G12C) in the *KRAS* gene.[17] *KRAS*(G12C) inhibitors aim to address a long-standing problem in cancer treatment, as *KRAS* mutations were previously considered "undruggable"[18] Sotorasib (AMG 510) was the first *KRAS*(G12C) inhibitor to receive accelerated FDA approval in May 2021 for the treatment of adult patients with NSCLC harboring the *KRAS*(G12C) mutation.[19] This marked a significant milestone in cancer therapy by introducing the first drug approved to target *KRAS* mutations. Sotorasib functions by irreversibly binding to the mutant *KRAS*(G12C), inhibiting its activity.[20] Specifically, Sotorasib forms a covalent bond with cysteine 12 in the switch II pocket of *KRAS*(G12C),[21] preventing its ability to activate downstream signaling pathways that promote cell growth and proliferation.[13]

Although Sotorasib has been approved as a drug for targeting *KRAS*(G12C),[19] there are still a limited number of virtual screening approaches to identify potential inhibitors for the *KRAS*(G12C).[22-24] Given this urging need to identify novel pharmacophores targeting the *KRAS*(G12C), we performed a covalent docking-based virtual screening in conjunction with molecular dynamics (MD) simulations to investigate the stability of *compounds* binding to the target *protein*. Based on scaffold-hopping approach with Sotorasib as a template, we identified 174 molecules exhibiting structural similarities to Sotorasib among billions of compounds. Our covalent docking-based virtual screening revealed potential putative binders of *KRAS*(G12C), namely, C01, C02a, C02b, and C03, with notable Glide score similarities to Sotorasib. MD simulations further demonstrated that the *KRAS*(G12C)-C02b complex exhibited stable binding interactions with a low root-mean-square deviation (RMSD) value, comparable to the reference FDA-approved drug Sotorasib. Moreover, the calculations of binding free energy indicated that these favorable interactions were highly preserved from the last 50 nanoseconds (ns) up to 500 ns of simulation time. To further investigate the dynamics of the *KRAS*(G12C)-Sotorasib and *KRAS*(G12C)-C02b complexes, we conducted post-MD analyses, including dynamic cross-correlation matrix (DCCM) and principal component analysis (PCA). These analyses aimed to elucidate the movements of the two switches and identify the flexible regions within the complexes. These findings suggest that C02b may serve as a promising inhibitor candidate for the *KRAS*(G12C) mutation, and further advanced studies are needed to validate its potential.

2. Materials and methods

2.1. Similarity search

Sotorasib served as a template to identify similar compounds within various chemical spaces including eXplore, Freedom Space, GalaXi, and KnowledgeSpace. The scaffold-hopping search was conducted using the infiniSee software (version 5.0.1; BioSolveIT GmbH, Sankt Augustin, Germany, 2023, www.biosolveit.de/infiniSee). The chemical spaces encompassing both building blocks and reactions allowed for the creation of a maximum number of virtual product molecules. Scaffold Hopper was executed using the infiniSee software.[25] The acrylamide warhead was selected from Sotorasib[22] (Figure A1 in Appendix). The target similarity was set at 1.00 to ensure a precise match, while a minimum similarity threshold was established at 90% to guarantee a significant level of similarity. Furthermore, a total diversity score of 1.00 was ensured to uncover compounds with a high degree of similarity with Sotorasib. As a result, 174 compounds containing the acrylamide warhead and showing structural similarities to Sotorasib were identified and saved in the structure-data file (SDF) format.

2.2. Protein and ligand preparation

The 174 compounds containing the acrylamide warhead were prepared and optimized for molecular docking using LigPrep.[26,27] Compounds were obtained using Epik at pH 7.0 ± 1.0.[28] The protein structure of the inactive GDP-bound state of human *KRAS*(G12C) (PDB ID: 6OIM) was retrieved from the Protein Data Bank, with a resolution of 1.65 Å.[29] Protein preparation was performed using the protein preparation wizard in Maestro,[30] which fixed all the missing residues and atoms of the protein structure. Several crucial steps in the preparation included introducing hydrogen atoms, refining hydrogen bond networks, and addressing missing side chains through the utilization of Prime. Subsequently, water molecules were eliminated, and a restrained minimization employing the OPLS3e force field was carried out.[31] The resulting optimized protein structure was then utilized for the subsequent docking process.

2.3. Covalent dockingbased virtual screening

A set of 174 pharmacophores similar to Sotorasib was employed in covalent docking-based virtual screening (CovDock-VS)[32] against the *KRAS*(G12C) target (PDB ID: 6OIM). The reactive residue Cys12 was specifically selected, and grid generation for docking was created as a cubic grid box centered at the coordinates (x, y, and z) of Sotorasib, with a length of 12 Å. The acrylamide electrophilic warhead group, identical to the template drug Sotorasib, was selected for the Michael addition – a reaction responsible for the covalent bonding of ligands featuring an acrylamide functional group with the side chain of nucleophilic protein residues. Subsequently, the ligands were selected and prioritized based on their Glide scores for the binding modes of pre-reactive complexes.[33,34]

2.4. MD simulations

MD simulations were conducted on the protein-ligand complexes derived from the covalent docking-based virtual screening using GROningen MAchine for Chemical Simulations (GROMACS) 2023.1.[35] The CHARMM-GUI (Chemistry at HARvard Macromolecular Mechanics Graphical User Interface) server was used for solution building. The CHARMM36 force field[36] was applied to define atom types and assign atomic partial charges for both the ligand and protein. Furthermore, the CHARMM-GUI server was employed to generate the topology for both the *KRAS*(G12C) receptor and the promising inhibitors. For the setup of MD simulations, the protein-ligand complexes were centered in an octahedral box and solvated with the TIP3P water model. Neutralization of the complexes charge was achieved by adding 150 mM ions. The minimization process concluded when the maximum force was reduced to <10.0 kJ/mol, with a maximum duration of 100 picoseconds (ps) for the minimization steps. Subsequently, a 100 ps equilibration phase with position restraints on the protein and ligand molecules was conducted using NVT (Number of Volume and Temperature) and NPT (Number of Pressure and Temperature) ensembles. During this period, the systems were heated to 303.15 K using a modified Berendsen thermostat (V-rescale),[37,38] and pressure was maintained using C-rescale pressure coupling set to a reference pressure of 1 bar.[39] For energy minimization, NVT, and NPT relaxation simulations, short-range interactions utilized a smooth force switch with a cut-off of 1.2 nm, and long-range electrostatics were computed using the Particle-Mesh-Ewald (PME) algorithm.[40] In addition, hydrogen bonds were constrained by employing the Linear Constraint Solver algorithm.[41] Finally, MD simulations were conducted for durations of 100 ns, 300 ns, and 500 ns (Table 1) without applying restraints. A 2 femtoseconds integration time step was employed, and trajectory snapshots were captured at 1 ps intervals.

MDS analyses were performed using GROMACS, a Linux-×86_64-multicore CUDA-enabled program on an NVIDIA GeForce GTX 1060 GPU and an 8-core Intel Core i5-7000 central processing infrastructure. The computation of a single complex required approximately 100 ns of processing time per simulation day. GROMACS modules were used to analyze MD trajectories, focusing on parameters such as root RMSD/mean-square deviation fluctuation (RMSF) and hydrogen bond (H-bond) analysis. Graphs were generated using XMGrace[42] and visual molecular dynamics (VMD).[43]

The calculation of binding free energy was performed using the gmx_MMPBSA tool, which is based on Assisted Model Building with Energy Refinement MM/PB(GB)SA approach. This tool is designed for conducting end-state free energy calculations using GROMACS trajectory files. Specifically, the final 50 ns of the simulation trajectory were considered for the molecular mechanics with Generalized Born surface area (MM/GBSA) calculation.

Table 1. *KRAS*(G12C) complexes feature with their atom numbers for duration 500 ns

Atoms	*KRAS*(G12C)-Sotorasib	*KRAS*(G12C)-C02b
Protein	2,704	2,704
Ligand	72	87
Na^+	14	13
Cl^-	7	7
Water (TIP3P)	7,914	7,923

Abbreviations: TIP3P: Transferable intermolecular potential three-point; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog.*

2.5. Post-MD simulations analyses

The analysis employed Bio3D,[44] an R library-based tool, to generate time-correlated DCCM for various ligand-bound forms of *KRAS*(G12C). This approach aimed to discern the localization and nature of fluctuations in C^{α} atoms within residues crucial to *KRAS*(G12C) function. Given the limitations of DCCM, which may overlook correlated but perpendicular motions, the analysis was complemented with PCA. In PCA, simulation frames are organized based on their principal components (eigenvectors), which represent the key directions capturing the most significant variability in the data. This approach allows for effective separation of structures by emphasizing dominant patterns and variations within the dataset. For both analyses, the final 400 ns of all MD simulations were utilized at 20 ps intervals per Bio3D file size constraints. This comprehensive approach provides a nuanced understanding of the dynamic behavior of *KRAS*(G12C) ligand-bound forms, integrating both normal mode analysis and PCA techniques to capture a broad spectrum of structural fluctuations.

3. Results

3.1. Covalent docking-based virtual screening

It is widely accepted that molecules with comparable chemical structures will have similar pharmacological properties. In ligand-based studies, Sotorasib, an FDA-approved drug for treating *KRAS*(G12C) protein in NSCLC, was reported to be a reliable template. To evaluate ligand interactions, we used the InfiniSee, a platform capable of examining billions of compounds across chemical spaces, including eXplore, Freedom Space, GalaXi, CHEMriya, and REAL, to identify potential chemical scaffolds similar to the template. This platform also provides access to KnowledgeSpace, a literature-based virtual chemical space with a strong emphasis on synthetic accessibility. These technological advancements offer great potential in the fields of drug discovery and chemical research due to its capacity to provide a greater spectrum of molecules, allowing the design of more effective and innovative therapeutic solutions. As a result, 174 molecules exhibiting structural similarities to Sotorasib were found from billions of compounds in the chemical spaces (Table A1 in Appendix). Subsequently, the selected compounds were further evaluated using a covalent docking-based virtual screening approach.

3.2. Covalent docking-based virtual screening identifies four compounds

For the covalent-based virtual screening approach, the X-ray structure of the *KRAS*(G12C) (PDB ID: 6OIM), forming a complex with Sotorasib, was selected.[29] The COVDOCK protocol was utilized to investigate the redocking capabilities as well as the binding of the co-crystallized Sotorasib to the *KRAS*(G12C) with GDP. As a result, we observed a similar binding mechanism to that observed in the crystallized structure, validating the covalent docking protocol (Figure 1A). Next, the binding modes of these four compounds were compared with the co-crystallized Sotorasib according to their 2-D interaction diagram (Figure 1B). The precision of the COVDOCK protocol in reproducing the binding mechanism of the co-crystallized ligand with the *KRAS*(G12C) protein has been demonstrated by prior researchers.[22,45]

The set of 174 selected compounds was then subjected to a covalent docking-based virtual screening, with their binding energies ranked according to their Glide scores. While none of the ranked compounds displayed a higher estimated binding energy compared to Sotorasib, which possesses a value of −8.1 kcal/mol, four compounds namely C01, C02a, C02b, and C03 exhibited closer and notable Glide scores of −7.8 kcal/mol, −7.5 kcal/mol, −7.3 kcal/mol, and −7.0 kcal/mol, respectively (Table 2 and Figure A1).

The comparison of binding modes for the four compounds with co-crystalized Sotorasib exhibited Pi-Pi stacking interactions with the residue TYR96. Similarly, compounds C01, C02a, C02b, and C03 exhibited similar interactions (Figure 2). Furthermore, these compounds formed hydrogen bonds with additional residues (LYS16, ALA59, and GLN61) in the binding pocket of *KRAS*(G12C) (Figure 2). Taken together, the four candidates show promise as covalent binders of *KRAS*(G12C) in its inactive GDP-bound configuration.

3.3. MD simulations of C01, C02a, C02b, C03, and Sotorasib with *KRAS*(G12C) protein

The root RMSD analysis allows us to monitor the fluctuation in the three-dimensional structure over time, offering valuable insight into the mobility of binding pocket residues during the MD simulation. RMSD is a measure used in computer simulations to determine how far a molecule or part of it has moved from its initial position. To assess possible fluctuations in the 3D structure over time, we ran simulations on various *KRAS*(G12C) protein complexes, including *KRAS*-C01, *KRAS*-C02a, *KRAS*(G12C)-C02b, *KRAS*-C03, and *KRAS*(G12C)-Sotorasib. Initially, we ran simulations for all complexes for 100 ns. Following that, we increased the simulation time for *KRAS*-C01, *KRAS*(G12C)-C02b, and *KRAS*(G12C)-Sotorasib to 300 ns. Finally, we ran simulations for *KRAS*(G12C)-C02b and *KRAS*(G12C)-Sotorasib for 500 ns. This enabled us to compare the structural stability of potential therapeutic

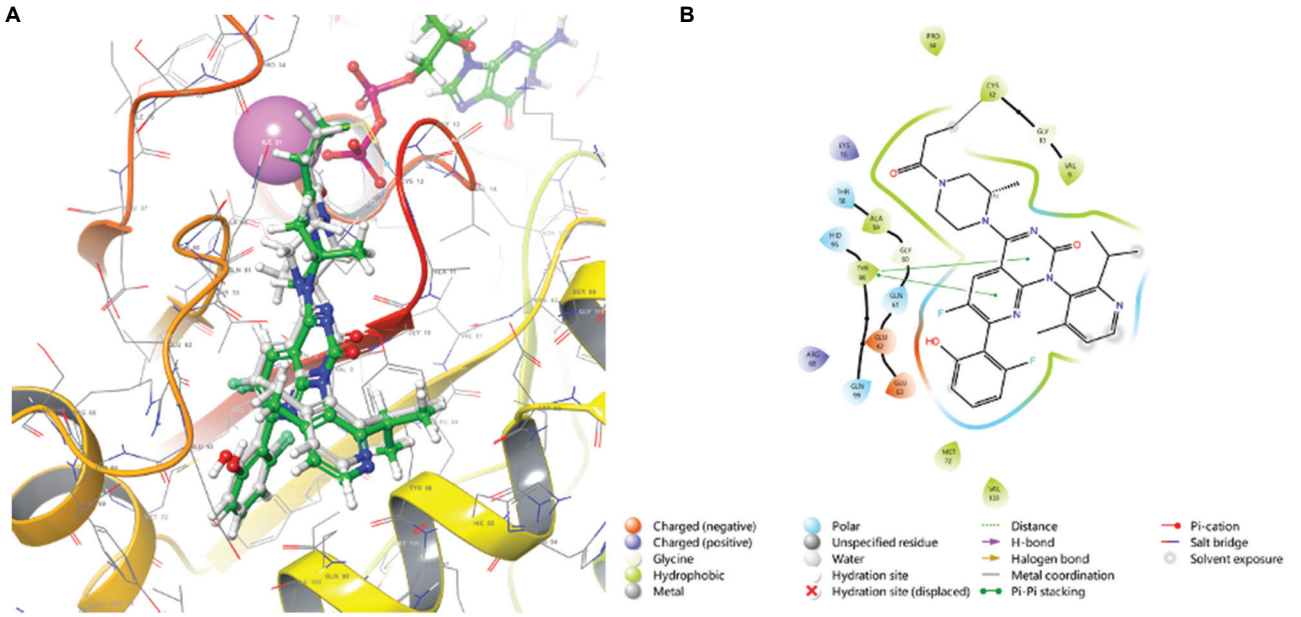


Figure 1. (A) The binding modes of the X-ray-bound Sotorasib and the re-docked Sotorasib were superimposed. (B) Amino acid interactions of Sotorasib with *Kirsten rat sarcoma viral oncogene homolog*(G12C) and compound prioritization based on molecular dynamics simulations and binding free energy calculations.

Table 2. The binding affinities and physicochemical properties of *KRAS*-C01, *KRAS*-C02a, *KRAS*(G12C)-C02b, and *KRAS*-C03 resembling the chemical scaffold of Sotorasib. TPSA: vdW polar surface area, LogP: Predicted octanol/water partition coefficient

Compound number	Compound name	Similarity	Glide score (kcal/mol)	Prime energy	Molecular weight	TPSA	LogP
	Sotorasib	1.00	−8.157	−19.6	560.6	102	4
C01	Shen_92298609_93443091_95741143	0.917	−7.801	−78.3	605.770	95.660	7.178
C02a	Shen_92298609_44473994_95741143	0.916	−7.596	−16.7	600.731	108.550	6.512
C02b	Shen_92298609_36250864_95741143	0.916	−7.347	−22.2	600.731	108.550	6.512
C03	Shen_92298609_50382533_8869305	0.916	−7.074	−5.6	611.675	130.480	4.944

Abbreviations: TPSA: Topological polar surface area.

candidates with the clinically approved drug Sotorasib using metrics such as RMSD and MM/GBSA to assess their binding strength. The 100 ns MD simulations revealed that, apart from *KRAS*-C02a and *KRAS*-C03, all complexes maintained an average RMSD value ranging from 0.2 to 0.4 nm, indicating a relatively stable simulation completion that was closely matched with their initial conformations (Figure 3). Furthermore, *KRAS*-C01 was unable to maintain its stability in subsequent 300 ns MD simulations (Figure 3). Complexes such as *KRAS*(G12C)-C02b and *KRAS*(G12C)-Sotorasib, on the other hand, successfully completed the simulation with RMSD values ranging from 0.2 to 0.4 nm (Figure 3). Continuing the 500 ns MD simulation, *KRAS*(G12C)-C02b maintained its stability until around 450 ns when its RMSD ranged from 0.3 to 0.6 nm near the end of the simulation at 500 ns (Figure 3). Of note, *KRAS*(G12C)-Sotorasib maintained an RMSD value ranging from 0.2 to 0.4 nm throughout the entire simulation. Thus, based on the RMSD of the ligand after least square fitting to the protein backbone, *KRAS*(G12C)-C02b exhibits a therapeutic candidate attribute comparable to the reference complex, *KRAS*(G12C)-Sotorasib, in that it maintains binding stability and shows no evidence of dissociation from the binding site.

To understand the fluctuations at the per residue level upon ligand binding, root mean-square fluctuation (RMSF) calculations were performed. Consistent with previously published findings,[26,31,46] it was observed that, in contrast to the entire protein, fluctuations were predominantly concentrated around switch-I (residues 30 – 38) and switch-II (residues 60 – 76). In Figure 4A, high peaks were noted in the RMSF graph within the region encompassing switch-I and switch-II loops in both complexes. Furthermore, hydrogen bond occupancy,

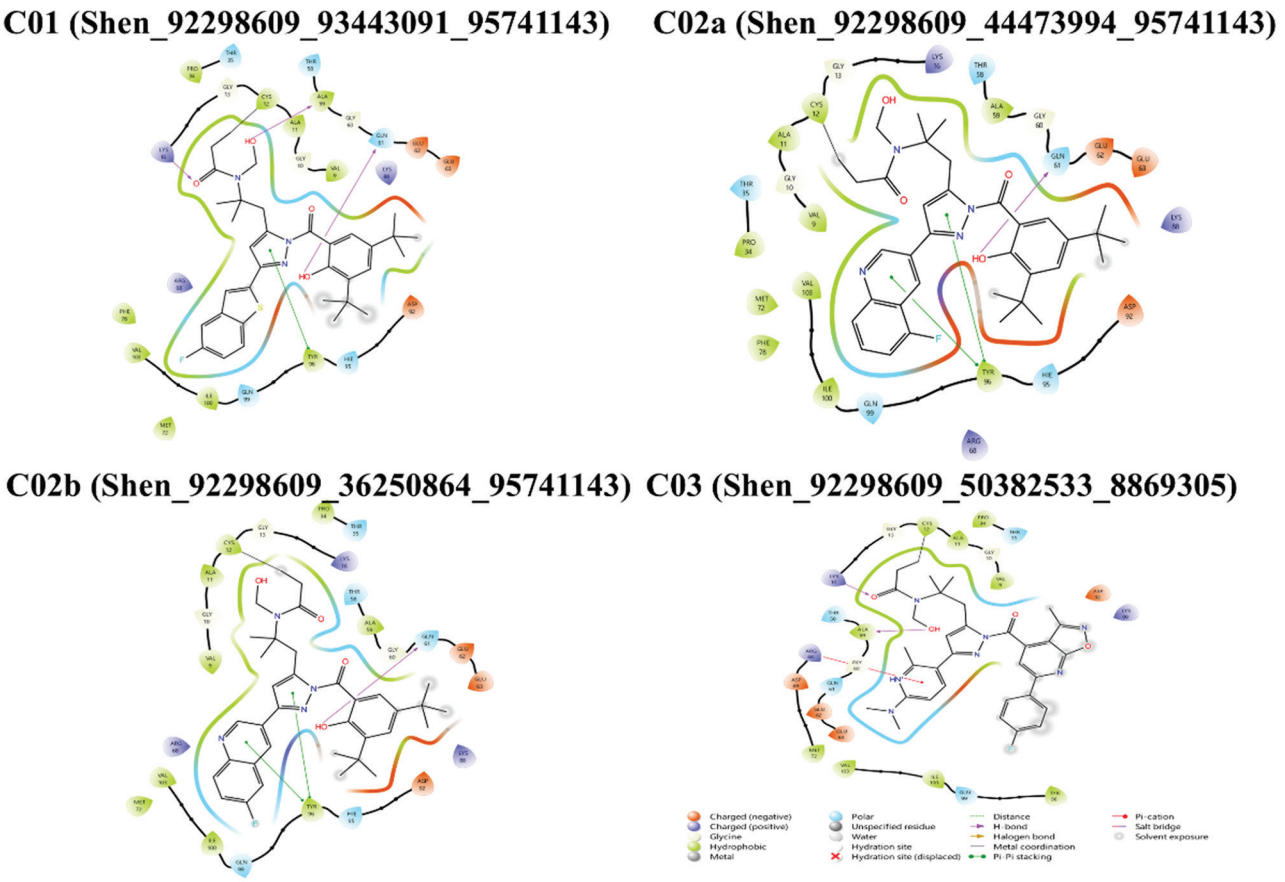


Figure 2. Amino acid interactions of the top-ranking compounds with *Kirsten rat sarcoma viral oncogene homolog*(G12C).

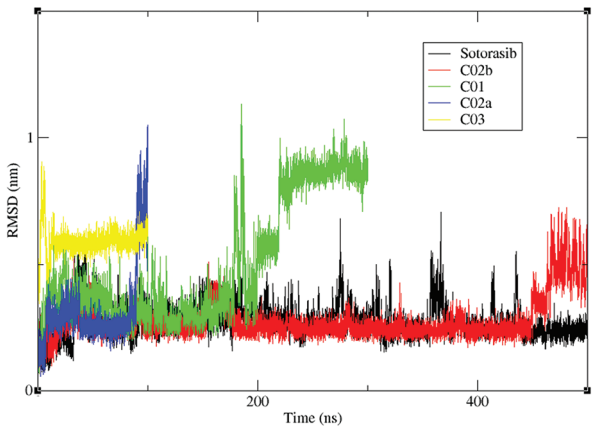


Figure 3. Lig-fit-prot RMSD graph spectrums of Sotorasib, *KRAS*-C01, *KRAS*-C02a, *KRAS*(G12C)-C02b, *KRAS*-C03, and *KRAS*(G12C)-Sotorasib in complex with *KRAS*(G12C).
Abbreviations: RMSD: Root-mean-square deviation; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog.*

quantified as the average number of hydrogen bonds per time frame, was also computed for these key residues. The percentages of hydrogen bond occupancies are detailed in Figures 4B and C. The elevated rate of hydrogen bond formation between C02b and GLN61, as well as Sotorasib and ASP69, signifies a robust and dependable hydrogen-bonding profile for these interactions.

3.4. MM/GBSA calculation for evaluating binding free energy (dG)

Afterward, binding free energy (dG) was calculated using gmx_MMPBSA, and the MM/GBSA dG distribution graph was generated for the last 50 ns of the 500 ns MD simulation. For *KRAS*(G12C)-Sotorasib, the average dG was −41 kcal/mol (Figure 5A), whereas for *KRAS*(G12C)-C02b, this value was −44 kcal/mol (Figure 5B). In addition, the Delta TDC (Total DeComposition) values, illustrating the residues contributing to the binding free energy, were investigated. During the interaction with *KRAS* G12C-Sotorasib, TYR96, and GLN99 were found to contribute significantly to the binding energy, followed by ARG68 and MET72 (Figure 5C). Examination of the Delta TDC values for *KRAS*(G12C)-C02b revealed that, like in the reference complex, TYR96 and GLN99 contributed most significantly to the binding free energy, followed by GLN61 and MET72,

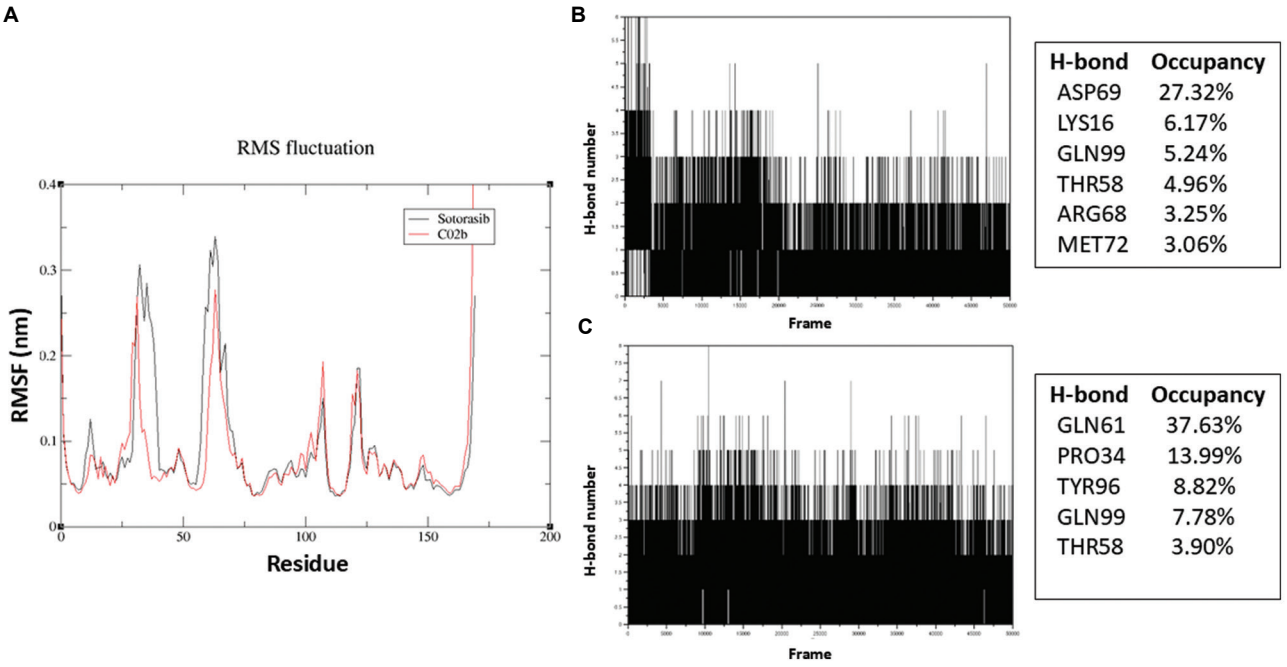


Figure 4. (A) RMSF profiles of *KRAS*(G12C) in complex with reference and hit compounds. H-bonding interactions of (B) *KRAS*(G12C)-Sotorasib and (C) *KRAS*(G12C)-C02b.
Abbreviations: RMSF: Root-mean-square fluctuation; H-bond: Hydrogen bond; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog.*

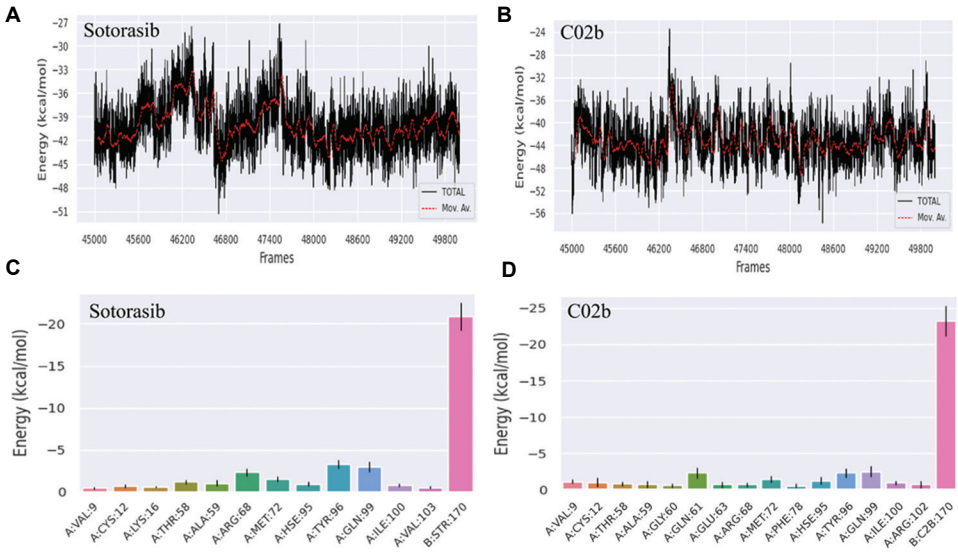


Figure 5. MM/GBSA dG distribution graph obtained for the last 50 ns of the 500 ns MD simulation for (A) Sotorasib, (B) C02b; and Delta total decomposition values for (C) STR: Sotorasib, (D) C2B: C02b.
Abbreviations: MM/GBSA: Molecular mechanics with generalized born surface area; MD: Molecular dynamics.

with a minor contribution from HIS95 (Figure 5D). It is worth reiterating that the interaction with GLN61 observed in the covalent docking results persisted throughout the 500 ns simulation. Both the *KRAS*(G12C)-C02b (Figure 6A) and *KRAS*(G12C)-Sotorasib (Figure 6B) complexes remained largely stable throughout the 500 ns trajectory, with minimal deviation from the targeted binding site, maintaining their initial positions. These findings suggest that C02b exhibits binding interactions similar to the reference FDA-approved drug Sotorasib, with highly preserved interactions. Further research into C02b, which shares similarities with Sotorasib, could potentially serve as an alternative.

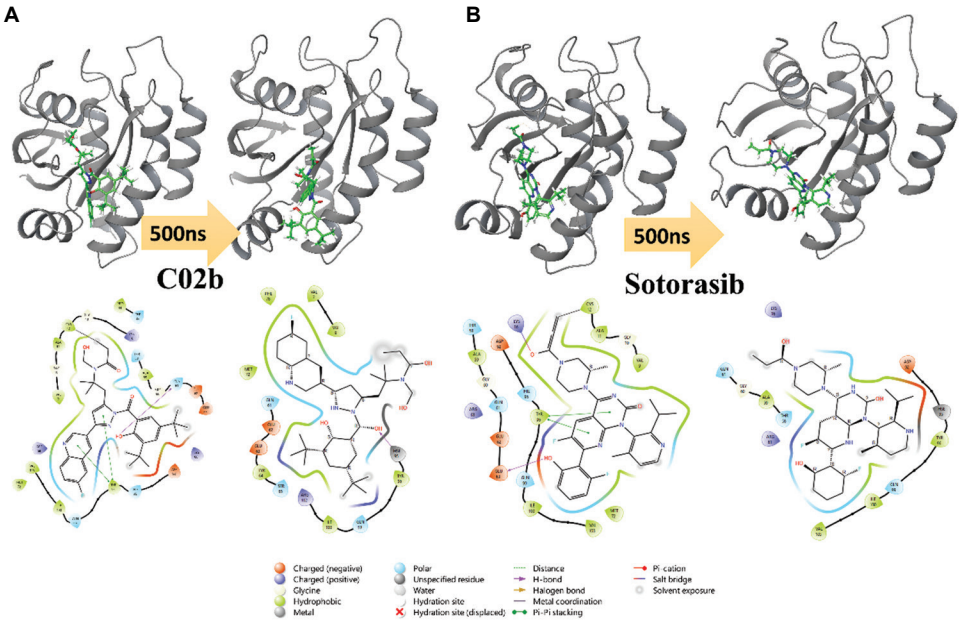


Figure 6. A schematic 2-D and 3-D representation of detailed ligand atom interactions with the protein residues at 0 ns and 500 ns. (A) C02b and (B) Sotorasib.

3.5. PCA and DCCM analysis

PCA was performed on the atomic backbone (C_α position) using three configurations: First principal component (PC1), second principal component (PC2), and third principal component (PC3). The three most representative principal components cumulatively accounted for approximately 50% of all variances among the least correlated components across all 2000 frames analyzed for *KRAS*(G12C)-ligand complexes (Figures 7-10).

In the PCA, the PC1 dominated the overall variance, accounting for more than a third of the total variance (25.45% for Sotorasib and 20.01% for C02b). PC2 accounted 12.21% for Sotorasib and 14.1% for C02b, while PC3 exhibited the lowest variability, with 5.67% for Sotorasib and 6.28% for C02b. Together, the first three components represented 43.3% and 40.5% of the total variance for Sotorasib and C02b, respectively (Figures 7 and 9).

To further elucidate the impact of Sotorasib on the structure of *KRAS*(G12C), RMSF analysis was performed to compare the flexibility of the two systems. Higher fluctuation peaks were observed in the switch-I (residues 30 – 38) and switch-II (residues 60 – 76) regions of the *KRAS*(G12C)-Sotorasib complex (Figure 8), consistent with previous observations of dynamic structural features in those regions.[46] However, for *KRAS*(G12C)-C02b, RMSF values showed a significant decrease, especially in the switch-II region (Figure 10). The 3D structures of *KRAS*(G12C)-Sotorasib and *KRAS*(G12C)-C02b revealed a noticeable reduction in flexibility in the two switch regions for the structure of *KRAS*(G12C)-C02b compared to the reference.

The correlated conformational motions of the *KRAS*(G12C)-Sotorasib and *KRAS*(G12C)-C02b complexes were examined through DCCM analysis (Figure 11). In this analysis, regions displaying high positive values (depicted in red) indicate a strong correlation in the movement of residues in the same direction. Conversely, negative regions (depicted in blue) signify robust anti-correlated motion, where residues move in opposite directions. The color intensity along the diagonal reflects the degree of movement for the corresponding atoms.

Within the reference and C02b hit complex systems, the *KRAS*(G12C)-Sotorasib complex exhibits relatively stronger correlated motions compared to the *KRAS*(G12C)-C02b complex (Figure 11). Specifically, in the DCCM of the *KRAS*(G12C)-Sotorasib complex, the flexible region is observed moving in an anti-correlated manner with both switch-I (residues 30 – 38) and switch-II (residues 60 – 76), as well as displaying anti-correlated motion with the C12 residue and switch-I. This suggests that Sotorasib brings the binding pocket closer through these anti-correlated motions of the flexible regions, including switch-I, switch-II, and the C12 residue.

On the other hand, the DCCM of the *KRAS*(G12C)-C02b complex reveals the flexible regions moving in

an anti-correlated manner with switch-I and switch-II, although the dynamic movement is less pronounced compared to the *KRAS*(G12C)-Sotorasib complex. Notably, there is no observed movement of the C12 residue in the *KRAS*(G12C)-C02b complex. These findings suggest a distinctive dynamic behavior in the correlated motions of the flexible loops and associated residues between the two complexes, highlighting the unique characteristics of Sotorasib and C02b interactions with *KRAS*(G12C).

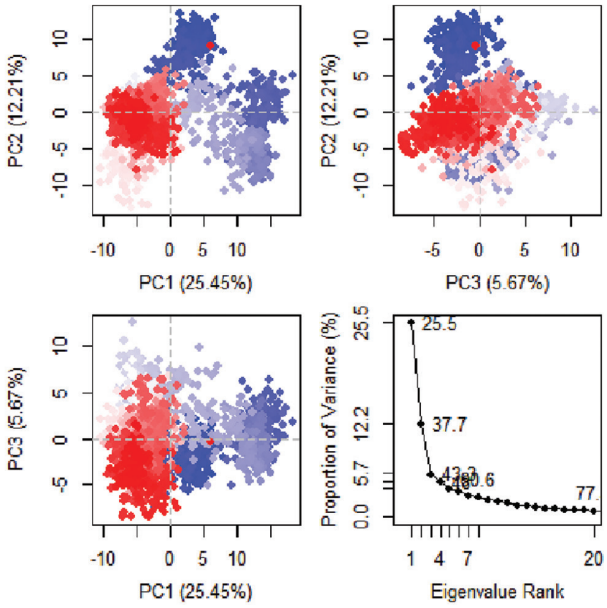


Figure 7. PCA dimension plots illustrating the first three principal components and displaying the cumulative percentages of variance covered by all eigenvectors for the *KRAS*(G12C)-Sotorasib structure. The color spectrum, transitioning from blue to red, delineates conformational alterations across the simulation. Blue dots signify the initial timesteps, white dots depict intermediate stages, and red dots denote the concluding timesteps. Abbreviations: PCA: Principal component analysis; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog*.

The interaction analysis revealed notable differences between C02b and Sotorasib in their binding to the *KRAS*(G12C). C02b forms a strong and persistent interaction with GLN61 within the switch-II region, a key flexible loop implicated in *KRAS* conformational dynamics. In contrast, Sotorasib predominantly interacts with ARG68 and ASP69 in the same region. These interactions were validated through hydrogen bond occupancy analysis, which highlighted the stability and significance of the C02b-GLN61 interaction throughout the MD simulations. This distinct binding pattern of C02b suggests a potential alteration in the modulation of *KRAS*'s switch regions compared to Sotorasib, which may influence downstream signaling pathways.

Furthermore, PCA and DCCM analyses demonstrated unique dynamic behavior for the *KRAS*(G12C)-C02b complex. While Sotorasib induces a broader range of flexibility in the switch-II region, C02b exhibited reduced fluctuations in this region, potentially stabilizing the inactive GDP-bound state of *KRAS*. This stabilization is further supported by MM/GBSA binding free energy

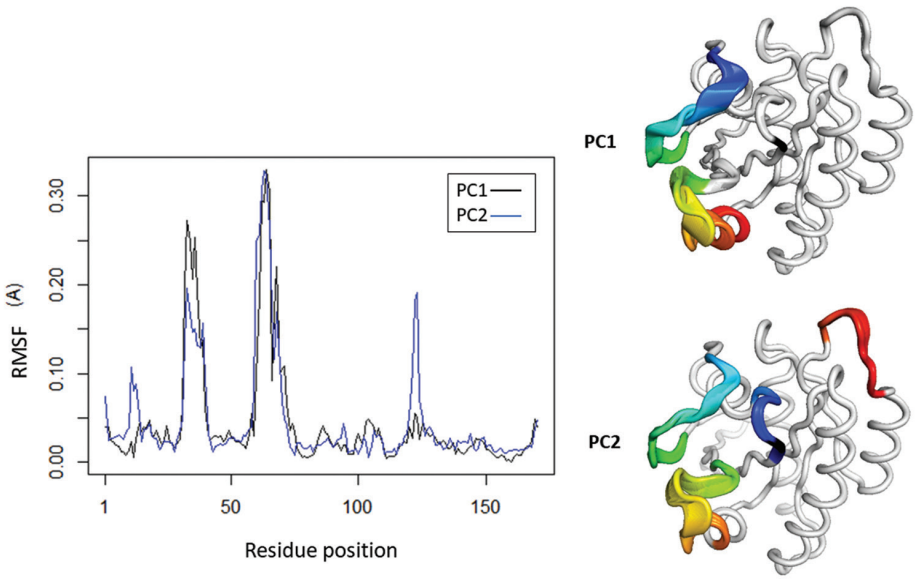


Figure 8. Flexibility analysis of the *KRAS*(G12C)-Sotorasib structure. The left side presents the RMSF values of the C^{α} atoms within the system, while the right side displays ribbon representations of the PC1 and PC2 structures. Flexible regions are highlighted with a rainbow color scheme, providing a visual cue to the dynamic nature of these segments. Notably, the C12G residue is depicted in black, and the structures were visualized using PyMOL.
Abbreviations: RMSF: Root mean-square fluctuation; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog*; PC1: First principal component; PC2: Second principal component.

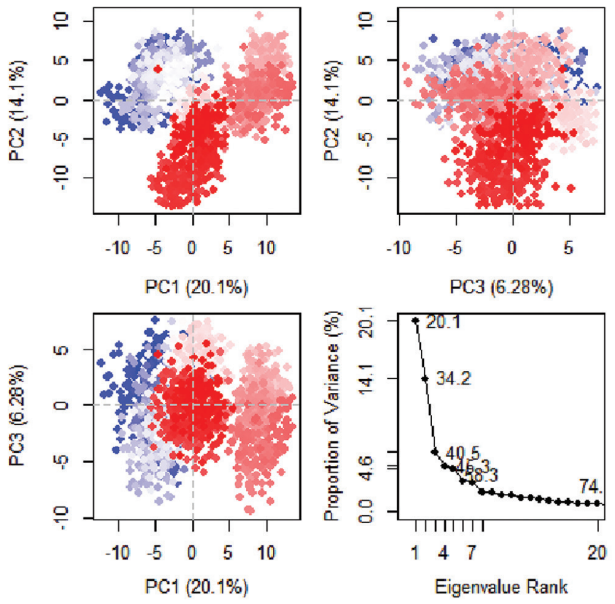


Figure 9. PCA dimension plots illustrating the first three principal components and displaying the cumulative percentages of variance covered by all eigenvectors for the *KRAS*(G12C)-C02b structure. The color spectrum, transitioning from blue to red, delineates conformational alterations across the simulation. Blue dots signify the initial timesteps, white dots depict intermediate stages, and red dots denote the concluding timesteps.
Abbreviations: PCA: Principal component analysis; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog.*

calculations, which confirm the preservation of strong binding interactions for both compounds but with distinct residue-specific contributions.

4. Discussion

In this study, we present potential *KRAS*(G12C) inhibitors identified through a comprehensive approach involving covalent docking-based virtual screening and MD simulations. Our strategy started with a similarity search based on the scaffold of Sotorasib, which successfully identified 174 molecules from a vast pool of compounds. Subsequently, covalent docking-based virtual screening was performed, revealing promising *KRAS*(G12C) protein binders, namely, C01, C02a, C02b, and C03. To assess the stability of the identified complexes and putative binders, MD simulations were employed. Significantly, among the candidates, C02b demonstrated exceptional stability as an inhibitor targeting the *KRAS*(G12C) during the 500 ns MD simulation, comparable to the reference.

In a previous study, it was found that Sotorasib (AMG 510) bound to the HIS95 groove and created a connected network of 25 ligand-protein interactions through van der Waals forces.[29] This network spanned from the backbone of helix 2 (involving HIS95 and TYR96) to the backbone of the flexible switch-II loop. According to the hydrogen bond analysis performed with VMD and decomposition analysis

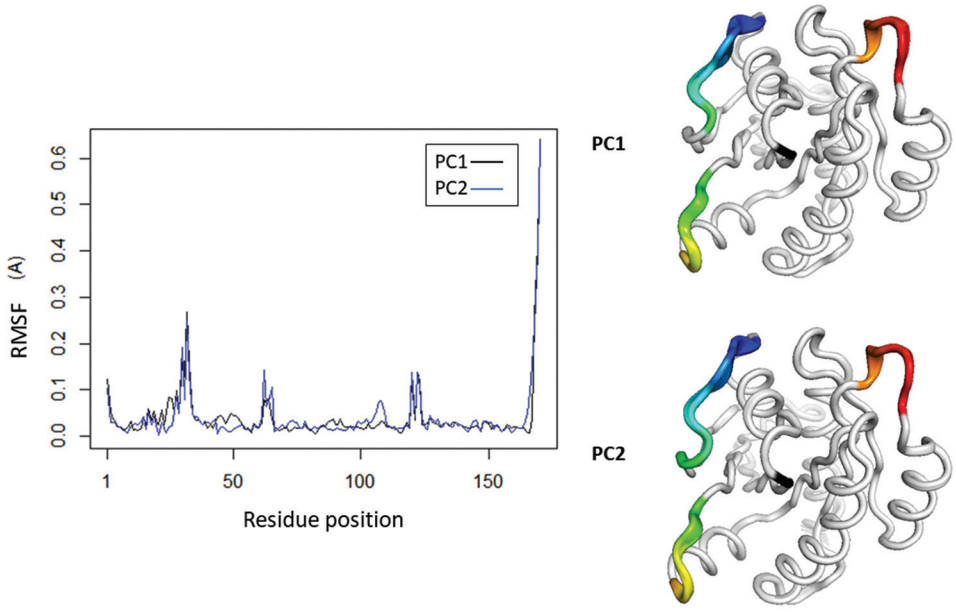


Figure 10. Flexibility analysis of *KRAS*(G12C)-C02b structure. The left side presents the RMSF values of C^{α} atoms within the system, while the right side displays ribbon representations of the PC1 and PC2 structures. Flexible regions are highlighted with a rainbow color scheme, providing a visual cue to the dynamic nature of these segments. Notably, the C12G residue is depicted in black, and the structures were visualized in PyMOL.
Abbreviations: RMSF: Root mean-square fluctuation; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog*; PC1: First principal component; PC2: Second principal component.

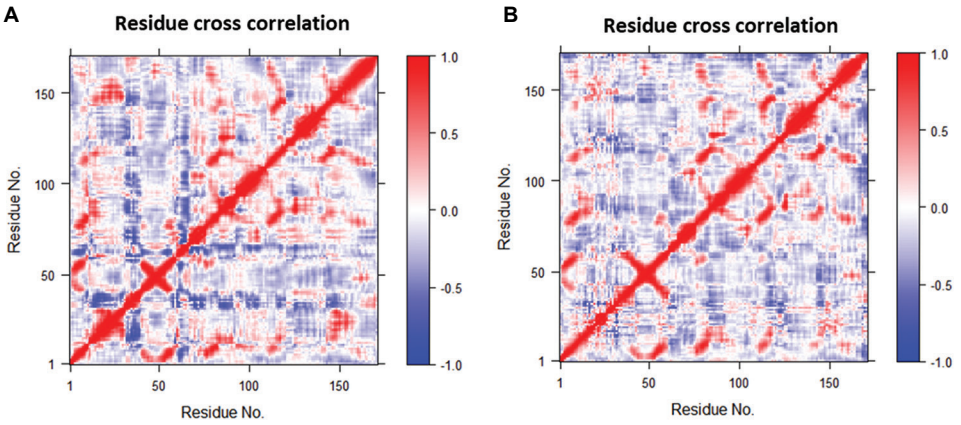


Figure 11. The post-MD simulation time-correlated DCCM plot for the *KRAS*(G12C) (A) Sotorasib and (B) *KRAS*(G12C)–C02b complexes. Blue color indicates anti-correlation and red positive correlation.
Abbreviations: DCCM: Dynamic cross-correlation matrix; MD: Molecular dynamics; *KRAS*: *Kirsten rat sarcoma viral oncogene homolog.*

using Delta TDC values in our study, the key interaction in C02b binding was observed with the GLN61 residue located in the switch-II loop. For Sotorasib, interactions were observed with ARG68 and ASP69 residues in the same loop. In addition, it was found that the TYR96 residue was crucial for both the reference and C02b interactions.

The distinct interaction of C02b with GLN61 in the switch-II region, compared to Sotorasib's interactions with ARG68 and ASP69, likely contributes to the observed differences in conformational dynamics and flexibility of the *KRAS*(G12C) complexes. Switch-II is a critical region for nucleotide exchange and downstream effector binding, and its modulation directly influences *KRAS* signaling activity. By stabilizing GLN61, C02b may better inhibit *KRAS* activation by preventing the conformational changes required for GTP loading and effector interactions. These molecular differences may translate into unique cellular outcomes. Reduced switch-II flexibility, as observed with C02b, could lead to more robust inhibition of *KRAS*-driven downstream signaling pathways, potentially altering cellular proliferation, survival, and apoptotic processes. This distinct mechanism of action highlights C02b as a promising candidate for further preclinical investigation, offering a potential advantage in overcoming resistance mechanisms associated with existing *KRAS* inhibitors like Sotorasib. Further experimental validation will be essential to confirm these findings. Comparative studies on the effects of C02b and Sotorasib in cellular models of *KRAS*(G12C)-driven cancers could provide critical insights into their impact on downstream signaling and therapeutic efficacy.

Moreover, the structural diversity of *KRAS*(G12C) inhibitors remains narrow, leaving substantial room for the development of alternative scaffolds. Such scaffolds could offer improved pharmacokinetics, enhanced binding interactions, and broader application to different tumor types harboring the *KRAS*(G12C) mutation. For instance, novel compounds with distinct binding modes may reduce side effects, enhance tolerability, or bypass resistance mechanisms arising in *KRAS* proteins with secondary mutations or altered conformations. These challenges highlight the urgent need for novel inhibitors that can either supplement or outperform existing treatments. Compounds like C02b, which demonstrate stable binding and distinct interaction profiles in computational analyses, hold promise for addressing these gaps. These novel compounds could expand the arsenal of therapeutic options, potentially improving outcomes when used alone or in combination with existing therapies.

While the current study primarily focuses on computational and simulation-based approaches, we have outlined the necessity of such biochemical analyses in the discussion section. Future work will include experimental assessments of compound selectivity using GTP-loaded *KRAS*(G12C) and wild-type *KRAS* proteins. These experiments will involve evaluating GDP/GTP exchange inhibition, covalent binding to Cys12, and downstream signaling modulation. We appreciate the reviewer's suggestion and recognize its importance for validating the therapeutic potential of C02b.

The catalytic domain of the RAS family maintains a relatively conserved conformation, apart from the two switch regions (switch-I and switch-II).[47] The conformations of these switch regions may play a crucial role in mediating interactions with other proteins.[47] Therefore, post-MD simulations analyses were carried out to evaluate the stability and flexibility of the two switches in the Sotorasib and C02b complexes targeting *KRAS*(G12C). The RMSD analysis determined the structural stability, while RMSF focused on flexibility, particularly in critical

functional regions. Throughout the 500 ns MD simulation, the *KRAS*(G12C)-C02b complex maintained low RMSD values and stability, similar to the reference complex. High peaks were observed in RMSF graph for the region containing switch-I and switch-II loops in both complexes. On the other hand, PCA identified dominant modes of motion, capturing a substantial portion of the total variance and DCCM elucidated correlated conformational motions within the complexes. The dynamic interaction of *KRAS*(G12C) and Sotorasib through a 10 microseconds MD simulation found that Sotorasib increased flexibility in the switch-I and Switch-II regions,[46] However, a previous study observed a decrease in the fluctuation of the switch-I region during an 800 ns MD simulation.[48] The disparity in these simulation results may be attributed to different simulation settings. In our RMSF analysis with PCA (Figures 8 and 10), through a 500 ns MD simulation of the dynamic interaction of *KRAS*(G12C) and Sotorasib, it was observed that Sotorasib did not stabilize *KRAS* switches in the complex.[49]

5. Conclusion

Taken together, our findings emphasize the distinctive stabilizing effects of C02b compared to Sotorasib on the conformational dynamics of *KRAS*(G12C), which may contribute to the development of potential inhibitors and highlight the importance of C02b in the development of *KRAS*(G12C) inhibitors in the preclinical setting.

Acknowledgments

None.

Funding

Ahmet Acar would like to acknowledge Republic of Türkiye The Council of Higher Education Research Universities Support Program (Grant number: ADEP-108-2022-11202).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Safiye Merve Bostancioglu, Ahmet Acar
Formal analysis: Safiye Merve Bostancioglu
Investigation: Safiye Merve Bostancioglu, Ahmet Acar
Methodology: Safiye Merve Bostancioglu, Ahmet Acar
Writing – original draft: Safiye Merve Bostancioglu, Ahmet Acar
Writing – review & editing: Safiye Merve Bostancioglu, Ahmet Acar

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Data are available from the corresponding author upon reasonable request.

References

1. Hanahan D. Hallmarks of cancer: New dimensions. *Cancer Discov*. 2022;12:31-46.
   doi: 10.1158/2159-8290.CD-21-1059
2. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*. 2011;144:646-74.
   doi: 10.1016/j.cell.2011.02.013
3. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481:306-313.
4. Gatenby RA, Brown JS. Integrating evolutionary dynamics into cancer therapy. *Nat Rev Clin Oncol*. 2020;17:675-86.
   doi: 10.1038/s41571-020-0411-1
5. Ermini L, Mallo D, Kleftogiannis D, *et al.* Editorial: Cancer evolution. *Front Genet*. 2023;14:1187687.
   doi: 10.3389/fgene.2023.1187687
6. Merlo LMF, Pepper JW, Reid BJ, *et al.* Cancer as an evolutionary and ecological process. *Nat Rev Cancer*. 2006;6:924-935.
   doi: 10.1038/nrc2013
7. Turajlic S, Sottoriva A, Graham T, *et al.* Resolving genetic heterogeneity in cancer. *Nat Rev Genet*. 2019;20:404-416.
   doi: 10.1038/s41576-019-0114-6
8. Yalcin GD, Yilmaz KC, Dilber T, *et al.* Investigation of evolutionary dynamics for drug resistance in 3D spheroid model system using cellular barcoding technology. *PLoS One*. 2023;18:e0291942.
   doi: 10.1371/journal.pone.0291942
9. Danisik N, Yilmaz KC, Acar A. Identification of collateral sensitivity and evolutionary landscape of chemotherapy-induced drug resistance using cellular barcoding technology. *Front Pharmacol*. 2023;14:1178489.
   doi: 10.3389/fphar.2023.1178489
10. Acar A, Nichol D, Fernandez-Mateos J, *et al.* Exploiting evolutionary steering to induce collateral drug sensitivity in cancer. *Nat Commun*. 2020;11:1923.
    doi: 10.1038/s41467-020-15596-z

11. Smalley KS, Lioni M, Noma K, *et al. In vitro* three-dimensional tumor microenvironment models for anticancer drug discovery. *Expert Opin Drug Discov*. 2008;3:1-10.

    doi: 10.1517/17460441.3.1.1

12. Downward J. Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer*. 2003;3:11-22.

    doi: 10.1038/nrc969

13. Kim HJ, Lee HN, Jeong MS, *et al.* Oncogenic *KRAS*: Signaling and drug resistance. *Cancers* (*Basel*). 2021;13:5599.

    doi: 10.3390/cancers13225599

14. Singh G, Thakur N, Kumar U. RAS: Circuitry and therapeutic targeting. *Cell Signal*. 2023;101:110505.

    doi: 10.1016/j.cellsig.2022.110505

15. Ferrer I, Zugazagoitia J, Herbertz S, *et al. KRAS*-Mutant non-small cell lung cancer: From biology to therapy. *Lung Cancer*. 2018;124:53-64.

    doi: 10.1016/j.lungcan.2018.07.013

16. Ciardiello D, Maiorano BA, Martinelli E. Targeting *KRAS*^{G12C} in colorectal cancer: The beginning of a new era. *ESMO Open*. 2023;8:100745.

    doi: 10.1016/j.esmoop.2022.100745

17. Naim N, Moukheiber S, Daou S, *et al. KRAS*-G12C covalent inhibitors: A game changer in the scene of cancer therapies. *Crit Rev Oncol Hematol*. 2021;168:103524.

    doi: 10.1016/j.critrevonc.2021.103524

18. Molina-Arcas M, Samani A, Downward J. Drugging the undruggable: Advances on RAS targeting in cancer. *Genes* (*Basel*). 2021;12:899.

    doi: 10.3390/genes12060899

19. Skoulidis F, Li BT, Dy GK, *et al.* Sotorasib for lung cancers with *KRAS* p.G12C mutation. *N Engl J Med*. 2021;384: 2371-2381.

    doi: 10.1056/nejmoa2103695

20. Lanman BA, Parsons AT. Sotorasib (LUMA*KRAS*), an ırreversible covalent Inhibitor of *KRAS* ^{G12C}. İn: *Current Drug Synthesis*. United States: Wiley; 2022. p. 183-199.

    doi: 10.1002/9781119847281.ch10

21. Saleh K, Kordahi M, Felefly T, *et al. KRAS*-targeted therapies in advanced solid cancers: Drug the undruggable? *Pharmacogenomics*. 2021;22:587-90.

    doi: 10.2217/pgs-2021-0045

22. Kim HJ, Lee HN, Jeong MS, *et al.* Oncogenic *KRAS*: Signaling and drug resistance. *Cancers* (*Basel*). 2021;13:5599.

    doi: 10.3390/cancers13225599

23. Oyedele AQK, Ogunlana AT, Boyenle ID, *et al.* Pharmacophoric analogs of sotorasib-entrapped *KRAS* G12C in its inactive GDP-bound conformation: Covalent docking and molecular dynamics investigations. *Mol Divers*. 2023;27:1795-1807.

    doi: 10.1007/s11030-022-10534-1

24. Mortier J, Friberg A, Badock V, *et al.* Computationally empowered workflow identifies novel covalent allosteric binders for *KRAS*^{G12C}. *ChemMedChem*. 2020;15:827-832

    doi: 10.1002/cmdc.201900727

25. Nnadi CI, Jenkins ML, Gentile DR, *et al.* Novel K-Ras G12C switch-II covalent binders destabilize ras and accelerate nucleotide exchange. *J Chem Inf Model*. 2018;58:464-471.

    doi: 10.1021/acs.jcim.7b00399

26. Boehm M, Wu TY, Haussen H, *et al.* Similarity searching and scaffold hopping in synthetically accessible combinatorial chemistry spaces. *J Med Chem*. 2008;51:2468-2480.

    doi: 10.1021/jm0707727

27. Madhavi Sastry G, Adzhigirey M, Day T, *et al.* Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 2013;27:221-34.

    doi: 10.1007/s10822-013-9644-8

28. *Schrödinger Release 2023-3: LigPre*. New York,: Schrödinger, LLC; 2023.

29. Shelley JC, Cholleti A, Frye LL, *et al.* Epik: A software program for pK(a) prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des*. 2007;21:681-691.

    doi: 10.1007/s10822-007-9133-z

30. Canon J, Rex K, Saiki AY, *et al.* The clinical *KRAS*(G12C) inhibitor AMG 510 drives Anti-tumour immunity. *Nature*. 2019;575:217-223.

    doi: 10.1038/s41586-019-1694-1

31. *Schrödinger Release 2023-3: Protein Preparation Wizard; Epik*. New York: Schrödinger, LLC; 2023.

32. Roos K, Wu C, Damm W, *et al.* OPLS3e: Extending force field coverage for drug-like small molecules. *J Chem Theory Comput*. 2019;15:1863-1874.

    doi: 10.1021/acs.jctc.8b01026

33. Zhu K, Borrelli KW, Greenwood JR, *et al.* Docking covalent inhibitors: A parameter free approach to pose prediction and scoring. *J Chem Inf Model*. 2014;54:1932-1940.

    doi: 10.1021/ci500118s

34. Friesner RA, Banks JL, Murphy RB, *et al.* Glide: A New Approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem*. 2004;47:1739-1749.

    doi: 10.1021/jm0306430

35. Halgren TA, Murphy RB, Friesner RA, *et al.* Glide: A new approach for rapid, accurate docking and scoring. 2. enrichment factors in database screening. *J Med Chem*. 2004;47:1750-1759.

doi: 10.1021/jm030644s

36. Van Der Spoel D, Lindahl E, Hess B, *et al.* GROMACS: Fast, flexible, and free. *J Comput Chem*. 2005;26:1701-1718.

doi: 10.1002/jcc.20291

37. Best RB, Zhu X, Shim J, *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ1 and χ2 Dihedral Angles. *J Chem Theory Comput*. 2012;8:3257-3273.

doi: 10.1021/ct300400x

38. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys*. 2007;126:014101.

doi: 10.1063/1.2408420

39. Berendsen HJC, Postma JPM, Van Gunsteren WF, *et al.* Molecular dynamics with coupling to an external bath. *J Chem Phys*. 1984;81:3684-3690.

doi: 10.1063/1.448118

40. Bernetti M, Bussi G. Pressure control using stochastic cell rescaling. *J Chem Phys*. 2020;153:114107.

doi: 10.1063/5.0020514

41. Darden T, York D, Pedersen L. Particle mesh ewald: An N·log(N) method for ewald sums in large systems. *J Chem Phys*. 1993;98:10089-100920.

doi: 10.1063/1.464397

42. Hess B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J Chem Theory Comput*. 2008;4: 116-122.

doi: 10.1021/ct700200b

43. Turner P. *XMGRACE, Version 5.1. 19. Center for Coastal and Land-Margin Research*. Beaverton, OR: Oregon Graduate Institute of Science and Technology; 2005.

44. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph*. 1996;14:33-8.

doi: 10.1016/0263-7855(96)00018-5

45. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LS. Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics*. 2006;22:2695-2696.

doi: 10.1093/bioinformatics/btl461

46. Scarpino A, Ferenczy GG, Keserü GM. Comparative evaluation of covalent docking tools. *J Chem Inf Model*. 2018;58:1441-1458.

doi: 10.1021/acs.jcim.8b00228

47. Pantsar T. *KRAS*(G12C)-AMG 510 interaction dynamics revealed by all-atom molecular dynamics simulations. *Sci Rep*. 2020;10:11992.

doi: 10.1038/s41598-020-68950-y

48. Milburn MV, Tong L, DeVos AM, *et al.* Molecular switch for signal transduction: Structural differences between active and inactive forms of protooncogenic ras proteins. *Science (1979)*. 1990;247:939-945.

doi: 10.1126/science.2406906

49. Li Y, Han L, Zhang Z. Understanding the influence of AMG 510 on the structure of *KRAS*G12C empowered by molecular dynamics simulation. *Comput Struct Biotechnol J*. 2022;20:1056-1067

doi: 10.1016/j.csbj.2022.02.018

Appendices

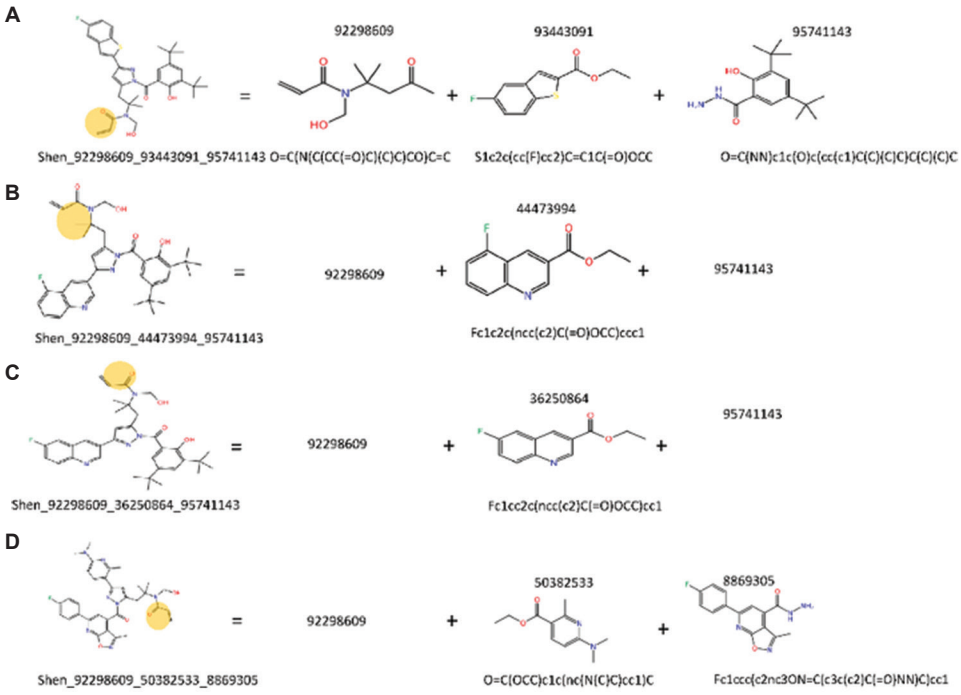


Figure A1. The acrylamide warheads and building blocks of compounds. (A) C01, (B) C02a, (C) C02b, and (D) C03.

Table A1. 174 compounds that have structural similarities to Sotorasib

No	Name
1	rxn501b____EMOL916035____EMOL4353066____EMOL45478818
2	rxn501b____EMOL485662____EMOL4353066____EMOL45478818
3	rxn508____EMOL916035____EMOL37154489____EMOL724744
4	rxn508____EMOL916035____EMOL37154489____EMOL4854098
5	rxn508____EMOL916035____EMOL37154489____EMOL300131641
6	rxn508____EMOL916035____EMOL37154489____EMOL9779925
7	rxn508____EMOL916035____EMOL37154489____EMOL2863515
8	rxn501b____EMOL882112____EMOL4353066____EMOL45478818
9	rxn508____EMOL916035____EMOL50157604____EMOL724744
10	rxn508____EMOL916035____EMOL2862529____EMOL4854098
11	rxn508____EMOL916035____EMOL50157604____EMOL2863515
12	rxn508____EMOL916035____EMOL50157604____EMOL300131641
13	rxn508____EMOL916035____EMOL2862529____EMOL300131641
14	rxn508____EMOL916035____EMOL50157604____EMOL9779925
15	rxn508____EMOL916035____EMOL2862529____EMOL2863515
16	rxn508____EMOL916035____EMOL50157604____EMOL4854098
17	rxn508____EMOL916035____EMOL2862529____EMOL9779925
18	rxn508____EMOL916035____EMOL2862529____EMOL724744
19	rxn508____EMOL916035____EMOL37154489____EMOL30070316
20	rxn508____EMOL916035____EMOL37154489____EMOL36725275
21	rxn508____EMOL916035____EMOL50157604____EMOL36725275
22	rxn508____EMOL916035____EMOL2862529____EMOL36725275
23	rxn508____EMOL916035____EMOL50157604____EMOL30070316
24	rxn508____EMOL916035____EMOL2862529____EMOL30070316
25	rxn501b____EMOL916035____EMOL4353066____EMOL48597142
26	rxn501b____EMOL916035____EMOL4353066____EMOL45478663
27	rxn508____EMOL916035____EMOL37154489____EMOL44474549
28	rxn508____EMOL916035____EMOL50157604____EMOL44474549
29	rxn508____EMOL916035____EMOL2862529____EMOL44474549
30	rxn508____EMOL916035____EMOL37154489____EMOL43633603
31	rxn508____EMOL916035____EMOL2862529____EMOL43633603
32	rxn508____EMOL916035____EMOL50157604____EMOL43633603
33	rxn508____EMOL916035____EMOL37154489____EMOL49844674
34	rxn508____EMOL916035____EMOL37154489____EMOL1522728
35	rxn508____EMOL916035____EMOL37154489____EMOL11554196
36	rxn508____EMOL916035____EMOL37154489____EMOL1885674
37	rxn508____EMOL916035____EMOL37154489____EMOL42785649
38	rxn508____EMOL916035____EMOL37154489____EMOL11488490
39	rxn508____EMOL916035____EMOL89945076____EMOL724744

(Cont'd...)

Table A1. (*Continued*)

No	Name
40	rxn508____EMOL916035____EMOL89945076____EMOL9779925
41	rxn508____EMOL916035____EMOL89945076____EMOL4854098
42	rxn508____EMOL916035____EMOL89945076____EMOL2863515
43	rxn508____EMOL916035____EMOL89945076____EMOL300131641
44	rxn508____EMOL916035____EMOL37154489____EMOL904614
45	rxn508____EMOL916035____EMOL2862529____EMOL42785649
46	rxn508____EMOL916035____EMOL50157604____EMOL49844674
47	rxn508____EMOL916035____EMOL50157604____EMOL1885674
48	rxn508____EMOL916035____EMOL2862529____EMOL11488490
49	rxn508____EMOL916035____EMOL50157604____EMOL42785649
50	rxn508____EMOL916035____EMOL50157604____EMOL11488490
51	rxn508____EMOL916035____EMOL2862529____EMOL49844674
52	rxn508____EMOL916035____EMOL2862529____EMOL1885674
53	rxn508____EMOL916035____EMOL2862529____EMOL1522728
54	rxn508____EMOL916035____EMOL50157604____EMOL11554196
55	rxn508____EMOL916035____EMOL2862529____EMOL11554196
56	rxn508____EMOL916035____EMOL50157604____EMOL1522728
57	rxn508____EMOL916035____EMOL50157604____EMOL904614
58	rxn508____EMOL916035____EMOL2862529____EMOL904614
59	rxn502b____EMOL916035____EMOL4353066____EMOL2251614
60	rxn508____EMOL916035____EMOL89945076____EMOL30070316
61	rxn508____EMOL916035____EMOL89945076____EMOL36725275
62	rxn508____EMOL485662____EMOL37154489____EMOL2863515
63	rxn508____EMOL485662____EMOL37154489____EMOL4854098
64	rxn508____EMOL485662____EMOL37154489____EMOL300131641
65	rxn508____EMOL485662____EMOL37154489____EMOL724744
66	rxn508____EMOL485662____EMOL37154489____EMOL9779925
67	rxn508____EMOL882112____EMOL37154489____EMOL2863515
68	rxn508____EMOL882112____EMOL37154489____EMOL724744
69	rxn508____EMOL882112____EMOL37154489____EMOL4854098
70	rxn508____EMOL882112____EMOL37154489____EMOL300131641
71	rxn508____EMOL882112____EMOL37154489____EMOL9779925
72	rxn508____EMOL485662____EMOL50157604____EMOL300131641
73	CSSS00161188889
74	WXDL_15A____SB1376____LN00004982____BT6378
75	Shen____92298609____48620564____44152703
76	Shen____92298609____48620700____44152703
77	Shen____92298609____93443091____44152703
78	Shen____92298609____48552545____29552264

(Cont'd...)

Table A1. (*Continued*)

No	Name
79	Shen____92298609____106083795____44152703
80	Shen____92298609____106083765____44152703
81	Shen____92298609____49239582____44152703
82	Shen____92298609____49239580____44152703
83	Shen____92298609____49239556____44152703
84	Shen____92298609____36250864____44152703
85	Shen____92298609____44473994____44152703
86	Shen____92298609____44473403____44152703
87	Shen____92298609____44473381____44152703
88	Shen____92298609____48573899____44152703
89	Shen____92298609____48552545____107053254
90	Shen____92298609____48552545____81813113
91	Shen____92298609____32456150____44152703
92	Shen____92298609____36745965____44152703
93	Shen____92298609____830425____44152703
94	Shen____92298609____42824050____44152703
95	Shen____92298609____42817358____44152703
96	Shen____92298609____25041346____44152703
97	Shen____92298609____97662126____29552264
98	Shen____92298609____36249205____29552264
99	Shen____92298609____48552545____8869305
100	Shen____92298609____36249205____107053254
101	Shen____92298609____97662126____107053254
102	Shen____92298609____56970423____44152703
103	Shen____92298609____49823880____44152703
104	Shen____92298609____36249205____81813113
105	Shen____92298609____97662126____81813113
106	Shen____92298609____50382533____29552264
107	Gray____107373053____6886343____107509511
108	Shen____92298609____53885480____44152703
109	Shen____92298609____53885502____44152703
110	Shen____92298609____44473387____44152703
111	Shen____92298609____53885486____44152703
112	Shen____92298609____42816206____44152703
113	Shen____92298609____36828355____44152703
114	Shen____92298609____53890255____44152703
115	Shen____92298609____50382533____107053254
116	Shen____92298609____50272909____29552264
117	Shen____92298609____50382533____81813113

(Cont'd...)

Table A1. (*Continued*)

No	Name
118	Shen____92298609____38501832____44152703
119	Shen____92298609____44848277____44152703
120	Shen____92298609____12814542____44152703
121	Shen____92298609____980919____44152703
122	Shen____92298609____44391072____44152703
123	Shen____92298609____520177____44152703
124	Shen____92298609____50272909____107053254
125	Shen____92298609____106068876____44152703
126	Shen____92298609____106068894____44152703
127	Shen____92298609____81802460____44152703
128	Shen____92298609____43045611____44152703
129	Shen____92298609____50272909____81813113
130	Shen____92298609____48620700____95741143
131	Shen____92298609____48620564____95741143
132	Shen____92298609____48552545____42793511
133	Shen____92298609____48620564____96426646
134	Shen____92298609____48620700____96426646
135	Shen____92298609____36249205____8869305
136	Shen____92298609____97662126____8869305
137	Shen____92298609____81802090____29552264
138	Shen____92298609____76750201____29552264
139	Shen____92298609____106978508____29552264
140	Shen____92298609____95759057____29552264
141	Shen____92298609____48552545____50495850
142	Shen____92298609____93443091____95741143
143	Shen____92298609____93443091____96426646
144	Shen____92298609____106083765____95741143
145	Shen____92298609____106083795____95741143
146	Shen____92298609____106083795____96426646
147	Shen____92298609____106083765____96426646
148	Shen____92298609____81802090____107053254
149	Shen____92298609____106978508____107053254
150	Shen____92298609____76750201____107053254
151	Shen____92298609____49253559____44152703
152	Shen____92298609____1451473____44152703
153	Shen____92298609____49253379____44152703
154	Shen____92298609____48552545____2066545
155	Shen____92298609____95759057____107053254
156	Shen____92298609____50382533____8869305

(Cont'd...)

Table A1. (*Continued*)

No	Name
157	Shen____92298609____48552545____50658478
158	Shen____92298609____49239556____95741143
159	Shen____92298609____49239582____95741143
160	Shen____92298609____49239580____95741143
161	Shen____92298609____44473381____95741143
162	Shen____92298609____44473403____95741143
163	Shen____92298609____36250864____95741143
164	Shen____92298609____44473994____95741143
165	Shen____92298609____48573899____95741143
166	Shen____92298609____76750201____81813113
167	Shen____92298609____81802090____81813113
168	Shen____92298609____106978508____81813113
169	Shen____92298609____95759057____81813113
170	Shen____92298609____49239556____96426646
171	Shen____92298609____49239580____96426646
172	Shen____92298609____49239582____96426646
173	Shen____92298609____84829545____44152703
174	Shen____92298609____44473389____44152703