ISTANBUL MEDENIYET UNIVERSITY
INSTITUE OF GRADUATE STUDIES
DEPARTMENT OF BIOENGINEERING

Target-Based Inhibitor Discovery: Computational Strategies and Analog Inhibitors against KRAS Mutants

Master's Thesis

Ecem Turhan

January 2026

ISTANBUL MEDENIYET UNIVERSITY
INSTITUE OF GRADUATE STUDIES
DEPARTMENT OF BIOENGINEERING

Target-Based Inhibitor Discovery: Computational Strategies and Analog Inhibitors against KRAS Mutants

Master's Thesis

Ecem Turhan

Supervisor

Assoc. Prof. Saliha Ece Acuner Zorluuysal

January 2026

THESIS JURY APPROVAL

This PhD thesis titled "Target-Based Inhibitor Discovery: Computational Strategies and Analog Inhibitors against KRAS Mutants" written by Ecem Turhan at the Department of Bioengineering was accepted by our jury.

JURY MEMBERS **SIGNATURE**

Supervisor:

Assoc. Prof. Saliha Ece Acuner Zorluuysal

Istanbul Medeniyet University

Other Jury Members:

Prof. Dr. Pemra Özbek Sarıca

Marmara University

Assoc. Prof. Elif Kubat Öktem

Istanbul Medeniyet University

Date of Defense: 12/01/2026

STATEMENTS

Style and Reference Manual Statement

Having reviewed this thesis written under my supervision, I confirm that it has been written in accordance with APA Manual of Style and used its in text reference format consistently throughout the entire text.

Signature

Assoc. Prof. Saliha Ece Acuner Zorluuysal

Declaration of Originality

I hereby declare that all information in this dissertation has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conducts, I have fully cited and referenced all material and results that are not original to this work.

Signature

Ecem Turhan

GENİŞ ÖZET

Hedefe Yönelik İnhibitör Keşfi: Mutant KRAS'lara Yönelik Hesaplamalı Stratejiler ve Analog İnhibitörler

Turhan, Ecem

Yüksek Lisans Tezi, Biyomühendislik Anabilim Dalı

Danışman: Doç. Dr. Saliha Ece Acuner Zorluuysal

Ocak 2026

Yaşamın temel birimleri olan hücreler, canlılıklarını, işlevselliklerini sağlayabilmek ve vücut içerisindeki görevlerini yerine getirebilmek için oldukça organize sinyal ağlarına sahiptirler. Bu organize sinyal ağları aracılığıyla hücreler, gerekli durumlarda dış uyaranlara yanıt verebilir, büyüyebilir, farklılaşabilir ve programlanmış hücre ölümünü (apoptoz) gerçekleştirebilirler. Ras protein ailesi, hücrenin kilit moleküler düzenleyicilerinden olan küçük GTP (Guanozin trifosfat) bağlayıcı proteinlerin bir alt ailesidir ve hücre içi sinyal yollarında merkezi bir görev almaktadır. Ras proteinleri, hücre iskeleti organizasyonu, hücre içi sinyal iletimi, gen ekspresyon düzenlenmesi, hücre döngüsünün ilerlemesi, çoğalma ve farklılaşma dahil olmak üzere çeşitli hücresel fonksiyonlarda rol oynamaktadır.

Ras proteinleri, aktif ve inaktif durumlar arasında geçiş yapan moleküler anahtarlar olarak işlev görmektedir. İnaktif formda Ras, GDP ile (Guanozin difosfat) bağlıyken, aktif formda GTP ile (Guanozin trifosfat) bağlıdır. Bu nükleotid değişim döngüsü, GDP'nin GTP ile değişimine yardımcı olan guanin nükleotid değişim faktörleri (GEF'ler) ve GTP hidrolizini hızlandıran GTPaz aktive edici proteinler (GAP'ler) tarafından sıkı bir şekilde düzenlenmektedir. Ras, GTP nükleotidi ile bağlanıp aktif duruma geçtikten sonra çeşitli efektör proteinlerle (Raf, PI3K, RALGDS) etkileşime girmektedir ve sinyalleri aşağı akış yoluna iletmektedir. Böylelikle Ras, büyümeyi düzenleyen hayati hücresel yolları kontrol eder. Dolayısıyla Ras'ın aktif/inaktif döngüsündeki bir bozulma, yalnızca tek bir enzimin problemi olmanın ötesine geçerek karmaşık bir ağ bozulması şeklinde değerlendirilmelidir.

Ras protein ailesi, üç homolog izoform gen tarafından kodlanan proteinlerden oluşmaktadır: bunlar, KRAS (Kirsten Rat Sarcoma Virus), HRAS (Harvey Rat Sarcoma Virus) ve NRAS (Neuroblastoma RAS Viral Oncogene) genleridir. KRAS izoformunun, KRAS4A ve KRAS4B olmak üzere iki ekleme varyantı vardır. GTP bağlanma bölgesini içeren G-alanının (G-domain) ilk 80 amino asidi, tüm Ras homolog proteinlerinde neredeyse aynıdır ve bu durum da korunmuş

bir işlevsel çekirdeğe işaret etmektedir. Aynı zamanda, bu homolog proteinlerin arasındaki C-terminal bölgelerinin farklılıkları, membran lokalizasyonu ve post-translasyonel modifikasyonlar gibi belirgin fonksiyonel çeşitlilik oluşturmaktadır.

Yapısal çalışmalar, G-alanının GDP-GTP değişimi sırasında önemli konformasyonel değişikliklere uğradığını ve MAPK (Mitojenle Aktifleştirilen Protein Kinaz), PI3K (Fosfoinozitid 3-Kinaz), Raf (Rapidly Accelerated Fibrosarcoma) ve Ral (Ras-like) gibi spesifik efektör proteinlerin bağlanıp çoğalma sinyallerini iletmesine olanak sağladığını göstermektedir.

Ras proteini yapısal olarak iki ana bölgeye ayrılmaktadır: Aktif (anahtar) bölge ve Allosterik bölge. Ras, yapısal aktivitesini belirleyen birkaç temel yapısal motif içermektedir: P-döngüsü (Fosfat bağlama döngüsü, aa 10-17), Anahtar I bölgesi (aa 30-40) ve Anahtar II bölgesi (aa 58-72). Bu bölgeler, GTP veya GDP'nin bağlı olup olmamasına göre konformasyonel değişikliklere uğramaktadır. Özellikle Anahtar I ve II bölgeleri, efektör proteinler için ana bağlanma cebini oluşturmaktadır ve aynı zamanda küçük molekül inhibitörleri için birincil hedef bölgelerdir. Allosterik bölgeler ise (özellikle aa 87-171), çekirdek yapıyı stabilize eder, GTPaz aktivitesinin düzenlenmesinde kritik bir rol üstlenir ve konformasyonel dinamikleri etkileyerek efektör etkileşimlerinin sürekliliğini belirlemektedir. Bu bölgeler, GTP hidrolizinin katalitik çevresini dolaylı olarak şekillendirmektedir ve aynı zamanda allosterik inhibitörler için potansiyel hedef bölgeleri oluşturmaktadır. Son yıllarda X-ışını kristalografisi, NMR ve kriyo-elektron mikroskobi gibi tekniklerin sağladığı yüksek çözünürlüklü yapısal veriler, bu allosterik bölgelerde geçici olarak oluşan veya yalnızca belirli konformasyonlarda ortaya çıkan "geçici ceplerin" de ilaç tasarımı açısından değerlendirilebileceğini göstermektedir. Böylece Ras proteini, yalnızca klasik nükleotid bağlanma bölgesine sahip küçük bir GTPaz olmaktan çıkıp, yapısı ilaç tasarımıyla dinamik bir hedef olarak yeniden tanımlanmaya başlanmıştır.

Normal fizyolojik koşullar altında Ras aktivasyonu geçicidir ve GEF ile GAP proteinleri tarafından sıkı bir şekilde düzenlenmektedir. Ancak, Ras genindeki nokta mutasyonları, proteini GTP'ye bağlı, sürekli aktif formunda kilitleyebilir. Bu durum, MAPK ve PI3K gibi yollar aracılığıyla sürekli aşağı akış sinyalizasyonuna yol açarak kontrolsüz hücre büyümesi sonuçlu tümöre neden olur. Raf protein ailesi (ARAF, BRAF, CRAF/RAF1), bir serin/treonin kinaz ailesidir. Raf proteinin N-terminalinde Ras bağlanma bölgesi (RBD) mevcuttur. Ras bağlanması ile Raf'ı plazma membranına toplar, Raf'ın konformasyonunu değiştirerek dimerleşmesini ve fosforilasyonunu kolaylaştırır. Raf proteini de bu sayede MEK1/2 (MAP2K1/2) - ERK1/2 (MAPK1/3) hattını durmadan fosforlamaktadır. PI3K kinazının ise p110 alt biriminde RBD mevcuttur. PI3K

kinazının sürekli aktif edilmesiyle PIP3 (Phosphatidylinositol (3,4,5)-trisphosphate) düzeyi yükselir ve AKT proteini (Protein Kinase B; PKB) etkilenerek membrana gelir. PDK1 (3-phosphoinositide-dependent protein kinase-1) ve mTORC2 (Mechanistic Target of Rapamycin Complex 2) tarafından fosforlanıp aktifleşerek apoptozu baskılar ve hücre büyümesini artırmaktadır.

Kapsamlı genomik analizler, Ras gen ailesindeki mutasyonların tüm insan kanserlerinin yaklaşık %30'unda görüldüğünü ve bu nedenle Ras'ın en sık mutasyona uğrayan onkogenlerden biri olduğunu ortaya koymaktadır. Bu ailede yer alan başlıca üç homolog gen (KRAS, NRAS ve HRAS) onkojenik dönüşümde farklı roller üstlenmektedir. Bunlar arasında KRAS mutasyonları, Ras kaynaklı mutasyonların yaklaşık %85'ini oluşturur ve özellikle pankreas adenokarsinomlarının %90'ından fazlasında, kolorektal kanserlerin yaklaşık %40'ından ve akciğer adenokarsinomlarının yaklaşık %30'undan sorumludur. NRAS mutasyonları ise toplam Ras mutasyonlarının yaklaşık %10-15'ini oluşturur. Bu mutasyonlar özellikle melanom, akut miyeloid lösemi (AML), akut lenfoblastik lösemi (ALL) ve bazı tiroid kanseri alt tiplerinde daha sık gözlenir. HRAS mutasyonları ise diğer homolog mutasyonlarına kıyasla yaklaşık %1-5'ini oluşturarak daha nadirdir, ancak mesane kanseri, baş-boyun skuamöz hücreli karsinomları ve servikal kanser gibi belirli tümörlerde anlamlı bir şekilde rapor edilmektedir.

En yaygın nokta mutasyonlar, Ras'ın GTPaz aktivitesini bozan ve inaktivasyonunu engelleyen 12., 13. ve 61. kodonlarda meydana gelmektedir. Bu mutasyonlar GTP'nin hidrolizini bloke ederek proteini sürekli aktif GTP-bağlı durumda tutar ve böylece hücresel proliferasyon, farklılaşma ve hayatta kalma sinyallerini kesintisiz hale getirir. KRAS'ta özellikle G12D, G12V, G12C gibi mutasyonlar öne çıkarken, NRAS'ta Q61R ve Q61K, HRAS'ta ise G12S ve G13R mutasyonları sık görülmektedir. Çeşitli KRAS mutantları arasında, yüksek sıklıkları ve klinik önemleri nedeniyle en çok G12D ve G12C mutasyonları ön plandadır. KRAS G12D mutasyonu, 12. pozisyonda glisin amino asidinin aspartik asit amino asidi ile yer değiştirmesinden kaynaklanır ve bu mutasyon tipi pankreas duktal adenokarsinom (PDAC) vakalarının yaklaşık %45'inde görülmektedir. Öte yandan, glisin amino asidinin sistein amino asidi ile yer değiştirdiği KRAS G12C mutasyonları, akciğer adenokarsinomlarında daha yaygın görülmektedir.

KRAS'ın G12C mutantını hedef alan kovalent inhibitörlerin keşfi, onkolojide büyük bir gelişmeyi temsil etmektedir. Bu kovalent inhibitörler, Switch II cebinin yakınında bulunan 12. pozisyondaki sistein rezidüsüne kovalent olarak bağlanarak KRAS'ı inaktif GDP-bağlı durumunda kilitler. Kovalent inhibitör örnekleri arasında, önemli klinik etkinlik gösteren ve hedefli KRAS tedavilerinin

önünü açan Sotorasib (AMG 510), Adagrasib (MRTX 849) ve ARS-1620 inhibitörleri bulunmaktadır. Ancak, KRAS G12D için inhibitörlerin geliştirilmesi çok daha zorlu olmuştur. G12D mutasyonunda reaktif bir sistein rezidüsü bulunmadığından, geleneksel kovalent inhibitör stratejileri uygulanamamaktadır. Glisinin aspartik asitle yer değiştirmesi, potansiyel bir kovalent bağlanma bölgesini ortadan kaldırmaktadır ve aspartik asidin karboksilat yan zinciri, nükleofilik saldırı için kimyasal olarak uygun değildir. Bu nedenle, KRAS G12D'nin inaktif formuna seçici olarak bağlanıp stabilize edebilen kovalent olmayan inhibitörlerin tasarımı, son araştırmaların odak noktası haline gelmiştir. Son çalışmalar, KRAS'ın G12D mutantını hedef alan birkaç umut verici kovalent olmayan inhibitör tanımlamıştır. MRTX-1133, TH-Z827 ve TH-Z835 gibi kovalent olmayan moleküllerin, KRAS'ın inaktif GDP-bağlı formunu stabilize ederek ve Raf, PI3K gibi aşağı akış sinyal yolu efektörlerinin aktivasyonunu önleyerek KRAS'ı inhibe ettiği gösterilmektedir. Bu bileşikler, Anahtar II cebini yüksek afiniteyle işgal etmektedir. Sinyal yayılımı için gerekli etkileşimleri yer doldurarak önlemektedir ve böylece tümör hücresi çoğalmasını baskılamaktadır.

Kovalent ve kovalent olmayan inhibitörlerin geliştirilmesine rağmen KRAS zorlu bir ilaç hedefi olmaya devam etmektedir çünkü Ras proteinin yüzeyinde küçük molekülleri barındırabilecek derin hidrofobik ceplerin bulunmaması nedeniyle, Ras onlarca yıl boyunca "ilaçla tedavi edilemez" olarak sınıflandırılmaktaydı. Aynı zamanda, Ras, GTP ve GDP'ye pikomolar afiniteyle bağlanarak potansiyel inhibitörlerin etkili bir şekilde rekabet etmesini zorlaştırmaktadır. Bütün bunlara ek olarak, birçok hastada tedaviye başlangıçta iyi yanıt alınsa da zamanla kazanılmış direnç mekanizmaları gelişmekte ve hastalığın yeniden ilerlemesi gözlenmektedir. Benzer şekilde, son yıllarda KRAS G12D mutasyonuna yönelik inhibitörlerin geliştirilmesinde önemli ilerlemeler kaydedilse de bu varyant için onaylı bir klinik ilaç henüz bulunmamaktadır. G12D inhibitör adaylarının preklinik ve erken faz klinik çalışmalarda umut verici sonuçlar göstermesine rağmen, direnç gelişimi riski bu mutasyon için de önemli bir endişe kaynağıdır. Dolayısıyla, KRAS mutasyonlarını hedefleyen mevcut tedavilerin ötesine geçilmesi, kombinasyon tedavi stratejilerinin geliştirilmesi, aktif veya yeni bağlanma bölgelerini hedefleyen alternatif inhibitörlerin tasarlanması ve allosterik mekanizmaları kullanan ilaç adaylarının klinik değerlendirmeye alınması büyük önem taşımaktadır. Bu yaklaşım, KRAS mutant tümörlerde daha kalıcı ve direnç gelişimine karşı daha dayanıklı yanıtların elde edilmesi açısından kritik bir adım olarak görülmektedir. Hesaplamalı ilaç keşfi, yapı tabanlı tasarım ve moleküler dinamik simülasyonlarındaki gelişmeler, Ras inhibitörleri arayışına metodsal yenilik getirmektedir. Bu hesaplamalı yaklaşımlar, geçici

bağlanma ceplerini belirlenmesine ve proteinin dinamik doğasından yararlanan ligandlar tasarlanmasına olanak tanımaktadır.

Bu tezin amacı, KRAS G12D ve G12C mutant varyantlarını hedef alan yeni aday inhibitörleri hesaplamalı yaklaşımlarla belirlemektir. Bu amaçla, çalışma grubumuz bünyesinde 2023 yılında yayınlanan 813197 numaralı yüksek lisans tez çalışmasında, G12D mutasyonu için küçük moleküllerden oluşan bir sanal kütüphane, en yüksek öngörülen bağlanma afinitesine sahip bileşikleri belirlemek için sanal tarama teknikleri kullanılarak KNIME programı ile taranmıştı. Önce ZINC20 veritabanı kullanılarak bir molekül kütüphanesi oluşturulmuş ve bu kütüphane KNIME programında Lipinski'nin Beş Kuralı ile filtrelenerek kurala uymayan moleküller elenmişti. Literatürde bildirilen bilinen inhibitörler pozitif kontrol inhibitörleri olarak tercih edilmiş ve bu moleküllere yapısal benzerlik gösteren ve Lipinski kuralına uyan moleküller, kovalent olmayan kenetlenme çalışmaları için AutoDock Vina yazılımı kullanılarak daha ileri analizlere tabi tutulmuştu. Sanal kütüphanenin oluşturulması aşamasında, ilaç benzeri özelliklere sahip bileşiklerin seçilebilmesi için Lipinski kuralları ve ilgili diğer farmakokinetik filtreler dikkate alınmıştı. Böylece tarama adımı, yalnızca teorik olarak güçlü bağlanma afinitesi sunan değil, aynı zamanda potansiyel klinik aday olma niteliği taşıyan molekülleri ön plana çıkarmayı hedeflemişti.

Sanal tarama aşaması sonrasında, 2023 yılında çalışma grubumuzdan yayınlanan 813197 numaralı tezin devamı olarak, bu tez kapsamında en iyi kenetlenme skoru veren molekülün, moleküler dinamik simülasyonlarının ve enerji hesaplamalarının gerçekleştirilmesi amaçlandı. Kenetlenme aşamalarının ardından, G12D mutasyonu için seçilen umut verici ligand-protein kompleksi, zaman içindeki kararlılığını, bağlanma etkileşimlerini, davranışlarını ve konformasyonel esnekliklerini değerlendirmek için moleküler dinamik (MD) simülasyonlarına tabi tutuldu. Bu hesaplamalı analizler, proteinin aktif bölgesindeki hidrojen bağlarının, hidrofobik temasların kalıcılığının yanı sıra protein-ligand kompleksinin genel yapısal stabilitesi hakkında değerli bilgiler sağladı. Moleküler dinamik (MD) simülasyonların sonucunda daha ileri hesaplamalı analizler için serbest enerji hesaplamaları ve çapraz korelasyon analizleri gerçekleştirildi. Aynı zamanda yüksek bağlanma afinitesine sahip analog alternatif moleküller geliştirmek için, aday bileşiğin fizikokimyasal özellikleri ve hedef bağlanma cebi ile etkileşimleri değerlendirildi. Bu değerlendirme, hidrojen bağı potansiyeli, hidrofobik etkileşimler ve ligand ile KRAS G12D proteininin bağlanma bölgesi arasındaki elektrostatik uyumluluk gibi temel faktörlerin değerlendirilmesini içermektedir. Bu analizlerin sonuçlarına dayanarak, aday bileşiğin moleküler yapısını optimize etmek için

rasyonel fonksiyonel grup modifikasyonları uygulandı. Bu modifikasyonlar uygulanırken 3 farklı yöntem; ChEMBL veritabanında mevcut deneysel olarak denenmiş modifikasyon kuralları, Tanimoto metodu ile benzerlik araştırması ve kullanılan aracın 1,4 milyon kimyasal kuralın öğretildiği yapay zekasının ürettiği moleküller kullanıldı. Aktif bölge ve allosterik bölgeye göre molekül üzerinde farklı fonksiyonel gruplar modifiye edildi. İki ayrı bağlanma bölgesine göre oluşturulan bu analog moleküller, kenetlenme aşaması gerçekleştirilerek protein ile bağlanma afinitileri değerlendirildi. Molekül kütüphanelerinde en iyi ilk 10 skoru veren moleküller bir sonraki aşamaya taşındı. Modifiye edilmiş moleküller daha sonra, kararlılıklarını, bağlanma davranışlarını ve zaman içindeki potansiyel konformasyonel değişimlerini incelemek ve inhibitör etkinlikleri hakkında daha derinlemesine bilgi alabilmek için KRAS G12D yapısıyla kenetlenme analizi ve moleküler dinamik simülasyonlarına tabi tutuldu. Moleküler dinamik simülasyonlar sonucunda simülasyon boyunca aktif bölgeye bağlı kalan analog moleküllerin enerji hesaplamaları gerçekleştirildi.

KRAS G12C mutasyonu için yeni kovalent inhibitörlerinin keşfedilmesi için sanal tarama aşamaları KNIME programında gerçekleştirildi. ZINC20 veritabanında molekül kütüphanesi oluşturulurken reaktif moleküller tercih edildi. Pozitif kontrol inhibitörleri olarak Sotorasib, Adagrasib ve ARS-1620 inhibitörleri seçildi ve Cresset Flare programı ile kovalent kenetlenme yöntemi uygulandı. En iyi skoru veren moleküller, proteinin bağlanma bölgesindeki etkileşime girdiği rezidülere göre değerlendirildi.

Bu tez çalışmasının genel amacı, KRAS mutantları ile potansiyel inhibitörler arasındaki moleküler etkileşimleri açıklığa kavuşturarak hedef bazlı tedavi stratejilerinin geliştirilmesine katkıda bulunmaktır. Hesaplamalı modellemeyi moleküler yerleştirme ve dinamiklerle bütünleştirerek, bu araştırma gelecekteki deneysel doğrulamalar için temel olabilecek ilaç benzeri molekülleri belirlemeyi amaçlamaktadır. Ayrıca, bu çalışma, modern tıbbi kimyada hesaplamalı ilaç keşif metodolojilerinin artan önemini vurgulamaktadır. Hesaplamalı araçlar, erken aşama ilaç tasarımında avantajlı yaklaşımlardır ve geleneksel yüksek verimli tarama yöntemlerine kıyasla hem zamandan hem de maliyetten önemli ölçüde tasarruf sağlamaktadır. Sanal tarama, kenetlenme ve moleküler dinamik simülasyonlarının birleşimi sayesinde araştırmacılar, bağlanma afinitelerini tahmin edebilir, moleküler etkileşimleri optimize edebilir ve biyolojik testler için en umut verici adaylara öncelik verebilirler.

Özetle, Ras protein ailesi, özellikle KRAS, birçok insan kanserinde önemli bir onkojenik etkendir. Bu proje, potansiyel inhibitör adaylarını belirlemek ve analiz etmek için hesaplamalı yaklaşımlar kullanarak KRAS'ı hedeflemeyle ilgili zorlukların üstesinden gelmeye odaklanmaktadır. Sonuçların, KRAS

inhibisyonunun yapısal temelinin daha iyi anlaşılmasına katkıda bulunması ve mutant KRAS izoformlarını seçici olarak hedeflemeyi amaçlayan yeni tedavi stratejilerinin geliştirilmesini desteklemesi beklenmektedir.

Anahtar Kelimeler: KRAS G12D, KRAS G12C, Sanal Tarama, Moleküler Dinamik Simülasyonlar, Kanser

ABSTRACT

Target-Based Inhibitor Discovery: Computational Strategies and Analog Inhibitors against KRAS Mutants

Turhan, Ecem

Master's Thesis, Department of Bioengineering

Supervisor: Assoc. Prof. Saliha Ece Acuner Zorluuysal

January 2026

The Ras protein family, encoded by NRAS, HRAS, and KRAS homolog genes, comprises small GTP-binding proteins that function as switches, alternating between active GTP-bound and inactive GDP-bound states to regulate signaling, cell division, and proliferation. Their conserved G-domain, the first 80 residues of which are highly conserved across homologs, contains P-loop (aa 10-17), Switch I (aa 30-40), and Switch II (aa 58–72) motifs. Oncogenic mutations, particularly at codons 12, 13, and 61, hold Ras in the active conformation and maintain constant downstream signaling. RAS mutations are found in about 30% of human cancers, with KRAS representing 85% of them and causing colorectal, pancreatic, and lung adenocarcinomas. Although RAS was considered "undruggable," the presence of a pocket between Switch I/II has enabled direct inhibition; covalently binding small molecules for KRAS G12C (Sotorasib and Adagrasib) and non-covalent for G12D (MRTX-1133) bind to the pocket near the mutated 12th amino acid and stabilize the inactive conformation. Patients developed resistance to known inhibitors over time, and molecules targeting G12C mutant cannot benefit patients with G12D mutation, leading to a search for new inhibitors. The aim of this thesis is to discover a new candidate drugs for the G12C mutant using virtual screening and covalent docking, and to examine binding behavior of the molecule discovered in thesis number 813197 for the G12D mutant, by performing molecular dynamics simulations and energy calculations. Furthermore, the aim is to create more stable analogs of the G12D mutant inhibitor by modification. This work aims to guide future experimental validation efforts.

Keywords: KRAS G12D, KRAS G12C, Virtual Screening, Molecular Dynamics Simulations, Cancer

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my dear advisor, Assoc. Prof. Dr. Saliha Ece Acuner Zorluuysal, for generously sharing her valuable knowledge, insights, and time throughout this study. I feel truly fortunate to have been her student, not only because of her exceptional academic guidance but also for the kindness, care, and support she has shown like a mother.

I would also like to thank my professors at Istanbul Medeniyet University and all the SEALab members, especially Hacı Aslan Onur İşcil, and Damla Nur Çamlı, whom I feel very fortunate to have met through this opportunity. I am deeply grateful to them for embracing me like their youngest sibling and for brightening my life.

I extend my deepest gratitude to my beloved husband, Ufuk Turhan, whose unconditional love I have felt at every moment of my life. Since our childhood, almost like a scene from a film, we unknowingly walked through the same schools, gathered with friends in the same places, and perhaps passed each other countless times. At some point, life brought us to the same place at the same time, and we finally met; I am profoundly grateful for that moment when our paths crossed and for all the coincidences that made today possible.

I am deeply grateful to my family, especially my mother Özlem Akyar and my brother Alp Ömer Bulut, for always making me feel their unwavering support and belief in me. My thanks also go to my little one, Bambi, the joy of our family. I wish you were here today; we would have celebrated in our own way, your favorite treats for you and cake for me. I will never forget how, in every difficult moment, you sensed my struggle, curled up in my arms, and brought me calm. I could not have done this without you.

TABLE OF CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

ΔG: Gibbs Free Energy Change

μs: Microsecond

Å: Ångstrom

aa: Amino acid

AKT: Protein Kinase B; PKB

ALL: Acute lymphoblastic leukemia

AML: Acute myeloid leukemia

DMSO: Dimethyl sulfoxide

FTMap: Computational Solvent Mapping

GAP: GTPase-Activating Protein

GEF: Guanine Nucleotide Exchange Factor

GDP: Guanosine diphosphate

GTP: Guanosine triphosphate

HCC: Hepatocellular Carcinoma

HRAS: Harvey Rat Sarcoma Virus

K:Kelvin

kcal: Kilocalorie

KRAS: Kirsten Rat Sarcoma Virus

MAPK: Mitogen-activated Protein Kinase

MD: Molecular Dynamics

MSCS: Multiple Solvent Crystal Structures

mTORC2: Mechanistic Target of Rapamycin Complex 2

NSCLC: Non-small-cell Lung Cancer

NRAS: Neuroblastoma RAS Viral Oncogene

ns: Nanosecond

PBC: Periodic Boundary Condition

PDB: Protein Data Bank

PDK1: 3-phosphoinositide-dependent protein kinase-1

PI3K: Phosphoinositide 3-kinase

PIP3: Phosphatidylinositol (3,4,5)-trisphosphate

P-loop: Phosphate-binding loop

pMD: Probe-based Molecular Dynamics Analysis

PMF: Potential of Mean Force

Raf: Rapidly Accelerated Fibrosarcoma

Ral: Ras-like

RCC: Renal Cell Carcinoma

Rg: Radius of Gyration

RMSD: Root Mean Square Deviation

RMSF: Root Mean Square Fluctuation

SMD: Steered Molecular Dynamics

TIP3P: Three-point Water Model

WHAM: Weighted Histogram Analysis Method

1.INTRODUCTION

The Ras protein family was first identified in the 1980s as being encoded by genes that are frequently activated through mutations in human cancers. Since then, numerous studies have shown that RAS genes (*KRAS*, *HRAS*, and *NRAS*) represent the oncogene family with the highest mutation frequency. RAS genes are particularly involved in pancreatic, colorectal, and lung cancers (Cox et al., 2014)(Moore et al., 2020). Among the regulators of signal transduction, Ras proteins belong to the superfamily of small proteins (GTPases) capable of binding and hydrolyzing GTP (guanosine triphosphate) (Moore et al., 2020). The Ras superfamily consists of small guanosine triphosphatase enzymes that play a role in a variety of cellular functions in response to extracellular signals. The Ras protein family, a subgroup of G proteins, plays a role in critical cellular processes such as intracellular signal transduction, regulation of gene expression, cell proliferation, division, and mutation (Reedquist et al., 2012). Ras is a proto-oncogene that becomes active in its GTP-bound state and switches off when bound to GDP (guanosine diphosphate). This nucleotide exchange cycle functions through a molecular "on/off" transition mechanism (Kim et al., 2021).

The Ras protein family consists of three homologous proteins KRAS (KRAS4A and KRAS4B), HRAS and NRAS (Muñoz-Maldonado et al., 2019). The three homologous Ras proteins share same conservation within the first 80 amino acids of the G domain, where two key regions, Switch I and Switch II (aa 30-40 and 58-72), undergo the principal structural changes between GTP and GDP-bound states (Figure 1.1.A). These regions constitute binding sites for Ras effectors such as MAPK, PI3K, Raf and Ral, and this region is also observed as a binding pocket for potential small molecules (McCormick, 2016). Guanine Nucleotide Exchange Factors (GEFs) and GTPase-Activating Proteins (GAPs) regulate the activity of small GTPases by controlling the guanine nucleotide cycle. While GEFs promote the dissociation of GDP from Ras, allowing its replacement by GTP, GAPs stimulate the intrinsic GTPase activity of Ras, accelerating GTP hydrolysis and contributing to signal termination (Figure 1.1.B) (Bos et al., 2007). Disruption of this cycle and the locking of Ras in its GTP-bound active state lead to persistent activation of downstream signaling pathways, particularly MAPK and PI3K, ultimately driving tumor initiation and progression (Yang and Wu, 2024).

The functional architecture of the Ras protein is defined through its GTPase domain, which can be divided into two main regions: the effector lobe and the allosteric lobe (Figure 1.1C). The effector lobe comprises, the Switch I and Switch II regions (aa 30-40 and 58-72) and is responsible for specific interactions with effector proteins such as RAF kinases, PI3K and MAPK, thereby enabling the GTP-bound active conformation of Ras to initiate downstream signaling

pathways. The allosteric lobe, on the other hand, represents a more distal region that governs intramolecular conformational rearrangements and membrane interactions; subtle structural changes within this lobe indirectly modulate the dynamics of the effector lobe and thus regulate the activity state of Ras (Pálfy et al., 2020), (Menyhárd et al., 2020).

Current data indicate that the Ras superfamily is found to be mutated in nearly 30% of human cancers, with a prominent role in colorectal (52.2%), pancreatic (97.7%) and lung (32.2%) cancers (Moore et al., 2020). In the most frequently mutated oncogene KRAS, alterations occurring at codons 12, 13 and 61 account for about 85% of Ras mutations (Prior et al., 2012). For this reason, intensive research efforts continue to focus on the development of effective therapeutic strategies targeting cancers driven by mutant KRAS.

Despite this, Ras was considered an "undruggable" target for more than 30 years due to its structural features and picomolar affinity for GTP/GDP. The absence of a well-defined drug binding site on the Ras protein surface limited the development of small molecules directly targeting Ras for many years (Wang et al., 2025). However, the discovery of small molecules capable of binding to a pocket located between the Switch I and Switch II regions (aa 30-40 and 58-72) led to an evaluation of Ras druggability and an increase in research in this field (Kessler et al., 2019). The inhibitors first developed by Shokat and colleagues, which inhibit SOS-driven nucleotide exchange and consequently disrupt Ras-Raf binding, followed by structural modifications leading to more potent molecules such as ARS-853 and ARS-1620, represented an important turning point in direct Ras targeting (McCormick, 2019). Ongoing studies have demonstrated that it is possible to design inhibitors specific to certain mutant forms of Ras.

One of the most notable advances in this regard has centered on designing covalent compounds against the G12C mutant KRAS. Sotorasib (AMG 510), the first FDA-approved KRAS inhibitor, targets the KRAS G12C mutation in non-small cell lung carcinoma. Ras inhibitors lock Ras in its inactive, GDP-bound form, preventing uncontrolled activation of downstream signaling pathways (Nakajima et al., 2022). Similarly, Adagrasib (MRTX 849) has shown significant anticancer efficacy and a favorable safety profile in patients with non-small cell lung cancer and colorectal cancer harboring the KRAS G12C mutation (Zhang et al., 2022). KRAS G12C inhibitors are designed to create a covalent bond with the cysteine residue at position 12 by targeting the pocket neighboring to the Switch II region (Kwan et al., 2022). However, considering the mutational variations observed in the clinic, KRAS G12C represents only a limited subgroup of patients. The KRAS G12D mutation, which is predominant in pancreatic ductal adenocarcinoma, has emerged as a more critical therapeutic target (Bannoura et

al., 2022). In the KRAS G12D mutation, codon 12 is mutated from glycine to aspartic acid, making it the most common KRAS mutation and accounting for approximately 45% of KRAS mutations, thereby representing a large patient population that does not benefit from G12C targeted therapies (Suyal et al., 2025). The lack of a suitable reactive cysteine in the Switch II pocket of KRAS G12D for covalent inhibition and the unfavorable nature of aspartic acid as a covalent binding target make the development of inhibitors specific to this mutant form more challenging. Consequently, current efforts focus on next-generation inhibitors such as MRTX-1133, TH-Z827 and TH-Z835, which are designed to selectively target KRAS G12D and stabilize KRAS in its GDP-bound inactive state; these molecules are still in preclinical stages of development (Hallin et al., 2022).

All these findings indicate that, despite being historically labeled as "undruggable," Ras is in fact a druggable oncogene at the molecular level. Nevertheless, a substantial proportion of patients fail to respond to KRAS inhibitor therapies due to intrinsic or acquired resistance mechanisms (Cox and Der, 2025). The successes achieved over the past decade with small molecule targeted therapies have further underscored the importance of better understanding the Ras oncogene and developing novel inhibitors. Since the first reports of small molecules directly targeting KRAS, the increasing number of clinical candidates and the accelerated FDA approval of Sotorasib have strongly supported the feasibility of direct Ras targeted drug development strategies (Luo et al., 2025).

RAS dimerization refers to two RAS proteins (HRAS/NRAS/KRAS) coming into close proximity on the plasma membrane to form transient dimers or nanoclusters. This arrangement is proposed to facilitate RAF activation and dimerization, thereby strengthening MAPK/ERK signaling. Oncogenic mutations such as KRAS G12D prolong the GTP-bound active state, which can increase the likelihood of such clustering and amplify downstream signaling. Consequently, peptides or small molecules that disrupt RAS dimerization interfaces are being explored. Although approved KRAS G12C inhibitors (and experimental G12D inhibitors) do not directly target the dimer interface, they weaken dimerization-dependent signaling by locking RAS in its inactive (GDP-bound) form, preventing activation and attenuating downstream pathways (MEK/ERK, Ral). Altough, RAS dimerization has been proposed to enhance RAF activation, current evidence indicates that RAS is predominantly monomeric, with transient dimers/nanoclusters forming under specific membrane conditions. Thus, dimerization can potentiate signaling but is not strictly

required, particularly for constitutively active oncogenic RAS mutants (Nabet and Gray, 2018)(Kim et al., 2020).

Within the scope of this project, a comprehensive library of KRAS G12C specific inhibitors was constructed based on the most recent FDA approved therapeutics and structurally related molecules reported in the literature. This library was subjected to virtual screening, followed by molecular docking studies. In addition, for the KRAS G12D mutation, the candidate molecule previously identified by virtual screening in the thesis numbered 813197, published in 2023 by a former M.S. student in our research group, was examined in detail at the molecular level using molecular dynamics simulations in order to characterize its interactions with the target. Furthermore, for KRAS G12D, alternative analog molecules were generated by systematic chemical group modifications guided by the binding site, and molecular dynamics simulations of these newly designed compounds were also performed. This integrated computational approach was expected to contribute to the development of more effective and selective inhibitors for Ras-driven tumors, to propose new solutions to current drug resistance, and to provide a reference inhibitor candidate for future experimental and clinical studies in the field of KRAS-targeted drug discovery.

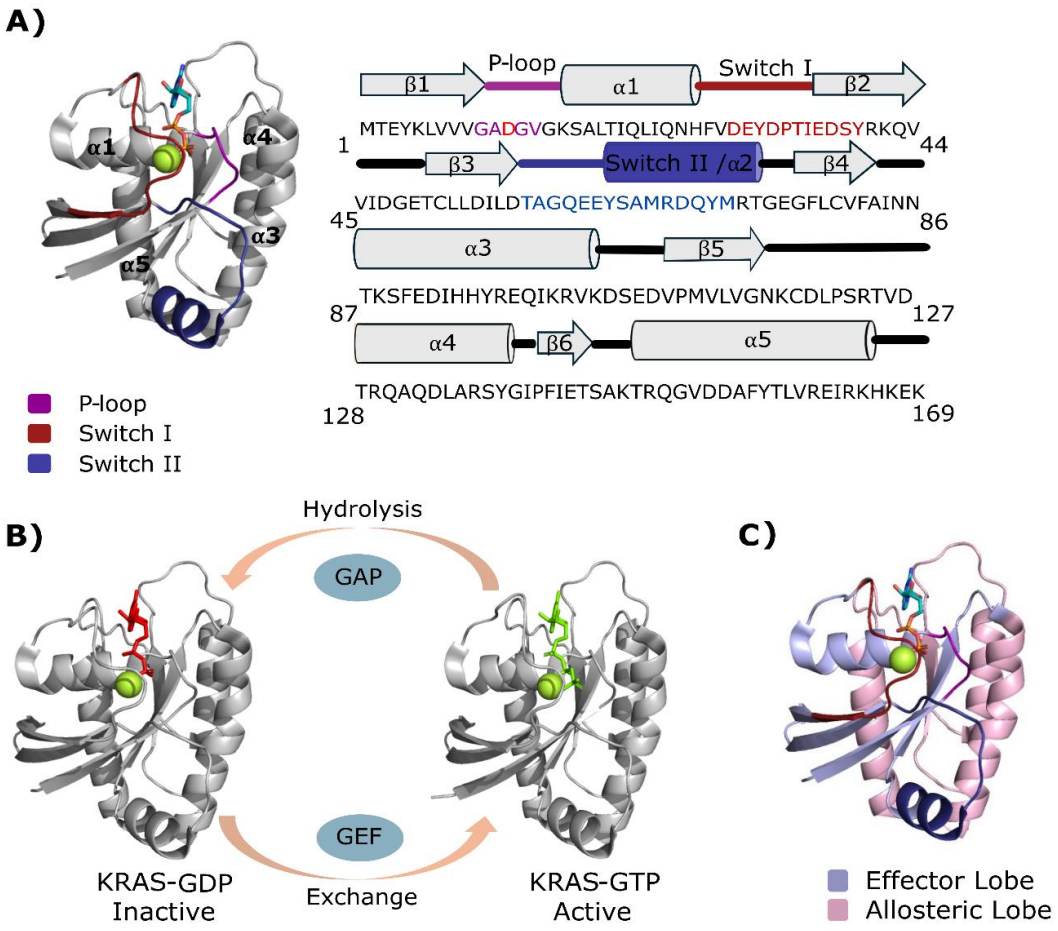

Figure 1.1 The structure, sequence, and working mechanism of the KRAS G12D. A) Crystal structure of GDP-bound KRAS G12D (PDB ID: 7RPZ) and two-dimensional representation of the secondary structure. B) GDP-GTP nucleotide exchange cycle and domains of KRAS. C) Effector and allosteric lobes of KRAS.

2.LITERATURE REVIEW

The KRAS protein is a membrane-linked small GTPase that acts as an on/off regulator for a wide range of signaling processes in cells. Active KRAS abundance in cells depends on the equilibrium between nucleotide hydrolysis and nucleotide exchange. KRAS bound to GDP is inactive, while KRAS bound to GTP is active (Hofmann et al., 2022).

RAS is among the most commonly mutated oncogenes in human cancer. Within this family, KRAS alterations occur most often, with NRAS mutations next in frequency, particularly in hematologic cancers, followed by HRAS mutations. The earliest Ras-associated mutation identified in human cancer tissue was a KRAS mutation detected in a patient's lung tumor in 1984, and this development was later followed by the discovery of a KRAS-mutation related colon cancer. This finding has encouraged researchers to search for RAS mutations in other

tumor types. To date, more than 20,000 different KRAS mutations have been found. Certain residues (G12, G13, and Q61) on RAS isoform proteins are more predominantly mutated in human cancers. Somatic activating mutations in the KRAS protein prevent the hydrolysis of GTP, which disrupts the normal balance between the GTP and GDP-bound state of KRAS, creating a constantly active KRAS molecule (Román et al., 2018). Over the past few years, a number of allele-specific mutated Ras inhibitors have been developed. In particular, molecules targeting tumors with KRAS G12C mutations have shown significant benefits across a variety of tumor types. In 2018, just a few years after the initial discovery of a druggable pocket in KRAS G12C, three small molecule (AMG 510, MRTX849, and JNJ74699157) reached human clinical trial testing. Today, there are at least five KRAS G12C inhibitors undergoing clinical trials (Krishnan et al., 2022). The mutant cysteine of KRAS G12C is located adjacent to a pocket found in the GDP-bound inactive form of KRAS (Liu et al., 2022). A common tactic in drug development is to design covalent compounds that react with active site cysteines.

Shokat's group was the first to describe a new allosteric binding site in KRAS G12C, referred to as the p2 pocket. They used a disulfide fragment-based screening approach to identify a chemical starting point. They screened a library of 480 molecules against KRAS G12C in the GDP state. Their work ultimately led to the development of a number of inhibitors, Shokat's inhibitor. The close proximity of the pocket and the mutant cysteine led to an extensive search for covalent inhibitors, eventually resulting in the identification of the inhibitor ARS-1620. After the discovery of the ARS-1620 inhibitor, the use of an alternative orientation of the inhibitor in the Switch II pocket of His95 and the addition of aromatic rings improved protein interactions with KRAS G12C and led to the development of Sotorasib (AMG 510) (Ostrem et al., 2013). Structure-guided drug design and subsequent optimization efforts resulted in the development of MRTX849 (Hallin et al., 2020).

New inhibitor discovery methods for KRAS-targeted computational studies include virtual screening and docking programs. Computational studies targeting KRAS, using new inhibitor discovery methods, demonstrate the use of virtual screening and docking programs. In their study on KRAS WT, KRAS G12D, KRAS G12V, KRAS G13D, and KRAS Q61H, Gupta and colleagues selected the ZINC database and removed compounds containing functional groups associated with adverse effects such as mutagenicity and tumorigenicity from the selected ligand set. These compounds were then filtered to meet the following values: Molecular weight = 150-500 Da, HBA ≤ 10, HBV ≤ 5, TPSA ≤

160, LogP ≤ 5, Number of rotatable bonds ≤ 7, Drug score ≥ 0.30, and Molecular complexity ≥ 0.40.

Protein structures were prepared for docking using the Protein Preparation Wizard in Schrödinger by assigning bond orders, adding hydrogen atoms, and removing water molecules 5 Å from heterogroups. Ligand libraries were prepared using the LigPrep program at pH 7.0. Pocket-specific libraries were constructed using Schrödinger's Glide docking program, which docked 8–12 representative conformers of each KRAS protein into the corresponding pocket. This study demonstrated that the E22 inhibitor has an effect on GEF-mediated GTP-GDP exchange (Gupta et al., 2019). In their studies targeting KRAS WT and KRAS G12D, Pathan and colleagues selected ligands from many databases (Asinex, BindingDB, DrugBank, National Cancer Institute (NCI), PubChem and Chembridge) and after applying the Lipinski test to these ligand sets, reactive groups were removed from the data set (Pathan et al., 2016). In a recent study conducted in 2022, compounds retrieved from the ZINC and IBS databases were subjected to ADMET filtering after the Lipinski test. After filtering, the number of these compounds decreased. The CDOCKER program was used to dock the target KRAS G12D structure (PDB ID: 6GJ8), assuming the target binding site as of the reference ligand's. The four compounds with the best scores were selected and molecular dynamics simulations were performed. After examining the hydrogen bonds, van der Waals, and pi interactions formed with the binding site, the number of selected compounds was reduced to two. As a result of the study, the inhibition properties of molecules 21 and 22 were discovered (Kulkarni et al., 2022).

In another study conducted for KRAS G13D, a molecular library containing 602,618 molecules was generated from the MCULE, ChemSpace, ChemDiv, CHEMBL30, ZINC, the NCI Open Chemistry Repository, WuXi LabNetwork, MolPort, and MCULE-ULTIMATE databases. This library was screened using Lipinski's rule, MM/GBSA, and ADMET. A 250 nanosecond molecular dynamics simulation of KRAS G13D was performed using the molecule CSC01 selected from the screening, and umbrella sampling (ΔG = -7 kcal/mol) was used for free energy calculations. This study has led to the discovery of a promising new inhibitor candidate for the G13D mutation, and its experimental evaluation is crucial (Durojaye et al., 2023).

In this thesis, we aimed to discover molecules similar to approved drugs in the literature by subjecting the molecular library generated in the ZINC20 database (Irwin et al., 2020) for KRAS G12C to filtering with Lipinski's Rule of Five, followed by fingerprint and 3D similarity searches. After the virtual screening,

we proposed the most potentially promising molecule among the obtained compounds through their covalent docking for further studies.

For KRAS G12D, we examined the interaction of the molecule (Discovered in thesis number 813197) with KRAS G12D and the behaviors of the complex structure through molecular dynamics simulation and post-MD analyses. With the cross-correlation analysis of these simulations, we observed the correlation between the critical regions of the protein (P-loop, Switch I, and Switch II) when the molecule is bound to the protein, and we calculated the free binding energy between the ligand and the protein using the umbrella sampling method. After this stage, we used the MolOpt tool (Shan and Ji, 2020) to generate more effective analog molecules than the original molecule. Then, we subjected the molecules to molecular dynamics simulations based on their binding affinities with AutoDock Vina (Trott and Olson, 2010), (Eberhardt et al., 2021).

The importance of analog molecules obtained through bioisosteric transformations is very clearly seen in the study by Karmacharya and colleagues (Karmacharya et al., 2021). In this study, they aimed to develop a better Cabozantinib analog targeting hepatocellular carcinoma (HCC) and non-small-cell lung cancer (NSCLC) that would show anticancer activity by inhibiting the receptor tyrosine kinase. The central benzene ring of the Cabozantinib molecule was replaced with 2,4,5-trimethylpyridine (3) and pyridine (4) using a bioisosteric approach. Among these synthesized analog molecules, 4 was found to show strong inhibition (~94%), and the IC_{50} of 4 was very close to Cabozantinib (4.9 nM vs. 5.4 nM). In the docking results, it was observed that 4 possesses a suitable conformation for binding, and in cell experiments 4 showed better antiproliferative effects than Cabozantinib in HCC (Hep3B, Huh7) and NSCLC (A549, H1299) lines and provided higher tumor selectivity especially against HCC. This was supported by the CAM xenograft model where 4 inhibited Hep3B tumor growth and angiogenesis more clearly than Cabozantinib. In conclusion, compound 4 was suggested to be a more promising therapeutic candidate for HCC.

In another study conducted in 2023 (Sharma et al., 2023), they aimed to develop multifunctional Chalcone derivatives against Alzheimer's. The researchers designed analogs by adding different amine groups to the chalcone scaffold via a two-carbon spacer, generated 826 candidates by making bioisosteric changes at 5 regions with MolOpt, proceeded to docking after filtering with synthetic accessibility and SwissADME, and synthesized and tested the PS(1–10) series. In vitro, PS10 (N-methylpiperazine-substituted) stood out as the strongest AChE inhibitor (IC_{50} = 15.3 nM) and was found to be slightly better than donepezil (IC_{50} = 15.68 nM). In the antioxidant test, PS10 again was the strongest candidate and

showed performance close to the standard with EC_{50} = 22.83 nM (EC_{50} = 21.7 nM). Selected as the most promising candidate, PS10 improved memory in rats with Alzheimer-like dementia induced by STZ i.c.v. 3 mg/kg (days 1 and 3) at doses of 2, 5, and 10 mg/kg p.o., significantly reversed the increased brain AChE activity and oxidative stress imbalance (TBARS↑, GSH↓), and 100 ns molecular dynamics simulations also showed that the PS10-AChE complex is stable.

3.MATERIALS AND METHODS

3.1 Targeting the G12D Mutant KRAS

In this study, various analyses were conducted for the molecule ZINC000008758797, which was identified as a potential drug candidate against KRAS G12D, with the best docking score determined in the master's thesis numbered 813197 (published by our research group).

3.1.1.Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were carried out with GROMACS 2023.5 software package (Abraham et al., 2024) to investigate the interaction and binding behavior between ZINC000008758797 molecule and both the G12D mutant and Wild Type KRAS proteins.

Initially, the complex structure obtained by molecular docking in thesis study No. 813197, was used. The coordinates of the ligand and the protein were separately extracted from the docked complex. The CHARMM36 force field (July 2022 version) (Huang and MacKerell, 2013) was added to the working directory to parameterize the system. The pdb2gmx command was used to generate the protein structure file in .gro format, employing the TIP3P water model. The terminal groups were defined as NH_3^+ for the N-terminus and COO^- for the C-terminus.

To generate the topology file of the ligand, hydrogen atoms were added to the ligand structure using the Avogadro software (Hanwell et al., 2012)(Avogadro Project, 2016), and the resulting file was saved in .mol2 format. To ensure proper bond order, the mol2 file was corrected using a Perl script named sort_mol2_bonds.pl (Lemkul, 2018). The corrected mol2 file was uploaded to the CGenFF (Vanommeslaeghe et al., 2010), (Vanommeslaeghe and MacKerell, 2012), (Vanommeslaeghe et al., 2012) server to obtain the corresponding .str file of the molecule. Subsequently, the cgenff_charmm2gmx_py3_nx2.py Python script (Lemkul, 2018) was employed to generate the topology and pdb files for the ligand. This process utilized Python 3.5.2 and NetworkX 2.3 versions.

Using the editconf command, the ligand's .gro file was prepared. The .gro files of the protein and the ligand were then merged to form the complex structure. The

ligand topology and parameter files were manually included in the topol.top file. A cubic simulation box was defined around the complex and solvated using the SPC216 water model. The system was then neutralized by adding K^+ and Cl^- ions with a concentration of 0.15 M.

Following energy minimization, a position restraint file was added to the topol.top file to constrain the ligand before the equilibration phase. The system was equilibrated at 310 K temperature and 1 atm pressure using NVT and NPT ensembles. After equilibration, the position restraints were removed, and a 1000 nanosecond production MD simulation was performed. For both the mutant and wild type complexes, three independent replicates of the 1000 nanosecond simulation were conducted to ensure statistical reliability.

Following the completion of the molecular dynamics simulations, a series of post-simulation analyses were conducted using the GROMACS 2023.5 software package (Abraham et al., 2024) to evaluate the structural stability, flexibility, and interaction dynamics of the protein-ligand complexes.

An index file was first generated using the make_ndx command to define atom groups required for subsequent analyses. The simulation trajectories were then processed to remove periodic boundary condition (PBC) artifacts. This was achieved using the trjconv command, initially with the -pbc nojump option to correct for molecular discontinuities, and subsequently with the -pbc mol -center option to re-center the protein-ligand complex within the simulation box. The resulting .xtc file, free of PBC effects, was used for all further analyses.

To evaluate the structural stability of the system, the root mean square deviation (RMSD) of the protein backbone atoms was calculated using rms command. The radius of gyration (Rg) of the protein, representing the overall compactness of the structure, was determined with the gyrate command based on the atom groups defined in the index file.

The root mean square fluctuation (RMSF) values were computed for the Cα atoms of each residue using the rmsf module to quantify local flexibility and identify regions of significant dynamic variation within the protein structure. Additionally, hydrogen bond (H-bond) analysis was performed with the hbond command to determine the number and persistence of hydrogen bonds between the protein and the ligand throughout the simulation.

3.1.2 Cross-Correlation Analysis

The GROMACS 2023.5 software package (Abraham et al., 2024) covar module was utilized to perform the covariance analysis of the molecular dynamics trajectories. The resulting .dat file obtained from this step served as the input for

the subsequent cross-correlation analysis. Cross-correlation calculations were conducted using a Jupyter Notebook script named covar_to_crosscorr (Barlas et al., 2022).

To capture the dynamic relationship between protein residues during different ligand binding events, separate cross correlation analyses were performed for each region (active and allosteric site) where the ligand interacted with the protein. For each simulation, only the trajectory segments corresponding to the periods in which the ligand remained bound to a specific region were extracted and analyzed independently.

Following the calculation of cross correlation matrices for each segment, average correlation maps were generated across the simulation replicas to ensure statistical consistency. The averaged matrices were then compared to identify variations in residue-residue correlation patterns between the mutant and wild type systems, as well as between distinct ligand binding regions. This approach enabled the assessment of how ligand binding modulates correlated motions within the protein over the course of the simulations.

3.1.3 Umbrella Sampling

Umbrella sampling simulations were performed to elucidate the ligand dissociation pathways from both the active and allosteric binding sites of the G12D mutant and wild type KRAS. All simulations were conducted with GROMACS 2023.5 software package (Abraham et al., 2024) at 310 K and 1 atm, using TIP3P water, adding K^+ and Cl^- ions with a concentration of 0.15 M, LINCS bond constraints, and a 2 fs time step. Prior to system setup, ligand parameters were generated with CGenFF (Vanommeslaeghe et al., 2010), (Vanommeslaeghe and MacKerell, 2012), (Vanommeslaeghe et al., 2012) and the corresponding topology/parameter files were prepared.

For each system, a representative structure was extracted from the equilibrated portion of the respective MD trajectory. The ligand interacting with the active site of G12D mutant structure was taken from replicate 2 at 917.4 ns, ligand interacting with the active site of wild type structure was taken from replicate 3 at 538 ns. The ligand interacting with the allosteric site of G12D mutant structure was taken from replicate 1 at 700 ns and the corresponding wild type complex was taken from replicate 1 at 700 ns.

A rectangular simulation box was constructed for each system with dimensions sufficient to allow full ligand dissociation along the pulling axis. The box sizes were 7.000 × 14.000 × 7.000 nm for the active site G12D mutant, 9.000 × 8.500 × 14.000 nm for the active site wild type, and 12.000 × 5.570 × 5.230 nm for both

allosteric systems. Each system was solvated with TIP3P water. Systems were neutralized by adding K^+ and Cl^- to a final concentration of 0.15 M. Energy minimization and equilibration were then performed in NVT and NPT ensembles at 310 K and 1 atm.

Steered molecular dynamics (SMD) simulations were executed to gradually pull the ligand from the binding pocket into bulk solvent using a harmonic force constant of 1000 kJ·mol^{-1}·nm^{-2} and a constant pulling rate of 0.01 nm·ps^{-1} for 500 ps. The reaction coordinate was defined normal to the binding site and aligned along the Y axis for the active site G12D mutant, the Z axis for the active site wild type, and the X axis for both allosteric systems. Each pulling simulation continued until complete dissociation, and snapshots were saved at regular intervals along the reaction coordinate.

From the SMD trajectories, equally spaced configurations were selected as umbrella sampling windows: 26 windows for the active site of G12D mutant, 25 for the active site of wild type, and 15 windows for each allosteric counterpart system. Each window was equilibrated under NPT conditions for 100 ps before production MD. Production umbrella simulations were then performed for 10 ns per window. The resulting biased trajectories were analyzed with WHAM to obtain the potential of mean force (PMF) profiles.

In summary, umbrella sampling simulations were carried out for four systems corresponding to ligand interactions with both the active and allosteric binding sites of the G12D mutant and the wild type KRAS. The PMF curves provided quantitative insight into the relative binding strength and dissociation energetics of the ligand across these distinct structural contexts.

3.1.4 Active and Allosteric Site Analysis with the Selected Inhibitor

We analyzed three independent simulations of the mutant-ligand complex and mapped, in nanoseconds, the intervals during which the ligand occupied each binding site. For every replicate, we extracted the protein conformation corresponding to the moment when the ligand at both the active and allosteric sites formed the maximum number of hydrogen bonds. The ligand interacting with the active site of the G12D mutant KRAS structure was taken from replicate 2 at 917.4 ns. The ligand interacting with the allosteric site of the G12D mutant KRAS structure was taken from replicate 1 at 700 ns. Using these representative structures, we characterized interactions at both sites and identified the ligand functional groups or atoms that engage the binding pockets. Interaction profiles were computed with PLIP (Schake et al., 2025) and Discovery Studio's Protein-Ligand Interactions tool (BIOVIA, Dassault Systèmes, 2021).

We prioritized modifying the non-interacting portions of the active site ligand, specifically, the disubstituted diazine ring (two nitrogens) bearing two carbon substituents, the N-substituted indole, carbon-nitrogen main scaffold and the chlorine (Cl) atom. For modification of the non-interacting portions of the allosteric site ligand, specifically, the disubstituted diazine ring (two nitrogens) bearing two carbon substituents, O-linked phenyl (phenoxy) group and the carbon-nitrogen main scaffold were selected based on the results in Section 4.1.4.

3.1.5 Preparation of Analog Molecules

Using the MolOpt web server (Shan and Ji, 2020), bioisosteric analogs were generated for the molecule with the identifier ZINC000008758797. Residues involved in protein-ligand interactions were examined separately for the active and allosteric binding sites. Based on these interaction profiles, specific atoms or functional groups of the ligand were selected for substitution to better accommodate each site. For the active site workflow, the ligand was uploaded in SMILES format to MolOpt (Shan and Ji, 2020) and Rules From Data Mining, Similarity Comparison, and Deep Generative Model options were applied. Individual regions targeted for modification were specified, and the resulting bioisosteric candidates were generated and downloaded. For the allosteric site, regions to be modified were likewise delineated, and bioisosteres were produced using Rules From Data Mining, Similarity Comparison, and the Deep Generative Model modules, after which all candidate molecules were downloaded.

Rules from Data Mining: This method extracts bioisosteric transformation rules from literature data. Wet lab level analog sets from ChEMBL (Gaulton et al., 2012) are processed via Matched Molecular Pairs Analysis using the Hussain-Rea algorithm (Hussain and Rea, 2010). Compounds are iteratively fragmented at all acyclic single bonds, complementary fragments are indexed, and matched pairs sharing a common scaffold with a single localized substitution are identified.

Similarity Comparison: In this method, interchangeable fragments are identified by similarity comparison rather than historical optimization data. Each fragment is represented by a simple profile: attachment points, H-bond donors/acceptors, hydrophobic atoms, and other heavy (non-hydrogen) atoms. Distances among these features are encoded as a topological pharmacophore fingerprint. Fragment pairs are then compared using the Tanimoto coefficient on these fingerprints; high-scoring pairs are taken as candidate bioisosteres, meaning fragments likely to substitute for one another without losing key interactions.

Deep Generative Model: This method learns to generate chemically meaningful transformation rules as SMIRKS strings using a sequence model. Bioisosteric transformations (1.4 million SMIRKS rules obtained from MMP analysis) are

tokenized with a predefined SMIRKS/SMARTS vocabulary and used to train a character-level Recurrent Neural Network implemented in PyTorch. This setup enables MolOpt (Shan and Ji, 2020) to propose syntactically valid, data-driven transformations conditioned on a user specified fragment while remaining compatible with standard cheminformatics toolchains.

3.1.6 Molecular Docking

Molecular docking of the designed analogs to KRAS G12D was performed using AutoDock Vina (Trott and Olson, 2010), (Eberhardt et al., 2021). To enable docking of multiple ligands, we followed the batch-docking procedure described by Li and Shash (Li and Shah, 2017). For the active site, the grid box was defined as: center (Å) x = 34.974, y = 26.7239, z = 26.6453; size (Å) x = 25.4255, y = 22.8467, z = 25.0023. For the allosteric site, the grid box was defined as: center (Å) x = 26.4158, y = 29.0933, z = 38.6318; size (Å) x = 23.8892, y = 28.4253, z = 25.8266. Docking for both the active and allosteric sites was carried out in three independent replicates per molecule, and for each compound the mean of the three Vina scores was used for downstream analyses.

3.1.7 Molecular Dynamics Simulations for Analog Inhibitors

Three methods were used to create analog molecules: Data Mining, Similarity Comparison, and Deep Generative. The docking scores of the molecules generated by these three methods for both active and allosteric regions were combined into two tables. The molecules with the top 10 scores were selected for the molecular dynamics simulation. The molecular dynamics simulations of the selected molecules, chosen based on the docking scores of all generated analogs, were carried out as described in Section 3.1.1.

3.1.8 Umbrella Sampling for Analog Inhibitors

Umbrella sampling simulations for the KRAS G12D-bound analog molecules were performed as described in Section 3.1.3. In these systems, ligand dissociation was sampled using steered molecular dynamics (SMD) along a reaction coordinate defined on a single Cartesian axis. Specifically, analogs 8929323, 3302325, and 5179173 were pulled along the Z axis, whereas analog 5576702 was pulled along the X axis. For the generation of umbrella sampling starting configurations, the initial bound structures were selected from the MD simulations by choosing the frames corresponding to the maximum hydrogen-bonding between the ligand and KRAS. The selected time points were 46.5 ns for analog 8929323, 692.8 ns for analog 3302325, 709.8 ns for analog 5179173, and 168.5 ns for analog 5576702. When selecting umbrella windows from the SMD trajectories, the window spacing was set to 0.05 nm for analog 8929323 due to its relatively long dissociation distance; using 0.2 nm would have provided fewer

configurations across the full separation range. For the other three analogs (3302325, 5179173, and 5576702), a 0.2 nm window spacing was used. With this setup, 6 umbrella windows were generated for analog 8929323, whereas 26, 25, and 26 windows were generated for analogs 3302325, 5179173, and 5576702, respectively. For each window, an NPT equilibration of 100 ps (0.1 ns) was followed by a 10 ns production run, resulting in total umbrella sampling simulation times (including equilibration) of 60.6 ns for analog 8929323, 262.6 ns for analog 3302325, 252.5 ns for analog 5179173, and 262.6 ns for analog 5576702.

3.2 Targeting the G12C Mutant KRAS

3.2.1 Molecular Library Preparation and Virtual Screening

For the construction of the molecular library to be used in the virtual screening phase, KRAS G12C targeted covalent inhibitors; Sotorasib, Adagrasib, and ARS-1620 were selected as positive control molecules, while dimethyl sulfoxide (DMSO) was employed as a negative control.

The molecular library was generated using the ZINC20 database (Irwin et al., 2020). Within the Tranches section of ZINC20 database (Irwin et al., 2020), the following filters were applied to define the compound set: 3D structure: *enabled*, Reactivity: *Reactive*, Purchasability: *In-stock*, pH: *Reference*, Charge categories: -2 (L), -1 (M), 0 (N), +1 (O), and +2 (P). The filtered compounds were downloaded in .sdf format. Compressed .gz files within the dataset were extracted, and all individual .sdf files were subsequently merged into a single .sdf file using the Open Babel software (O'Boyle et al., 2011).

The KNIME Analytics Platform (Berthold et al., 2007) was employed to perform the virtual screening workflow (Figure 3.1). The combined .sdf file was first imported into the SDF Reader node, which was then connected to the Lipinski Rule of Five node to assess the drug-likeness of each compound. Using the Row Filter node, compounds that fully satisfied all five of Lipinski's criteria (value = 0) were selected. The Column Filter node was applied to retain only the columns containing molecular information. For the calculation of molecular fingerprints, the Fingerprint node was used with the standard fingerprinting option. A 3D-D Similarity node was then utilized to perform three-dimensional similarity comparisons between compound sets. Two datasets were included in the analysis: The filtered ZINC20 (Irwin et al., 2020) compound set (after Lipinski filtering), and The positive control dataset containing the molecular fingerprints of Sotorasib, Adagrasib, and ARS-1620. For the positive control molecules, a parallel workflow was applied, starting from the SDF Reader node and proceeding through the Fingerprint node to ensure methodological consistency. The similarity thresholds were defined as 0.85 for Sotorasib, 0.85 for Adagrasib,

and 0.90 for ARS-1620. Compounds meeting or exceeding these threshold values were selected using additional Row Filter and Column Filter nodes. The final subset containing the structural and identification information of the selected compounds was exported in .sdf format using the SDF Writer node.

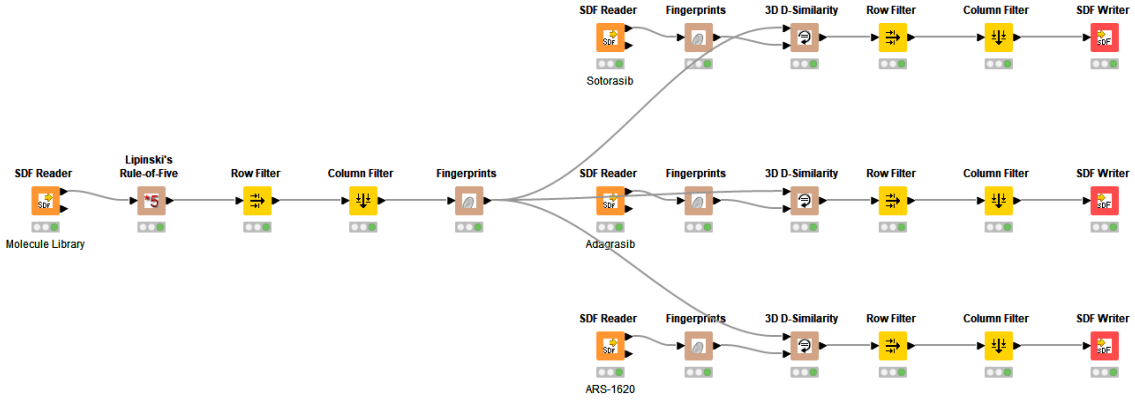

Figure 3.1 Virtual Screening Workflow.

3.2.2 Preparation of the Receptor Structure for Docking

To obtain the physiologically relevant conformation of the receptor to be used in the docking studies, the structure with PDB ID: 6OIM (Canon et al., 2019) was subjected to molecular dynamics (MD) simulation based minimization using the GROMACS 2023.5 software package (Abraham et al., 2024). Prior to the simulation, all water molecules and the co-crystallized Sotorasib inhibitor were removed from the structure. The CHARMM36 force field (July 2022 version) (Huang and MacKerell, 2013) was selected for system parameterization. A cubic simulation box was defined around the protein, and the system was solvated with SPC216. To ensure overall charge neutrality, appropriate counter ions (K^+ and Cl^-) were added to the system. The system underwent energy minimization, followed by temperature (NVT) and pressure (NPT) equilibration phases.

For the production MD simulation, a total simulation time of 200 nanoseconds (ns) was defined. Upon completion of the simulation, the root mean square deviation (RMSD) plot was analyzed to assess system stability. Based on the RMSD profile, the structure at 130 ns was identified as a stable representative conformation. The coordinates of this frame were extracted and saved in PDB format (Bernstein et al., 1977). The resulting structure was then converted to PDBQT format and subsequently used as the receptor input for the molecular docking studies.

3.2.3 Covalent Docking

The molecules received from virtual screening in molecular docking were subjected to covalent docking using the Cresset Flare software (Cresset, 2024), (Cheeseright et al., 2006), (Bauer and Mackey, 2019), (Kuhn et al., 2020). Within the program, the *Dock* option was selected, and the docking type was set to *Covalent*. The *Residue* parameter was defined as *CYS 12*, and the docking process was subsequently performed.

3.2.4 Active Site Analysis with the Selected Covalent Inhibitors

Discovery Studio's Protein-Ligand Interaction tool (BIOVIA, Dassault Systèmes, 2021) was used to determine the residues that interact between the protein and ligands in the complex KRAS G12C structure obtained after docking.

4.RESULTS

4.1 Targeting the G12D Mutant KRAS: Structural and Dynamic Insights

4.1.1 Molecular Dynamics Simulations

We performed molecular dynamics (MD) simulations of complexes of ZINC000008758797 with the G12D mutant and wild type KRAS to examine the binding behavior of the protein and the ligand, through three independent replicate 1-μs (1000 ns) runs for each case.

For the mutant-complex simulations, in the first trajectory, the ligand remained in the active site until 10 ns, after which it dissociated and subsequently associated with the region between Helix 3 and Helix 4 from approximately 210 to 750 ns. In the second trajectory, the ligand stayed bound within the active site for the entire 1-μs simulation. In the third trajectory, the ligand occupied the active site from the beginning of the simulation until 190 ns, then departed and re-bound at the Helix 3/Helix 4 interface from 450 to 1000 ns.

For the wild type protein and ligand complex, in the first trajectory, the ligand remained in the active site from the start until 160 ns, then dissociated and associated with the region between Helix 3 and Helix 4 during 210-740 ns. In the second trajectory, the ligand occupied the active site only up to 20 ns and then resided at the Helix 3/Helix 4 interface for the remainder of the simulation. In the third trajectory, the ligand stayed in the active site until 650 ns, followed by binding at the Helix 3/Helix 4 interface between 730 and 820 ns (Table 4.1 and Figure 4.1).

After analyzing the trajectories, the Root Mean Square Deviation (RMSD) profiles suggest that the mutant-ligand complex is generally more stable than the wild type-ligand complex. In particular, during the second mutant-ligand run, the

ligand stayed bound to the active site for the entire 1-μs simulation, the most stable behavior we observed. The third mutant-ligand run showed the next best stability: after 450 ns, the ligand moved to the allosteric site and remained bound there for the rest of the simulation (Figure 4.2.A).

Table 4.1 Simulations of the Top Candidate Molecule with Mutant and Wild Type KRAS and Its Binding Behavior During the Simulation.

Runs / **Sites**	**Mutant**			**Wild Type**		
	Active Site	Allosteric Site	Unbound	Active Site	Allosteric Site	Unbound
Run 1	0 - 10 ns	210 - 750 ns	750 - 1000 ns	0 - 160 ns	210 - 740 ns	-
Run 2	0 - 1000 ns	-	-	0 - 20 ns	20 - 1000 ns	-
Run 3	0 - 190 ns	450 - 1000 ns	-	0 - 650 ns	730 - 820 ns	650 - 730, 820 - 1000 ns

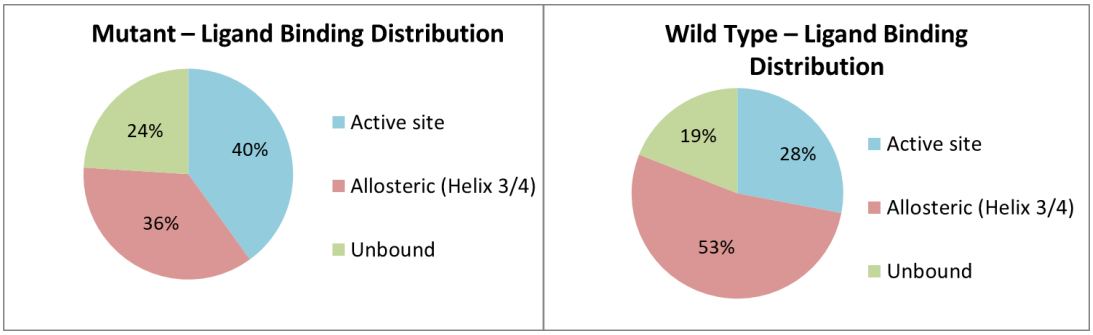

Figure 4.1 Regional Distribution of the Top Candidate Molecule Occupancy in Mutant and Wild Type.

Based on the Radius of Gyration (Rg) analysis, the mutant–ligand complex generally maintains a more compact conformation than the wild type-ligand complex across simulations. Among the mutant-ligand replicates, simulation 3 showed the greatest compactness, followed by simulation 2 (Figure 4.2.B).

Accounting for the fact that the RMSF traces were averaged over frames in which the ligand was bound either to the active site or to the allosteric pocket. With

active site binding, the mutant-ligand complex shows lower fluctuations across the P-loop (aa 10-17) and Switch I/II (aa 30-40 and 58-72) than the wild type. When the ligand occupies the allosteric site, fluctuations around residues 90-94 and 133-137 are likewise reduced in the mutant, importantly, under this allosteric bound condition the active site regions exhibit higher RMSF than in the active bound condition. (Figure 4.2.C).

From the hydrogen bond analysis, when the ligand is bound at the active site, the time-averaged number of hydrogen bonds formed with the mutant is greater than that with the wild type. The same trend holds when the ligand bound to the allosteric site: the mutant-ligand pair exhibits a higher hydrogen-bond count than the wild type-ligand pair (Figure 4.2.D).

4.1.2 Cross-Correlation Analysis

In MD simulations with the ligand bound to the active pocket of the mutant KRAS, cross correlation analysis shows positive correlations of the P-loop region (aa 10-17) with the active site (Figure 4.3.A). In contrast, there are negative correlations between the P-loop and Switch II (aa 58-72).

In the wild type with the ligand bound to the active pocket, cross correlation analysis shows positive correlations around the P-loop (aa 10-17) are short-range and weak (Figure 4.3.B). Also, negative correlation between P-loop (aa 10-17) and Switch II (aa 58-72) is much less pronounced.

When the cross-correlation analyses with the ligand bound to the active site are compared between the mutant and the wild type, the ligand shows a greater stabilizing effect in the mutant than in the wild type (Figure 4.3.A).

In MD simulations with the ligand bound to the allosteric pocket, cross correlation analysis shows positive correlations between the P-loop (aa 10-17) and the allosteric residues (F90-H94, L133-Y137). In contrast, the allosteric residues display negative correlation with Switch II (aa 58-72).

In the wild type with the ligand bound to the allosteric pocket, cross correlation analysis shows positive correlation between the allosteric residues (F90-H94, L133-Y137) and the P-loop (aa 10-17) but the correlation is weaker than in the mutant. By contrast, correlations between the allosteric residues and Switch II (aa 58-72) are mostly neutral to weakly negative (Figure 4.3.B).

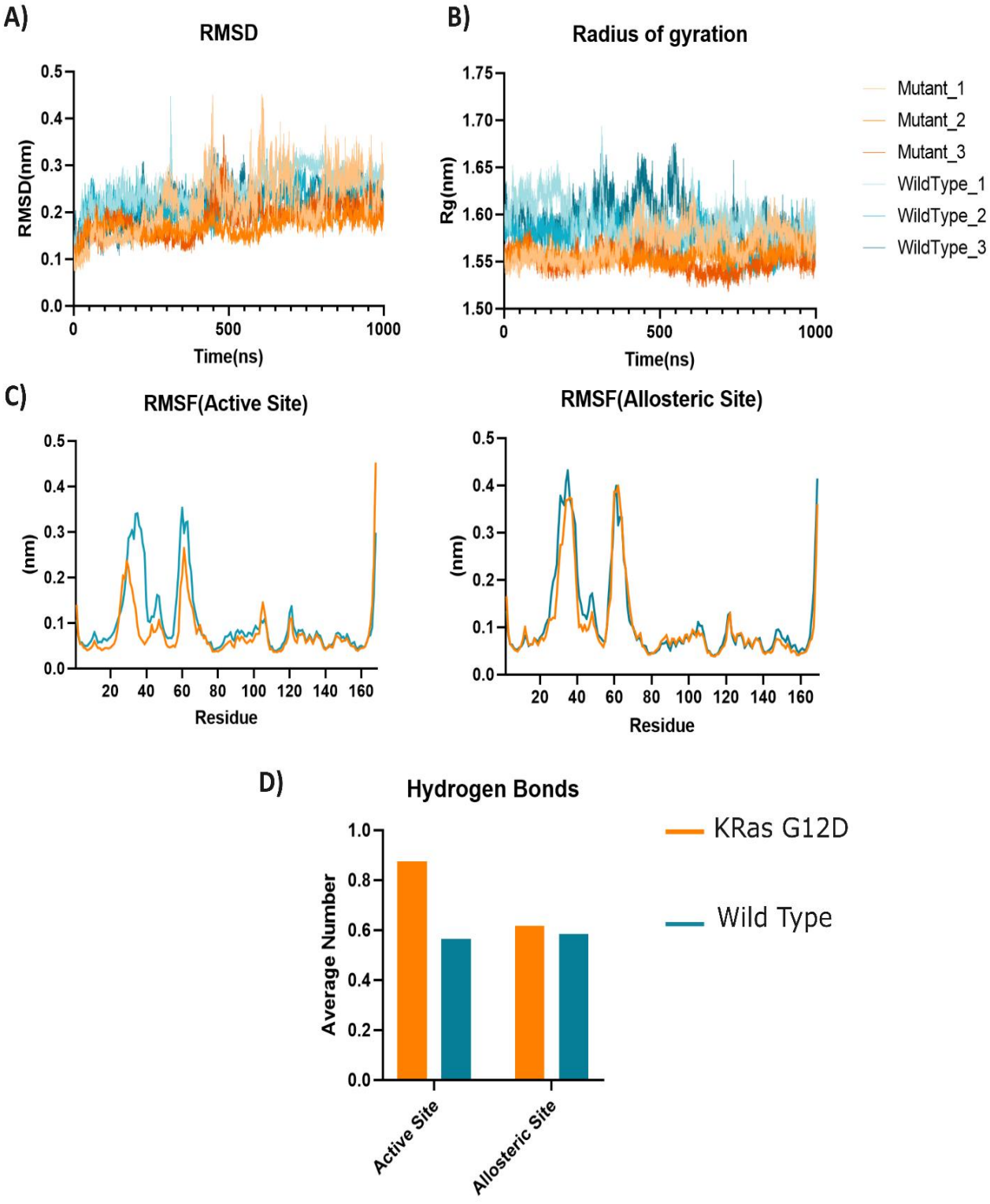

Figure 4.2 A) RMSD, B) Radius of Gyration C) RMSF, and D) Hydrogen Bond graphs from the molecular dynamics simulations of KRAS G12D and Wild Type.

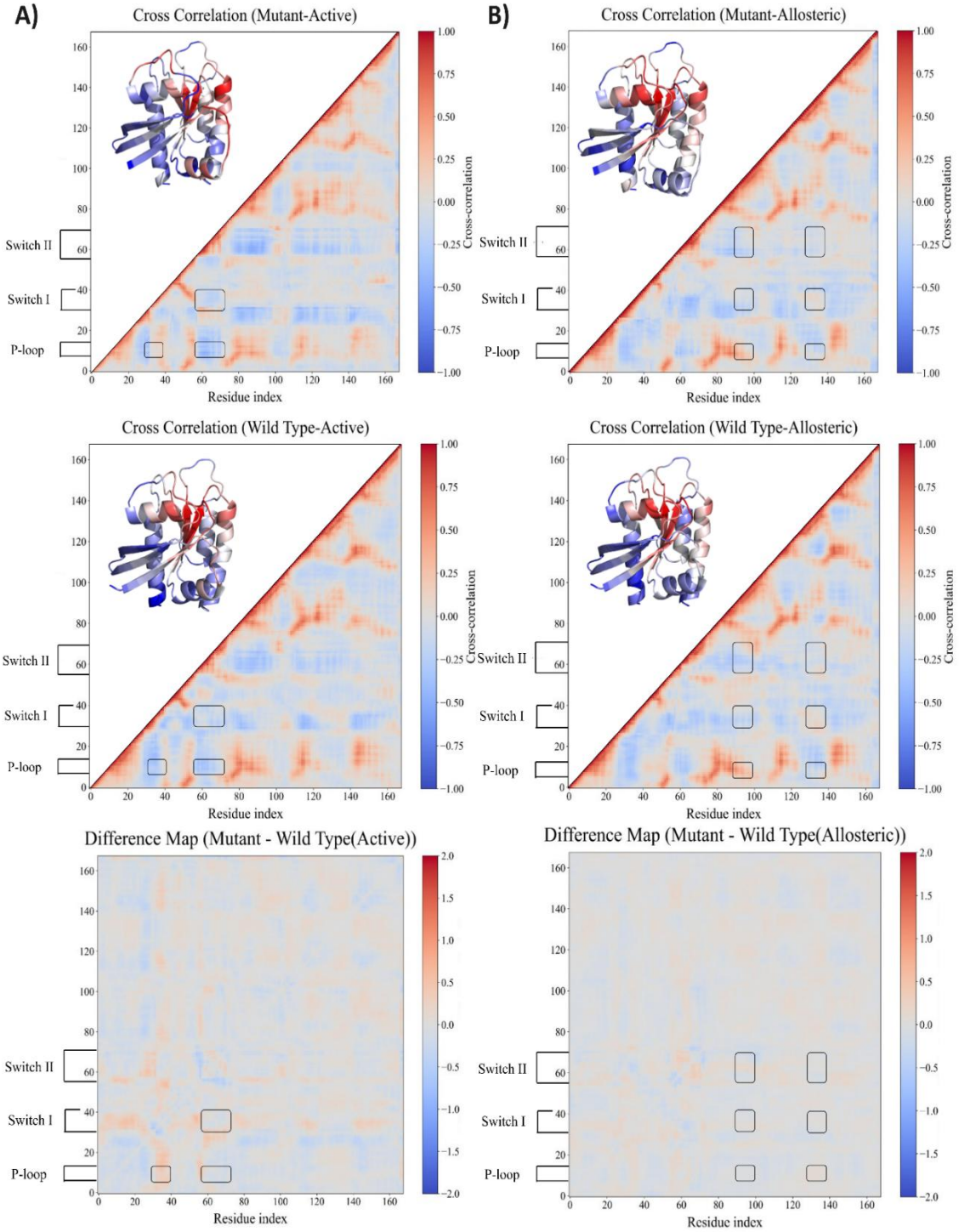

Figure 4.3 Cross-Correlation and Difference (Mutant-Wild Type) Maps Between Mutant and Wild Type with Ligand Bound to A) Active Site and B) Allosteric Site.

4.1.3 Umbrella Sampling

In this thesis, we define the binding free energy landscape along a relevant reaction coordinate by $\Delta G = F(max) - F(min)$, which quantifies the barrier height

between the most stable basin (at F(min)) and the transition region (at F(max)). This expression captures the kinetic aspect of the process: larger barriers imply slower transitions, even when the overall binding is favorable.

When the ligand was bound to the active site in the mutant structure, analysis of the umbrella sampling results showed a minimum value of -1.55 kcal/mol and a maximum value of 8.92 kcal/mol. According to the Gibbs free energy calculation, $\Delta G = 8.92 - (-1.55) = 10.47$ kcal/mol. When the ligand was bound to the active site in the wild type structure, the umbrella sampling results showed a minimum of -2.67 kcal/mol and a maximum of 16.33 kcal/mol. Based on the Gibbs free energy calculation, $\Delta G = 16.33 - (-2.67) = 19.00$ kcal/mol.

When the ligand was bound to the allosteric site in the mutant structure, analysis of the umbrella sampling results showed a minimum of -2.60 kcal/mol and a maximum of 5.06 kcal/mol. According to the Gibbs free energy calculation, $\Delta G = 5.06 - (-2.60) = 7.66$ kcal/mol. When the ligand was bound to the allosteric site in the wild type structure, the umbrella sampling results showed a minimum of -2.80 kcal/mol and a maximum of 4.64 kcal/mol. Based on the Gibbs free energy calculation, $\Delta G = 4.64 - (-2.80) = 7.44$ kcal/mol (Figure 4.4).

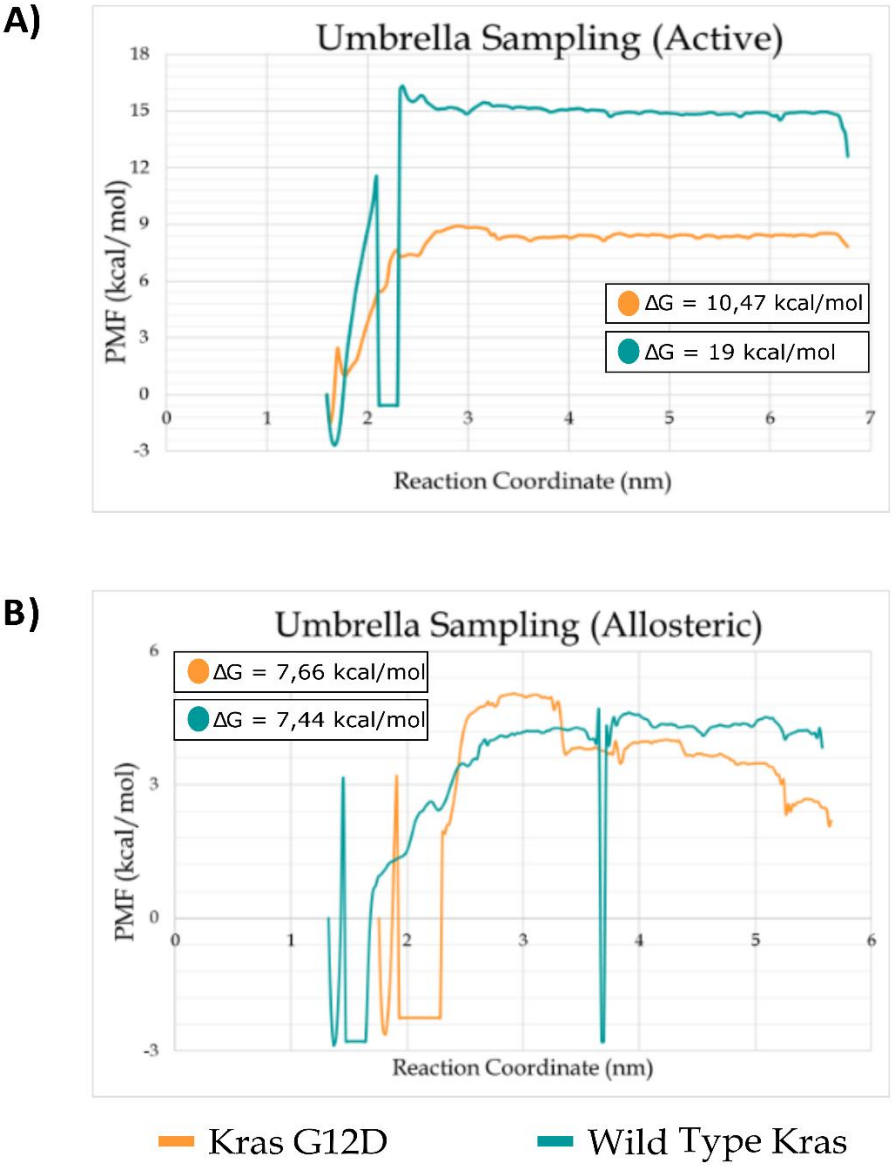

Figure 4.4 Potential of Mean Force (PMF) from Umbrella Sampling. A) Active site: Mutant vs. Wild Type, B) Allosteric site: Mutant vs. Wild Type.

4.1.4 Active and Allosteric Site Analysis with the Selected Inhibitor

For the interface analysis, residues involved in the mutant/wt-ligand interactions were identified using the PLIP web server (Schake et al., 2025) and the Protein-Ligand Interactions tool in Discovery Studio (BIOVIA, Dassault Systèmes, 2021).

According to PLIP, the ligand interacts with active site residues: GLU62, TYR64, SER65, ARG68, ASP69, TYR96, GLN99, ARG102, and VAL103. Discovery Studio reports interactions with GLN61, GLU62, TYR64, ARG68, MET72, TYR96, and ARG102. The residues common to both tools are GLU62, TYR64, ARG68, TYR96,

and ARG102 (Table 4.2). For the allosteric site, according to PLIP, the ligand interacts with GLU91, ASP132 and TYR137. Discovery Studio reports interactions with PHE90, HSD94, ARG97, GLU98, ASP132, LEU133 and SER136. The residue common to both tools are ASP132 (Table 4.3).

Table 4.2 Residue Level Interaction for the G12D mutant KRAS with the Ligand Bound in the Active Site.

Interacting Residues	**PLIP**	**Discovery Studio**
GLN61		X
GLU62	X	X
TYR64	X	X
SER65	X	
ARG68	X	X
ASP69	X	
MET72		X
TYR96	X	X
GLN99	X	
ARG102	X	X
VAL103	X	

Table 4.3 Residue Level Interaction for the G12D mutant KRAS with the Ligand Bound in the Allosteric Site.

Interacting Residues	**PLIP**	**Discovery Studio**
PHE90		X
GLU91	X	
HSD94		X
ARG97		X
GLU98		X
ASP132	X	X
LEU133		X
SER136		X
TYR137	X	

4.1.5 Analog Molecule Library

For constructing bioisosteric analog molecules with respect to the active and allosteric sites, the rules from data mining method, deep generative model and similarity comparison methods were used to select modification regions, different functional groups or atoms. We thus prioritized modifying the non-interacting portions of the active site ligand, specifically, the disubstituted diazine ring (two nitrogens) bearing two carbon substituents, the N-substituted indole, and the chlorine (Cl) atom. For modification of the non-interacting portions of the allosteric site ligand, specifically, the disubstituted diazine ring (two nitrogens) bearing two carbon substituents, O-linked phenyl (phenoxy) group and the carbon-nitrogen main scaffold were selected.

According to these selected regions (Tables 4.2 and 4.3, residues indicated as bold), a total of 5646 molecules were generated for the active site inhibitor analogs using the MolOpt server: 393 molecules from the data mining method, 2875 molecules from the deep generative model, and 2378 molecules from the similarity comparison method (Figures 4.5 - 4.7). On the other hand, a total of 4142 molecules were generated for the allosteric site inhibitor analogs with the MolOpt server: 342 molecules from the data mining method, 1741 molecules from

the deep generative model, and 2059 molecules from the similarity comparison method (Figures 4.8 - 4.10).

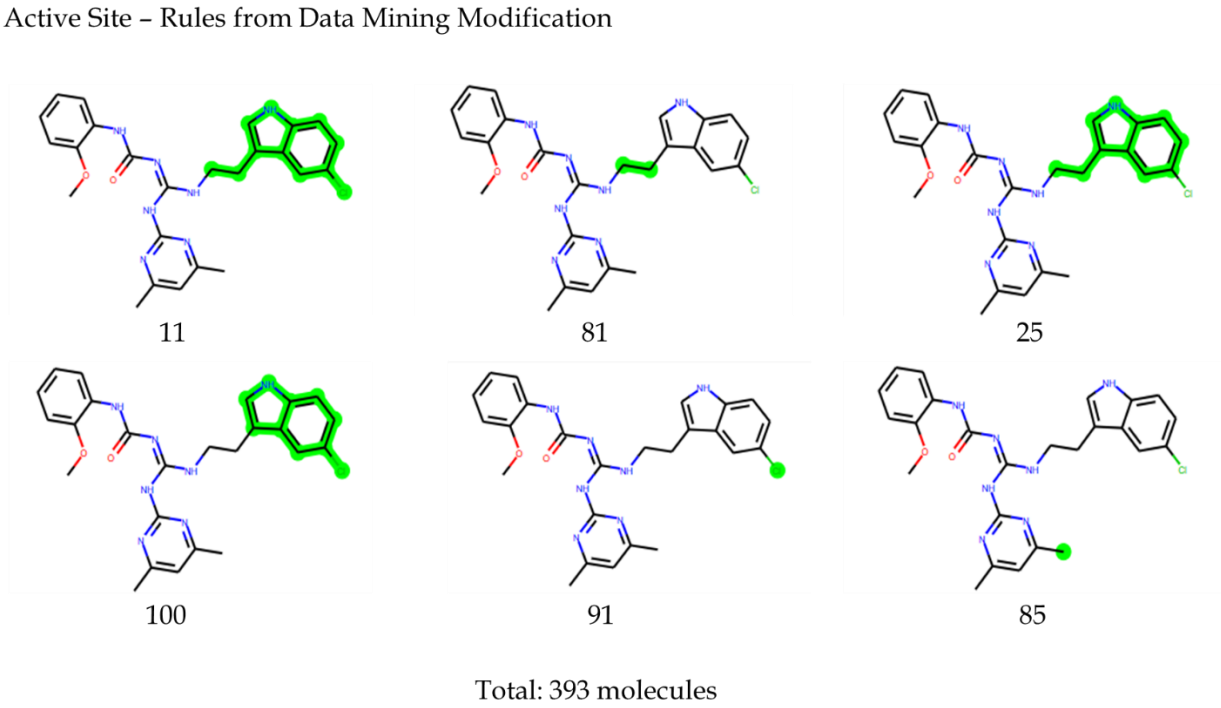

Figure 4.5 Regions of the Ligand Modified for Active Site Interacting Functional Groups / Atoms Revealed by Data Mining.

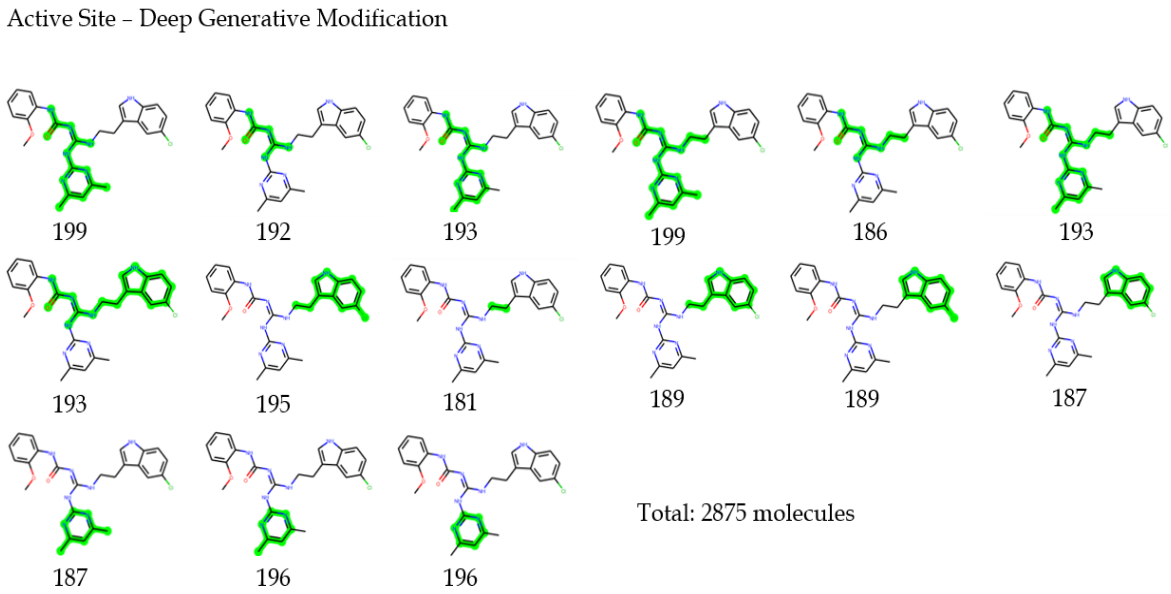

Figure 4.6 Regions of the Ligand Modified for Active Site Interacting Functional Groups / Atoms Revealed by Deep Generative Modification.

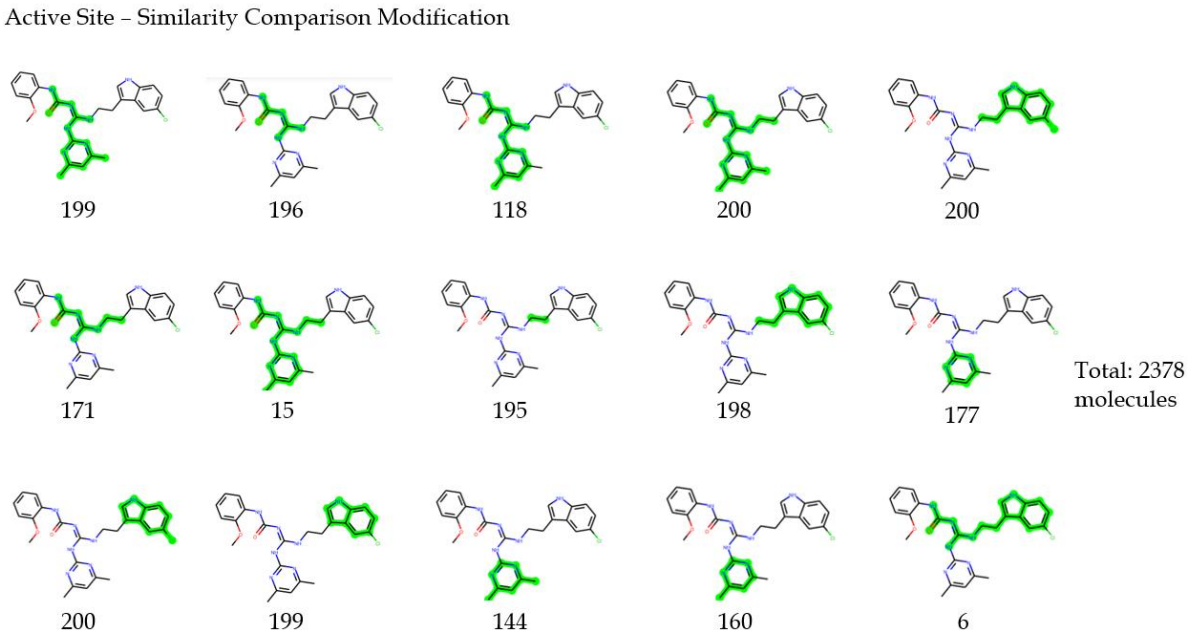

Figure 4.7 Regions of the Ligand Modified for Active Site Interacting Functional Groups / Atoms Revealed by Similarity Comparison.

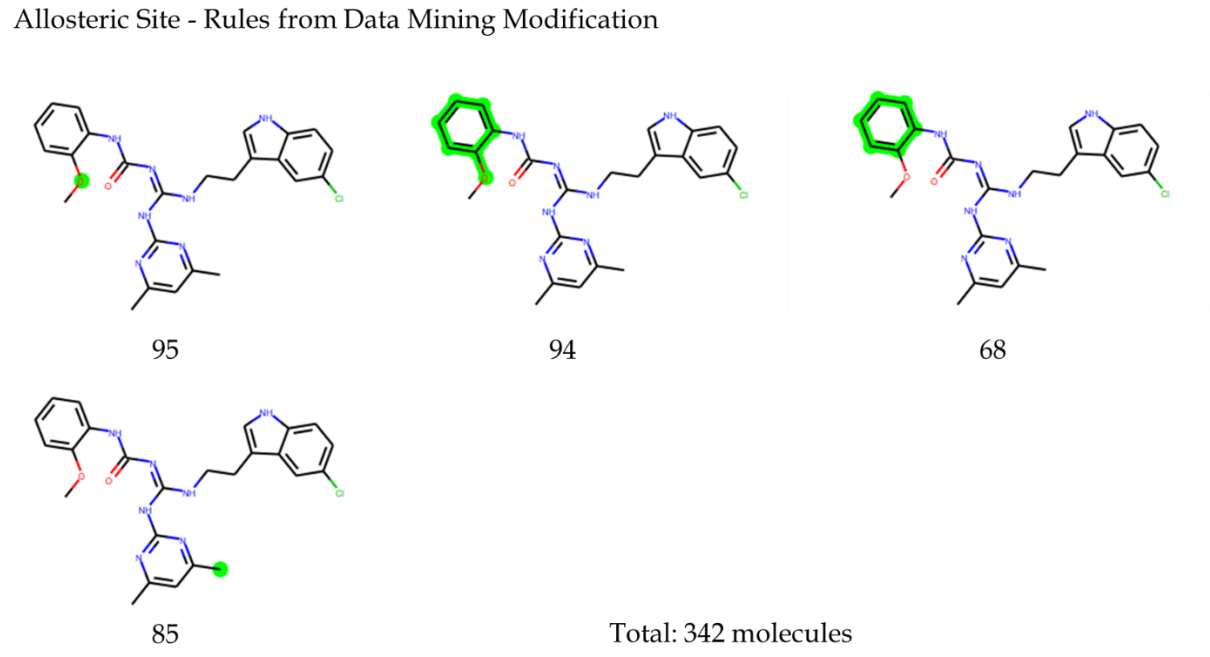

Figure 4.8 Regions of the Ligand Modified for Allosteric Site Interacting Functional Groups / Atoms Revealed by Data Mining.

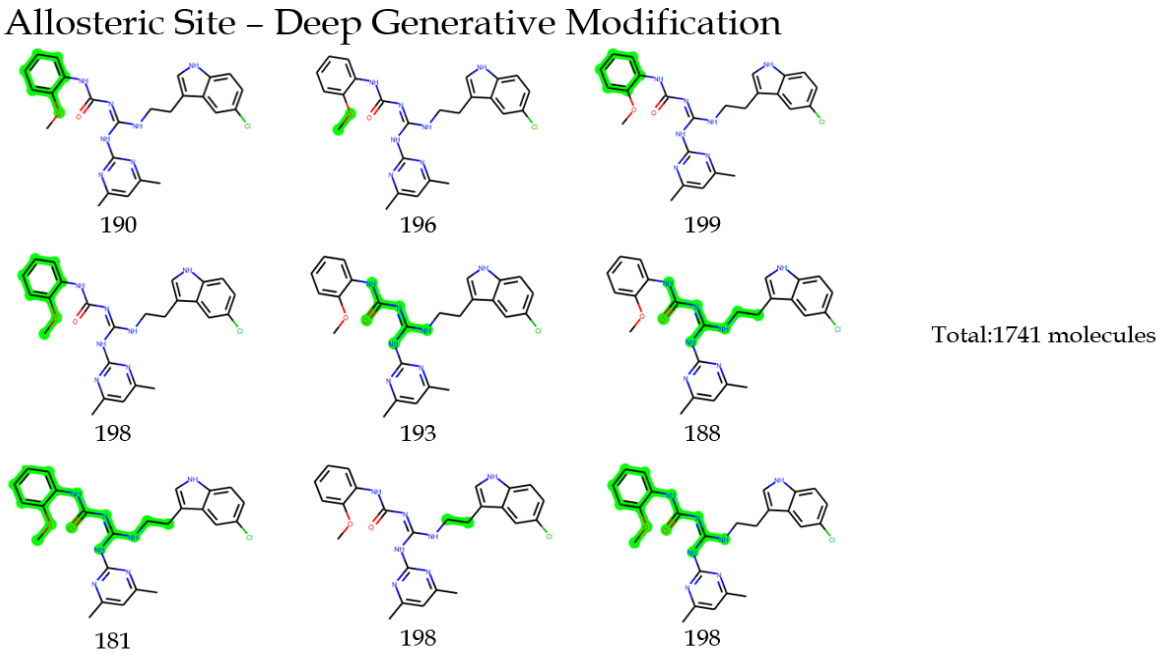

Figure 4.9 Regions of the Ligand Modified for Allosteric Site Interacting Functional Groups / Atoms Revealed by Deep Generative Modification.

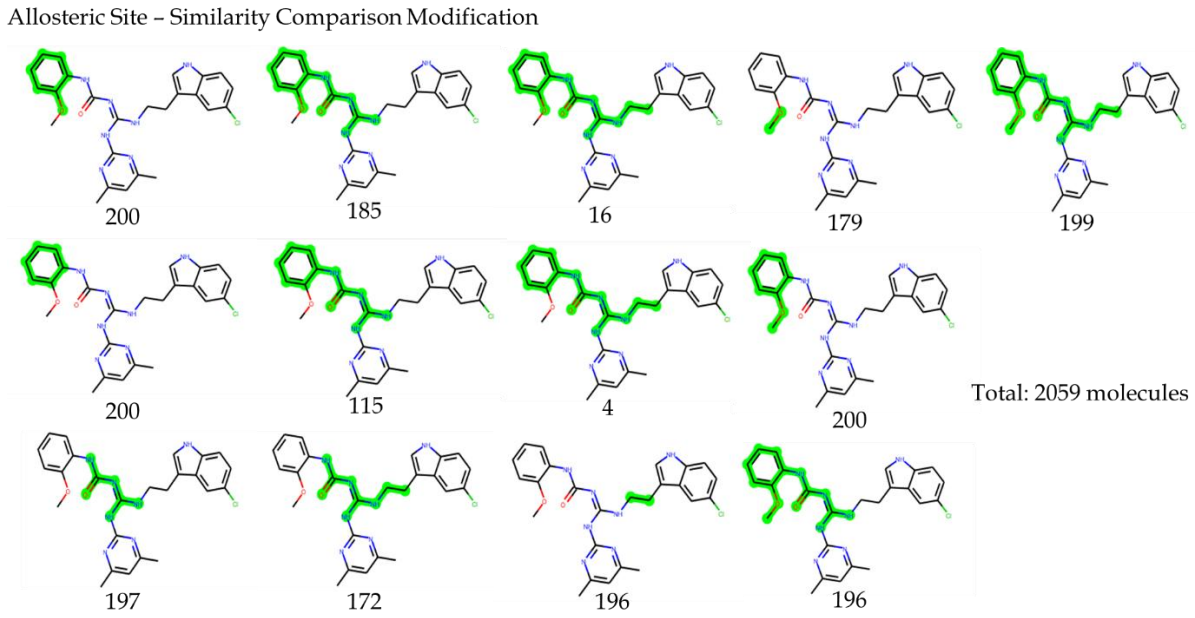

Figure 4.10 Regions of the Ligand Modified for Allosteric Site Interacting Functional Groups / Atoms Revealed by Similarity Comparison.

4.1.6 Molecular Docking

For KRAS G12D, the docking scores of the analog molecules of ZINC000008758797, which were generated by the Data Mining method based on the active site, are presented in Table 4.4. Among the analog molecules generated using the Data Mining method, the best docking score (-7.85233) was obtained for the analog with molecule ID 1099339.

Table 4.4 Active Site Analog Molecules (from Data Mining): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
1099339	-7.85233	5	8	525.57	170.78	2.71	3.91
5872702	-7.38133	4	5	612.16	116.32	4.79	6.62
5013692	-7.339	4	8	525.57	159.92	2.71	3.9
5236032	-7.31633	6	6	507.98	154.37	2.83	3.77
5170273	-7.27767	6	6	519.99	166.19	2.29	3.78
1291715	-7.193	4	6	556.01	125.55	3.97	4.2
5490605	-7.18633	4	8	525.56	155.24	2.44	3.98
3103294	-7.17033	4	6	485.84	133.39	1.97	3.65
2531283	-7.16133	4	5	612.16	116.32	4.79	6.62
3770040	-7.15533	5	6	544.01	145	2.58	3.91

The docking scores of the analog molecules generated by the Deep Generative method based on the active site are presented in Table 4.5. Among the analog molecules generated using the Deep Generative method, the best docking score (-7.72433) was obtained for the analog with molecule ID 5052871.

Table 4.5 Active Site Analog Molecules (from Deep Generative): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
5052871	-7.72433	3	5	498.96	104.82	3.17	3.2
7446167	-7.70933	4	4	512.99	103.43	3.87	3.75
4302027	-7.606	6	5	467.91	139.45	2.8	3.61
4443359	-7.57433	4	6	529.98	129.46	3.35	3.98
3974494	-7.48767	4	7	460.49	138.59	2.14	4.11
2990178	-7.46067	3	3	461.94	83.22	4.02	3.07
4487108	-7.459	4	5	506	116.32	2.87	3.91
6257233	-7.45167	4	6	495.94	116.32	3.24	3.65
6939128	-7.451	5	4	501.97	119.22	3.81	3.74

The docking scores of the analog molecules generated by the Similarity Comparison method based on the active site are presented in Table 4.6. Among the analog molecules generated using the Similarity Comparison method, the best docking score (-8.35933) was obtained for the analog with molecule ID 8929323.

Table 4.6 Active Site Analog Molecules (from Similarity Comparison): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
8929323	-8.35933	1	2	482	34.25	5.51	3.4
3302325	-7.73933	2	4	510	75	3.85	3.85
4609934	-7.66067	1	4	510	67	4.17	3.58
5576702	-7.659	2	4	512	71	4.24	4.17
5179173	-7.586	3	5	565.01	86.46	4.79	4.78
3109754	-7.58533	5	5	497.93	139.95	2.56	3.72
6772048	-7.576	1	4	522.04	68.62	4.23	3.59
2218051	-7.56767	2	3	470.99	66.59	3.44	3.94
5289413	-7.56333	1	7	527.96	103.4	3.76	3.6
5791629	-7.56267	2	5	531.96	53.7	6.37	3.46
7427027	-7.551	1	4	498.02	60.03	4.59	3.64
2347817	-7.518	1	4	512.04	60.03	4.19	4.19

The docking scores of the analog molecules generated by all three methods based on the active site are presented in Table 4.7. In the docking results, the analog with molecule ID 8929323 ranks first with the most favorable docking score of −8.35933, followed by the analogs with molecule IDs 1099339, 3302325, 5052871, 7446167, 4609934, 5576702, 4302027, 5179173, and 3109754, which exhibit docking scores of −7.85233, −7.73933, −7.72433, −7.70933, −7.66067, −7.659, −7.606, −7.586, and −7.58533, respectively.

Table 4.7 Active Site Analog Molecules (from all methods): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
8929323	-8.35933	1	2	482	34.25	5.51	3.4
1099339	-7.85233	5	8	525.57	170.78	2.71	3.91
3302325	-7.73933	2	4	510	75	3.85	3.85
5052871	-7.72433	3	5	498.96	104.82	3.17	3.2
7446167	-7.70933	4	4	512.99	103.43	3.87	3.75
4609934	-7.66067	1	4	510	67	4.17	3.58
5576702	-7.659	2	4	512	71	4.24	4.17
4302027	-7.606	6	5	467.91	139.45	2.8	3.61
5179173	-7.586	3	5	565.01	86.46	4.79	4.78
3109754	-7.58533	5	5	497.93	139.95	2.56	3.72

The docking scores of the analog molecules generated by the Data Mining method based on the allosteric site are presented in Table 4.8. Among the analog molecules generated using the Data Mining method, the best docking score (-7.772) was obtained for the analog with molecule ID 3243164.

Table 4.8 Allosteric Site Analog Molecules (from Data Mining): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
3243164	-7.772	4	7	545	147.03	2.71	3.96
2531283	-7.37	4	5	612.16	116.32	4.79	6.74
4462927	-7.354	4	4	552.07	107.09	4.55	4.11
4344757	-7.35067	5	5	491.97	127.32	3.08	3.63
1182517	-7.318	5	5	492.96	142.34	2.78	3.66
5795557	-7.25767	5	5	491.97	127.32	3.08	3.63
5067776	-7.19933	4	4	554.87	107.09	3.34	3.68
5721677	-7.19467	4	7	562.05	170.34	2.39	4.1
3789213	-7.11967	5	5	519	136.19	2.84	3.83
8275954	-7.11933	4	5	542.03	120.23	2.61	4.23

The docking scores of the analog molecules generated by the Deep Generative method based on the allosteric site are presented in Table 4.9. Among the analog molecules generated using the Deep Generative method, the best docking score (-7.60733) was obtained for the analog with molecule ID 8912677.

Table 4.9 Allosteric Site Analog Molecules (from Deep Generative Model): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
8912677	-7.60733	6	5	503.99	156.96	2.24	3.77
8075047	-7.57233	4	5	482.97	127.4	1.9	3.59
7919766	-7.50733	5	6	517.97	145.14	2.75	3.98
7850335	-7.396	4	5	531.01	127.4	2.77	3.93
6451109	-7.39067	4	6	572.03	116.32	3.67	4.03
8991732	-7.387667	4	6	502.96	133.12	2.17	3.89
5211816	-7.377	4	6	513.98	132.87	1.79	3.77
6116459	-7.36767	2	5	470.96	105.35	2.44	3.84
5654502	-7.33967	4	6	477.95	132.87	1.63	3.69
6401231	-7.32067	3	4	418.88	95.06	2.67	3.3

The docking scores of the analog molecules generated by the Similarity Comparison method based on the allosteric site are presented in Table 4.10. Among the analog molecules generated using the Similarity Comparison method, the best docking score (-7.961) was obtained for the analog with molecule ID 4083211.

Table 4.10 Allosteric Site Analog Molecules (from Similarity Comparison): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
4083211	-7.961	1	7	597.01	84.25	4.56	4.45
5267697	-7.85267	2	5	519.98	92.79	3.42	3.5
3810718	-7.80133	1	4	521.01	78.95	4.78	3.46
4621950	-7.73867	4	4	486.95	115.56	2.43	3.99
5733578	-7.685	1	3	522.08	50.8	5.17	3.89
4195506	-7.66167	3	6	532.95	82.28	4.76	3.48
1342773	-7.642	1	5	561.03	84.25	4.11	4.35
5532744	-7.59967	1	5	566.04	50.8	6.13	3.9
6956576	-7.597	3	4	484	91.41	3.35	3.6
2984330	-7.56233	2	4	517.04	96.12	3.8	3.55

The docking scores of the analog molecules generated by all three methods based on the allosteric site are presented in Table 4.11. In the docking results, the analog with molecule ID 8929323 ranks first with the most favorable docking score of −7.961, followed by the analogs with molecule IDs 5267697, 3810718, 3243164, 4621950, 5733578, 4195506, 1342773, 8912677 and, 5532744, which exhibit docking scores of -7.961, -7.85267, -7.80133, -7.772, -7.73867, -7.685, -7.66167, -7.642, -7.60733, -7.59967, respectively.

Table 4.11 Allosteric Site Analog Molecules (from all methods): Docking Score and Physicochemical Descriptors.

Molecule ID	Binding Affinity (kcal/mol)	Num Lipinski HB Donor	Num Lipinski HB Acceptor	Mol Wt (g/mol)	Topological Polar SA (Å²)	LogP	Synthetic Accessibility
4083211	-7.961	1	7	597.01	84.25	4.56	4.45
5267697	-7.85267	2	5	519.98	92.79	3.42	3.5
3810718	-7.80133	1	4	521.01	78.95	4.78	3.46
3243164	-7.772	4	7	545	147.03	2.71	3.96
4621950	-7.73867	4	4	486.95	115.56	2.43	3.99
5733578	-7.685	1	3	522.08	50.8	5.17	3.89
4195506	-7.66167	3	6	532.95	82.28	4.76	3.48
1342773	-7.642	1	5	561.03	84.25	4.11	4.35
8912677	-7.60733	6	5	503.99	156.96	2.24	3.77
5532744	-7.59967	1	5	566.04	50.8	6.13	3.9

4.1.7 Molecular Dynamics Simulations for Analog Inhibitors

Molecular dynamics simulations of 100 nanoseconds were performed for the top 10 best-scoring molecules from the library generated for the active site. Since the analog molecules with ID 8929323, 3302325, 5179173 and 5576702 remained in the active site, the simulations was extended to 1000 nanoseconds, and molecules did not leave the active site throughout the simulation. After the analog molecules with ID 1099339 and 4302027 stayed in the active site, the simulation was extended to 200 nanoseconds for 1099339 and 700 nanoseconds for 4302027; however, during the continuation of the simulation, the molecule moved out of the binding pocket. In the 100-nanosecond simulations of the analog molecules with IDs 5052871, 7446167, 4609934 and 3109754 the molecules dissociated from the protein. (Table 4.12).

Table 4.12 Molecular Dynamics Simulation Duration and Binding Behavior of the Top 10 Active Site Analog Molecules.

Molecule ID	Simulation Time	Binding Behavior
8929323	1000 ns	✓
1099339	200 ns	**x**
3302325	1000 ns	✓
5052871	100 ns	**x**
7446167	100 ns	**x**
4609934	100 ns	**x**
5576702	1000 ns	✓
4302027	700 ns	**x**
5179173	1000 ns	✓
3109754	100 ns	**x**

When the post-MD analysis results of the molecule with ID 8929323 were examined, the RMSD values were found to be generally stable. The RMSD values started around 0.08-0.15 nm and increased during the simulation, reaching higher levels of 0.30-0.40 nm in the middle stages, and then fluctuating mostly around 0.22-0.30 nm toward the end. The Radius of Gyration graph indicated that the complex structure remained compact throughout the simulation. The Rg values were generally within 1.53-1.60 nm, with brief rises up to 1.62 nm. In the RMSF graph, a significant fluctuation was observed particularly in the Switch I region (aa 30-40). The average hydrogen bond number was 0.27 (Figure 4.11).

For molecule with ID 3302325, the RMSD values were mostly in the 0.18-0.35 nm range, with higher peaks up to 0.45 nm, followed by values of 0.22-0.30 nm in the later stages. The Radius of Gyration fluctuated mainly between 1.56 and 1.60 nm, with a transient increase up to 1.63 nm. The RMSF analysis showed the highest peak around Switch I (aa 30-40) and, another peak around Switch II (aa 60-70). The Hydrogen Bond analysis showed an average of 0.58 hydrogen bonds (Figure 4.11).

For molecule with ID 5179173, the RMSD values were relatively stable. The RMSD started around 0.10-0.16 nm and mostly fluctuated in the 0.14-0.25 nm range throughout the simulation, indicating a generally equilibrated complex.

The Radius of Gyration profile suggested that the complex largely preserved a compact structure, with Rg values mainly within 1.55-1.60 nm, while showing a transient increase near the end. In the RMSF graph, the most pronounced flexibility was again observed in the Switch I region (aa 30-40), with additional fluctuations around Switch II (aa 60-70), whereas the remaining regions stayed comparatively stable. The average hydrogen bond number was 0.75 (Figure 4.11).

For molecule with ID 5576702, the RMSD values showed fluctuations with an overall stable trend. The RMSD values began around 0.12-0.18 nm and increased during the simulation, generally remaining in the 0.20-0.35 nm range, with occasional peaks approaching 0.38-0.40. The Radius of Gyration values were largely confined to 1.56-1.61 nm, pointing to a compact complex, with only minor transient deviations. The RMSF analysis indicated the highest flexibility in Switch I (aa 30-40) and a noticeable secondary fluctuation around Switch II (aa 60-70). The Hydrogen Bond analysis showed an average of 1.09 hydrogen bonds (Figure 4.11).

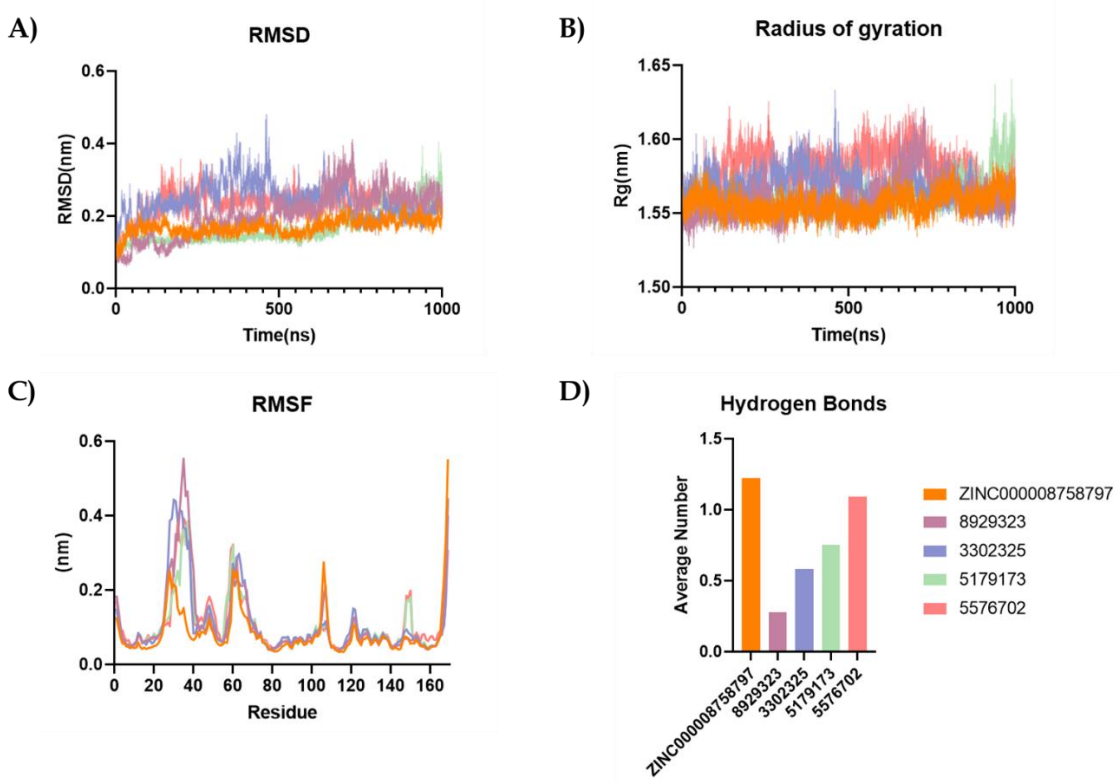

Figure 4.11 A) RMSD, B) Radius of Gyration C) RMSF, and D) Hydrogen Bond graphs from the molecular dynamics simulations of KRAS G12D with molecule ZINC000008758797 (second MD run), 8929323, 3302325, 5179173 and 5576702.

Molecular dynamics simulations of 100 nanoseconds were performed for the top 10 best-scoring molecules from the library generated for the allosteric site. During the 100 ns simulations of compounds with IDs 4083211 and 5267697, the

molecules did not remain in the allosteric site and were observed to drift away from this region. Since compound 3810718 remained stable in the allosteric site, its simulation was extended to 400 ns. For compound 4621950, the simulation had previously been extended to 700 ns because the molecule initially remained in the allosteric site; however, it was later observed to move away from the allosteric site. The 100 ns simulations of compounds with IDs 3243164, 5733578, 4195506, 1342773, 8912677, and 5532744 are still ongoing (Table 4.13).

Table 4.13 Molecular Dynamics Simulation Duration and Binding Behavior of the Top 10 Allosteric Site Analog Molecules.

Molecule ID	Simulation Time	Binding Behavior
4083211	100 ns	**x**
5267697	100 ns	**x**
3810718	400 ns	continues
3243164	100 ns	continues
4621950	700 ns	**x**
5733578	100 ns	continues
4195506	100 ns	continues
1342773	100 ns	continues
8912677	100 ns	continues
5532744	100 ns	continues

4.1.8 Umbrella Sampling for Analog Inhibitors

Umbrella sampling simulations were performed for the four active site analog compounds that remained in the active site throughout the entire simulation period. The 60 nanosecond umbrella sampling simulation of analog molecule 8929323 has been completed, whereas the umbrella sampling simulations of the other three compounds are still ongoing (Table 4.14). For analog molecule 8929323, the free energy difference was calculated using the equation $\Delta G = F_{max} - F_{min}$, yielding $\Delta G = 2.92 - (-4.01) = 6.93$ kcal/mol.

Table 4.14 Umbrella Sampling Simulation Time and Status of Active Site Analog Molecules.

Molecule ID	Simulation Time	Simulation Status
8929323	60 ns	finished
3302325	260 ns	continues
5179173	250 ns	continues
5576702	260 ns	continues

4.2 Targeting the G12C Mutant KRAS: Structural and Dynamic Insights

4.2.1 Molecular Library and Virtual Screening

Filtering in the Tranches section of the ZINC20 database (Irwin et al., 2020) resulted in the download of 11,856,000 protomers. After extracting the files and combining the molecules into a single sdf file using Open Babel (O'Boyle et al., 2011), 11,429,395 molecules were obtained. Applying Lipinski's rule of five in KNIME (Berthold et al., 2007) resulted in 8,720,782 molecules. Using the 3D-D Similarity node with positive control molecules, 78 molecules were obtained with the Sotorasib inhibitor, 16 with the Adagrasib inhibitor, and 35 with the ARS1620 inhibitor. At the end of the virtual screening phase, 129 molecules were obtained for the docking phase, and a total of 133 molecules were obtained with the control molecules.

4.2.2 Covalent Docking

As a result of docking 133 molecules with the Cresset Flare program software (Cresset, 2024), (Cheeseright et al., 2006), (Bauer and Mackey, 2019), (Kuhn et al., 2020), only 26 of the 129 molecules, excluding control molecules, were covalently bound to KRAS G12C. The docking scores for the positive control inhibitors Sotorasib, Adagrasib, and ARS-1620 were calculated as -4.767, -5.902, and -5.111, respectively. The docking score for the negative control DMSO molecule could not be calculated by the program because it could not form a covalent bond. The scores of the top 10 molecules with the best scores are summarized in the Table 4.15.

Table 4.15 Docking Score of Covalent and Positive Control Molecules.

Covalent and Control Molecules	Binding Affinity (kcal/mol)
ZINC000000626790	-5.972
ZINC000057607951	-5.946
Adagrasib	-5.902
ZINC000003005730	-5.726
ZINC000100738897	-5.431
ZINC000004760237	-5.423
ARS-1620	-5.111
ZINC000009850516	-5.074
ZINC000020740022	-5.054
ZINC000033551403	-5.012
ZINC000027017866	-4.973
ZINC000012056017	-4.902
Sotorasib	-4.767

4.2.3 Active Site Analysis with the Selected Covalent Inhibitors

When the amino acids interacting with the best-scoring molecules, ZINC000000626790 and ZINC000057607951, other than CYS 12 in the active pocket of the protein, are examined, the molecule named ZINC000000626790 interacted with the amino acids ALA59, GLU62, GLN61, MET67, and GLY60, while the molecule named ZINC000057607951 interacted with the amino acids PRO34, GLY60, TYR96, ASN86, ALA11, and TYR32 (Figure 4.12). Among the positive control molecules, Sotorasib interacted with GLU62, GLN61, and GLU63, while Adagrasib interacted with GLU62, GLN61, and GLU63, HSD95, ASP92, and LYS88, and ARS-1620 interacted with ALA59, PRO34, GLU62, THR35, GLU37, and MET67. Among the amino acids interacted with by the molecule ZINC000000626790, the following amino acids were also observed in the positive control molecules: GLU62 and GLN61 in the Sotorasib inhibitor, GLU62 and GLN61 in the Adagrasib inhibitor, and ALA59, GLU62, and MET67 in the ARS-1620 inhibitor. Among the amino acids that the ZINC000057607951

molecule interacts with, the ARS-1620 inhibitor interacts with the PRO34 amino acid (Table 4.16).

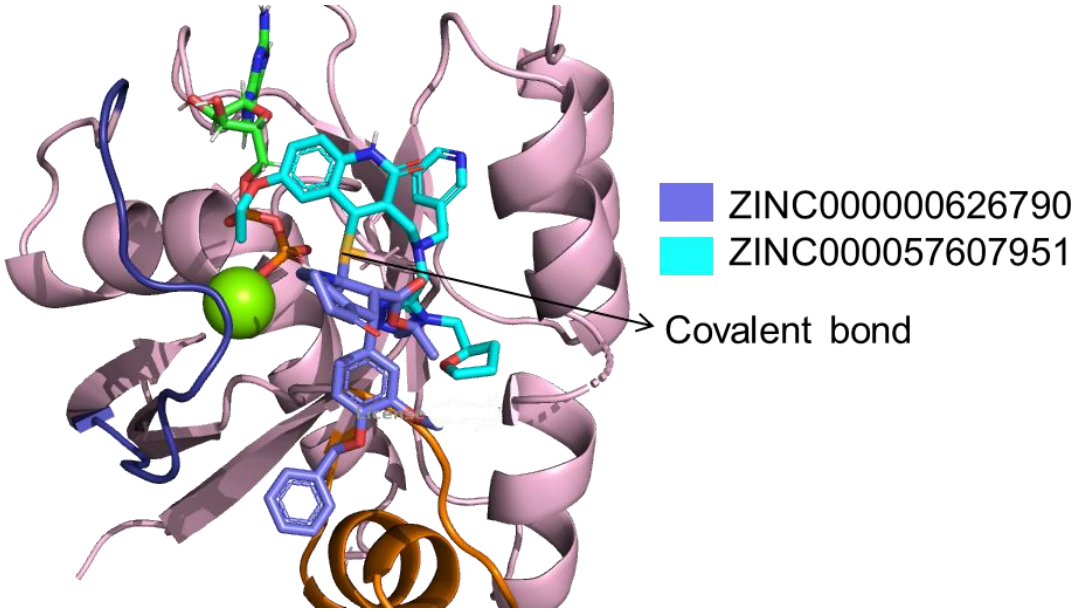

Figure 4.12 Complex Structure of the Best Candidate (ZINC000000626790 and ZINC000057607951) that Covalently Bind to the G12C Mutant KRAS.

Table 4.16 The Residues Interacting with the Best Covalent Inhibitor Candidates and Positive Controls.

Best Covalent Inhibitor Candidates and Positive Control Molecules	**Interacting Residues**
Sotorasib	GLU62, GLN61, GLU63
Adagrasib	GLU62, GLN61, GLU63, HSD95, ASP92, LYS88
ARS-1620	ALA59, PRO34, GLU62, THR35, GLU37, MET67
ZINC000000626790	ALA59, GLU62, GLN61, MET67, GLY60
ZINC000057607951	PRO34, GLY60, TYR96, ASN86, ALA11, TYR32

5.DISCUSSION

5.1 Targeting the G12D Mutant KRAS

When three replicate simulations of the G12D mutant KRAS structure with the molecule ID ZINC000008758797 were examined, it was observed that in two replicates, the molecule exited the active site and bound to an allosteric site between Helix 3 and Helix 4, while in the other replicate, it remained in the active site throughout the simulation. Similarly, in three replicate simulations of the

molecule with wild type KRAS, identified as ZINC000008758797, the ligand moves away from the active site and binds to the allosteric site in all replicate simulations. The fact that this situation was encountered repeatedly strengthened the possibility that this molecule could be an allosteric inhibitor.

In a 2011 study of binding sites on Ras, Buhrman et al. (Buhrman et al., 2011) suggested that this region between Helix 3 and Helix 4 (F90, E91, I93, H94, L133, S136, Y137) and interacting with the membrane could be an allosteric pocket. In this study, a hotspot was identified in the allosteric site between Helix 3 and Helix 4 to disrupt interactions of Ras with membranes and effector proteins. The conformational change in the effector domain extends deep into the structure and into the allosteric region, suggesting that these distant hotspots may be involved in allosteric regulation of the effector lobe or inhibition of its function. Two methods were used in the study: Multiple Solvent Crystal Structures (MSCS) and Computational Solvent Mapping (FTMap). MSCS is based on crystallizing the same protein in different solvents and analyzing these crystals using high-resolution X-ray crystallography. The target protein is crystallized in different organic solvents or solvent mixtures. These solvents localize to potential binding pockets on the protein surface. FTMap is a computational technique used to identify binding hotspots on the surface of biomolecules. Various small probe molecules (water, methanol, acetonitrile) are placed on the protein surface. These probe molecules are used to explore potential binding sites on the protein surface. Energy calculations identify the regions where the probe molecules bind most strongly. These regions are called binding hotspots and define potential ligand binding pockets on the protein surface. This pocket, located between Helix 3 and Helix 4, was also observed in a 2015 study by Prakash et al. (Prakash et al., 2015), which identified binding sites on KRAS using probe-based molecular dynamics analysis (pMD). Despite analyses suggesting that this pocket, which they named S4, was unsuitable for ligand binding, it was determined that small organic molecules could bind. S4 is considered a region that may be important for interactions with other proteins or molecules. It is thought to play a role in dimerization and interactions with biomolecules. When the HRAS complex crystal structure was used as a reference, some B2-B3 contact points appeared to coincide with a segment of the Sos interface involved in Ras binding. However, these contacts seem secondary because Ras-Sos association is driven largely by the Switch I and II regions.

Molecular dynamics simulations of the G12D mutant and Wild Type KRAS structures using the candidate molecule, revealed that the Switch I domain moved away from the active pocket with respect to to the starting structure, disrupting the structure of this pocket, which is the binding site for effector

proteins. In the mutant structure, the Switch II domain appeared to approach Helix 3 more closely, closing the pocket, with respect to to the starting structure (PDB ID: 7RPZ) (Figure 5.1).

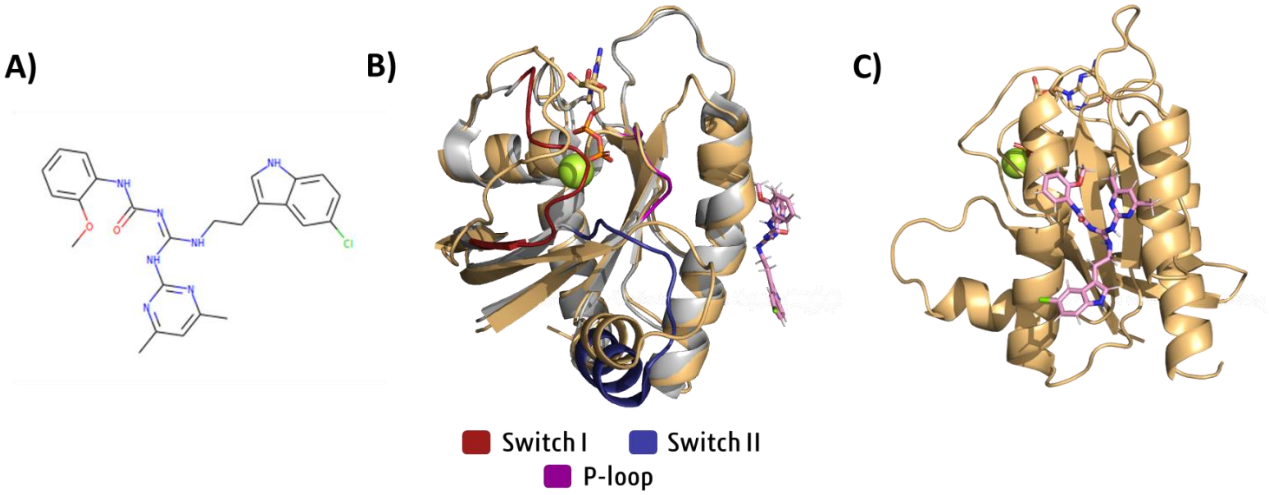

Figure 5.1 A) Two-dimensional chemical structure of ZINC000008758797. B) Alignment of the KRAS G12D structure obtained at the 700th nanosecond of the molecular dynamics simulation (first MD run) with the reference structure (PDB ID: 7RPZ). C) Ninety-degree rotated view of the KRAS G12D-ZINC000008758797 complex obtained at the 700th nanosecond, showing the drug candidate bound in the allosteric pocket formed between Helices α3 and α4.

The simulations of the mutant-ligand complex structure are generally more stable and compact than those of the wild type complex structure according to the RMSD and Rg analysis results. This suggests that the ligand's affinity for the mutant KRAS is greater than the wild type KRAS. In both complex structures, stability fluctuates as the ligand transitions from the active site to the allosteric site.

When the ligand is bound to the active site in the mutant-ligand complex, the P-loop (aa 10–17) and Switch I/II regions (aa 30-40 and 58-72) exhibit lower RMSF values compared with the wild type-ligand complex, which is consistent with tighter local stabilization. When the ligand is bound to the allosteric site, the fluctuations of the allosteric site residues are also reduced in the mutant complex relative to the wild type, again indicating a more compact allosteric pocket. At the same time, when the ligand is bound to the allosteric site, the RMSF values of the active site residues are higher than when the ligand is bound to the active site. This increase in active site mobility is consistent with allosteric coupling, suggesting that allosteric binding modulates active site flexibility.

In the hydrogen-bond analyses, the mutant complex forms a greater number of hydrogen bonds with the ligand than the wild type, both when the ligand is bound to the active site and when it is bound to the allosteric site, suggesting a higher affinity of the ligand for the mutant structure. Considered together with the other post-MD analyses, these results indicate that the ligand shows greater specificity for the mutant and forms a more favorable complex relative to the wild type (Figure 4.2).

In the cross-correlation analysis of the mutant complex with the ligand bound to the active site, the P-loop (aa 10–17) shows a positive correlation with the active-site region and a negative correlation with Switch II (aa 58-72). The positive correlation suggests same-phase motion and a more ordered phosphate-binding conformation, whereas the negative correlation points to opposite-phase behavior. This pattern indicates that Switch II (aa 58-72) tends to close over the active site, producing a more compact, catalytically ready state. In the wild type complex, the positive correlation involving the P-loop (aa 10-17) is weaker, indicating a less stabilized phosphate-binding pose. The negative correlation between the P-loop (aa 10-17) and Switch II (aa 58-72) is also weaker than in the mutant, which is consistent with the absence of a robustly stabilized closure over the active site. When the ligand is bound to the allosteric site, a positive correlation between the allosteric site residues and the P-loop (aa 10–17) is observed in the mutant complex. This indicates same-phase motion and a reduction in the independent opening/closing amplitude of the P-loop (aa 10–17), thereby stabilizing a more ordered phosphate-binding conformation. By contrast, a negative correlation between the allosteric residues and Switch II (aa 58-72) is consistent with an opposite-phase relationship and with the observed closure of Switch II (aa 58-72) over the active site in structural inspection. In the wild type complex, these correlations are weaker or closer to neutral, again indicating reduced stabilization and a less pronounced tendency of the active site to adopt a closed, binding-competent conformation. Together, these results suggest that allosteric binding in the mutant stabilizes the P-loop (aa 10-17) while driving Switch II (aa 58-72) toward a closed state, thereby reducing active-site accessibility (Figure 4.3).

The interaction energy analysis through umbrella sampling reveals that when the ligand is bound to the active site, the ΔG value for the mutant structure is 10.47 kcal/mol, whereas for the wild type it is 19 kcal/mol. The higher ΔG value in the wild type, in contrast to the post-MD simulation and cross-correlation analysis results, indicates that the ligand dissociates more reluctantly from the wild type structure and is therefore more strongly bound than in the mutant. When the ligand is bound to the allosteric site, the ΔG value is 7.66 kcal/mol in the mutant

and 7.44 kcal/mol in the wild type. Although this difference is small, it suggests that the ligand binds slightly more strongly to the mutant structure at the allosteric site (Figure 4.4).

When the docking of analog molecules was first evaluated, it was observed that compound 8929323 stood out with the best docking score among the molecules generated for the active site. As shown in Table 4.12, in the 1000 ns simulations performed with KRAS G12D for four compounds (8929323, 3302325, 5179173 and 5576702), each ligand was observed to remain bound to the active site throughout the entire simulation. When the docking poses of these analogs were compared with the docking poses of the original molecule and the positive control compounds, it was found that all molecules commonly interacted with the amino acid GLU62. According to Table 5.1, the original molecule (ZINC000008758797) interacted with five active site residues (VAL9, GLU62, TYR64, ARG68, and MET72). The ligand with ID 8929323 interacted with four active site residues (VAL9, GLU62, ARG68, and MET72), the ligand with ID 3302325 interacted with six active site residues (VAL9, ALA11, GLU62, TYR64, ARG68, and MET72), the ligand with ID 5179173 interacted with seven active site residues (ALA11, ASP12, GLU62, TYR64, ARG68 and MET72) and the ligand with ID 5576702 interacted with five active site residues (VAL9, GLU62, TYR64, ARG68, MET72). At first glance, these results suggest that compound 5179173 forms more interactions with KRAS G12D than the other molecules and may therefore be more effective in comparison (Table 5.1 and Figure 5.2).

Table 5.1 Residues Interacting with the Best Analog Candidates, Original Molecule and Positive Controls.

Analog Molecules and Positive Control Inhibitors	**Interacting Residues**
ZINC000008758797	VAL9, GLU62, TYR64, ARG68, MET72, HIS94, HIS95, TYR96, GLN99
MRTX-1133	ALA11, ASP12, GLU62, ASN86, LYS88, SER89
TH-Z827	GLU62, TYR64, ARG68, MET72, HIS95
TH-Z835	ALA11, ASP12, GLU62, LYS88, ASP92, HIS95
8929323	VAL9, GLU62, ARG68, MET72, ASP92, HIS95, TYR96
3302325	VAL9, ALA11, GLU62, TYR64, ARG68, MET72, HIS95, TYR96
5179173	ALA11, ASP12, GLU62, TYR64, ARG68, ASP69, MET72, ASP92, HSD95, TYR96
5576702	VAL9, GLU62, TYR64, ARG68, MET72, ASP92, HSD95, TYR96

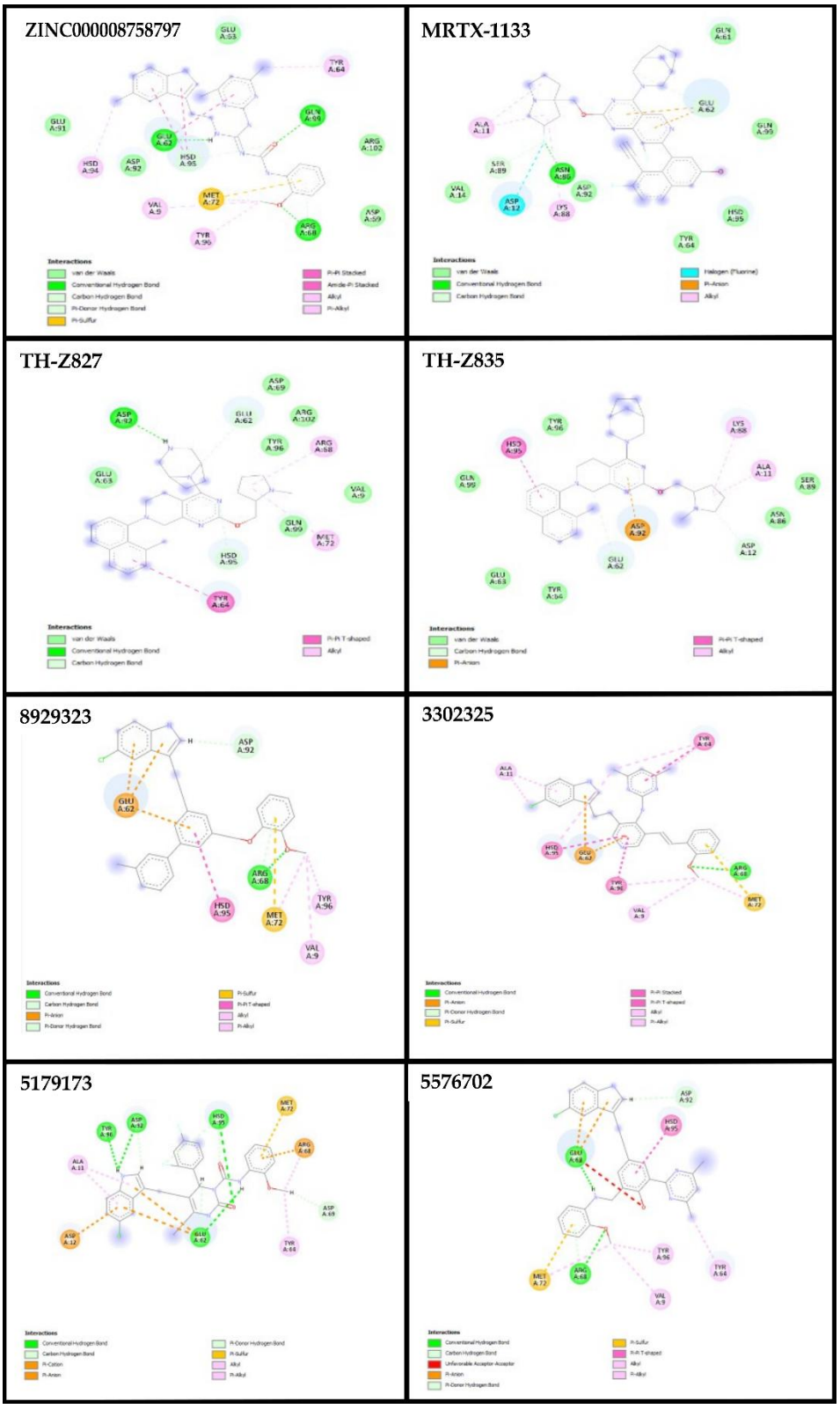

Figure 5.2 Two-dimensional Chemical Structure Representations and Docking Poses of Original Molecule (ZINC000008758797), Analog Molecules (8929323, 3302325, 5179173 and 5576702), and Positive Controls in the KRAS G12D Active Site.

When the simulations of the complexes formed by the analog molecules with KRAS G12D are compared, it is generally observed that all systems remain stable over the 1000 ns timescale (Figure 4.11). In terms of the RMSD profiles, the lowest values belong to ZINC000008758797 and 5179173, which may suggest that binding of the ligand induces less conformational fluctuation in the protein and that the complex exhibits a more settled behavior. In contrast, the occasional increases in RMSD observed for the 8929323, 3302325 and 5576702 systems indicate that conformational rearrangements within the active site are more pronounced in the presence of these analogs. Between the four analogs, although 3302325 shows higher RMSD fluctuations in certain intervals, the trajectory continues within a more stable band in the later stages of the simulation, suggesting re-stabilization of the system. The Radious of gyration values, remain within similar ranges for all ligands; therefore, the overall folded state of KRAS G12D appears to be preserved regardless of ligand type. The RMSF patterns of the four systems are largely similar; however, the higher RMSF observed in the Switch I region for 8929323, 3302325 and 5179173 compared to the original ligand suggests that these analogs may require greater local adaptation around the pocket during binding. While the original ligand ZINC000008758797 stands out with a higher average number of hydrogen bonds, followed by 5576702, 3302325 and 5179173 shows an intermediate level and 8929323 exhibits the lowest average H-bond value. This comparison suggests that binding of the original ligand may be stabilized predominantly through of polar/ionic interactions, whereas the analogs, particularly 8929323, may maintain binding stability more through hydrophobic interactions and van der Waals contacts.

When the simulations of the KRAS G12D complexes with the two analog molecules (8929323 and 3302325) are visually inspected, the molecules may appear to be leaving the pocket when comparing the initial and final structures. However, this observation can be explained by the gradual closure of Switch II over the binding pocket over time. Focusing on the loop region of Switch II, as the loop moves in the direction of closing the pocket, the molecules also seem to shift in the same direction due to their interactions with this region. In fact, the ligands moving together with the binding site may indicate that they form favorable interactions and remain coupled to the pocket dynamics rather than dissociating. In the other two analog molecules (5179173 and 5576702), conversely, the initial structure has a more closed pocket, while the final structure has molecules positioned further inside the active pocket. In addition, the region marked with an ellipse in the visualization corresponds to Switch I, where fluctuations are also observed in the RMSF profiles (Figure 5.3).

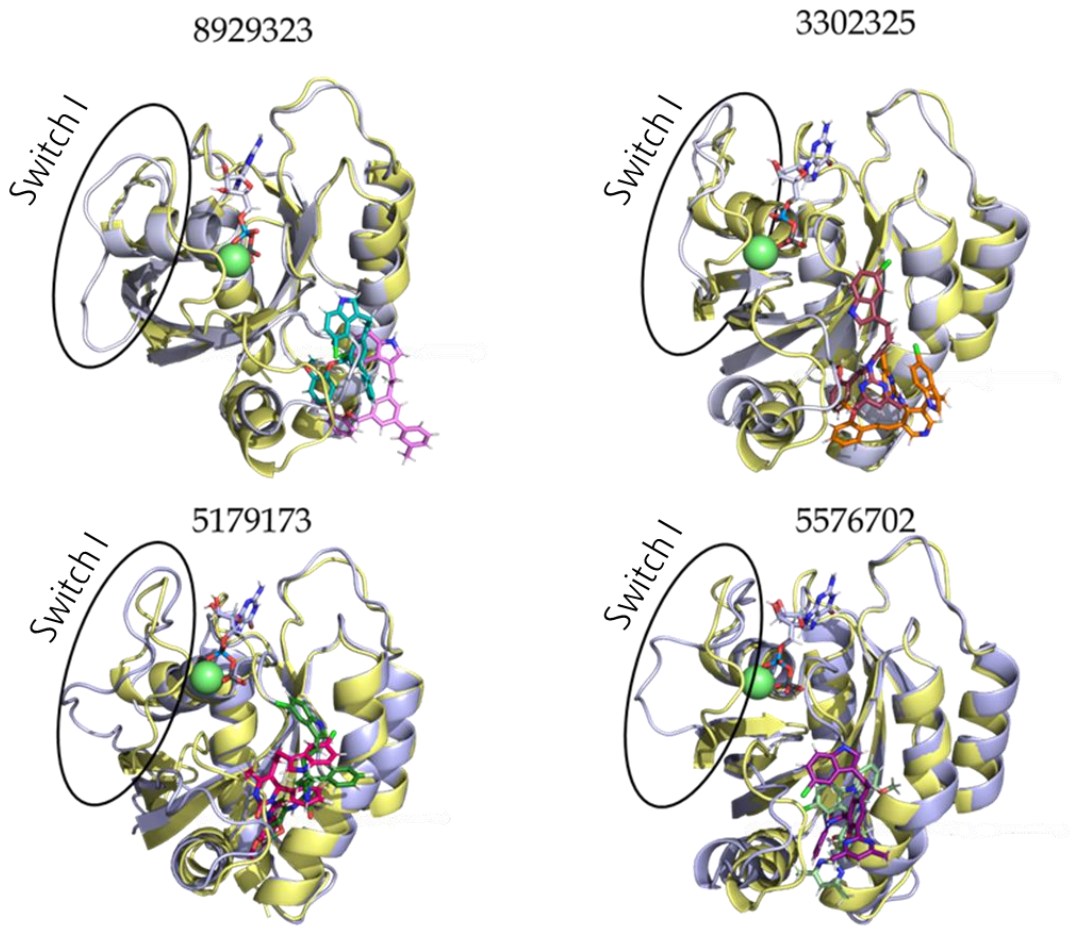

Figure 5.3 Alignment-based Comparison of KRAS G12D Conformations upon Binding of Molecules 8929323, 3302325, 5179173 and 5576702: Initial (KRAS: Yellow; Molecules: 3302325 burgundy, 8929323 cyan, 5179173 dark pink, 5576702 purple) and Final (KRAS: light lilac; Molecules: 3302325 orange, 8929323 magenta, 5179173 dark green, 5576702 light green) MD Frames.

5.2 Targeting the G12C Mutant KRAS

In the discovery of covalent inhibitors for KRAS G12C, only 26 of the molecules obtained from the virtual screening were able to covalently bind to the 12th Cysteine. Docking of these molecules and the positive control molecules revealed that two molecules had better docking scores than all of the positive control molecules (Table 4.15). This finding increases the likelihood that they represent potential covalent drug candidates. When the residues with which the two molecules interact in the active site of KRAS G12C are examined, the interactions of the molecule ZINC000000626790 with GLU62 and GLN61 are also observed for the molecules Sotorasib and Adagrasib. In addition, the residues that are common to its interactions with the inhibitor ARS-1620 are ALA59, GLU62 and MET67. This molecule and the three positive control inhibitors all interact with GLU62. This makes our potential inhibitor, which exhibits common interactions

with the drugs Sotorasib and Adagrasib that have been approved by the FDA for KRAS G12C, particularly promising. For the molecule ZINC000057607951, only the residue PRO34 appears to be shared with ARS-1620. Therefore, the molecule ZINC000000626790 has a higher potential as a drug candidate compared with the other molecule.

When the 2D structure of the molecule ZINC000000626790 is examined, its IUPAC name can be given as *isopropyl 4-[4-(benzyloxy)-3-methoxyphenyl]-2-methyl-5-oxo-1,4,5,6,7,8-hexahydro-3-quinolinecarboxylate*. This molecule can be evaluated in terms of two key functional features: the quinoline/quinolinone scaffold and the Michael acceptor (α,β-unsaturated carbonyl) moiety. In the literature, the tyrosine kinase inhibitor Lenvatinib, which contains a quinoline core, is of particular interest. Lenvatinib (*4-{3-chloro-4-[(cyclopropylcarbamoyl)amino]phenoxy}-7-methoxyquinoline-6-carboxylic acid*) is an FDA approved drug for the treatment of thyroid cancer, renal cell carcinoma (RCC), endometrial cancer and hepatocellular carcinoma. It targets the VEGF receptors VEGFR1 (FLT1), VEGFR2 (KDR) and VEGFR3 (FLT4), thereby suppressing the signaling required for uncontrolled cell proliferation and growth through downstream MAPK and PI3K pathways (Scott, 2015). Similarly, when the α,β-unsaturated carbonyl (Michael acceptor) group is considered, the approved KRAS G12C inhibitors Sotorasib and Adagrasib also contain an acrylamide warhead that forms a covalent bond with Cys12 (Figure 5.4) (Lee and Park, 2022).

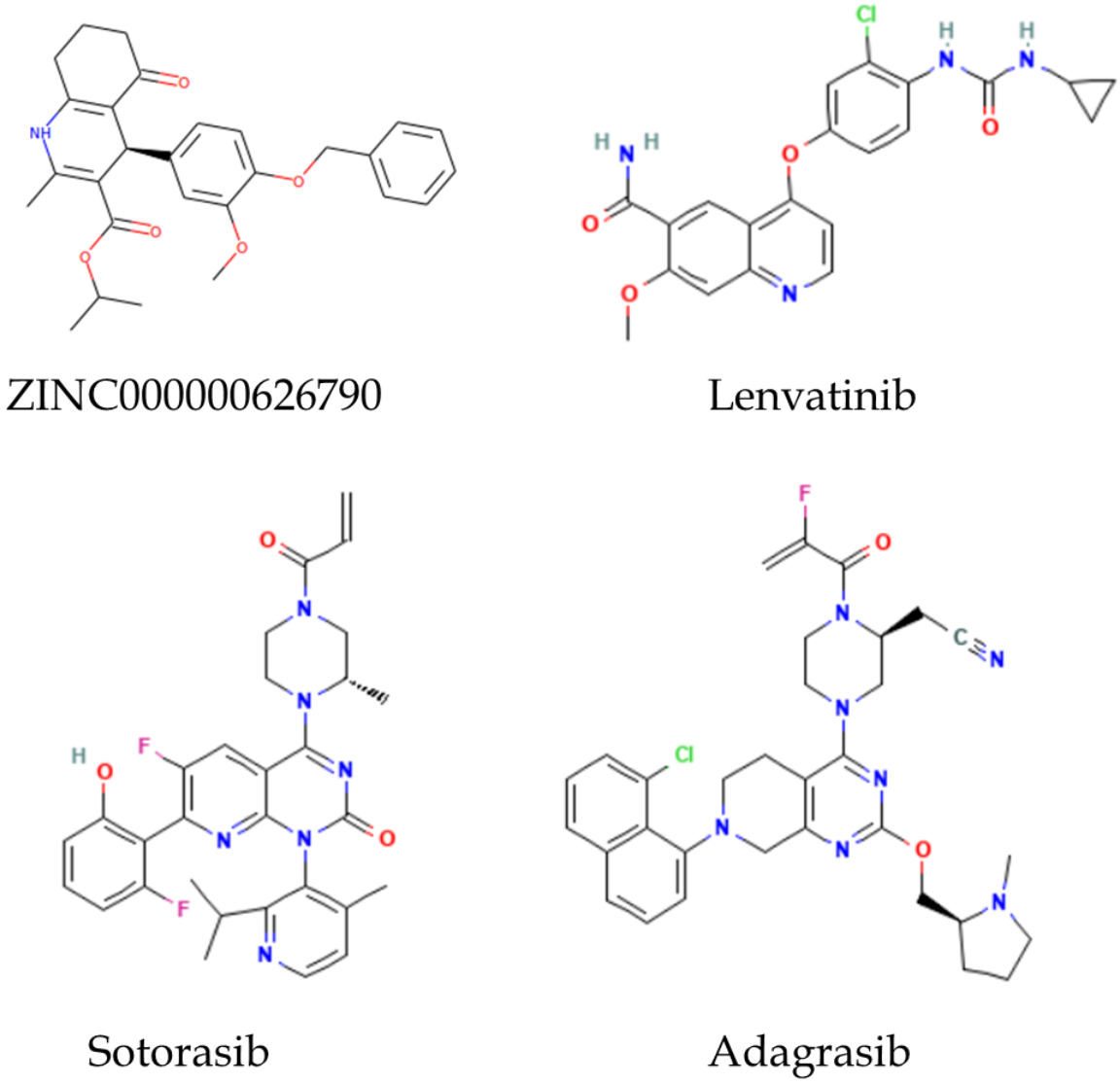

Figure 5.4 Two-dimensional Molecular Structures of the Hit Compound ZINC000000626790 and the Approved Anticancer Drugs Lenvatinib, Sotorasib and Adagrasib.

6.CONCLUSION & FUTURE WORK

In this thesis, which aims to target KRAS mutants and identify novel alternative inhibitors, our studies on the KRAS G12D mutant led to the design of four promising active site analogs (8929323, 3302325, 5179173, and 5576702). Based on the post-MD analyses, compound 5576702 emerged as the most prominent candidate, as it exhibited low structural fluctuations, the highest number of hydrogen bonds among the analogs, and a strong interaction profile in docking, reflected by the number of active site residues involved in binding. However, to ensure reproducibility and reliability, each of these four analogs should be evaluated with three independent simulation replicates. Although our

conclusions may be refined as additional umbrella sampling results become available, the current findings indicate that 5576702 has produced the most favorable outcomes so far. In parallel, the ongoing MD simulations and subsequent free energy calculations of the designed allosteric analogs may enable the discovery of an allosteric inhibitor targeting KRAS G12D.

For the KRAS G12C mutant, the compound ZINC000000626790 stood out due to its best docking score and its shared interaction residues with the FDA-approved drugs Sotorasib and Adagrasib. Nevertheless, performing covalent MD simulations for this compound is of critical importance to validate its binding stability and inhibitory potential.

Overall, these results provide a structured framework for prioritizing the most promising candidates and demonstrate that combining rational design with MD-based validation and free energy calculations is a robust strategy for advancing the development of next-generation KRAS inhibitors.

REFERENCES

Mark Abraham, Andrey Alekseenko, Cathrine Bergh, Christian Blau, Eliane Briand, Mahesh Doijade, Stefan Fleischmann, Vytautas Gapsys, Gaurav Garg, Sergey Gorelov, Gilles Gouaillardet, Alan Gray, M. Eric Irrgang, Farzaneh Jalalypour, Joe Jordan, Christoph Junghans, Prashanth Kanduri, Sebastian Keller, Carsten Kutzner, … Erik Lindahl. (2024). GROMACS 2023.5 Source code (2023.5). Zenodo. https://doi.org/10.5281/zenodo.11104849

Avogadro (Version 1.2.0) [Computer software]. https://avogadro.cc/

Bannoura, S. F., Khan, H. Y., & Azmi, A. S. (2022). KRAS G12D targeted therapies for pancreatic cancer: Has the fortress been conquered?. *Frontiers in oncology*, *12*, 1013902. https://doi.org/10.3389/fonc.2022.1013902

Barlas, A. B., Savaş, B., & Karaca, E. (2022). CSB-KaracaLab/gmx_corr: Initial version (Version 1.0.0). Zenodo. https://doi.org/10.5281/zenodo.7180338

Bauer, M. R., & Mackey, M. D. (2019). Electrostatic Complementarity as a Fast and Effective Tool to Optimize Binding and Selectivity of Protein-Ligand Complexes. *Journal of medicinal chemistry*, *62*(6), 3036–3050. https://doi.org/10.1021/acs.jmedchem.8b01925

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. Journal

of molecular biology, 112(3), 535–542. https://doi.org/10.1016/s0022-2836(77)80200-3

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., & Wiswedel, B. (2007). KNIME: The Konstanz Information Miner. *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)*. Springer.

Discovery Studio (Version 2021) (Computer software). Dassault Systèmes. https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-discovery-studio/

Bos, J. L., Rehmann, H., & Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. *Cell*, *129*(5), 865–877. https://doi.org/10.1016/j.cell.2007.05.018

Buhrman, G., O'Connor, C., Zerbe, B., Kearney, B. M., Napoleon, R., Kovrigina, E. A., Vajda, S., Kozakov, D., Kovrigin, E. L., & Mattos, C. (2011). Analysis of binding site hot spots on the surface of Ras GTPase. Journal of molecular biology, 413(4), 773–789. https://doi.org/10.1016/j.jmb.2011.09.011

Canon, J., Rex, K., Saiki, A. Y., Mohr, C., Cooke, K., Bagal, D., Gaida, K., Holt, T., Knutson, C. G., Koppada, N., Lanman, B. A., Werner, J., Rapaport, A. S., San Miguel, T., Ortiz, R., Osgood, T., Sun, J.-R., Zhu, X., McCarter, J. D., … Lipford, J. R. (2019). The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature, 575*(7781), 217–223. https://doi.org/10.1038/s41586-019-1694-1

Cheeseright, T., Mackey, M., Rose, S., & Vinter, A. (2006). Molecular field extrema as descriptors of biological activity: definition and validation. *Journal of chemical information and modeling*, *46*(2), 665–676. https://doi.org/10.1021/ci050357s

Cox, A. D., & Der, C. J. (2025). "Undruggable KRAS": druggable after all. *Genes & development*, *39*(1-2), 132–162. https://doi.org/10.1101/gad.352081.124

Cresset. (2024). *Flare* (Computer software). https://cresset-group.com/flare/

Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J., & Der, C. J. (2014). Drugging the undruggable RAS: Mission possible?. *Nature reviews. Drug discovery*, *13*(11), 828–851. https://doi.org/10.1038/nrd4389

Durojaye, O. A., Ejaz, U., Uzoeto, H. O., Fadahunsi, A. A., Opabunmi, A. O., Ekpo, D. E., Sedzro, D. M., & Idris, M. O. (2023). CSC01 shows promise as a potential inhibitor of the oncogenic G13D mutant of KRAS: an in silico

approach. *Amino acids*, *55*(12), 1745–1764. https://doi.org/10.1007/s00726-023-03304-2

Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of chemical information and modeling*, *61*(8), 3891–3898. https://doi.org/10.1021/acs.jcim.1c00203

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic acids research*, *40*(Database issue), D1100–D1107. https://doi.org/10.1093/nar/gkr777

Gupta, A. K., Wang, X., Pagba, C. V., Prakash, P., Sarkar-Banerjee, S., Putkey, J., & Gorfe, A. A. (2019). Multi-target, ensemble-based virtual screening yields novel allosteric KRAS inhibitors at high success rate. Chemical biology & drug design, 94(2), 1441–1456. https://doi.org/10.1111/cbdd.13519

Hallin, J., Bowcut, V., Calinisan, A. *et al.* Anti-tumor efficacy of a potent and selective non-covalent $KRAS^{G12D}$ inhibitor. *Nat Med* **28**, 2171–2182 (2022). https://doi.org/10.1038/s41591-022-02007-7

Hallin, J., Engstrom, L. D., Hargis, L., Calinisan, A., Aranda, R., Briere, D. M., Sudhakar, N., Bowcut, V., Baer, B. R., Ballard, J. A., Burkard, M. R., Fell, J. B., Fischer, J. P., Vigers, G. P., Xue, Y., Gatto, S., Fernandez-Banet, J., Pavlicek, A., Velastagui, K., Chao, R. C., … Christensen, J. G. (2020). The KRASG12C Inhibitor MRTX849 Provides Insight toward Therapeutic Susceptibility of KRAS-Mutant Cancers in Mouse Models and Patients. Cancer discovery, 10(1), 54–71. https://doi.org/10.1158/2159-8290.CD-19-1167

Hanwell, M.D., Curtis, D.E., Lonie, D.C. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminform 4, 17 (2012). https://doi.org/10.1186/1758-2946-4-17

Hofmann, M. H., Gerlach, D., Misale, S., Petronczki, M., & Kraut, N. (2022). Expanding the Reach of Precision Oncology by Drugging All KRAS Mutants. Cancer discovery, 12(4), 924–937. https://doi.org/10.1158/2159-8290.CD-21-1331

Huang J, MacKerell AD Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem. 2013 Sep 30;34(25):2135-45. https://doi.org/10.1002/jcc.23354

Hussain, J., & Rea, C. (2010). Computationally efficient algorithm to identify matched molecular pairs (MMPs) in large data sets. *Journal of chemical information and modeling*, *50*(3), 339–348. https://doi.org/10.1021/ci900450m

Irwin, J. J., Tang, K. G., Young, J., Dandarchuluun, C., Wong, B. R., Khurelbaatar, M., Moroz, Y. S., Mayfield, J., & Sayle, R. A. (2020). ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery. Journal of chemical information and modeling, 60(12), 6065–6073. https://doi.org/10.1021/acs.jcim.0c00675

Karmacharya, U., Guragain, D., Chaudhary, P., Jee, J. G., Kim, J. A., & Jeong, B. S. (2021). Novel Pyridine Bioisostere of Cabozantinib as a Potent c-Met Kinase Inhibitor: Synthesis and Anti-Tumor Activity against Hepatocellular Carcinoma. International journal of molecular sciences, 22(18), 9685. https://doi.org/10.3390/ijms22189685

Kessler, D., Gmachl, M., Mantoulidis, A., Martin, L. J., Zoephel, A., Mayer, M., Gollner, A., Covini, D., Fischer, S., Gerstberger, T., Gmaschitz, T., Goodwin, C., Greb, P., Häring, D., Hela, W., Hoffmann, J., Karolyi-Oezguer, J., Knesl, P., Kornigg, S., Koegl, M., … McConnell, D. B. (2019). Drugging an undruggable pocket on KRAS. Proceedings of the National Academy of Sciences of the United States of America, 116(32), 15823–15829. https://doi.org/10.1073/pnas.1904529116

Kim, D., Xue, J. Y., & Lito, P. (2020). Targeting KRAS(G12C): From Inhibitory Mechanism to Modulation of Antitumor Effects in Patients. Cell, 183(4), 850–859. https://doi.org/10.1016/j.cell.2020.09.044

Kim, H. J., Lee, H. N., Jeong, M. S., & Jang, S. B. (2021). Oncogenic KRAS: Signaling and Drug Resistance. *Cancers*, *13*(22), 5599. https://doi.org/10.3390/cancers13225599

Krishnan, T., Roberts-Thomson, R., Broadbridge, V., & Price, T. (2022). Targeting Mutated KRAS Genes to Treat Solid Tumours. Molecular diagnosis & therapy, 26(1), 39–49. https://doi.org/10.1007/s40291-021-00564-0

Kuhn, M., Firth-Clark, S., Tosco, P., Mey, A. S. J. S., Mackey, M., & Michel, J. (2020). Assessment of Binding Affinity via Alchemical Free-Energy Calculations. *Journal of chemical information and modeling*, *60*(6), 3120–3130. https://doi.org/10.1021/acs.jcim.0c00165

Kulkarni, A. M., Kumar, V., Parate, S., Lee, G., Yoon, S., & Lee, K. W. (2022). Identification of New KRAS G12D Inhibitors through Computer-Aided Drug Discovery Methods. *International journal of molecular sciences*, *23*(3), 1309. https://doi.org/10.3390/ijms23031309

Kwan, A. K., Piazza, G. A., Keeton, A. B., & Leite, C. A. (2022). The path to the clinic: a comprehensive review on direct $KRAS^{G12C}$ inhibitors. *Journal of experimental & clinical cancer research : CR*, *41*(1), 27. https://doi.org/10.1186/s13046-021-02225-w

Lee, J., & Park, S. B. (2022). Extended Applications of Small-Molecule Covalent Inhibitors toward Novel Therapeutic Targets. Pharmaceuticals (Basel, Switzerland), 15(12), 1478. https://doi.org/10.3390/ph15121478

Lemkul, J. A. (2018). From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package [Article v1.0]. *Living Journal of Computational Molecular Science*, *1*(1), 5068. https://doi.org/10.33011/livecoms.1.1.5068

Li Q., Shah S. (2017) Structure-Based Virtual Screening. In: Wu C., Arighi C., Ross K. (eds) Protein Bioinformatics. Methods in Molecular Biology, vol 1558. Humana Press, New York, NY https://doi.org/10.1007/978-1-4939-6783-4_5

Liu, J., Kang, R., & Tang, D. (2022). The KRAS-G12C inhibitor: activity and resistance. *Cancer gene therapy*, *29*(7), 875–878. https://doi.org/10.1038/s41417-021-00383-9

Luo, X., Zhou, F., Tang, Y., Liu, X., Xiao, R., Gu, M., Bai, J., Jiang, D., Yang, G., You, L., & Zhao, Y. (2025). Novel paradigms in KRAS targeting: Unveiling strategies to combat drug resistance. Chinese medical journal, 138(18), 2243–2267. https://doi.org/10.1097/CM9.0000000000003776

McCormick F. (2016). K-Ras protein as a drug target. *Journal of molecular medicine (Berlin, Germany)*, *94*(3), 253–258. https://doi.org/10.1007/s00109-016-1382-7

McCormick F. (2019). Progress in targeting RAS with small molecule drugs. *The Biochemical journal*, *476*(2), 365–374. https://doi.org/10.1042/BCJ20170441

Menyhárd, D. K., Pálfy, G., Orgován, Z., Vida, I., Keserű, G. M., & Perczel, A. (2020). Structural impact of GTP binding on downstream KRAS signaling. *Chemical science*, *11*(34), 9272–9289. https://doi.org/10.1039/d0sc03441j

Moore, A. R., Rosenberg, S. C., McCormick, F., & Malek, S. (2020). RAS-targeted therapies: is the undruggable drugged?. *Nature reviews. Drug discovery*, *19*(8), 533–552. https://doi.org/10.1038/s41573-020-0068-6

Muñoz-Maldonado, C., Zimmer, Y., & Medová, M. (2019). A Comparative Analysis of Individual RAS Mutations in Cancer Biology. *Frontiers in oncology*, *9*, 1088. https://doi.org/10.3389/fonc.2019.01088

Nabet, B., & Gray, N. S. (2018). It Takes Two To Target: A Study in KRAS Dimerization. Biochemistry, 57(16), 2289–2290. https://doi.org/10.1021/acs.biochem.8b00376

Nakajima, E. C., Drezner, N., Li, X., Mishra-Kalyani, P. S., Liu, Y., Zhao, H., Bi, Y., Liu, J., Rahman, A., Wearne, E., Ojofeitimi, I., Hotaki, L. T., Spillman, D., Pazdur, R., Beaver, J. A., & Singh, H. (2022). FDA Approval Summary: Sotorasib for KRAS G12C-Mutated Metastatic NSCLC. *Clinical cancer research : an official journal of the American Association for Cancer Research*, *28*(8), 1482–1486. https://doi.org/10.1158/1078-0432.CCR-21-3074

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, 33. https://doi.org/10.1186/1758-2946-3-33

Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A., & Shokat, K. M. (2013). K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*, *503*(7477), 548–551. https://doi.org/10.1038/nature12796

Pathan, A. A., Panthi, B., Khan, Z., Koppula, P. R., Alanazi, M. S., Sachchidanand, Parine, N. R., & Chourasia, M. (2016). Lead identification for the K-Ras protein: virtual screening and combinatorial fragment-based approaches. *OncoTargets and therapy*, *9*, 2575–2584. https://doi.org/10.2147/OTT.S99671

Pálfy, G., Menyhárd, D. K., & Perczel, A. (2020). Dynamically encoded reactivity of Ras enzymes: opening new frontiers for drug discovery. *Cancer metastasis reviews*, *39*(4), 1075–1089. https://doi.org/10.1007/s10555-020-09917-3

Prakash, P., Hancock, J. F., & Gorfe, A. A. (2015). Binding hotspots on K-ras: consensus ligand binding sites and other reactive regions from probe-based molecular dynamics analysis. Proteins, 83(5), 898–909. https://doi.org/10.1002/prot.24786

Prior, I. A., Lewis, P. D., & Mattos, C. (2012). A comprehensive survey of Ras mutations in cancer. *Cancer research*, *72*(10), 2457–2467. https://doi.org/10.1158/0008-5472.CAN-11-2612

Reedquist, K. A., & Tak, P. P. (2012). Signal transduction pathways in chronic inflammatory autoimmune disease: small GTPases. *The open rheumatology journal*, *6*, 259–272. https://doi.org/10.2174/1874312901206010259

Román, M., Baraibar, I., López, I., Nadal, E., Rolfo, C., Vicent, S., & Gil-Bazo, I. (2018). KRAS oncogene in non-small cell lung cancer: clinical perspectives on the

treatment of an old target. Molecular cancer, 17(1), 33. https://doi.org/10.1186/s12943-018-0789-x

Schake, P., Bolz, S. N., Linnemann, K., & Schroeder, M. (2025). PLIP 2025: introducing protein-protein interactions to the protein-ligand interaction profiler. Nucleic acids research, 53(W1), W463–W465. https://doi.org/10.1093/nar/gkaf361

Scott L. J. (2015). Lenvatinib: first global approval. Drugs, 75(5), 553–560. https://doi.org/10.1007/s40265-015-0383-0

Shan, J., & Ji, C. (2020). MolOpt: A Web Server for Drug Design using Bioisosteric Transformation. Current computer-aided drug design, 16(4), 460–466. https://doi.org/10.2174/1573409915666190704093400

Sharma, P., Singh, M., Singh, V., Singh, T. G., Singh, T., & Ahmad, S. F. (2023). Recent Development of Novel Aminoethyl-Substituted Chalcones as Potential Drug Candidates for the Treatment of Alzheimer's Disease. Molecules (Basel, Switzerland), 28(18), 6579. https://doi.org/10.3390/molecules28186579

Suyal, C., Shenoy, K. M., Kishore, A., & Kini, S. G. (2025). Structure-activity relationships of KRAS-G12D inhibitors for pancreatic cancer. *Drug discovery today*, *30*(7), 104396. https://doi.org/10.1016/j.drudis.2025.104396

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. https://doi.org/10.1002/jcc.21334

Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., & Mackerell, A. D., Jr (2010). CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. Journal of computational chemistry, 31(4), 671–690. https://doi.org/10.1002/jcc.21367

Vanommeslaeghe, K., & MacKerell, A. D., Jr (2012). Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. Journal of chemical information and modeling, 52(12), 3144–3154. https://doi.org/10.1021/ci300363c

Vanommeslaeghe, K., Raman, E. P., & MacKerell, A. D., Jr (2012). Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. Journal of chemical information and modeling, 52(12), 3155–3168. https://doi.org/10.1021/ci3003649

Wang, X., Wu, J., Xiao, A., Wang, J., & Tian, J. (2025). Evolution of direct RAS inhibitors: from undruggable target to clinical breakthroughs. Molecular cancer, 24(1), 229. https://doi.org/10.1186/s12943-025-02364-0

Yang, X., & Wu, H. (2024). RAS signaling in carcinogenesis, cancer therapy and resistance mechanisms. *Journal of hematology & oncology*, *17*(1), 108. https://doi.org/10.1186/s13045-024-01631-9

Zhang, Y., Li, C., Xia, C., Wah To, K. K., Guo, Z., Ren, C., Wen, L., Wang, F., Fu, L., & Liao, N. (2022). Adagrasib, a KRAS G12C inhibitor, reverses the multidrug resistance mediated by ABCB1 in vitro and in vivo. *Cell communication and signaling : CCS*, *20*(1), 142. https://doi.org/10.1186/s12964-022-00955-8