МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
імені В. Н. Каразіна

Кафедра неорганічної хімії

МІНІСТЕРСТВО ВИЩОЇ ОСВІТИ І НАУКИ ФРАНЦІЇ
УНІВЕРСИТЕТ ЛАЗУРНОГО БЕРЕГА

Інститут хімії

УДК 544.165+577.29

До захисту допускаю
_____ В.о. завідувача кафедри
«__» _____ 2025 р. к.х.н., доц. М.М. Волобуєв

КОМП'ЮТЕРНИЙ ДИЗАЙН НОВИХ ІНГІБІТОРІВ ТА ЛІГАНД-БІЛОК ВЗАЄМОДІЯ У ОНКОЛОГІЧНО ВАЖЛИВИХ БІЛКАХ

Кваліфікаційна робота магістра
ІІ курсу хімічного факультету
МОРОЗОВОЇ ВІРИ ОЛЕГІВНИ

Наукові керівники:

д.х.н., професор О.В. Кириченко
д-р, професор С. Ронко

ХАРКІВ 2025

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

V.N. KARAZIN KHARKIV NATIONAL UNIVERSITY

Department of Inorganic Chemistry

MINISTRY OF HIGHER EDUCATION AND SCIENCE OF FRANCE
UNIVERSITÉ CÔTE D'AZUR

Institute of Chemistry

UDK 544.165+577.29

Admit to defense
_____ Head of the Department
«___» _____ 2025 Ph.D. M. M. Volobuev

COMPUTER-AIDED DESIGN OF NOVEL INHIBITORS AND LIGAND-PROTEIN INTERACTIONS IN ONCOLOGY-RELEVANT PROTEINS

Master thesis of 2nd year student
of the School of Chemistry
MOROZOVA VIRA

Supervisors:

Professor, Doc. Sci. A.V. Kyrychenko
Professor, Ph.D. C. Ronco

KHARKIV 2025

ABSTRACT

The MSc thesis consists of 50 pages, 29 figures, 2 tables, and 50 references.

The object of the study is a targeted library of 1H-1,2,4-triazole derivatives, designed to inhibit the oncology-relevant DRAK1 enzyme.

The aim of this study is to interpret the structure–activity relationship (SAR) data for a ligand library by identifying key interactions within ligand–receptor complexes. This will help determine a promising direction for further structure optimization. To accomplish this, molecular docking and classical molecular dynamics simulations were performed using two pairs of regioisomers as model systems. This approach enables the assessment of how the spatial arrangement of functional groups influences binding within the protein's active site.

Research methods – molecular docking, classical molecular dynamics simulation, and binding free energy calculations with per-residue energy decomposition (MM-PBSA).

A bioinformatic analysis of the target protein DRAK1 was conducted using MEGA12 software. Molecular docking of 24 original ligands and 11 ligands derived from literature was performed using AutoDock Vina and Avogadro 1.2.0 for ligand preparation. Potential binding pockets were identified using the FTMove web service. A homology model of DRAK1 for molecular dynamics was constructed using the SWISS-MODEL platform. Following molecular dynamics simulations of the protein in OPLS-AA/TIP3P and CHARMM36/TIP3P (CHARMM-modified) environments, CHARMM36 was selected for further modeling based on analyses of RMSD, RMSF, and radius of gyration. Two pairs of regioisomers, which differed in biological activity, were investigated through classical molecular dynamics simulations of protein-ligand complexes. Binding free energies were calculated using the MM-PBSA method along with per-residue energy decomposition. The stability of the complexes was evaluated through RMSD analysis.

The results of this study may be useful for the further development of DRAK1 inhibitors as potential therapeutic agents for glioblastoma multiforme.

Keywords: DRAK1, glioblastoma multiforme, oncology-relevant proteins, molecular docking, molecular dynamics, MM-PBSA, per-residue decomposition, CHARMM36, OPLS-AA.

РЕФЕРАТ

Кваліфікаційна робота складається з 50 сторінок, 29 рисунків, 2 таблиці та 50 літературних посилань.

Об'єкт дослідження – фокусна бібліотека лігандів похідних 1H-1,2,4-триазолу, створена з метою інгібіювання онкологічно-важливого ензиму DRAK1.

Мета дослідження – інтерпретація отриманих даних структура–активність (SAR) для бібліотеки з 24 лігандів шляхом з'ясування ключових взаємодій у комплексах ліганд–рецептор та визначення перспективного напряму для подальшої оптимізації. Молекулярний докінг і молекулярно динамічне моделювання із використанням двох пар регіоізомерів як модельних систем, що дозволить оцінити вплив просторового розташування функціональних груп на характер зв'язування в активному сайті білка.

Методи дослідження – молекулярний докінг, класична молекулярна динаміка, енергетичний аналіз комплексів із амінокислотною декомпозицією (MM PBSA).

Проведено біоінформатичний аналіз таргетного білка DRAK1 із використанням серверу MEGA12. Молекулярний докінг 24 оригінальних та 11 літературних лігандів здійснено з використанням AutoDock Vina, а також Avogadro 1.2.0 для підготовки лігандів. Потенційні сайти зв'язування досліджено з використанням веб-сервісу FTMove. Гомологічну модель білка DRAK1 для молекулярної динаміки побудовано на платформі SWISS-MODEL. Після проведення молекулярної динаміки білка в системах OPLS-AA/TIP3P та CHARMM36/TIP3P (CHARMM-modified), на основі аналізу RMSD, RMSF і радіуса гірації обрано силове поле CHARMM36 для подальшого моделювання. Дві пари регіоізомерів, що відрізняються за біологічною активністю, були досліджені методом класичної молекулярної динаміки комплексів протеїн-ліганд. Розрахунок вільної енергії зв'язування здійснено за методом MM-PBSA.

Результати дослідження можуть бути корисними для подальшої розробки інгібіторів DRAK1 як потенційних терапевтичних засобів при гліобластомі мультиформній.

Ключові слова: DRAK1, гліобластома мультиформна, онкологічно-важливі білки, молекулярний докінг, молекулярна динаміка, MM-PBSA, поамінокислотна декомпозиція, CHARMM36, OPLS-AA.

CONTENT

GLOSSARY

Abbreviation	Definition
DRAK1	Death-associated protein kinase-related apoptosis-inducing protein kinase 1
DAPK	Death-associated protein kinase
STK17A	Serine/threonine kinase 17A
CADD	Computer-Aided Drug Design
SAR	Structure-Activity Relationship
FDA	Food & Drug Administration
SMKI	Small Molecule Kinase Inhibitor
MSK1	Mitogen- and stress-activated kinase 1
CAMK	Calcium/calmodulin-dependent protein kinase
MLCK	Myosin light chain kinase
NLS	Nuclear localization signals
PKC	Protein kinase C
TNF	Tumor necrosis factor
TRAF6	TNF receptor-associated factor 6
HNSCC	Head and neck squamous cell carcinoma
GBM	Glioblastoma multiforme
SBDD	Structure-Based Drug Design
LBDD	Ligand-Based Drug Design
MD	Molecular dynamics
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
Rg	Radius of gyration
GROMACS	GROningen MAchine for Chemical Simulation
MM-GBSA	Molecular Mechanics with Generalized Born and Surface Area
MM-PBSA	Molecular Mechanics Poisson–Boltzmann Surface Area
SASA	Solvent-Accessible Surface Area
MUSCLE	Multiple Sequence Comparison by Log-Expectation
OPLS-AA	Optimized Potentials for Liquid Simulations – All Atom
CHARMM36	Chemistry at HARvard Macromolecular Mechanics, version 36

INTRODUCTION

Serine/threonine kinase 17A (STK17A), also known as death-associated protein kinase-related apoptosis-inducing protein kinase 1 (DRAK1), is a member of the death-associated protein kinase (DAPK) family and is classified within the so-called "dark kinome." DRAK1 plays a significant role in the tumorigenesis of several cancers, particularly glioblastoma multiforme. Despite its importance in cancer biology, there are currently no selective inhibitors available for DRAK1 [1].

Modern computer-aided drug design (CADD) techniques, such as molecular docking and molecular dynamics simulations, significantly accelerate drug discovery. These methods allow researchers to concentrate their experimental efforts on the most promising compounds, thereby reducing the dependence on expensive high-throughput screening. In this study, we constructed a focused library of 24 ligands based on previous docking studies and structure-activity relationship (SAR) data. To further investigate the most promising candidates and gain deeper insight into their binding behavior, we conducted a computational analysis of two pairs of regioisomers, each consisting of an active compound and an inactive counterpart. This approach enables us to examine critical molecular interactions within the binding site and to identify a rational direction for optimizing lead compounds.

Objects of the research is a library of 24 ligands containing a 1H-1,2,4-triazole ring as the core scaffold, developed for binding to the DRAK1 enzyme.

The goal of the diploma work is to understand the structure-activity relationships observed within a library of 24 ligands that contain a 1H-1,2,4-triazole core scaffold, which have been developed to bind to the DRAK1 kinase. By analyzing the SAR differences between two pairs of regioisomers, this research aims to identify key molecular interactions, investigate the structural characteristics of the binding pocket, and propose new design strategies for developing improved compounds.

CHAPTER 1. MOLECULAR TARGETS IN ONCOLOGY: FOCUS ON KINASES

1.1 The role of kinases in oncology

Tumorigenesis is a multistep process driven by genetic alterations in key regulatory genes that control cellular homeostasis. When these genes are dysregulated, it results in uncontrolled cell proliferation and resistance to apoptosis, often due to the abnormal activation of kinases. Oncogenic signaling involves the transmission of signals from the extracellular environment to the nucleus, mediated by membrane receptors and cytoplasmic signaling cascades, which ultimately promote tumor growth. Cellular signaling depends on the phosphorylation carried out by kinases, which are located at the membrane, in the cytosol, or in the nucleus. In cancer, deregulated kinases disrupt normal signaling pathways by phosphorylating tyrosine residues or the serine and threonine residues of other kinases and proteins, thereby promoting cell proliferation and transformation [2].

The FDA's (Food & Drug Administration) approval of imatinib in 2001 represented a significant advancement in targeted cancer therapy and spurred the development of kinase inhibitors. Since then, protein kinases have become one of the most important classes of drug targets in modern medicinal chemistry (see Figure 1.1) [3].

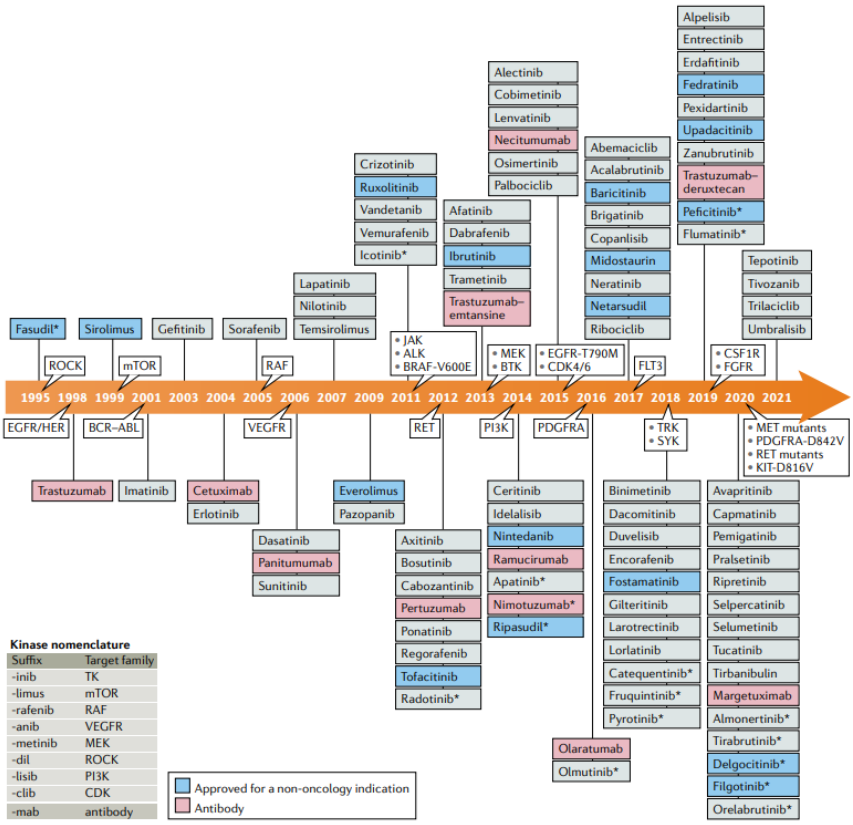


Figure 1.1 – Chronology of kinase inhibitor approvals [3].

As of 2025, the FDA has approved a total of 85 kinase inhibitors. Among these, 46 are designed to target receptor tyrosine kinases, 21 target non-receptor tyrosine kinases, 13 focus on serine/threonine kinases, and 5 target dual-specificity kinases. In 2024, four of these inhibitors received approval, with an additional one approved in 2025. While oncology remains the primary focus for these treatments, several small molecule kinase inhibitors (SMKIs) have also been authorized for non-oncological uses, including conditions like rheumatoid arthritis (notably upadacitinib, tofacitinib, and baricitinib) [3].

Future development will focus on optimizing classical chemotypes and expanding novel ones. This includes designing allosteric and irreversible kinase inhibitors, as well as developing combination strategies with immunotherapies, particularly for lung, renal, and breast cancers. Continued innovation in these areas is expected to enhance the therapeutic potential of kinase inhibitors [4].

1.2 Serine/Threonine kinase 17A (DRAK1): structure and function

Protein kinases are enzymes that mediate the transfer of the γ-phosphate group from a magnesium-coordinated ATP molecule ($MgATP^-$) to the hydroxyl group of a target protein substrate. This phosphorylation reaction yields a phosphoprotein (protein–O–PO_3^{2-}), magnesium-bound ADP (MgADP), and a proton (H^+):

$$MgATP^- + \text{Protein–OH} \rightarrow \text{Protein–O–}PO_3^{2-} + MgADP + H^+$$

By catalyzing the process of phosphorylation, kinases introduce covalent modifications to various proteins, thereby affecting many cellular processes, including proliferation, development, and survival [5].

Protein kinases possess a conserved bilobal structure consisting of a smaller N-terminal lobe and a larger C-terminal lobe, which are linked by a flexible hinge region. The active site, situated between these lobes, contains several conserved motifs, including the glycine-rich loop, DFG motif, catalytic loop, and activation segment. These motifs play crucial roles in coordinating ATP binding and catalysis. A conserved lysine residue helps stabilize the phosphate groups of ATP, while magnesium ions (Mg^{2+}), coordinated by residues in the DFG motif, aid in neutralizing the negative charge of ATP and ensure its proper orientation. Additionally, a distinct aspartate residue in the catalytic loop acts as a base by accepting a

proton from the hydroxyl group of the substrate, facilitating the transfer of the γ-phosphate to a serine, threonine, or tyrosine residue (Figure 1.2) [6].

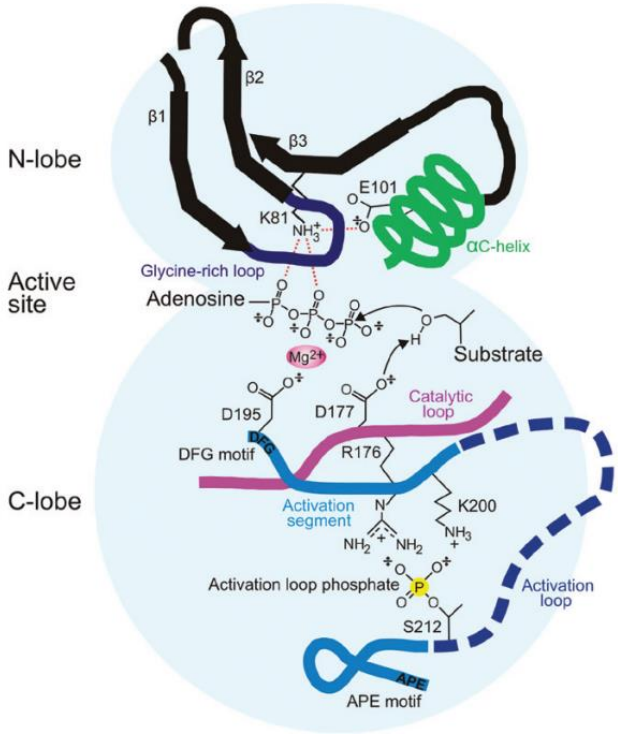


Figure 1.2 – Conserved architecture of a classical kinase active site, exemplified by human mitogen- and stress-activated kinase 1 (MSK1). Structural motifs such as the glycine-rich loop, catalytic loop, DFG motif, and activation segment coordinate ATP binding and phosphate transfer – elements also conserved in DRAK1 [6].

Protein kinases are typically categorized based on the type of hydroxyl-containing amino acid residue that they phosphorylate, leading to the formation of three main categories within this enzyme family:

- protein tyrosine kinases (90 members),
- protein tyrosine kinase-like enzymes (43 members),
- protein serine/threonine kinases (385 members) [7].

The death-associated protein kinase (DAPK) family is a specific subgroup of serine/threonine kinases (STKs) that play a crucial role in regulating various biological processes in human cells. This family includes several members: DAPK1, DAPK2, DAPK3, as well as DRAK1 and DRAK2, which are also known as DAPK-related apoptosis-inducing protein kinases 1 and 2, or STK17A and STK17B, as summarized in Figure 1.3 [5].

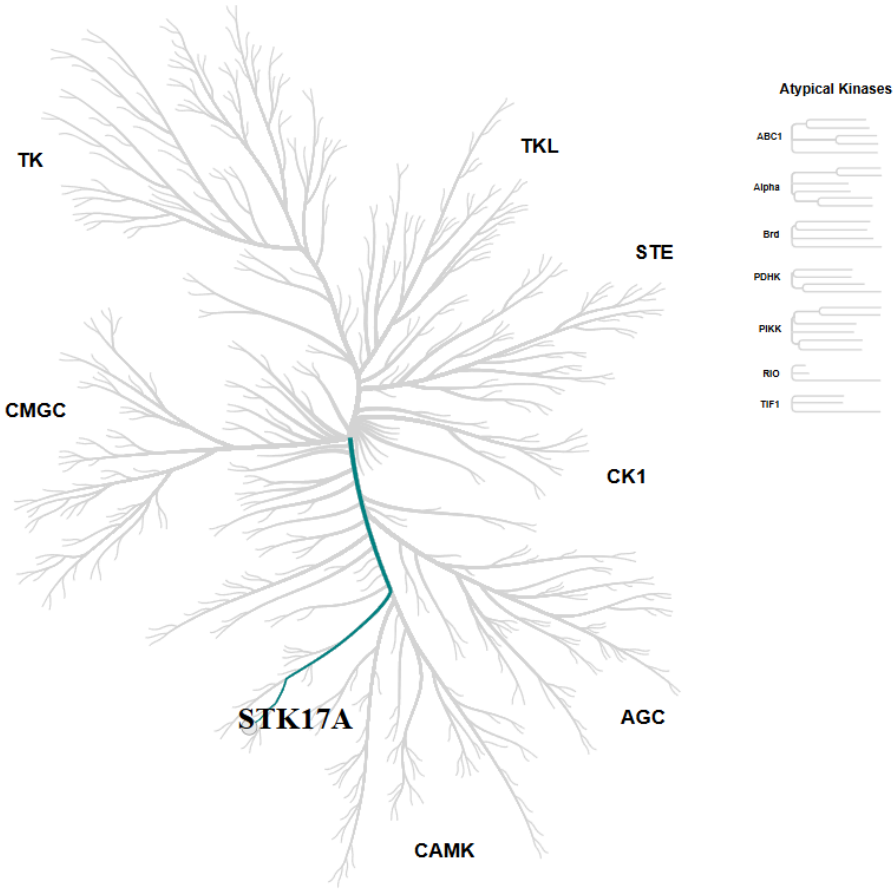


Figure 1.3 – Phylogenetic tree of the human kinome highlighting the position of STK17A (DRAK1). STK17A is classified within the CAMK (calcium/calmodulin-dependent protein kinase) group [8].

DRAK kinases were first identified in 1998 but have been studied less extensively than DAPK1, DAPK2, and DAPK3. DRAK1 and DRAK2, these kinases are encoded by the STK17A and STK17B genes, which are located on chromosomes 7 and 2, respectively. They have molecular weights of 46 kDa (414 amino acids) and 42 kDa (372 amino acids) (see Figure 1.4). Both kinases can undergo autophosphorylation and are capable of phosphorylating the myosin light chain (MLC). They exhibit a sequence identity of 67.1% in their kinase domains and 24.2% in their extra-catalytic C-terminal regions. The kinase domains of DAPK2 and DAPK3 share 80% and 83% identity with DAPK1, respectively, while DRAKs show approximately 50% similarity and partial homology with calcium/calmodulin-dependent protein kinase II (CAMK2) and myosin light chain kinase (MLCK). The five members of the DAPK family are grouped together due to their highly conserved catalytic domains, despite significant differences in their extra-catalytic structures, localization, and molecular weight [5].

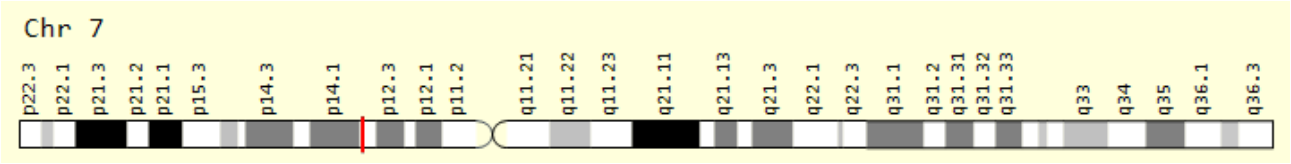
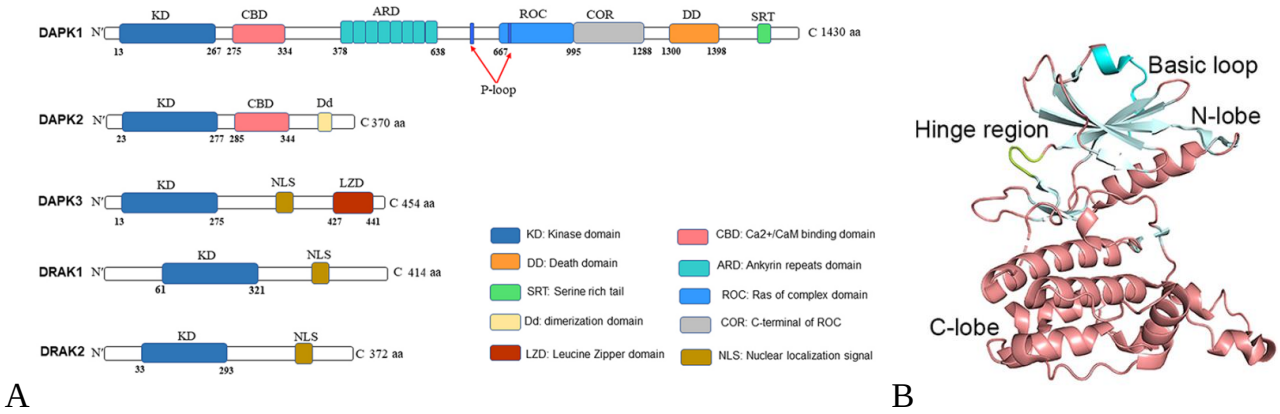


Figure 1.4 – Genomic context of STK17A gene on chromosome 7p13, based on UCSC Genome Browser data (GRCh38/hg38, GRCh37/hg19) and GeneCards custom tracks [9].

Members of the DAPK family possess a conserved catalytic kinase domain that consists of 263 residues in DAPK1-3 or 261 residues in DRAKs. This domain is organized into 11 subdomains and features a small N-terminal lobe, which includes five β-sheets and an αC-helix, and a larger C-terminal lobe predominantly composed of α-helices. These two lobes are connected by a flexible hinge region (see Figure 1.5) [10].

A conserved lysine in the ATP-binding site is essential for the enzyme's activity; mutating this residue in DRAK1 (K90A) eliminates its catalytic function. Another important feature is the basic loop, which consists of 12 basic residues that are crucial for dimerization and activation, though they are not necessary for substrate binding. The αC-β4 loop at the N-lobe terminus interacts with Hsp90, which helps stabilize DAPK1-3. However, the inhibition of Hsp90 by tanespimycin leads to the degradation of these proteins. The role of DRAKs in regulating Hsp90 is still unclear [10].



A B

Figure 1.5 – A) Schematic diagram of the DAPK structures. B) X-ray structure of DRAK1 (PDB code 7QUE). DRAK1 consists of the N-terminal lobe and C-terminal lobe, with the active site situated between them, accommodating the ATP molecule. The basic loop is highlighted in cyan, while the hinge region is shown in yellow [11].

The extracatalytic regions of DRAKs are distinctive and do not share homology with any known proteins. These regions contain nuclear localization signals (NLS) that direct

DRAK proteins to the nucleus, linking them to the regulation of apoptosis. DRAK1 predominantly resides in the nucleus but moves to the cytoplasm when phosphorylated at Ser395 by protein kinase C (PKC). This phosphorylation occurs near the C-terminal NLS. Additionally, an extra NLS within the kinase domain also aids in nuclear localization; mutations in either of these sequences hinder this process [5].

To gain a clearer understanding of the physiological significance of DRAK1, it's essential to evaluate its expression in various human tissues. Transcriptomic data indicate a widespread distribution of DRAK1, with particularly high levels found in blood, the small intestine, and certain regions of the brain as summarized in Figure 1.6 [12].

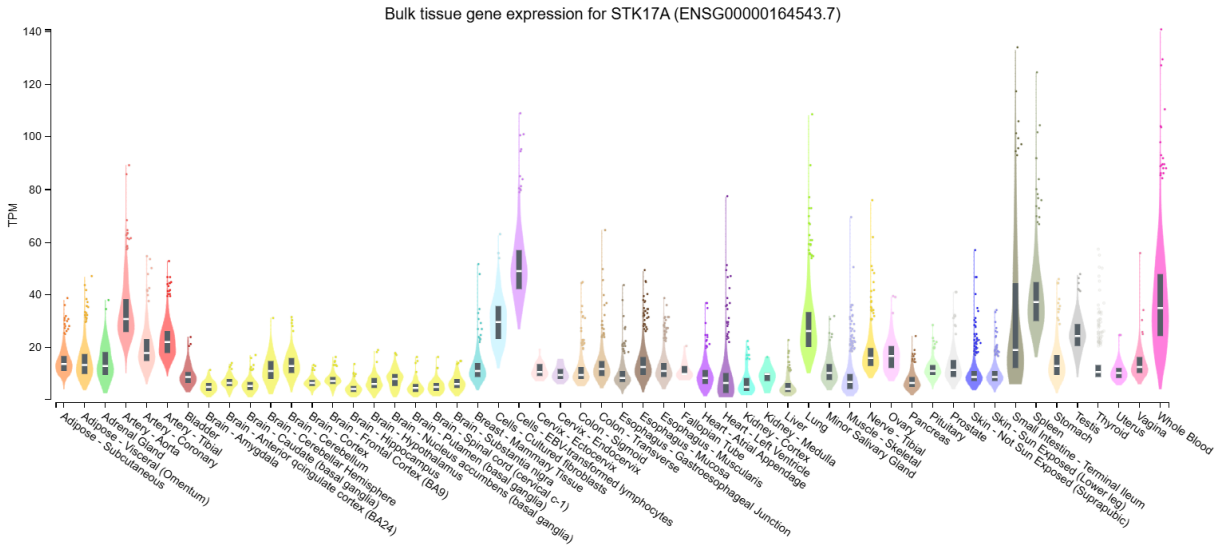


Figure 1.6 – Bulk tissue expression profile of STK17A across normal human tissues. Violin plots represent transcript levels (TPM) measured across tissue types [12].

DRAK1 plays a crucial role in regulating apoptosis and various downstream signaling pathways. Both DRAK1 and DRAK2 can be activated through phosphorylation mediated by PKC. Once activated, DRAK1 has a complex role in both inflammation and tumor progression. In the context of cervical cancer, DRAK1 inhibits pro-inflammatory signaling by promoting the autophagy-mediated degradation of TNF (tumor necrosis factor) receptor-associated factor 6 (TRAF6). This action prevents the activation of the IL-1β pathway and the production of inflammatory cytokines. A loss of DRAK1 is associated with increased levels of TRAF6 and enhanced tumor growth. Furthermore, Figure 1.7 shows that DRAK1 acts as a negative regulator of TGF-β signaling in head and neck squamous cell carcinoma (HNSCC)

by binding to Smad3 and blocking the formation of the Smad3/Smad4 complex. This suppression interferes with tumor suppressor signaling [11].

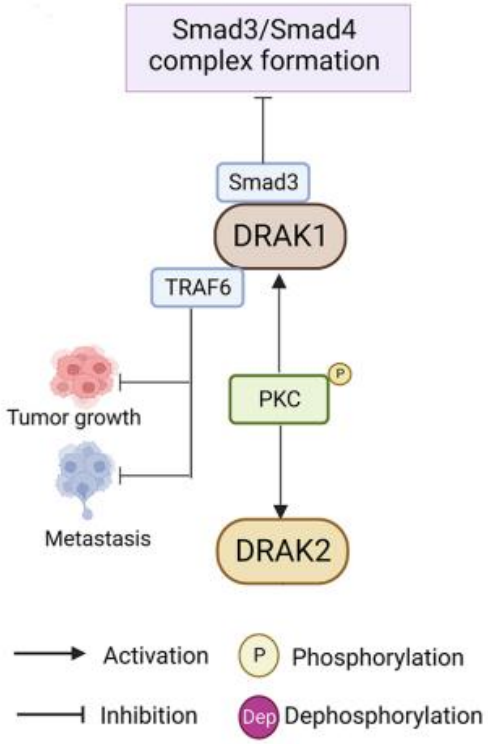


Figure 1.7 – DRAK1-centered regulatory network. PKC-mediated phosphorylation activates DRAK1 and DRAK2, while DRAK1 suppresses TRAF6-dependent inflammation and Smad3/Smad4 signaling, thereby inhibiting tumor growth and metastasis [11].

DRAK1 is known to exhibit oncogenic properties in gliomas, especially in glioblastoma multiforme (GBM), where its expression increases in a manner dependent on tumor grade compared to normal brain tissue. In this context, DRAK1 promotes tumor cell proliferation and enhances resistance to genotoxic stress, rather than showing the tumor-suppressive functions observed in other tissues. When DRAK1 is knocked down in GBM cells, it leads to changes in cellular phenotypes and increases sensitivity to genotoxic agents. Additionally, high levels of DRAK1 expression are associated with poor prognosis [11].

Thus, the combination of conserved kinase domain features and unique structural and biological properties differentiates DRAK1 from other DAPK family members, highlighting its potential as a selective therapeutic target, particularly for glioblastoma treatment [1, 11].

CHAPTER 2. COMPUTATIONAL CHEMISTRY FOR DRUG DESIGN

2.1 Molecular docking

Molecular docking is a widely used computational technique in the field of structure-based drug design (SBDD), with its origins tracing back to the early 1980s. As a component of the broader approach known as Computer-Aided Drug Design (CADD), SBDD utilizes structural information about biological targets to facilitate the rational design of bioactive compounds (see Figure 2.1) [13].

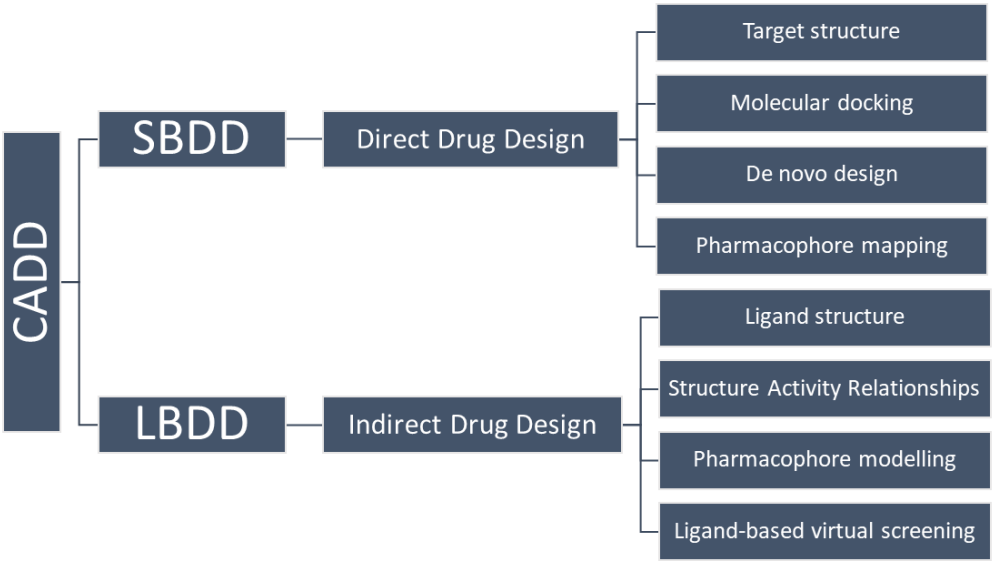


Figure 2.1 – The place of the molecular docking in CADD. SBDD is the Structure-Based Drug Design and LBDD is the Ligand-Based Drug Design [14].

Molecular docking has become a fundamental tool for *in silico* prediction of ligand-receptor interactions and for guiding lead optimization. The primary aim of molecular docking is to predict molecular recognition by identifying potential binding modes and estimating binding affinity. Although it was initially developed for ligand-protein interactions, its applications have now expanded to include protein-protein, nucleic acid-ligand, and nucleic acid-protein-ligand docking [15].

Molecular docking methods can be classified based on the flexibility of the receptor. There are two main categories: rigid docking, where the receptor structure remains fixed, and flexible docking, where the receptor structure can change. Additionally, some docking methods allow flexibility only in specific amino acid side chains. One example of software that supports this approach is AutoDock Vina [14, 16].

Various molecular docking programs utilize different algorithms, each offering unique advantages based on the system and application. This classification is based on search algorithms [16]:

- Systematic or direct algorithms investigate the conformational and positional space of the ligand in a structured and deterministic manner. They include:
    - Conformational search, which samples different torsional angles and spatial orientations of the ligand to identify low-energy conformers that may fit into the binding site.
    - Fragmentation breaks ligands into smaller fragments, which then dock individually and recombine in the binding site.
    - Database search that uses libraries of ligand conformations.
- Random algorithms, also known as stochastic algorithms, use randomized sampling along with heuristics to explore conformational space:
    - Monte Carlo algorithm generates random ligand poses followed by energy-based acceptance criteria. It uses probabilistic rules to accept or reject new conformations, allowing broad exploration of the conformational landscape.
    - Genetic algorithm evolves a population of ligand poses over generations. Operators such as mutation, crossover, and selection applied to generate new conformations with improved binding scores.
    - Tabu search is an iterative local search that keeps track of previously visited (tabu) states to avoid revisiting them. It helps escape local minima and encourages exploration of diverse binding poses [16].

Ligand binding affinity can be estimated using various scoring functions, generally classified into four categories [17]:

- Physics-based methods originated from classical force field models, which estimate binding energies through van der Waals and electrostatic interactions. Subsequent refinements incorporated solvent effects, polarization, and charge distribution.
- Empirical scoring functions estimate binding affinity as the free energy difference upon complex formation by summing weighted energy terms before and after binding. Derived from known 3D structures, they are computationally simpler and faster than physics-based methods.

- Knowledge-based (statistical potential) scoring functions derive interaction energies from pairwise atomic or residue distances, transforming them into potentials via Boltzmann inversion.
- Machine learning techniques analyze intricate relationships among interface features, energetic terms, and solvent-accessible surface areas to accurately predict binding affinity.

First three groups are classic. Hybrid scoring strategies have also been developed [17].

Molecular docking typically involves the following key steps (Figure 2.2) [14]:

1. Preparation of the 3D protein and ligand structures (removing water molecules and co-crystallized ligands from protein structure, adding hydrogen atoms and optimizing geometry of the ligand etc.).
2. Selection of a docking method. Rigid docking is suitable for large-scale virtual screening, while flexible is preferred for complex or adaptive binding sites.
3. Scoring function selection to rank ligand poses based on predicted affinity.
4. Conformational sampling. Prior to docking, ligand conformational space is explored to generate diverse low-energy conformers for input into docking algorithms.
5. Docking execution. Common tools include DOCK, Vina, Glide, and GOLD [18].
6. Refine the initial ligand poses obtained from docking to improve accuracy by using, for example, rescoring techniques.
7. Visualize and analyze to define poses and interactions within the binding site.
8. Validate the docking results by comparing them to experimental data. Interpret the binding poses and interactions [18, 19].

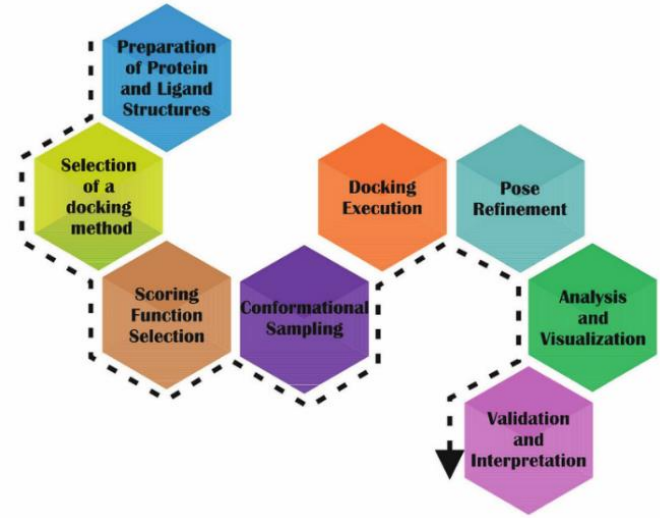


Figure 2.2 – Key steps involved in molecular docking [14].

Here are examples of common software used to perform molecular docking:

- AutoDock Vina – an open-source program for doing molecular docking. It was originally designed by Dr. Oleg Trott at The Scripps Research Institute. Mostly this tool is for rigid docking of ligands. Although some receptor side chains can be chosen to be treated as flexible. AutoDock Vina lacks support for modeling specific features such as macrocycles or explicit water molecules. AutoDock has three main search algorithms: Lamarkian, Simulated Annealing, Genetic Algorithm. It uses an empirical united-atom scoring function [20, 21].
- DiffDock (and DiffDock-L) – a diffusion-based machine learning (ML) model for the docking. DiffDock-L is the newer version of DiffDock [22].
- Boltz-1 – an open-source deep learning model incorporating innovations in model architecture, speed optimization, and data processing achieving AlphaFold3-level accuracy in predicting the 3D structures of biomolecular complexes [23].
- Maestro Schrödinger – is a multifunctional molecular modeling platform that includes tools for molecular docking, notably using the Glide docking engine. Glide uses an empirical scoring function. The software supports both rigid docking and induced-fit (Glide with Prime) docking protocols [24].

Molecular docking plays a versatile role in drug discovery, aiding in structure-activity relationship analysis, lead optimization, virtual screening, binding mode prediction for mutagenesis, and interpretation of X-ray or cryo-EM data through ligand fitting into electron density maps [15].

2.2 Molecular dynamics simulations with GROMACS

Molecular dynamics (MD) is a widely utilized computational method that simulates atomic motions in molecular systems. This technique offers detailed insights into the structure, dynamics, and interactions of these systems. MD enables the exploration of various aspects, such as membrane organization, protein-ligand interactions, and the mechanisms of drug action. Recent advancements in force fields and algorithms have established MD simulation as a vital tool for investigating biomolecular mechanisms, supporting experimental design, and accelerating early-stage drug discovery [25].

MD methods have significantly advanced computational chemistry by enabling detailed modeling of atomic-scale motions with high spatial and temporal resolution. With

improved force fields and simulation techniques, MD simulation can now capture long-timescale events such as protein folding. Widely used MD engines include GROMACS, AMBER, NAMD, CHARMM, LAMMPS, and Desmond [26]. MD simulations estimate the motion of individual atoms within proteins or other molecular systems over time, using a general physical model that describes interatomic forces. At each time step, the force acting on a given atom due to interactions with all surrounding atoms is computed, and then used to update the atom's position and velocity. Repeating this process produces a time-resolved atomic trajectory, essentially a three-dimensional «movie» of molecular behavior [27].

The interatomic forces described by molecular mechanics force fields, which include terms for electrostatic interactions between atoms, spring-like terms that model the preferred length of covalent bonds, as well as additional components for other types of interactions. Importantly, classical MD does not account for bond formation or cleavage. To capture such chemical transformations, quantum mechanics/molecular mechanics approaches are employed [28]. To maintain numerical stability, simulations must use very short time steps, usually a few femtoseconds (10^{-15} s). However, most biologically relevant phenomena, such as conformational rearrangements or ligand binding, occur on nanosecond to microsecond timescales. As a result, only recent advances in computational power and hardware have made such simulations broadly accessible and practical for complex biological systems [28].

GROMACS (GROningen MAchine for Chemical Simulation) originated in the early 1990s at the University of Groningen through a collaboration between the Department of Chemistry and the Department of Computer Science. It evolved from the earlier GROMOS package, initially developed by Wilfred van Gunsteren, who later continued its development at ETH Zürich [29].

GROMACS offers built-in support for a broad range of biomolecules, including proteins, nucleic acids, and lipids, along with commonly used force fields. While implicit solvent models are available, explicit solvent models such as TIP3P, SPC/E, and TIP4P are more commonly used due to their greater accuracy. The software also includes advanced free energy algorithms and employs multithreading for efficient parallel performance across platforms, including standard desktop systems. Combined with optimized assembly kernels and modern parallelization strategies, GROMACS achieves high computational efficiency suitable for both high-throughput and large-scale simulations [30, 31].

The trajectory analysis of a protein or protein-ligand complex typically begins with the evaluation of three parameters: RMSD (root mean square deviation), which tracks the overall

structural stability over time; RMSF (root mean square fluctuation), which reflects the flexibility of individual residues; and R_g (radius of gyration), which indicates the compactness of the molecular structure during the simulation [32].

To compute RMSD parameter is possible by equation 2.1:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}((x_i^m - x_i^1)^2 + (y_i^m - y_i^1)^2 + (z_i^m - z_i^1)^2)}, \qquad (2.1)$$

where x^m, y^m, z^m and x^1,y^1,z^1 represent the initial coordinates and trajectory coordinates at frame t, respectively, N is the number of atoms [32].

RMSF is calculated through the given equation 2.2:

$$RMSF = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(x_i - \bar{x})^2}, \qquad (2.2)$$

where T demonstrates trajectory frame numbers and $\bar{x}$ is the time-averaged position [32].

Lower Rg values indicate greater structural compactness during the simulation. Structural changes in complexes can be quantified using Rg, calculated by the following formula 2.3:

$$R_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|r(i) - r_{center}|^2}, \quad (2.3)$$

where r(i) displays the coordinates of the atom i and r_{center} is the center of mass, N refers to the number of protein atoms [32].

MD simulations go beyond just structural metrics; they also offer various methods to calculate binding and interaction energies. This provides deeper insights into molecular recognition and stability. Overall, MD serves as a powerful and versatile platform for understanding biomolecular behavior at the atomic level, making it an essential tool in modern computational biology and drug discovery.

2.3 Application of the gmx_MMPBSA

Binding free energy analysis highlights key residues relevant for inhibitor design. Stabilization of the inhibitor within the catalytic region is often reflected in its thermodynamic contribution to overall binding. To estimate binding affinity of top-ranked compounds, MM-GBSA (Molecular Mechanics with Generalized Born and Surface Area) and MM-PBSA

(Molecular Mechanics Poisson-Boltzmann Surface Area) are commonly used, offering more reliable predictions than many alternative approaches. These methods offer higher accuracy than most docking scoring functions while remaining less computationally intensive than alchemical free energy methods [33].

In the MMP(G)BSA the free-energy change for binding of a ligand (L) to a protein (P) to form a complex (PL) can be expressed as equation 2.4 [34]:

$$\Delta G_{bind} = \langle G_{PL} - G_P - G_L \rangle_{PLS}, \quad (2.4)$$

where ⟨⟩PLS shows the averaged ensemble from the simulation carried in protein-ligand-solvent system. Therefore, the formula 2.5 for binding free energy is:

$$\Delta G_{bind} = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{solv} - T\Delta S, \quad (2.5)$$

where ΔE_{MM}, ΔG_{solv}, $-T\Delta S$ are the changes in the gas-phase molecular mechanics energy, solvation free energy, and conformational entropy changes upon ligand binding. ΔE_{MM} includes the energy from ΔE_{bonded} terms (bond, angle, dihedral) and $\Delta E_{non-bonded}$ which is decomposed into electrostatic and van der Waals interactions (see equation 2.6) [34]:

$$\Delta E_{MM} = \Delta E_{bonded} + \Delta E_{non-bonded} = \Delta E_{bonded} + \Delta E_{elelectrostatic} + \Delta E_{vdW}, \quad (2.6)$$

where $\Delta E_{bonded} = 0$ for non-covalent protein-ligand interaactions. The solvation term ΔG_{solv} can be decomposed into equation 2.7:

$$\Delta G_{solv} = \Delta G_{polar} + \Delta G_{non-polar} = \Delta G_{PB/GB} + \Delta G_{non-polar}, \quad (2.7)$$

where $\Delta G_{PB/GB} = \Delta G_{polar}$ and it is electrostatic (polar) contribution calculated by PB or GB model. Typically, the dielectric constants are set to 1 for the solute (interior) and 80 for the solvent (exterior). However, research suggests that for highly charged solutes, using a solute dielectric value greater than 1 can improve accuracy, and adjustments may be necessary depending on the system. The $\Delta G_{non-polar}$ is generally calculated based on the solvent-accessible surface area (SASA) or the volume enclosed by SASA (SAV) [34].

gmx_MMPBSA is a versatile tool for end-state free energy analysis using GROMACS MD trajectories. *gmx_MMPBSA* is a Python 3.8 based tool that integrates functionalities from GROMACS and AmberTools to facilitate binding free energy calculations. It supports binding free energy calculations with various solvation models (PB, GB, 3D-RISM), stability assessments, entropy corrections, per-residue energy decomposition, and computational

alanine scanning. Advanced options include interaction entropy and dielectric-adjusted alanine scanning. The package also includes auxiliary tools *gmx_MMPBSA_test* and *gmx_MMPBSA_ana* for streamlined testing and result analysis. Its workflow is divided into three main stages: system preparation, energy calculation, and result analysis and visualization. The core computational tasks are carried out using the MMPBSA.py engine, adapted for seamless use within the *gmx_MMPBSA* framework (Figure 2.3) [35, 36].

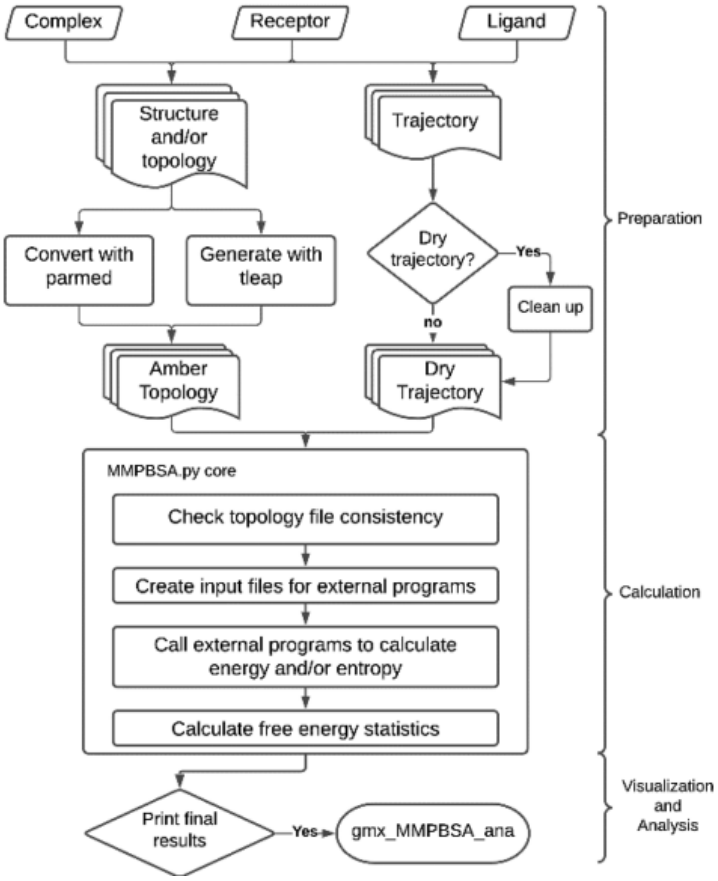


Figure 2.3 – General workflow of *gmx_MMPBSA* for end-state binding free energy calculations. Firstly, Amber topology files are prepared using parmed or LEaP, and GROMACS is used to generate cleaned PDB structures and trajectories. Then binding free energy is calculated using MMPBSA.py. The results are processed and visualized using *gmx_MMPBSA_ana* [35].

In summary, *gmx_MMPBSA* stands out as a powerful and flexible tool for end-state binding free energy calculations. Its ability to handle a wide range of biological systems including protein-ligand, protein-protein, protein-DNA, metalloprotein-ligand, protein-glycan, membrane proteins, and even complex multicomponent assemblies underscores its broad applicability and value in computational drug discovery [35].

CHAPTER 3. RESULTS AND DISCUSSION

3.1 Bioinformatic analysis and target justification for DRAK1 kinase

The DRAKs and DAPKs are part of a subfamily of serine/threonine kinases that have partially overlapping functions. They are classified as paralogs, which are genes or proteins that originate from gene duplication within the same organism and often retain similar functions. Both DRAKs and DAPKs are derived from ancestral DAPK-like kinases.

The closest paralogs of human DRAK1 are DRAK2, DAPK1, DAPK2, and DAPK3, all of which are part of the DAPK (death-associated protein kinase) family. These kinases share conserved domains and functional motifs, but they exhibit varying degrees of evolutionary divergence. The evolutionary relationships among them are depicted in the phylogenetic tree obtained from the KinBase database (Figure 3.1) [37].

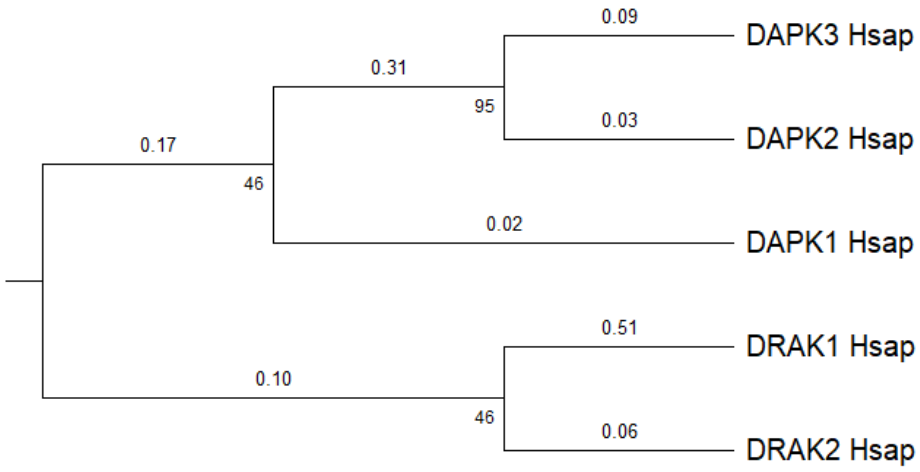


Figure 3.1 – Phylogenetic tree of human serine/threonine kinases family members (Hsap means human).

The branch lengths in the phylogenetic tree represent the estimated number of amino acid substitutions per site, derived from the aligned protein sequences. This metric reflects the evolutionary distance between the proteins: the higher the value, the more divergent the sequences are. Bootstrap support values at the internal nodes, calculated from 100 replicates, indicate the statistical confidence for each branching point [37].

The phylogenetic tree clearly separates the DRAK and DAPK branches within the DAPK kinase family. DRAK1 exhibits the greatest evolutionary distance from the others, indicating its potential for functional specificity.

To assess the sequence similarity between DRAK1 and DRAK2, and to further compare them with members of the DAPK subfamily, a multiple sequence alignment was performed using the MUSCLE algorithm (Multiple Sequence Comparison by Log-Expectation) in Jalview. The alignment was conducted with the BLOSUM62 substitution matrix, which defines substitution scores between amino acids based on evolutionary likelihood. The ClustalW weighting scheme was applied to both weighting steps to balance the contribution of closely related and divergent sequences during the alignment process (Figure 3.2) [38].

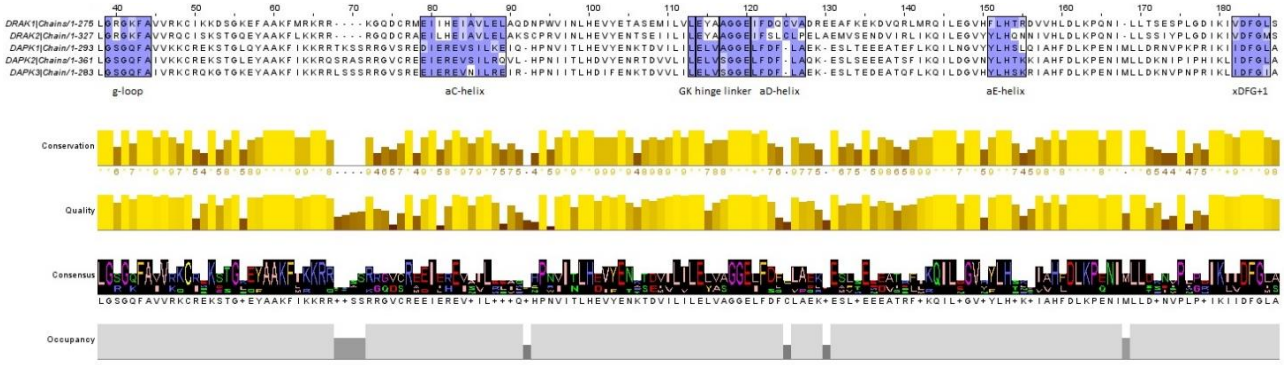


Figure 3.2 – Multiple sequence alignment of the kinase domains from DRAK and DAPK family members. Conserved regions such as the *g*-loop, αC-helix, hinge region, and xDFG motif are annotated. Divergences in sequence highlight structural differences that may underlie functional specificity within the family [1, 38].

There are several conserved motifs within the kinase domain that are essential for ATP binding and catalytic activity, including the G-loop, αC-helix, GK hinge linker, αD-helix, αE-helix, and the xDFG+1 motif. Since our goal is to inhibit selectively the catalytic site, it is particularly important to note that this region exhibits notable sequence divergence between DRAK1/DRAK2 and the DAPK1/2/3 subfamily.

In addition, differences in sequence are observed between DRAK1 and DRAK2, particularly within the catalytically relevant elements such as the αD- and αE-helices. This divergence is further supported by a full-length sequence alignment, which showed only 64.36% sequence identity between DRAK1 and DRAK2.

Multiple sequence alignment indicates that DRAK1 shows considerable divergence from both the larger DAPK1/2/3 subfamily and its nearest paralog, DRAK2. The differences

observed, particularly within important catalytic motifs, suggest that DRAK1 has unique sequence characteristics that could be leveraged for the development of selective inhibitors. These findings emphasize DRAK1 as a promising and distinct therapeutic target within the DAPK kinase family.

3.2 Molecular docking and model selection for MD simulations

A focused library of 24 ligands targeting DRAK1, all sharing a 1H-1,2,4-triazole ring as the core scaffold, was generated based on previous docking studies and structure-activity relationships analysis (Figure 3.3).

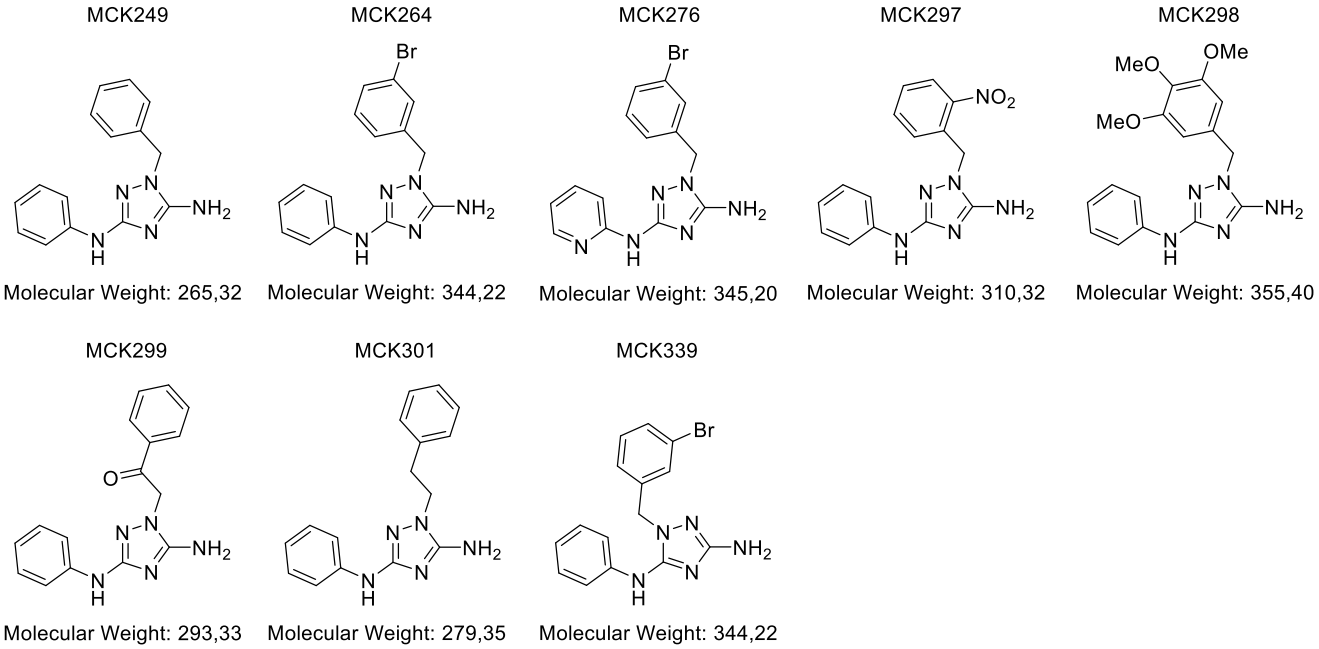


Figure 3.3 – Focused ligand library selected for investigation. Due to copyright issues and regulations from UNIVERSITÉ CÔTE D'AZUR, the structures of some ligands are currently undisclosed until the results will be published in the public domain.

In addition, some reference compounds from the literature were included for comparison during this study. Among these literature compounds, three are known inhibitors reported to bind DRAK1 as an off-target: Tomivosertib, AT9283, and BAY 985. Four are pyrazolo[1,5-a]pyrimidine-based macrocycles – CK23, CK156, CK228, and CKJB68 – reported in a recent study as selective and potent DRAK1 inhibitors [1, 39]. These

macrocycles were developed through structure-based optimization and showed high affinity and selectivity in kinome profiling. And the remaining are potent and selective STK17A/B inhibitors sharing pyrazolo[1,5-a]pyrimidine core scaffold, a fused bicyclic heterocycle commonly used in kinase inhibitors due to its ability to mimic the adenine moiety of ATP and form key hinge interactions in the kinase binding site (Figure 3.4) [40].

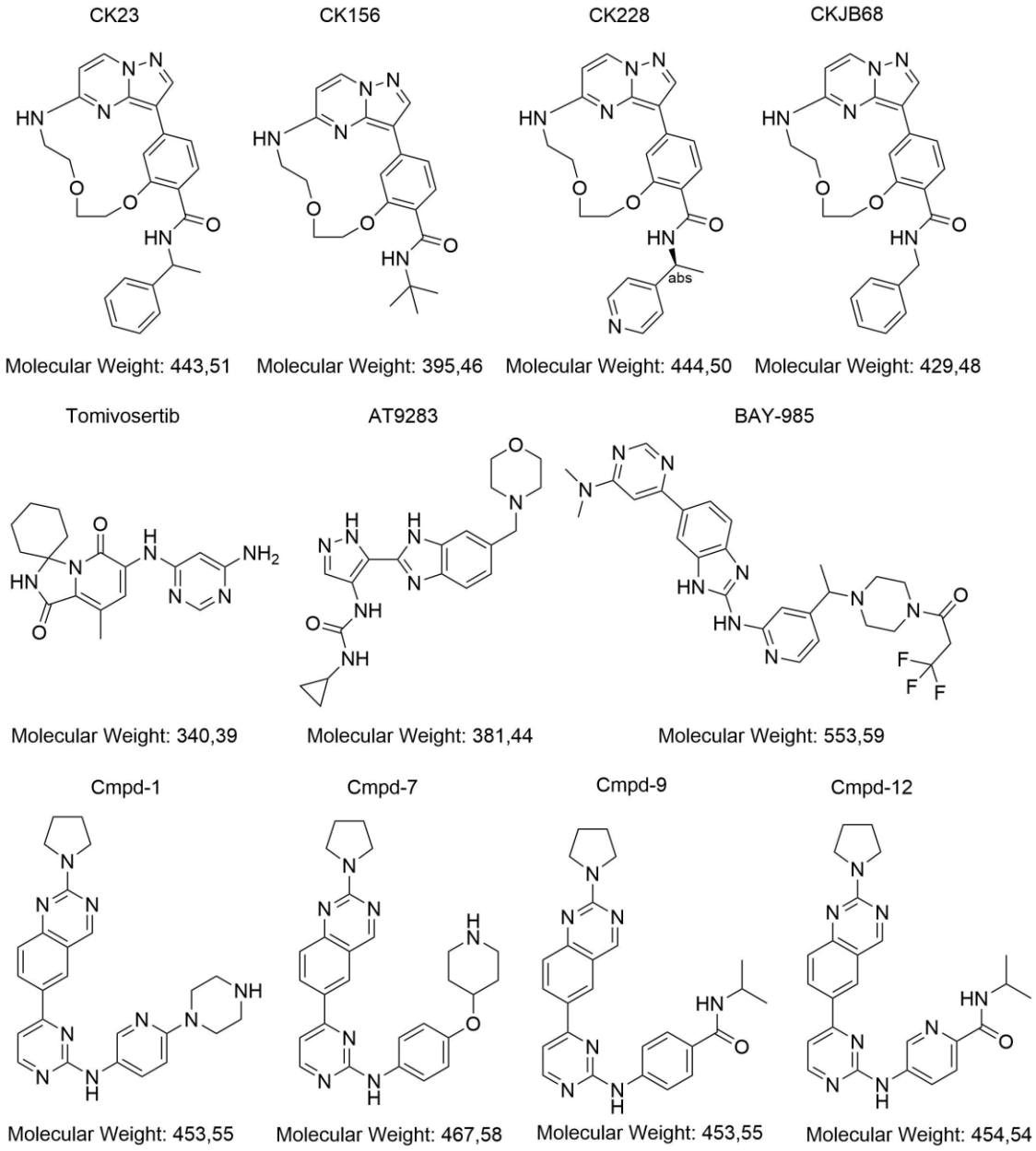


Figure 3.4 – Reference ligands from the literature included for comparative analysis.

Two X-ray structures of DRAK1 are available in the Protein Data Bank: 7QUE and 7QUF. They have resolutions of 2.40 Å and 2.60 Å, respectively. Both structures demonstrate

that DRAK1 exists as a dimer. Each structure exhibits a canonical kinase fold with a hinge region forming part of the ATP-binding pocket. Structural alignment revealed an RMSD of 0.485 Å, mainly resulting from minor differences in flexible loops. Upon comparison of missing residues (5 residues in chain A of 7QUF and 9 residues in chain A of 7QUE), the structure 7QUF was selected for work in FTMove and Boltz-1 [23, 41].

In this study, FTMove was employed to predict potential binding pockets on the DRAK1 protein using the 7QUF structure. FTMove is the novel computational tool that enables dynamic analysis of protein binding sites based on FTMap hotspot detection. FTMove facilitates the identification of potential allosteric sites and off-target binding of known bioactive molecules. In the case of our ligands, allosteric interactions were not detected (see Figure 3.5) [41].

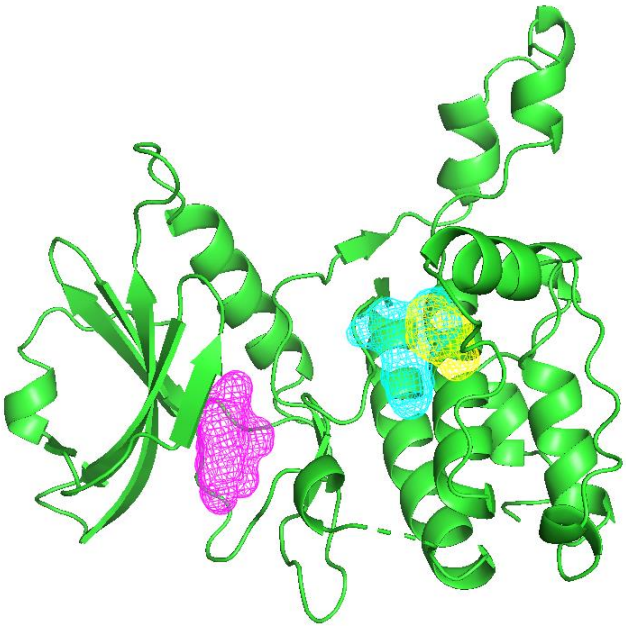


Figure 3.5 – Binding sites in DRAK1 (PDB: 7QUF).

Subsequent docking experiments showed that only the violet-colored pocket was suitable for our ligands, which corresponds to the ATP binding site. To determine if our ligands bind to the same region as ATP, we utilized Boltz-1, an open-source model designed to predict biomolecular complex structures, including combinations of proteins, RNA, DNA, and small molecules. To validate the accuracy of Boltz-1 in our study, we first tested its predictions using the 7QUF structure alongside a known binder, CK156, which is documented to bind at the same site as ATP (see Figure 3.6) [23].

The binding poses predicted by Boltz-1 for both ligands were consistent with experimental data. Our docking studies further confirmed that the ligands under investigation bind to the ATP site, indicating a competitive inhibition mechanism. No evidence was found for co-substrate interactions involving ATP, the ligands, and the protein. ATP forms several key interactions within the DRAK1 binding pocket. The phosphate groups engage in electrostatic and hydrogen bond interactions with positively charged residues, notably Lys43, Lys141, and Asp160, which may play a role in coordination. Additional polar contacts are observed with Gly23 and Gln143. The adenine moiety participates in π-stacking and hydrophobic interactions with Leu20, Ala41, Val28, and Leu146. Meanwhile, Glu92 and Glu98 are positioned to potentially form hydrogen bonds or electrostatic contacts with the ribose or base. These interactions collectively stabilize ATP within the active site and are characteristic of ATP-competitive kinase binding.

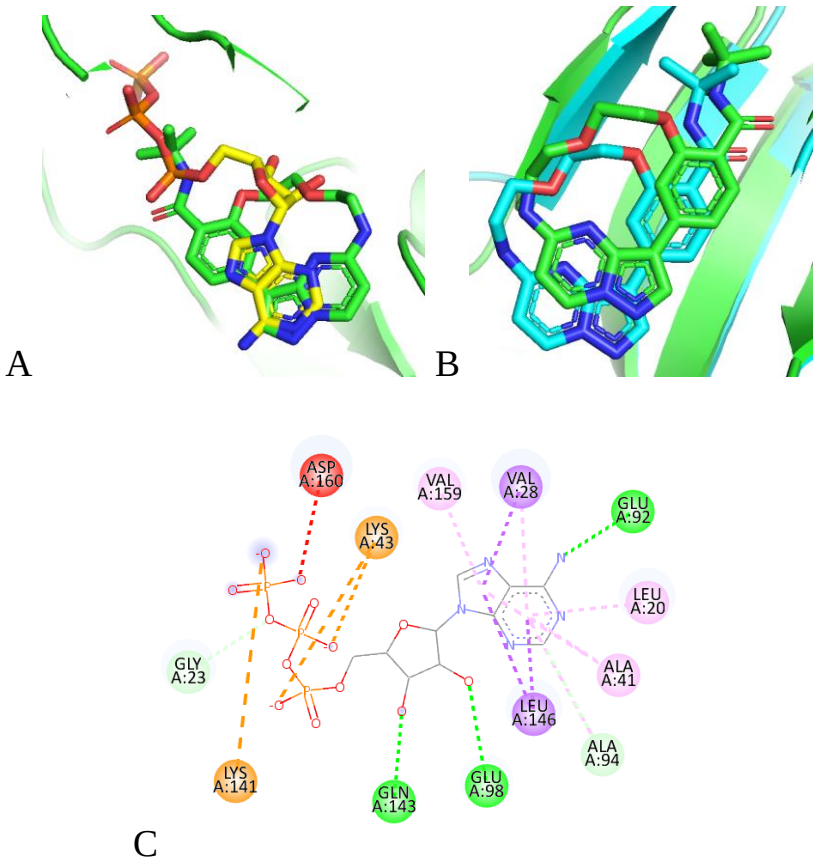


Figure 3.6 – (A) CK156 and ATP in binding site of DRAK1 (PyMOL 3.1). (B) Alignment by PyMOL 3.1 of the predicted complex 7QUF-CK156 with X-Ray structure of Chain A of 7QUF (RMSD = 0,599). (C) Representation of interactions in catalytic site of ATP in Discovery Studio 2024 [42, 43].

The initial rigid docking of our ligand library did not produce results that correlated with the biological assay data. One of possible reasons might be using the rigid receptor model. So, to consider the receptor flexibility and to gain deeper insights into the binding behavior of promising ligands, MD simulations are necessary. However, for accurate MD simulations, a complete protein structure without any missing residues is required. Therefore, the next step was to select an appropriate protein structure for these simulations. Since DRAK1 functions as a dimer, we first verified whether a monomeric form would suffice for the simulations. Rigid docking using AutoDock Vina showed no significant differences in docking scores or ligand poses between the dimeric and monomeric forms (see Table 3.1). Additionally, the docking results were consistent across both structures (7QUF and 7QUE). Consequently, we selected the monomeric form of 7QUF for further investigation [20, 21].

Chain A of the 7QUF structure has several missed residues: Ala152, Asp153, Arg154, Glu155, and Glu156. To address this, a complete structural model was created using the SWISS-Model server to fill in these gaps. Although these missing residues are located away from the binding site, their inclusion is essential for constructing the protein topology and conducting MD simulations.

Table 3.1 Molecular docking calculations using AutoDock Vina

Name	IC50, nM	Predicted binding affinity, kcal/mol			
		7QUF Monomer	7QUF Dimer	7QUE Dimer	SWISS-Model
AT9283	10-30	-8.3±0.1	-8.9±0.1	-8.1±0.1	-8.4±0.2
BAY 985	311	-9.3±0.1	-9.6±0.2	-9.1±0.2	-9.4±0.2
CK23	131	-10.7±0.1	-11.1±0.1	-10.1±0.1	-10.8±0.0
CK156	31	-10.4±0.1	-10.4±0.0	-9.8±0.0	-9.3±0.1
CK228	49	-11.2±0.1	-11.3±0.0	-9.8±0.0	-10.4±0.1
CKJB68	15	-10.7±0.0	-10.8±0.1	-10.7±0.1	-10.7±0.1
Tomivosertib	155	-9.6±0.0	-9.6±0.0	-9.4±0.0	-9.4±0.0
Cmpd-1	18	-10.0±0.2	-10.0±0.1	-9.9±0.2	-10.2±0.1
Cmpd-7	>10,000	-10.0±0.1	-10.0±0.1	-9.8±0.2	-10.0±0,1
Cmpd-9	23	-10.1±0.2	-10.4±0.0	-9.5±0.1	-10.2±0.1
Cmpd-12	12	-9.9±0.2	-9.9±0.2	-9.4±0.1	-10.1±0.1
ATP	-	-7.5±0.2	-7.5±0.1	-7.5±0.1	-7.6±0.1

Name	Binding Activity at 0.5μM, %	Predicted binding affinity, kcal/mol			
		7QUF Monomer	7QUF Dimer	7QUE Dimer	SWISS-Model
MCK249	3	-7.6±0.0	-7.6±0.0	-7.6±0.0	-7.8±0.0
MCK264	23	-7.8±0.0	-7.9±0.0	-8.1±0.0	-7.9±0.1
MCK276	0	-8.1±0.0	-8.1±0.0	-7.9±0.0	-7.8±0.1
MCK297	5	-7.9±0.0	-8.0±0.1	-7.8±0.1	-8.0±0.0
MCK298	0	-7.9±0.1	-8.0±0.2	-7.9±0.1	-8.2±0.2
MCK299	3	-8.3±0.0	-8.3±0.0	-8.0±0.1	-8.5±0.0
MCK301	3	-8.3±0.1	-8.2±0.2	-7.9±0.0	-8.3±0.0
MCK339	1	-7.9±0.0	-7.9±0.0	-7.6±0.2	-8.0±0.0
MCK340	53	-7.9±0.1	-8.0±0.0	-7.9±0.2	-8.1±0.0
MCK341	0	-8.0±0.0	-8.0±0.0	-8.2±0.0	-8.3±0.1
MCK342	73	-8.0±0.0	-8.1±0.1	-8.0±0.1	-8.2±0.0
MCK343	0	-7.8±0.0	-7.8±0.0	-8.0±0.1	-8.0±0.1
MCK344	78	-7.9±0.1	-8.1±0.1	-7.9±0.1	-8.1±0.1
MCK345	3	-7.7±0.0	-7.8±0.1	-7.9±0.1	-8.0±0.1
MCK346	80	-8.0±0.1	-8.0±0.1	-7.8±0.0	-8.1±0.0
MCK347	8	-8.1±0.1	-8.0±0.2	-8.2±0.1	-8.3±0.1
MCK348	65	-8.4±0.2	-8.6±0.1	-8.4±0.2	-8.8±0.0
MCK349	82	-8.1±0.0	-8.4±0.0	-8.4±0.0	-8.3±0.0
MCK350	7	-8.9±0.1	-9.0±0.0	-8.4±0.1	-8.8±0.1
MCK351	14	-7.8±0.1	-7.7±0.1	-7.7±0.0	-7.9±0.1
MCK352	78	-7.7±0.0	-7.7±0.0	-7.6±0.0	-7.8±0.0
MCK385	-	-8.2±0.1	-8.1±0.1	-8.4±0.1	-8.0±0.0
MCK411	100	-8.3±0.1	-8.5±0.0	-8.4±0.0	-8.7±0.1
MCK412	17	-8.4±0.1	-8.6±0.0	-8.2±0.1	-8.6±0.0

Molecular docking calculations using the modeled structure showed results consistent with the original crystallographic structure, and structural validation using PyMOL confirmed its reliability (Table 3.1). Specifically, alignment between the original chain A and the modeled structure yielded an RMSD of 0.092 Å [44 – 46].

It is important to note that the docking scores for the macrocyclic ligands are higher, likely because CK156 was co-crystallized with the DRAK1 structure used in this study. Nevertheless, the predicted binding affinities of our ligands are superior to that of ATP and only slightly lower than those for the macrocycles [1].

The relatively lower docking scores observed for AT9283, BAY 985, Tomivosertib, and others with a pyrazolo[1,5-a]pyrimidine scaffold can be partially attributed to their larger molecular size, as well as their higher number of hydrogen bond donors and acceptors. These factors may lead to suboptimal fitting within the binding pocket. This is especially justified for the latter compounds (Cmpd-1, Cmpd-7, Cmpd-9, Cmpd-12), which are indeed effective inhibitors with an IC_{50} in the nanomolar range [1, 40].

In conclusion, we have confirmed that the mechanism of action of our ligands involves competitive inhibition with ATP. Additionally, we established that simple rigid docking was inadequate to differentiate between the ligands, as correlation with biological activity was not observed. As a result, on the next step, we conducted MD simulations using the refined protein model.

3.3 MD force field selection

In this study, we evaluated two widely used force fields: OPLS-AA (Optimized Potentials for Liquid Simulations – All Atom) and CHARMM36 (Chemistry at HARvard Macromolecular Mechanics, version 36). CHARMM36 is a highly accurate and extensively validated force field designed for proteins and nucleic acids. It is often used in conjunction with a modified version of the TIP3P water model. This modified TIP3P is specifically tailored for CHARMM-based simulations and incorporates Lennard-Jones parameters for hydrogen atoms. This enhancement improves compatibility and structural fidelity during simulations. On the other hand, the OPLS-AA force field is a versatile all-atom parameter set optimized for both proteins and small organic molecules [47 – 49]. The TIP3P water model was chosen as the solvent environment for both force fields because of its common use and computational efficiency. TIP3P represents water as a rigid molecule with three sites, where each atom has assigned partial charges and Lennard-Jones parameters. Its simplicity and established effectiveness make it a suitable choice for comparative MD studies [50].

To evaluate the dynamic behavior of the protein under different force field conditions, MD simulations were performed. Firstly, for 10 ns in NVT conditions 200 K, then 50 ns NPT 250K, and productive MD for 400 ns at 300 K using both CHARMM36/TIP3P and OPLS-

AA/TIP3P systems in GROMACS 2024 [29 – 31]. To determine the most suitable setup for further analysis, key structural parameters: root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and radius of gyration (R_g) were calculated.

Figure 3.7 demonstrates that the OPLS-AA/TIP3P system reveals more stable backbone RMSD values during the initial phase of the simulation. However, over the long-term MD trajectory, the CHARMM36/TIP3P system exhibits slightly greater stability.

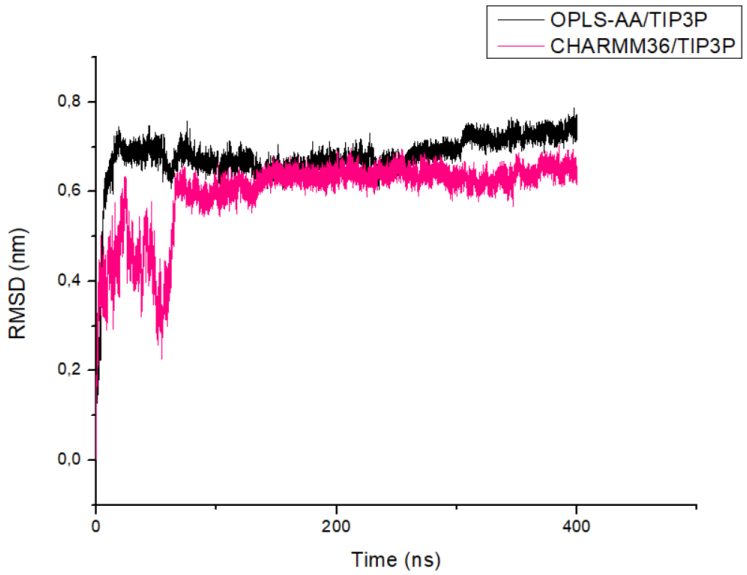
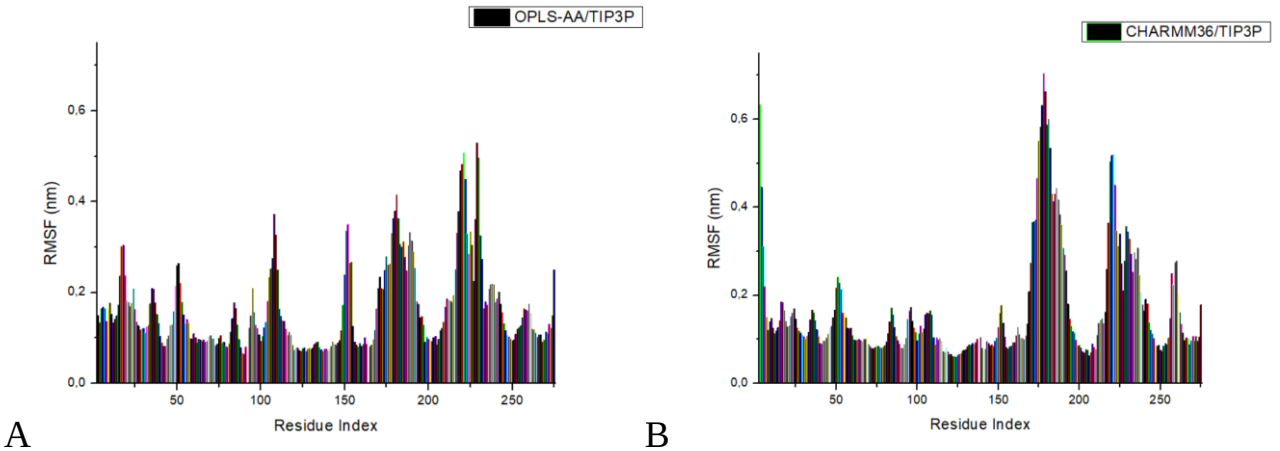


Figure 3.7 – Comparison of backbone RMSD values for OPLS-AA/TIP3P and CHARMM36/TIP3P systems with the protein alone.

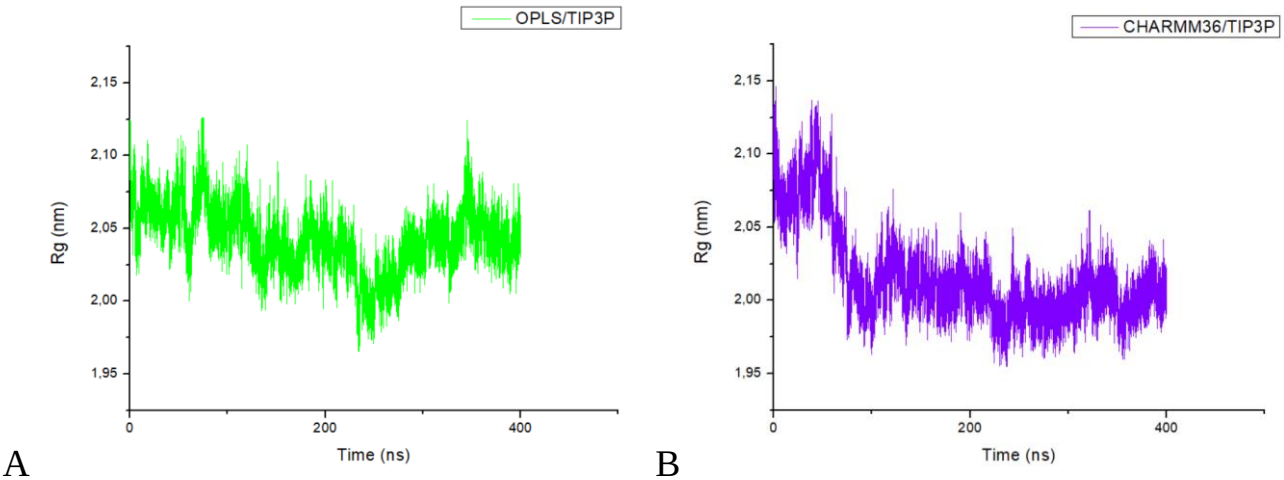
In terms of RMSF, the OPLS-AA/TIP3P system shows slightly lower residue-level fluctuations overall. This suggests better local stability and reduced flexibility throughout the protein structure. Although CHARMM36 model exhibits fluctuations in the terminal regions and in a flexible loop, which is located quite far from the catalytic domain. In the catalytic region, which includes residues such as Lys43, Asp160, Ala94, and Tyr93, the CHARMM36 model performs better, as seen in Figure 3.8.



A B

Figure 3.8 – (A) RMSF profile of the protein in the OPLS-AA/TIP3P system. (B) RMSF profile of the protein in the CHARMM36/TIP3P system.

As illustrated in Figure 3.9, the Rg analysis shows that the CHARMM36/TIP3P system reaches a stable plateau earlier than the OPLS-AA/TIP3P system and maintains more consistent values throughout the simulation. This indicates a more compact and structurally stable conformation.



A B

Figure 3.9 – (A) R_g of the protein in the OPLS-AA/TIP3P system. (B) R_g for the CHARMM36/TIP3P system.

While there can be significant differences between the two force fields, the variations in this case are relatively minor. As a result, both the OPLS-AA/TIP3P and CHARMM36/TIP3P force fields were utilized for a comparative analysis of protein-ligand

complexes. The choice of the most appropriate force field will ultimately depend on the results of these evaluations.

Therefore, productive MD simulations of the protein-ligand complex with MCK411 were performed for 150 ns at 300 K using the structure obtained from molecular docking with AutoDock Vina. The CHARMM36/TIP3P system showed greater structural stability throughout the simulation, as evidenced by consistently lower backbone RMSD values. Although the OPLS-AA/TIP3P system exhibited increasing in RMSD at the beginning, it also achieved a stable conformation toward the end of the trajectory, as seen in Figure 3.10.

To validate the selection of the CHARMM36/TIP3P setup for simulating various protein-ligand complexes, we conducted structural comparisons based on the last frames of the MD trajectories. Specifically, we aligned the initial protein structure prior to the MD simulations with the final frames from the MCK411-protein complex simulations, which utilized both the OPLS-AA/TIP3P and CHARMM36/TIP3P force fields. The resulting RMSD was 3.333 Å for the OPLS-AA/TIP3P system, while the CHARMM36/TIP3P system showed a significantly lower RMSD of 1.702 Å, indicating better structural preservation.

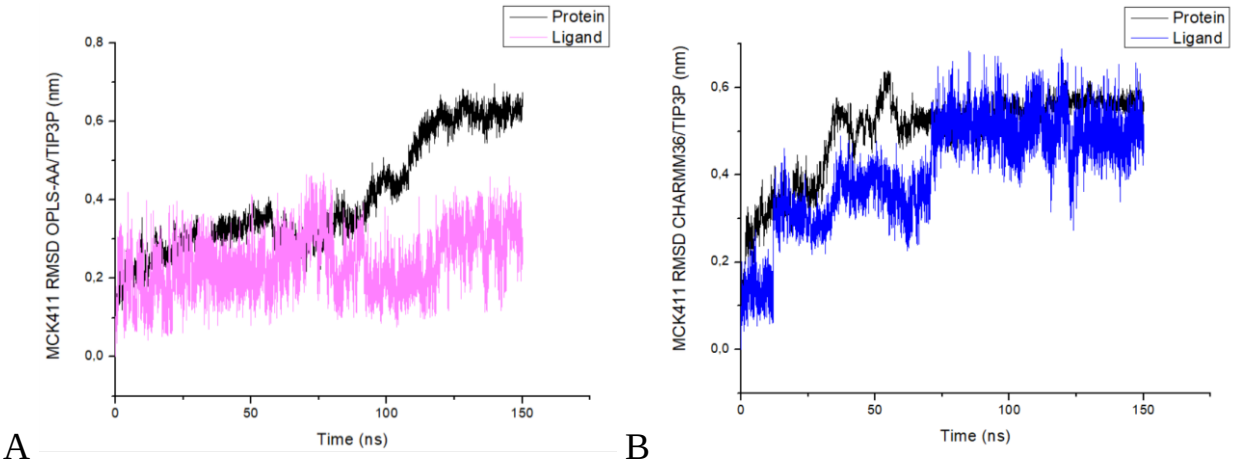


Figure 3.10 – (A) Backbone RMSD values for the protein and ligand in the MCK411–protein complex simulated with OPLS-AA/TIP3P. (B) Backbone RMSD values for the same complex simulated with CHARMM36/TIP3P.

Finally, based on the comparative analysis of structural stability and conformational integrity, the CHARMM36/TIP3P combination was selected as the force field-water model system for all subsequent MD simulations.

3.4 Comparative analysis of protein-ligand complex trajectories

Our ligand library consists of two main components, built around a constant core scaffold represented by the 1H-1,2,4-triazole ring. The first subset includes two hydrophobic moieties: bromobenzene and another benzene derivative. Among this series, MCK349 exhibited the most promising activity (see Table 3.1), while its regioisomer, MCK385, showed no activity in biological assays. This pair was selected for MD simulations to investigate key interactions, involved amino acid residues and potential binding mechanisms.

Additionally, MD simulations were conducted for MCK411 and MCK412, which are structural isomers that differ only in the position of two atoms. *MCK411 is our lead inhibitor*, while *MCK412 is biologically inactive*. The MD simulations aimed to explore two key aspects. First, binding free energy calculations were performed using the MMPBSA method (gmx_MMPBSA, excluding entropy) to determine if the results correlate with the experimental biological activity, thereby validating the MD approach for this system. Second, per-residue energy decomposition was carried out to identify specific interactions between the ligands and the DRAK1 monomer. This analysis aims to enhance our understanding of the binding mechanism and guide further structural optimization of MCK411.

The results of MD simulations can be influenced by the initial binding pose of the ligands. To address this, we considered various starting conformations for each ligand. These included poses generated by AutoDock Vina and alternative poses predicted by DiffDock-L, enabling a more comprehensive comparison of their binding behaviors.

The MD protocol consisted of 10 ns of equilibration under NVT conditions at a temperature of 200 K, followed by 50 ns of NPT equilibration at 250 K. This was followed by a production run of 200 ns under NPT conditions at 300 K. In the case of MCK385, when starting from the pose generated by AutoDock Vina (see Figure 3.11), the ligand dissociated from the binding pocket within the first nanoseconds of the production run, which is consistent with the structure-activity relationship data for this series. For MCK349, MCK411, and MCK412, complete 200 ns production simulations were conducted. The resulting trajectories were analyzed in terms of RMSD, per-residue energy decomposition, and overall binding energy using the *gmx_MMPBSA* tool.

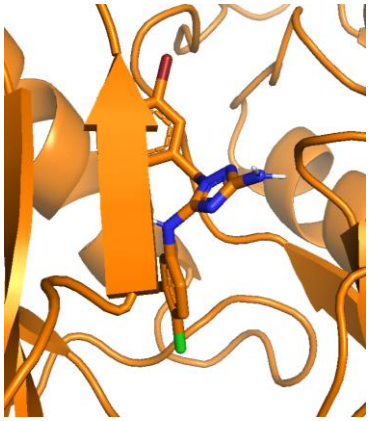


Figure 3.11 – (A) Initial binding pose of MCK385 obtained with Autodock Vina.

The interaction map generated by Discovery Studio 2024 using the final frame of the MCK349 MD trajectory revealed several key contacts.

It should be noted that the 1H-1,2,4-triazole core interacts with residues Ala94 and Tyr93 in the hinge region, which justifies the choice of this particular scaffold. Additionally, the dichlorobenzene group occupies the ATP-binding site, effectively blocking the catalytic region. This area includes the residues Lys43 and Asp160, which are essential for ATP coordination, as they typically interact with the phosphate groups of ATP.

For the active compounds MCK349 and MCK411, a stepwise insertion into the inner cavity was observed, as indicated by their RMSD profiles. Both ligands achieved stable conformations, with RMSD values consistently remaining below 2 Å throughout the simulation. The final binding poses were maintained for extended periods during the third stage of integration, approximately 75 ns for MCK349 and 100 ns for MCK411, demonstrating stable binding. In contrast, MCK412, which is biologically inactive, exhibited significant fluctuations in its RMSD trajectory, suggesting a less favorable and unstable interaction with the binding pocket, as summarized in Figure 3.13.

Binding free energy calculations were conducted using the *gmx_MMPBSA* tool, which employs the Poisson-Boltzmann (PB) model to estimate solvation energy. The PB method was chosen based on recommendations from the developers for use with CHARMM force fields, as it tends to provide more reliable results than the Generalized Born (GB) model in this context. While entropy was not included in the energy calculations, both gas-phase interaction energy, consisting of electrostatic and van der Waals contributions in a vacuum, and solvation free energy, comprising polar and nonpolar contributions, were taken into account. It is important to note that the calculated binding energies are not absolute values;

rather, they serve as relative estimates that are useful for comparing ligands within the same system.

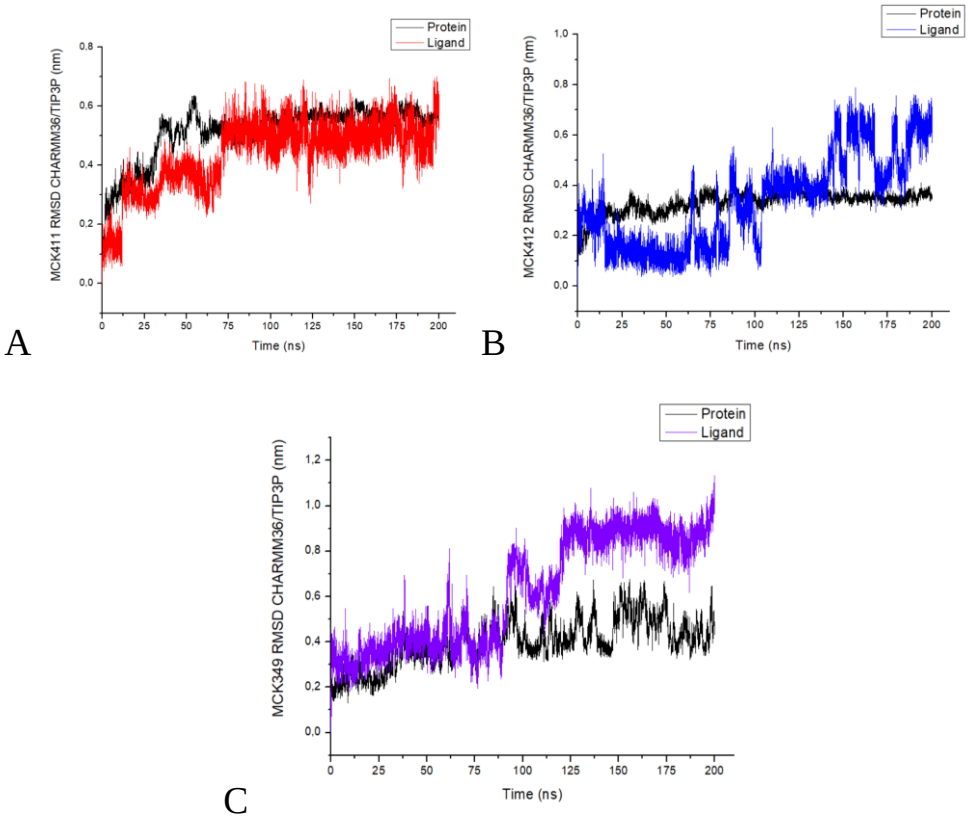


Figure 3.13 – (A) RMSD profile of the DRAK1-MCK411 complex over the 200 ns MD simulation. (B) RMSD profile of the DRAK1-MCK412 complex. (C) RMSD profile of the DRAK1-MCK349 complex.

For each ligand, binding energy was calculated during the portion of the trajectory where the complex achieved a relatively stable conformation, as indicated by the RMSD analysis. This analysis focused on the interval from 150 to 200 ns of the simulation (see Figure 3.14). The binding energy profile for the inactive MCK412 exhibited relative stability; however, this suggests that relying solely on energy calculations may not explain ligand-protein interactions. The RMSD trajectory showed noticeable fluctuations, indicating that binding stability should be evaluated using a combination of parameters to capture the complex behavior fully. In contrast, the results for MCK349 and MCK411 were consistent and aligned well with their biological activity, reinforcing the reliability of the combined analysis approach in these cases.

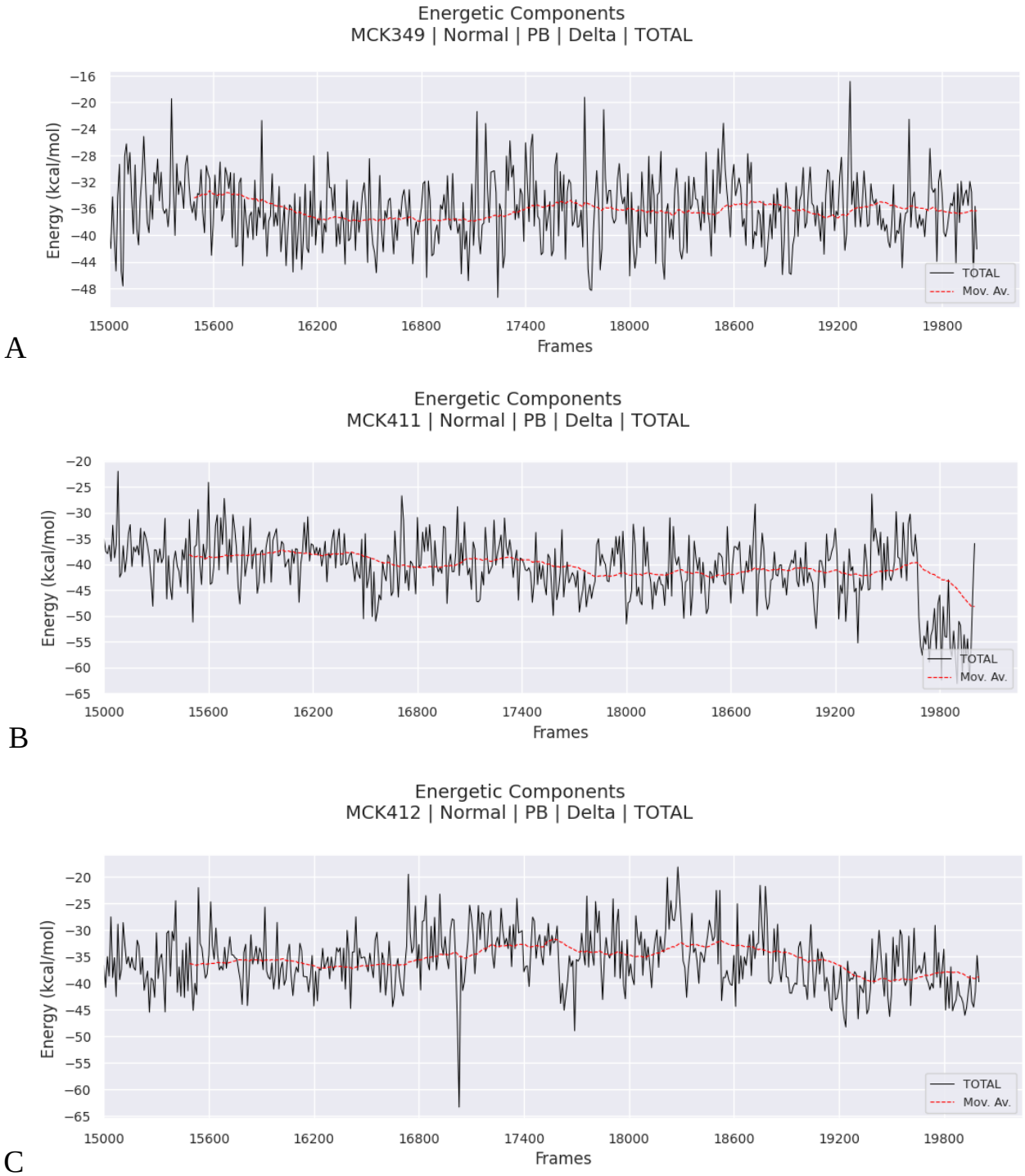


Figure 3.14 – (A) Binding energy profile over the last 50 ns of the MD trajectory for MCK349. (B-C) Binding energy profiles for MCK411 and MCK412, respectively.

Binding free energies were calculated for the 150-200 ns segment of the molecular dynamics trajectories, as shown in Table 3.2. Although a general trend can be observed between the predicted energies and the experimental activity, the correlation remains weak. Overall, strong predicted binding does not necessarily indicate high biological activity. In MCK411, we see an ideal case where both biological activity and predicted binding energy consistently indicate a strong inhibitor.

It is important to note that while MCK412 shows a moderately favorable binding energy, no inhibitory effect was observed in the biological assay. This indicates that *binding energy alone cannot fully explain the activity and should be considered alongside other structural and dynamic parameters*.

Table 3.2 Correlation between biological activity and calculated binding energy

Name	Binding Activity at 0.5μM, %	Predicted Binding Free Energy*, kcal/mol
MCK412	17	-35.84
MCK349	82	-36.18
MCK411	100	-40.78

*Calculated using *gmx_MMPBSA* over the 150-200 ns interval of the MD trajectory.

Per-residue energy decomposition highlights the key residues involved in ligand recognition. For MCK349, strong favorable contributions were noted from Leu146, Leu20, and Tyr93, which were also identified in the interaction maps generated using Discovery Studio 2024 (see Figure 3.15). In contrast, Lys43 consistently contributed unfavorable results for MCK349. As a result, this residue became a focal point for optimization in the MCK411 ligand series.

Current efforts are focused on the catalytic region, particularly on Asp160 and Lys143, to enhance binding in this area.

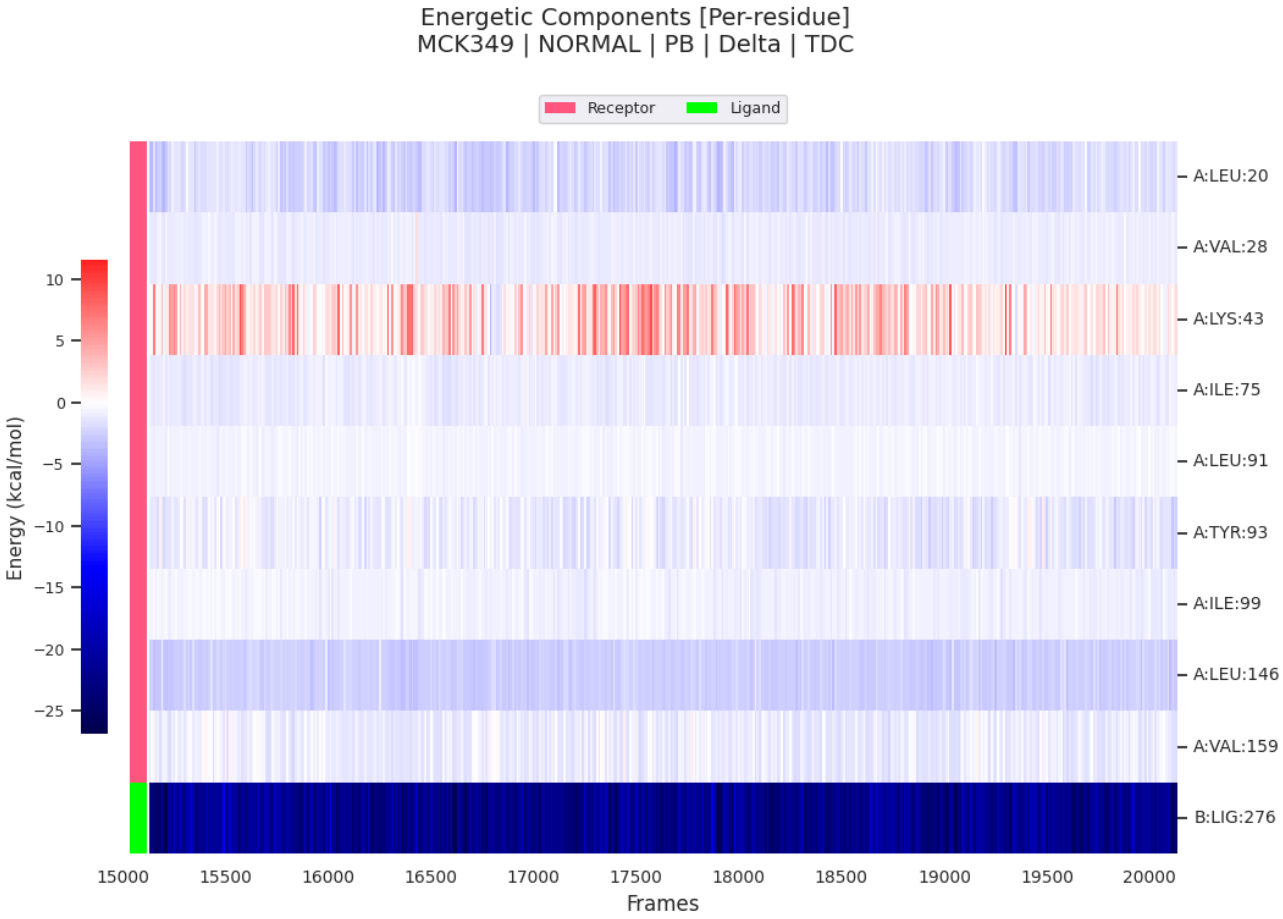


Figure 3.15 – Per-residue binding energy contributions for the DRAK1-MCK349 complex (150-200 ns, PB model).

Compared to MCK349, MCK411 shows a broader range of favorable interactions, particularly in the hinge region (Figure 3.16). Residues Glu92, Ala94, Tyr93, and Gly97 contribute significantly to ligand stabilization, reinforcing the role of the triazole core in this region. A strong favorable contribution is also observed from Lys43, which is especially important as it is located near the region where MCK411 and MCK412 differ by a single atom. In contrast, Asp160 shows a consistent unfavorable contribution, highlighting a potential area for structural refinement in future lead optimization efforts.

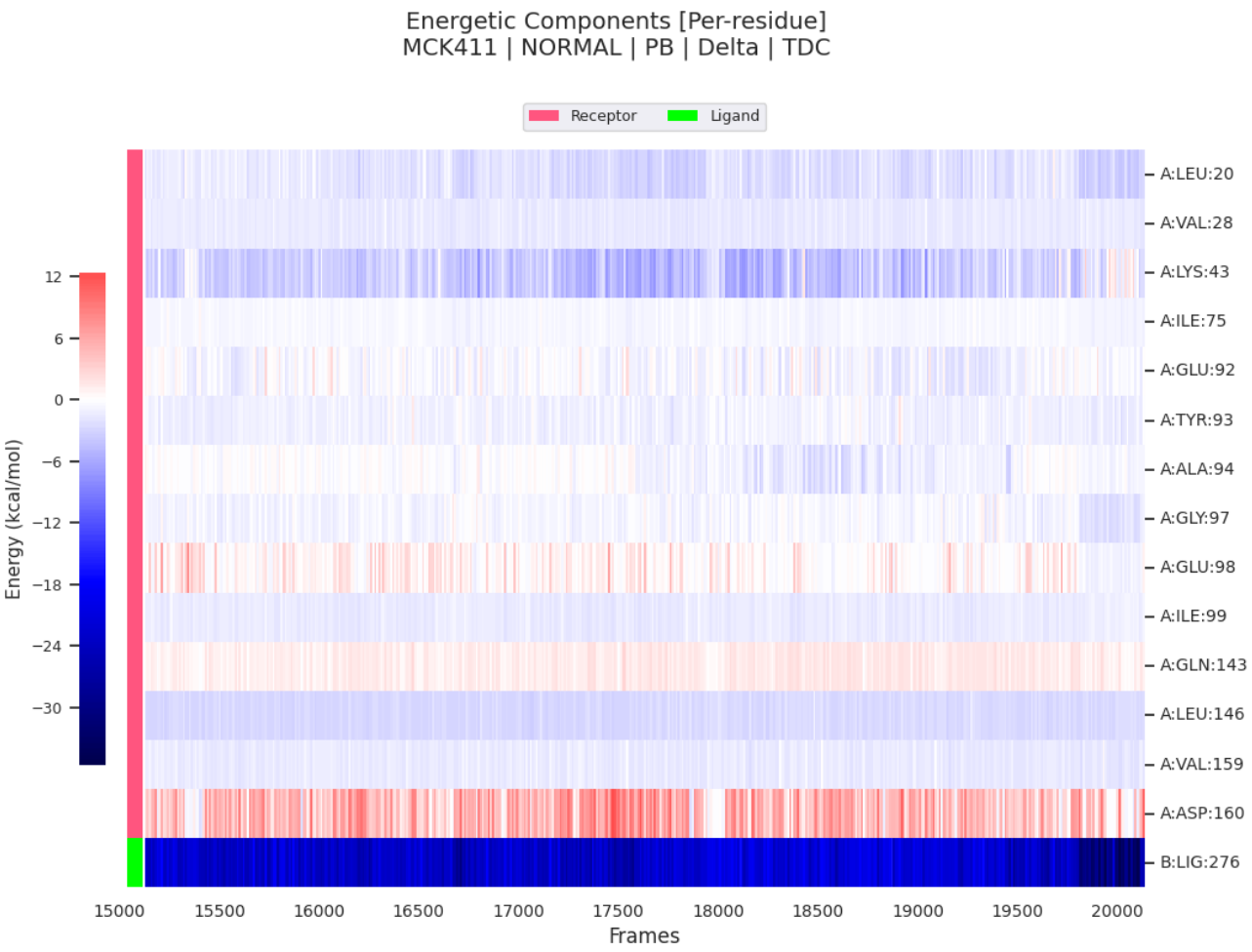


Figure 3.16 – Per-residue binding energy contributions for the DRAK1-MCK411 complex (150-200 ns, PB model).

In comparison to MCK411, the MCK412 ligand exhibits fewer favorable interactions overall (see Figure 3.17). Although the structural difference between MCK411 and MCK412 involves only two atoms, it has a significant impact on key interaction patterns. Notably, the interaction with Lys43 becomes strongly unfavorable in MCK412, while it was favorable for MCK411. Conversely, there is a slight improvement in the interaction with Asp160 in MCK412, which can be attributed to the second site of structural variation between these two ligands. However, this positive shift is not enough to offset the loss of stabilizing contacts, such as those with Ala94 in the hinge region, as well as the unfavorable interaction with Lys43.

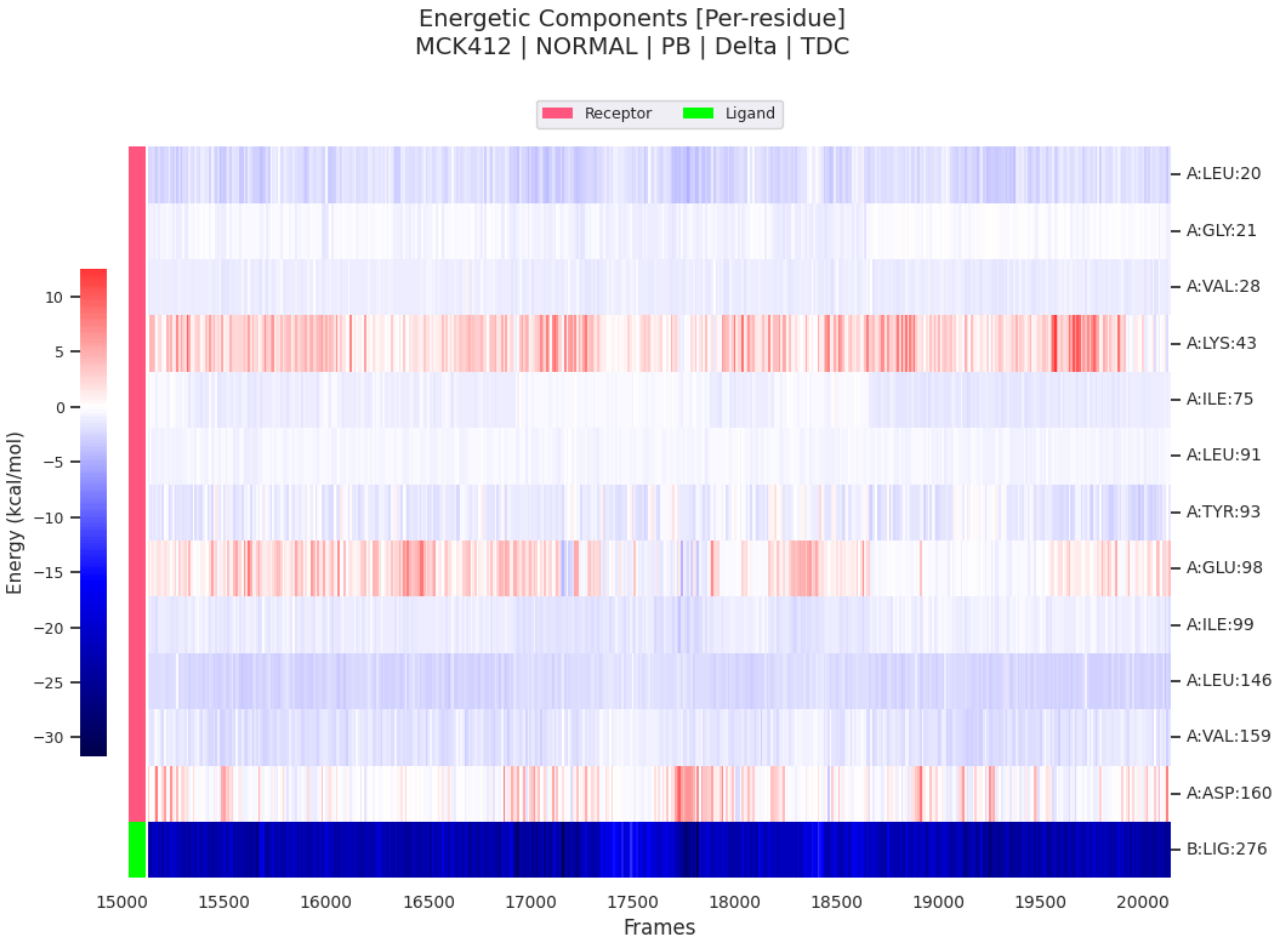


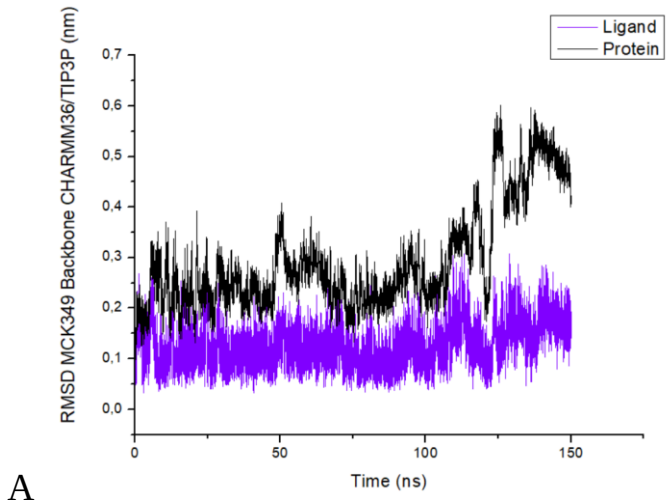
Figure 3.17 – Per-residue binding energy contributions for the DRAK1-MCK412 complex (150-200 ns, PB model).

Thus, MD simulations revealed a three-step integration pattern of ligands into the binding pocket, as shown by the RMSD profiles of all three compounds. The simulations also confirmed that MCK385 was inactive, as it rapidly dissociated from the pocket. Additionally, the per-residue energy analysis identified key residues responsible for ligand binding, which provided a clear explanation for the significant difference in activity between the structural isomers MCK411 and MCK412. Overall, these findings suggest a promising avenue for future ligand optimization by emphasizing a favorable direction for structural expansion.

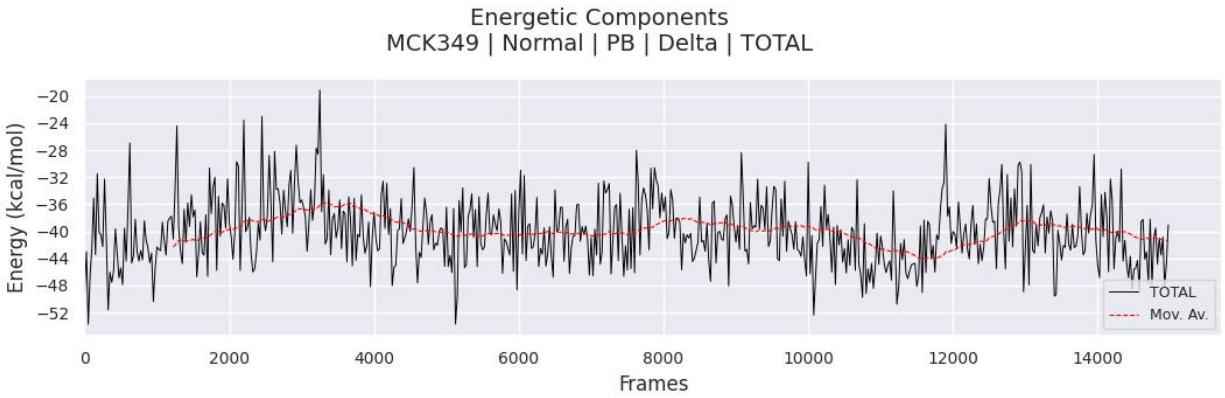
We initially conducted MD simulations based on docking poses obtained from AutoDock Vina. To enhance our analysis, we also tested alternative poses generated using DiffDock-L, a machine learning-based method that utilizes a diffusion model to predict ligand binding without the need to define a fixed binding site. Unlike rigid docking approaches such as Vina, DiffDock-L allows for some flexibility in both the ligand and the receptor by not assuming fixed atomic positions, which can better approximate aspects of the induced fit mechanism. However, a limitation of this method is that it does not include hydrogen atoms

during pose generation. As a result, the resulting complexes often lack chemical plausibility, particularly in cases where hydrogen bonding or precise geometry is critical. This limitation has been acknowledged in the original DiffDock implementation, which omits hydrogens to simplify modeling, based on the premise that they are frequently missing from PDB structures and were not essential for the model's training.

In the case of the MCK411 complex, the ligand gradually reoriented during the simulation, eventually adopting a pose similar to the original Vina-derived conformation. This behavior highlights the importance of hydrogen atoms in pose stability and suggests that DiffDock-L may not be suitable for our system. In contrast, for MCK349, the alternative pose remained stable throughout the simulation, with favorable RMSD and binding energy profiles, indicating a reliable binding mode (see Figure 3.18).



A



B

Figure 3.18 – (A) RMSD plot of the corresponding trajectory. (B) Binding free energy calculated with *gmx_MMPBSA* remained stable throughout the entire trajectory (average ΔG = –40.08 kcal/mol).

A per-residue energy decomposition analysis is also performed. While a comparable number of residues contribute favorably to the overall binding energy, often with stronger

individual contributions, several residues exhibit unfavorable interactions, which partially offset the energetic gain (Figure 3.19).

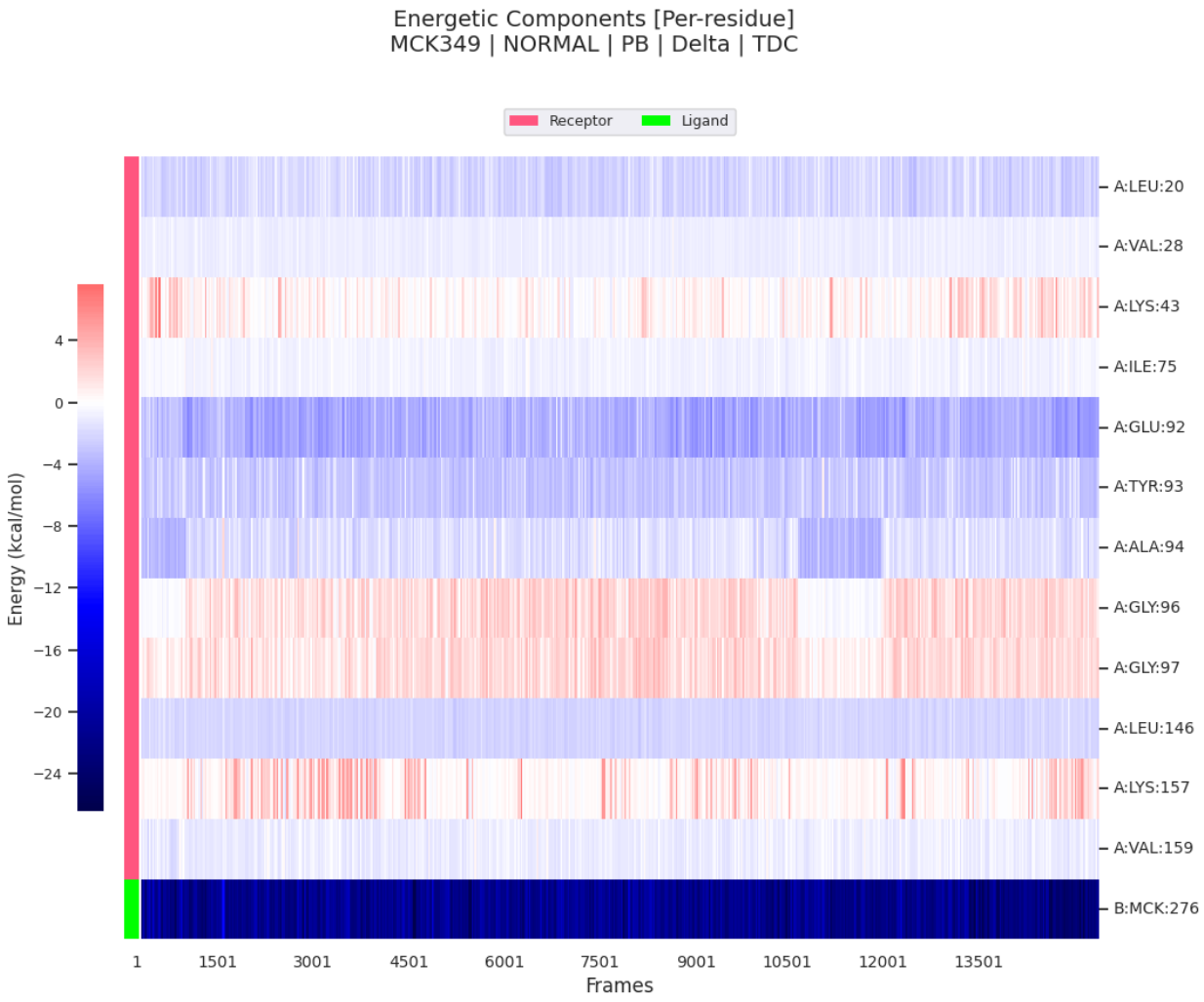


Figure 3.19 – Per-residue binding energy contributions for the DRAK1-MCK349 complex (150 ns, PB model).

Despite the promising calculation results, this simulation does not clarify the structure–activity relationship observed in this series of compounds. The dichlorobenzene group, which remains exposed to the solvent in this pose, has been substituted with various groups in other analogs, leading to different levels of biological activity. This situation underscores the importance of evaluating all available computational data in an integrated manner, rather than relying solely on a single parameter. In contrast, the molecular dynamics simulations based on AutoDock Vina poses show a better correlation with SAR trends and offer more consistent mechanistic insights. Notably, the DiffDock-L pose for MCK411, our lead compound, was

invalidated during MD simulations, which further reinforces our decision to prioritize Vina-derived conformations in subsequent analyses.

In summary, MD simulations provided valuable insights into the behavior of our ligand-protein complexes. Through the analysis of backbone RMSD profiles, binding energy calculations, and per-residue energy decomposition, we were able to explain the significant differences in biological activity between the structural isomers MCK411 and MCK412, as well as between MCK349 and MCK385. Based on these analyses, we identified key interactions and a promising pathway for further lead optimization, and we are currently developing a new series of analogs. Additionally, our computational results suggest that docking poses obtained with AutoDock Vina offer a more reliable foundation for MD simulations than those predicted by DiffDock-L.

CONCLUSIONS

In this MSc thesis, various computational methods were utilized to analyze the structure-activity relationships of the triazole-based ligand library targeting the DRAK1 kinase. The key findings are summarized below:

- To validate the receptor model, we conducted molecular docking on both the monomeric and dimeric forms of the DRAK1 structure (PDB ID: 7QUF) using reference inhibitors along with our local library of ligands. The similarity in the docking scores supported *the choice of the monomer* as a suitable basis for building a model for molecular dynamics simulations.
- A comparative analysis of OPLS-AA and CHARMM36 force fields, based on RMSD, RMSF, and radius of gyration (R_g), indicated that the CHARMM36/TIP3P combination provides enhanced stability for protein-ligand complexes; therefore, it was selected for all subsequent MD simulations.
- Molecular dynamics simulations were conducted on two pairs of regioisomers. These simulations enabled us to evaluate conformational stability, estimate binding free energies using MM-PBSA methods, and perform per-residue energy decomposition to identify crucial binding interactions and explain differences in structure-activity relationships.
- Our findings highlighted crucial ligand-receptor interactions and proposed new ligand growth vectors within the DRAK1 binding pocket, providing a solid foundation for future inhibitor design.

A new series of compounds is currently being developed based on the results obtained in this study.

REFERENCES

1. Kurz C. G., Preuss F., Tjaden A., Cusack M., Amrhein J. A., Chatterjee D., etc. Illuminating the dark: highly selective inhibition of serine/threonine kinase 17A with pyrazolo[1,5-a]pyrimidine-based macrocycles. J. Med. Chem. 2022, 65 (11), 7799–7817.

2. Tsatsanis C., Spandidos D. A. The role of oncogenic kinases in human cancer (Review). Int. J. Mol. Med. 2000, 5 (6), 583–590.

3. Attwood M. M., Fabbro D., Sokolov A. V., Knapp S., Schiöth H. B. Trends in kinase drug discovery: targets, indications and inhibitor design. Nat. Rev. Drug Discov. 2021, 20 (11), 839–861.

4. Roskoski R. Jr. Properties of FDA-approved small molecule protein kinase inhibitors: A 2025 update. Pharmacol. Res. 2025, 216, 107723.

5. Farag A. K., Roh E. J. Death-associated protein kinase (DAPK) family modulators: Current and future therapeutic outcomes. Med. Res. Rev. 2019, 39 (1), 349–385.

6. Duda K., Frödin M. Stimuli that activate MSK in cells and the molecular mechanism of activation. MSKs, Landes Bioscience, 2012, Chapter 2, 1–19.

7. Roskoski R. Jr. Properties of FDA-approved small molecule protein kinase inhibitors: A 2024 update. Pharmacol. Res. 2024, 200, 107059.

8. Berginski M. E., Moret N., Liu C., Goldfarb D., Sorger P. K., Gomez S. M. The Dark Kinase Knowledgebase: an online compendium of knowledge and experimental results of understudied kinases. Nucleic Acids Res. 2021, 49 (D1), D529–D535.

9. Barshir R., Fishilevich S., Iny-Stein T., Zelig O., Mazor Y., Guan-Golan Y., Safran M., Lancet D. GeneCaRNA: a comprehensive gene-centric database of human non-coding RNAs in the GeneCards Suite. J. Mol. Biol. 2021, 433 (11), 166913.

10. Velentza A. V., Schumacher A. M., Weiss C., Egli M., Watterson D. M. A protein kinase associated with apoptosis and tumor suppression: Structure, activity, and discovery of peptide substrates. J. Biol. Chem. 2001, 276, 38956–38965.

11. Zhang L., Luo B., Lu Y., Chen Y. Targeting death-associated protein kinases for treatment of human diseases: Recent advances and future directions. J. Med. Chem. 2023, 66 (2), 1112–1136.

12. Lonsdale J., Thomas J., Salvatore M., Phillips R., Lo E., Shad S., Hasz R., etc. The Genotype-Tissue Expression (GTEx) project. Nat. Genet. 2013, 45 (6), 580–585.

13. Kuntz I. D., Blaney J. M., Oatley S. J., Langridge R., Ferrin T. E. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. 1982, 161 (2), 269–288.

14. Bhagat R. T., Butle S. R., Khobragade D. S., Wankhede S. B., Prasad C. C., Mahure D. S., Armarkar A. V. Molecular docking in drug discovery. J. Pharm. Res. Int. 2021, 33 (30B), 46–58.

15. Stanzione F., Giangreco I., Cole J. C. Use of molecular docking computational tools in drug discovery. Prog. Med. Chem. 2021, 60, 273–343.

16. Agu P. C., Afiukwa C. A., Orji O. U., Ezeh E. M., Ofoke I. H., Ogbu C. O., Ugwuja E. I., Aja P. M. Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. Sci. Rep. 2023, 13, 13398.

17. Shirali A., Stebliankin V., Karki U., Shi J., Chapagain P., Narasimhan G. A comprehensive survey of scoring functions for protein docking models. BMC Bioinformatics 2025, 26, 25.

18. Mursal M., Ahmad M., Hussain S., Khan M. F. Navigating the computational seas: a comprehensive overview of molecular docking software in drug discovery. Unravelling Molecular Docking – From Theory to Practice. IntechOpen, 2024.

19. Raval K., Ganatra T. Basics, types and applications of molecular docking: A review. IP Int. J. Compr. Adv. Pharmacol. 2022, 7 (1), 12–16.

20. Eberhardt J., Santos-Martins D., Tillack A. F., Forli S. AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. J. Chem. Inf. Model. 2021, 61 (8), 3891–3898.

21. Trott O., Olson A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J. Comput. Chem. 2010, 31 (2), 455–461.

22. Corso G., Stärk H., Jing B., Barzilay R., Jaakkola T. DiffDock: diffusion steps, twists, and turns for molecular docking. arXiv preprint 2022, arXiv:2210.01776.

23. Wohlwend J., Corso G., Passaro S., Reveiz M., Leidal K., Swiderski W., Portnoi T., Chinn I., Silterra J., Jaakkola T., Barzilay R. Boltz-1: democratizing biomolecular interaction modeling. bioRxiv 2024.

24. Friesner R. A., Banks J. L., Murphy R. B., Halgren T. A., Klicic J. J., Mainz D. T., Repasky M. P., Knoll E. H., etc. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J. Med. Chem. 2004, 47, 1739–1749.

25. Filipe H. A. L., Loura L. M. S. Molecular dynamics simulations: advances and applications. Molecules 2022, 27 (7), 2105.

26. Abraham M. J., Murtola T., Schulz R., Páll S., Smith J. C., Hess B., Lindahl E. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015, 1–2, 19–25.

27. Karplus M., McCammon J. A. Molecular dynamics simulations of biomolecules. Nat. Struct. Biol. 2002, 9 (9), 646–652.

28. Hollingsworth S. A., Dror R. O. Molecular dynamics simulation for all. Neuron 2018, 99 (6), 1129–1143.

29. van der Spoel D., Lindahl E., Hess B., Groenhof G., Mark A. E., Berendsen H. J. C. GROMACS: fast, flexible, and free. J. Comput. Chem. 2005, 26 (16), 1701–1718.

30. Pronk S., Páll S., Schulz R., Larsson P., Bjelkmar P., Apostolov R., Shirts M. R., Smith J. C., Kasson P. M., van der Spoel D., Hess B., Lindahl E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 2013, 29 (7), 845–854.

31. Abraham M., Alekseenko A., Andrews B., Basov V., Bauer P., Bird H., Briand E., Brown A., Doijade M., etc. GROMACS 2025.1 Manual. Zenodo 2025.

32. Ghahremanian S., Rashidi M. M., Raeisi K., Toghraie D. Molecular dynamics simulation approach for discovering potential inhibitors against SARS-CoV-2: a structural review. J. Mol. Liq. 2022, 354, 118901.

33. Wang E., Sun H., Wang J., Wang Z., Liu H., Zhang J. Z. H., Hou T. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. Chem. Rev. 2019, 119 (16), 9478–9508.

34. Akkus E., Tayfuroglu O., Yildiz M., Kocak A. Revisiting MMPBSA by adoption of MC-based surface area/volume, ANI-ML potentials, and two-valued interior dielectric constant. J. Phys. Chem. B 2023, 127 (20), 4415–4429.

35. Valdés-Tresanco M. S., Valdés-Tresanco M. E., Valiente P. A., Moreno E. gmx_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS. J. Chem. Theory Comput. 2021, 17 (10), 6281–6291.

36. Miller B. R., McGee T. D., Swails J. M., Homeyer N., Gohlke H., Roitberg A. E. MMPBSA.py: an efficient program for end-state free energy calculations. J. Chem. Theory Comput. 2012, 8 (9), 3314–3321.

37. Manning G., Whyte D. B., Martinez R., Hunter T., Sudarsanam S. The protein kinase complement of the human genome. Science 2002, 298, 1912–1934.

38. Waterhouse A. M., Procter J. B., Martin D. M. A., Clamp M., Barton G. J. Jalview Version 2: a multiple sequence alignment editor and analysis workbench. Bioinformatics 2009, 25, 1189–1191.

39. Gao L.-J., Kovackova S., Sála M., Ramadori A. T., De Jonghe S., Herdewijn P. Discovery of dual death-associated protein related apoptosis inducing protein kinase 1 and 2 inhibitors by a scaffold hopping approach. J. Med. Chem. 2014, 57 (18), 7624–7643.

40. Chaudhry S., Castro J. R., Totiger T. M., Afaghani J., Khurshid R., Nicholls M., Zhang Z., Schürer S. C., Shah A., Taylor J., Feng Y. Potent, selective, and orally bioavailable quinazoline-based STK17A/B dual inhibitors. ACS Med. Chem. Lett. 2024, 15, 945–949.

41. Egbert M., Jones G., Collins M. R., Kozakov D., Vajda S. FTMove: a web server for detection and analysis of cryptic and allosteric binding sites by mapping multiple protein structures. J. Mol. Biol. 2022, 434 (11), 167587.

42. DeLano, W. L. The PyMOL Molecular Graphics System, DeLano Scientific, 2002.

43. BIOVIA, Dassault Systèmes. Discovery Studio Visualizer, Release 2024. San Diego: Dassault Systèmes, 2024.

44. Waterhouse A., Bertoni M., Bienert S., Studer G., Tauriello G., Gumienny R., Heer F. T., de Beer T. A. P., Rempfer C., Bordoli L., Lepore R., Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018, 46 (W1), W296–W303.

45. Bienert S., Waterhouse A., de Beer T. A. P., Tauriello G., Studer G., Bordoli L., Schwede T. The SWISS-MODEL Repository – new features and functionality. Nucleic Acids Res. 2017, 45 (D1), D313–D319.

46. Guex N., Peitsch M. C., Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 2009, 30 (S1), S162–S173.

47. Robertson M. J., Tirado-Rives J., Jorgensen W. L. Improved peptide and protein torsional energetics with the OPLS-AA force field. J. Chem. Theory Comput. 2015, 11 (7), 3499–3509.

48. Lopes P. E. M., Guvench O., MacKerell A. D. Jr. Current status of protein force fields for molecular dynamics simulations. In: Molecular Modeling of Proteins; Kukol A., Ed.; Methods in Molecular Biology; Vol. 1215; Springer: New York, 2015; pp. 47–71.

49. Huang J., Rauscher S., Nawrocki G., Ran T., Feig M., de Groot B. L., Grubmüller H., MacKerell A. D. Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat. Methods 2017, 14 (1), 71–73.

50. Kadaoluwa Pathirannahalage S. P., Meftahi N., Elbourne A., Weiss A. C. G., McConville C. F., etc. Systematic comparison of the structural and dynamic properties of commonly used water models for molecular dynamics simulations. J. Chem. Inf. Model. 2021, 61 (9), 4521–4536.