BEZMIALEM VAKIF UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

THE QUEST FOR LIGANDS AGAINST KINESIN MOTOR PROTEINS USING CADD

PhD THESIS

Ahmet Fatih SAHIN

Department of Drug Discovery and Development

Drug Discovery and Development PhD Program

Thesis Advisor: Prof. Dr. Atilla AKDEMIR

JULY 2025

BEZMIALEM VAKIF UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

THE QUEST FOR LIGANDS AGAINST KINESIN MOTOR PROTEINS USING CADD

PhD THESIS

Ahmet Fatih SAHIN
205332002

Department of Drug Discovery and Development

Drug Discovery and Development PhD Program

Thesis Advisor: Prof. Dr. Atilla AKDEMIR

JULY 2025

After fulfilling all the necessary conditions determined by the related regulations, PhD student Ahmet Fatih SAHIN (student ID number: 205332002) successfully presented his thesis titled as " THE QUEST FOR LIGANDS AGAINST KINESIN MOTOR PROTEINS USING CADD" in front of the jury members.

Thesis Advisor : **Prof. Dr. Atilla AKDEMIR** .....
Istinye University

Jury Members : **Prof. Dr. Vecdi Melih ALTAN** .....
Bezmialem Vakif University

Prof. Dr. Simone CARRADORI .....
Università degli Studi "G. d'Annunzio"

Prof. Dr. Aydan DAG .....
Bezmialem Vakif University

Asst. Prof. Dr. Dilek OZTURK CIVELEK .....
Istanbul University

Asst. Prof. Dr. Muhammed TRAWALLY .....
Istanbul University

Date of Submission : 27 June 2025
Date of defense Exam : 25 July 2025

To my family,

FOREWORD

The completion of this doctoral dissertation signifies the completion of a journey that has been academically demanding and transformational. This technique has broadened my scientific comprehension and influenced my character and viewpoint as a researcher. The progress made here would not have been attainable without the constant support of many people and institutions to whom I extend my heartfelt appreciation.

I would like to express my deepest appreciation to my academic advisor Prof. Dr. Atilla Akdemir for his constant mentoring, insightful guidance, and persistent support. His insightful comments and consistent backing were important for determining the scientific trajectory of our efforts. I express my gratitude to Prof. Dr. Özlen Güzel Akdemir for her kind sharing of skills and experiences throughout our collaboration in many research projects.

I express my gratitude to Dr. Serap İpek Dingiş Birgül, Dr. Kaan Birgül, and Dr. Muhammad Trawally for their important support, friendship, and company during our collaborations at Bezmialem Vakıf University and beyond. Their encouragement significantly influenced my academic growth and personal drives.

I wish to express my gratitude for the constructive feedback and academic guidance received from my thesis monitoring committee members, Prof. Dr. Vecdi Melih Altan and Prof. Dr. Simone Carradori, whose critical insights have enhanced the scientific rigor of this dissertation.

I am very appreciative of the Scientific and Technological Research Council of Türkiye (TÜBİTAK) for funding my PhD studies under the 2211-C National PhD Scholarship in Priority Fields of Science and Technology. I express my gratitude to the 2214-A International Research Fellowship for PhD Students, which facilitated my study at the University of Tübingen. The grants not only enabled the actual execution of this project but also enhanced it via international cooperation and exposure to other scientific contexts.

I am honored to have worked at the University of Tübingen under the guidance of Prof. Dr. Frank M. Boeckler. His guidance was essential and I am particularly grateful for his tough critiques and scientific insights. I express my heartfelt gratitude to the members of the Boeckler Study Group for their kind assistance and favorable environment over the six-month study period. Finally, this research was made possible by the computational resources provided by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG, grant INST 40/575-1 FUGG) for the JUSTUS 2 cluster.

I would also like to express my gratitude to Prof. Dr. Wolfgang Sippl and his research team for hosting me throughout my erasmus training. Their mentorship and sincerity enhanced my academic development and offered a distinctive multidisciplinary viewpoint. I would also like to acknowledge the personal support I have received throughout this endeavor.

I am very appreciative of my parents, Ayşe Şahin and Ülkü Şahin, as well as my sisters, Şeyda Şahiner and Büşra Şahin, for their unfailing affection, patience, and support. I extend my sincere gratitude to my lovely girlfriend Zuhal Gerede for her constant presence and emotional support during the most difficult stages of my journey. I am profoundly grateful to all my friends and colleagues—too many to name individually—whose generosity and support have supported me throughout my journey.

I like to underscore that while I accept complete responsibility for any deficiencies or mistakes in this work, the contributions of those recognized herein—and some who are unmentioned—have been indispensable. Their thoughts, support, and participation have significantly enhanced the scientific and intellectual qualities of this dissertation.

July 2025 Ahmet Fatih Sahin

DECLARATION

I declare that this thesis study is mine. I do not have any unethical behavior at all stages of the thesis. I have obtained all the information within academic and ethical rules. I have referred to all information and comments that have not been obtained through this thesis study, and I have included them in the list of references, and I have not violated any patent and copyright during the study and writing of this thesis.

Ahmet Fatih Sahin

TABLE OF CONTENTS

Page

ABBREVIATIONS

ABC **:**ATP-Binding Cassette
ADMET **:**Absorption, Distribution, Metabolism, Elimination (Excretion), and Toxicity
AAG **:**Alpha 1-acid Glycoprotein
AI **:**Artificial Intelligence
AI4DD **:**Artificial Intelligence for Drug Discovery
AIC **:**Active Inhibitor Concentration
AML **:**Acute Myeloid Leukemia
AUC **:**Area Under the Curve
ATP **:**Adenosine Triphosphate
AURKA **:**Aurora Kinase A
AURKB **:**Aurora Kinase B
BASS **:**Bipolar Assembly Module
BBB :Blood-Brain Barrier
BSA **:**Bovine Serum Albumin
CADD **:**Computer-Aided Drug Discovery
CIN :Chromosomal Instability
CMC **:**Critical Micelle Concentration
CNS :Central Nervous System
CPU :Central Processing Unit
CRISPR :Clustered Regularly Interspaced Short Palindromic Repeats
DD **:**DeepDocking
DL **:**Deep Learning
DNA **:**Deoxyribonucleic Acid
DMSO **:**Dimethyl Sulfoxide
DNN **:**Deep Neural Network
DOX **:**Doxorubicin
DUD-E **:**Directory of Useful Decoys: Enhanced
ETKDG **:**Experimental-Torsion Knowledge Distance-Geometry (conformer embedding)
FBDD **:**Fragment-Based Drug Discovery
FDA **:**U.S. Food and Drug Administration
FPDE **:**Full Predicted Database Enrichment
FPR **:**False Positive Rate
GI50 **:**50% Growth Inhibition
GNN **:**Graph Neural Network
GPU **:**Graphics Processing Unit
GUI **:**Graphical User Interface
HTS **:**High-Throughput Screening
HTVS **:**High-Throughput Virtual Screening (Glide mode)
IC50 **:**Half-Maximal Inhibitory Concentration
KSP **:**Kinesin Spindle Protein

LBDD :Ligand-Based Drug Design
LSTM :Long Short-Term Memory (recurrent neural network)
MD :Molecular Dynamics
ML :Machine Learning
MM/GBSA:Molecular Mechanics / Generalized Born Surface Area
MM/PBSA:Molecular Mechanics / Poisson-Boltzmann Surface Area
MMFF94 :Merck Molecular Force Field 1994
MPO :Multi-Parameter Optimization
MT :Microtubule
NCI :National Cancer Institute
OPLS4 :Optimized Potentials for Liquid Simulations v4
pROC :Predictive Receiver-Operating Characteristic
PD :Pharmacodynamics
PDI :Polydispersity
PEG :Poly(ethylene glycol)
PK :Pharmacokinetics
PLK1 :Polo-Like Kinase 1
QSAR :Quantitative Structure–Activity Relationship
RAFT :Reversible Addition–Fragmentation Chain-Transfer
RDKit :RDKit cheminformatics toolkit
RNA :Ribonucleic Acid
RMSD :Root-Mean-Square Deviation
SAR :Structure–Activity Relationship
SBDD :Structure-Based Drug Design
SDF :Structure Data File
SMILES :Simplified Molecular-Input Line-Entry **System**
SP :Standard-Precision (Glide mode)
SSH :Secure Shell
St :Styrene
STLC :***S*-**Trityl-*L*-cysteine
TI :Thermodynamic Integration
TPR :True Positive Rate
VSGB :Variable Solvent Generalized Born (implicit-solvent model)
WHO :World Health Organization
XP :Extra-Precision (Glide mode)

SYMBOLS

Å	:Angstrom (10^{-10} m)
nm	:Nanometer (10^{-9} m)
µm	: Micrometer (10^{-6} m)
°C	:Degree Celsius
bar	:Bar (pressure unit ≈ 10^5 Pa)
h	:Hour
K	:Kelvin (absolute temperature)
kcal mol^{-1}	:Kilocalories per mole (energy)
mM	:Millimolar (10^{-3} mol L^{-1})
µM	:Micromolar (10^{-6} mol L^{-1})
nM	:Nanomolar (10^{-9} mol L^{-1})
nm	:Nanometre (10^{-9} m)
ns	:Nanosecond (10^{-9} s)
ps	:Picosecond (10^{-12} s)
mg m^{-2}	: Milligrams per square meter (dose)
ρ	:Density
t	:Time (generic variable)

LIST OF TABLES

Page

LIST OF FIGURES

Page

THE QUEST FOR LIGANDS AGAINST KINESIN MOTOR PROTEINS USING CADD

SUMMARY

The kinesin motor protein Eg5 is a well-validated anticancer target due to its essential role in mitotic spindle formation. However, despite numerous clinical trials, no Eg5 inhibitor has yet achieved regulatory approval, largely due to challenges with modest efficacy, dose-limiting toxicities, and drug resistance. This highlights a critical need for novel chemical scaffolds with superior therapeutic profiles.

This dissertation addresses this challenge by employing a comprehensive, dual-track computational strategy to discover next-generation Eg5 inhibitors. The first track utilized a validated, structure-based virtual screening cascade on the ~11 million compound ZINC Anodyne library. The second, more exploratory track deployed a state-of-the-art DeepDocking active learning campaign to navigate the vast, 1-billion-compound ZINC-ML chemical space.

The conservative Track 1 successfully identified four novel, active chemotypes, with the lead candidate ZINC243914928 demonstrating an in vitro IC_{50} of 28.97 μM. This experimental validation of a computationally derived hit provided a crucial proof-of-concept for the entire workflow. More significantly, the AI-driven Track 2 discovered multiple, structurally diverse classes of inhibitors with superior in silico profiles. The top candidates from this track exhibited calculated binding free energies that were profoundly more favorable than both the reference inhibitor and the experimentally validated Track 1 hit. Crucially, the final hit compounds from both tracks exhibited favorable drug-like properties, conforming to Lipinski's rules, and demonstrated significant structural novelty with low Tanimoto similarity to known Eg5 chemotypes, confirming the success of the workflow in achieving genuine scaffold hopping.

In conclusion, this thesis not only delivers multiple novel and validated chemical scaffolds for Eg5 inhibition but also demonstrates a powerful, multi-faceted computational strategy for tackling challenging drug targets. The exceptional computational profiles of the Track 2 candidates, supported by the validated success of the overall methodology, strongly suggest their potential for high potency and mark them as highly promising starting points for the development of a clinically successful Eg5-targeted therapeutic.

Keywords: Cancer therapy, Kinesin Eg5, Mitotic arrest, Allosteric Inhibition, Eg5 inhibitors, Computer-Aided Drug Design, Virtual Screening, Active Learning.

CADD KULLANARAK KİNESİN MOTOR PROTEİNLERİNE KARŞI LİGAND ARAYIŞI

ÖZET

Kinesin motor proteini Eg5, mitotik iğ iplikçiği oluşumundaki temel rolü nedeniyle, kanser tedavisinde etkinliği kanıtlanmış bir hedeftir. Ancak, çok sayıda klinik denemeye rağmen, bugüne dek hiçbir Eg5 inhibitörü, sınırlı etkinlik, doz kısıtlayıcı toksisiteler ve ilaç direnci gibi zorluklar nedeniyle ruhsat onayı alamamıştır. Bu durum, üstün terapötik profillere sahip yeni kimyasal iskelet yapılarına yönelik kritik bir ihtiyacı ortaya koymaktadır.

Bu tez, yeni nesil Eg5 inhibitörleri keşfetmek amacıyla kapsamlı ve çift yönlü bir hesaplamalı strateji kullanarak bu ihtiyaca yanıt vermektedir. Birinci yol, ~11 milyon bileşik içeren ZINC Anodyne kütüphanesine yönelik, valide edilmiş, yapı-temelli bir sanal tarama iş akışını içermektedir. İkinci ve daha keşifsel olan yol ise, 1 milyar bileşiklik devasa ZINC-ML kimyasal uzayını taramak için en son teknoloji bir DeepDocking aktif öğrenme kampanyasını kullanmıştır.

Daha konservatif olan birinci yol, dört yeni ve aktif kemotipi başarıyla tanımlamış, lider aday ZINC243914928 in vitro olarak 28.97 μM'lik bir IC_{50} değeri sergilemiştir. Hesaplamalı olarak bulunan bir adayın deneysel olarak doğrulanması, tüm iş akışının geçerliliği için kritik bir kanıt sağlamıştır. Daha da önemlisi, yapay zeka destekli ikinci yol, çok daha üstün in silico profillere sahip, yapısal olarak çeşitli yeni inhibitör sınıfları keşfetmiştir. Bu yoldan çıkan en iyi adaylar, hem referans inhibitörden hem de deneysel olarak doğrulanmış ilk yol adayından çok daha elverişli olan, hesaplanmış bağlanma serbest enerjileri sergilemiştir. Daha da önemlisi, her iki yoldan elde edilen nihai aday bileşikler, Lipinski'nin kurallarına uygun, olumlu ilaç-benzeri özellikler sergilemiş ve bilinen Eg5 kemotiplerine düşük Tanimoto benzerliği ile önemli yapısal yenilik göstermiştir; bu da iş akışının yeni yapilar tespit elde etmedeki başarısını doğrulamıştır.

Sonuç olarak bu tez, Eg5 inhibisyonu için sadece çok sayıda yeni ve valide edilmiş kimyasal iskelet yapıları sunmakla kalmamış, aynı zamanda zorlu ilaç hedeflerinin üstesinden gelmek için güçlü ve çok yönlü bir hesaplamalı strateji ortaya koymuştur. Metodolojinin genel başarısıyla desteklenen ikinci yol adaylarının olağanüstü hesaplamalı profilleri, bu bileşiklerin yüksek bir etki potansiyeline sahip olduğunu kuvvetle düşündürmekte ve onları klinik olarak başarılı bir Eg5 hedefli terapötiğin geliştirilmesi için son derece umut verici başlangıç noktaları olarak işaretlemektedir.

Anahtar Kelimeler: Kanser tedavisi, Kinesin Eg5, Mitotik duraklama, Allosterik İnhibisyon, Eg5 inhibitörleri, Bilgisayar Destekli İlaç Tasarımı, Sanal Tarama, Aktif Öğrenme.

1. INTRODUCTION

Cancer is globally recognized as the second leading cause of mortality, responsible for nearly ten million deaths annually, and imposing a significant economic burden on healthcare systems [1]. Conventional anticancer agents that target cell division, primarily taxanes and vinca alkaloids, exert their effects by disrupting microtubule dynamics [2]. However, the efficacy of these drugs is constrained by significant challenges, including neurotoxicity, bone marrow suppression, and eventual development of drug resistance [3]. This situation underscores the urgent need for novel, targeted anticancer agents that can selectively affect dividing cells while sparing non-dividing normal tissues.

In this context, the mitotic motor protein Eg5, also known as KIF11 or kinesin spindle protein (KSP), is a promising therapeutic target [4]. Eg5 is a homotetramer that facilitates the formation of a bipolar spindle during the early stages of mitosis by separating the anti-parallel microtubules. When the activity of Eg5 is inhibited, proliferating cells are arrested with a monopolar spindle structure and subsequently undergo apoptosis [5]. In contrast, non-dividing cells such as neurons remain largely unaffected. Furthermore, it has been reported that Eg5 expression is elevated in many solid tumors and hematological malignancies compared to healthy tissues, and this increase is often associated with a poorer clinical prognosis [6]. These characteristics establish Eg5 as a physiologically significant and pharmacologically valid therapeutic target.

Structural analyses revealed multiple small compound-binding sites in the Eg5 motor domain. The most extensively studied and validated of these is Allosteric Site-I (AS-I), a druggable pocket formed at the interface between loop L5 and helices α2/α3 [7]. The first-in-class inhibitor, monastrol, and subsequent, more potent clinical

candidates, such as ispinesib and filanesib, all target this site, functioning by locking the motor in an inactive, ADP-bound conformation [8]. However, despite advances in late-stage clinical trials, these pioneering inhibitors have been consistently limited by a narrow therapeutic window, exhibiting rapid pharmacokinetic clearance, dose-limiting bone marrow toxicity, and modest efficacy in patients. Ultimately, none of them received regulatory approval [9]. This history of clinical disappointment does not invalidate the L5 pocket as a target; rather, it highlights a critical deficiency in the chemical nature of the inhibitors developed to date. The challenge lies in the discovery of novel chemical scaffolds that can occupy this validated pocket with superior affinity and improved pharmacological properties. Although other allosteric sites have been identified, such as the cavity between helices α4 and α6, the L5 pocket remains the most promising site for achieving highly selective inhibition, indicating that its potential for binding diverse and more effective chemotypes is far from exhausted [10].

This dissertation addresses the critical need for next-generation inhibitors by employing a comprehensive dual-track computational strategy to discover novel chemical scaffolds. The first track utilizes a validated, structure-based virtual screening cascade, integrating GPU-accelerated shape screening, pharmacophore filtering, and hierarchical docking on the ~11 million compound ZINC Anodyne library. In parallel, a more exploratory second track deploys a state-of-the-art DeepDocking active learning campaign to navigate the vast 1-billion-compound ZINC-ML chemical space [11–13]. For both tracks, the most promising candidates were subjected to rigorous post-screening validation, including extensive molecular dynamics simulations and binding free energy calculations, to ensure the identification of hits with high predicted affinities and stable binding modes.

The primary objective of this study was to leverage this multifaceted computational strategy to identify and validate novel, drug-like Eg5 inhibitors that can overcome the well-documented clinical limitations of previous compounds. By discovering structurally diverse chemotypes with superior in silico profiles and stable binding dynamics, this study aimed to provide a set of high-potential lead compounds, thereby offering superior starting points for the development of a clinically successful Eg5-targeted therapeutic

2. GENERAL INFORMATION AND LITERATURE REVIEW

2.1 Targeting Mitosis in Cancer Therapy

2.1.1 Global cancer burden and epidemiological trends

Cancer is a multifaceted disease characterized by unregulated cell proliferation and metastasis. This results from genetic mutations and epigenetic modifications [14]. Cancer is a major public health problem worldwide, ranking second in terms of mortality rates. As of 2022, approximately 20 million new cancer cases and 9.7 million cancer-related deaths have been reported (Figure 2.1) [15].

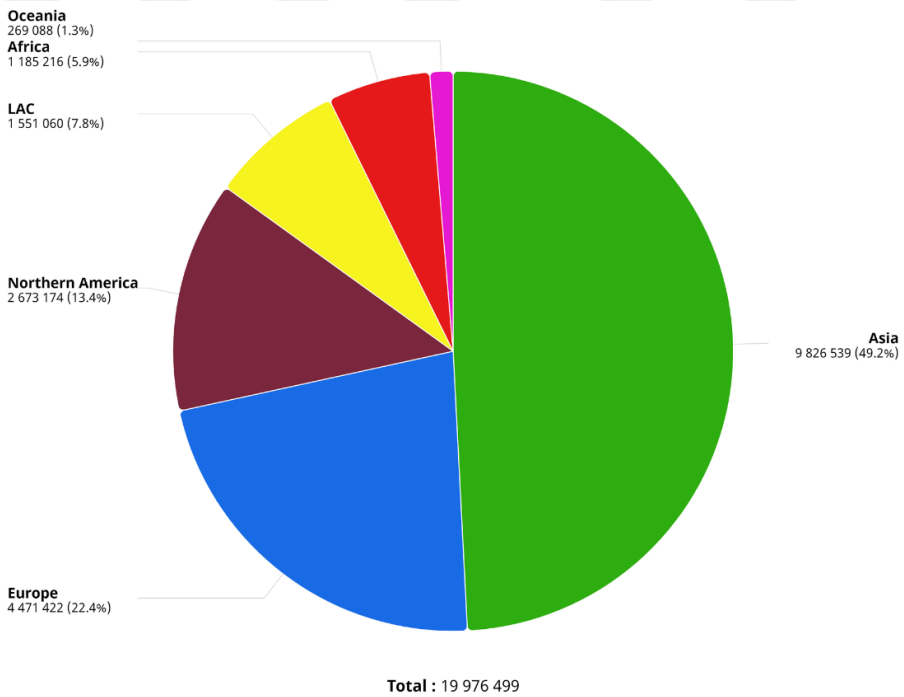

Figure 2.1 : Continental distribution of cancer incidence in 2022 (both sexes, absolute values) [15].

According to current estimates, approximately 20% of the world's population will develop cancer during their lifetime; 11% of men and 8% of women will die of cancer [15]. By 2050, the number of global cancer cases and deaths is expected to rise to 35 million and 18 million, respectively, annually. The main reasons for this dramatic increase are population growth, aging, and lifestyle risk factors (e.g., tobacco, alcohol, obesity, and infections) [16, 17].

2.1.2 Molecular roots of uncontrolled division

The molecular basis of cancer involves the dysregulation of cellular growth control pathways, often stemming from the activation of proto-oncogenes to oncogenes or inactivation of tumor suppressor genes. These alterations disrupt the crucial regulatory mechanisms governing cell proliferation, apoptosis, and DNA repair [18].

Among the hallmarks of cancer, uncontrolled mitosis plays a central role in promoting tumor progression. Mitosis, a tightly regulated phase of the cell cycle, ensures accurate segregation of replicated chromosomes into two daughter cells [19]. The mitotic spindle, a dynamic bipolar structure composed primarily of microtubules, is fundamental to this process. Microtubules extend from two opposing poles of the cell, defined by the microtubule organizing centers (centrosomes), and their plus ends can attach to kinetochores located at the centromeric regions of chromosomes. This precise arrangement ensures that chromosomes align correctly at the metaphase plate before sister chromatids segregate into opposite poles during anaphase (Figure 2.2) [20].

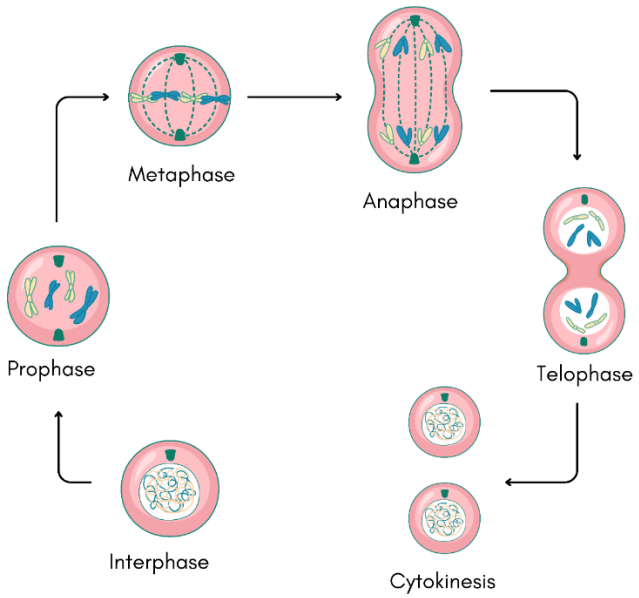

Figure 2.2 : Stages of the mitosis [21].

The proper assembly and function of the mitotic spindle, driven by the dynamic instability of microtubules, centrosome integrity, and accuracy of kinetochore-microtubule attachments, is essential for faithful chromosome segregation [22]. In cancer cells, which are characterized by rapid and often uncontrolled division, the spindle apparatus plays a critical role in sustaining mitotic progression (Figure 2.3). Consequently, any disruption in spindle formation or function can lead to chromosome missegregation, mitotic arrest, and ultimately trigger apoptotic cell death, which is a vulnerability frequently exploited in anticancer therapeutic strategies [23].

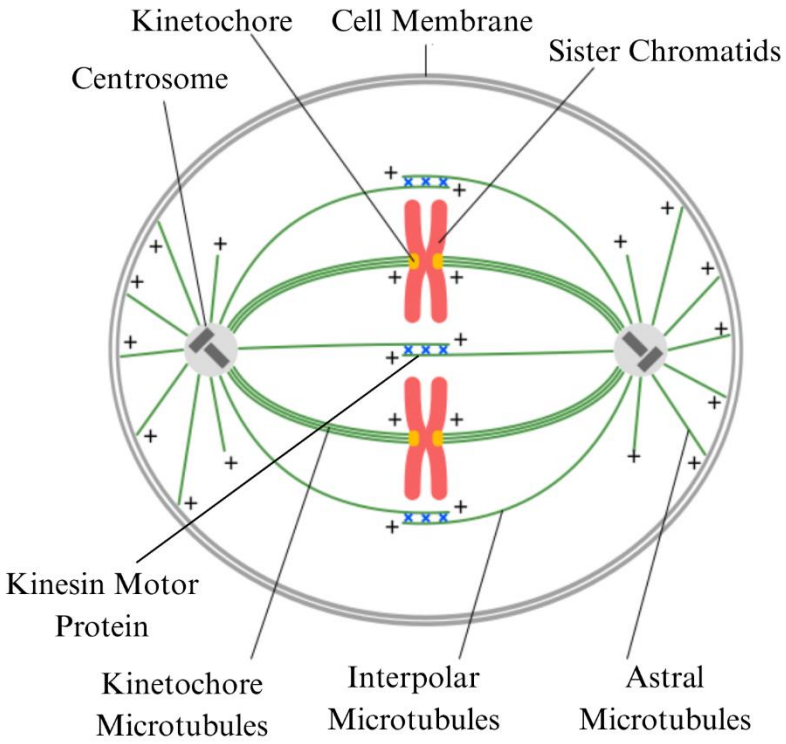

Figure 2.3 : Structure and components of the mitotic spindle apparatus during metaphase [24].

Targeting mitosis offers a therapeutically relevant approach to cancer treatment, because it is one of the most precisely orchestrated and vulnerable events in the cell cycle. The rapid proliferation of cancer cells makes them particularly susceptible to mitotic checkpoint activation and stresses induced by mitotic inhibitors. This characteristic allows for a degree of selectivity in targeting tumor cells over many normal, non-dividing, or slowly dividing host cells [25].

2.1.3 Conventional microtubule targeting agents: success and limitations

Conventional anticancer drugs that target the mitotic machinery, often referred to as microtubule-targeting agents (MTAs), interfere with microtubule dynamics. These agents typically arrest cells in mitosis by either stabilizing microtubules (e.g., taxanes such as paclitaxel; Figure 2.4) or by promoting their depolymerization (e.g., vinca alkaloids such as vinblastine and vincristine; Figures 2.5 and 2.6), thereby activating the spindle assembly checkpoint and inducing apoptosis [26]. Although MTAs have been mainstays in cancer chemotherapy for decades, their efficacy is often limited by two major factors: a lack of specificity, leading to significant off-target effects, and the frequent development of clinical drug resistance (e.g., through overexpression of drug efflux pumps such as P-glycoprotein/ABCB1 or mutations in tubulin itself) [27, 28]. Since microtubules are crucial for various cellular functions in non-dividing cells, particularly for axonal transport in neurons and for the division of hematopoietic progenitor cells, dose-limiting toxicities such as peripheral neuropathy and myelosuppression are common [29].

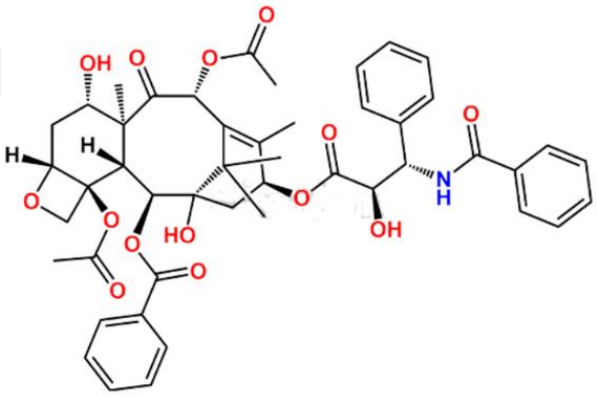

Figure 2.4 : Paclitaxel.

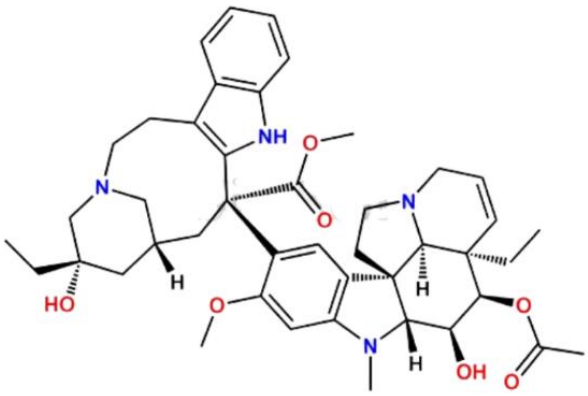

Figure 2.5 : Vinblastine.

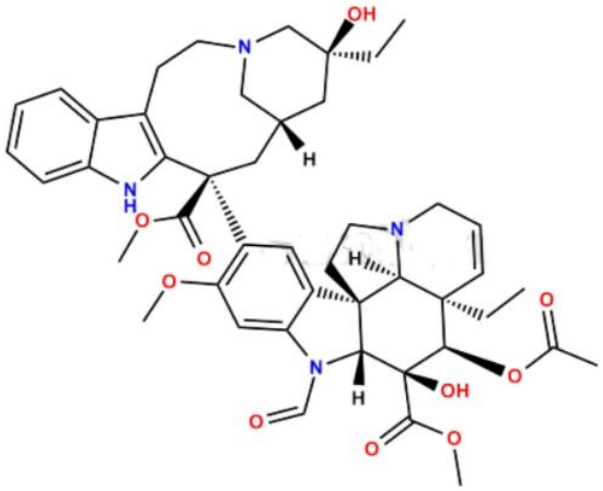

Figure 2.6 : Vincristine.

The limitations of conventional MTAs have driven research on more selective antimitotic targets. Recent studies have demonstrated that mitotic arrest, aneuploidy, and apoptosis can also be induced by inhibiting proteins specifically involved in mitosis, such as mitotic motor proteins (e.g., Eg5/KIF11, KIF18A), aurora kinases (e.g., Aurora A, Aurora B), and polo-like kinase 1 (PLK1) [30]. These spindle-specific targets offer the potential for improved therapeutic indices compared to broadly acting MTAs, primarily by selectively affecting dividing cells, thus aiming to spare quiescent normal cells to a greater extent [31].

Table 2.1 : Conventional vs. spindle-specific antimitotics: mechanism, dose-limiting toxicity, and resistance.

Drug class	Representative agent(s)	Primary mechanism of action	Major dose-limiting toxicity	Dominant resistance mechanism(s)	Key references
Conventional microtubule poisons—stabilizers	Paclitaxel (taxanes)	Binds β-tubulin and stabilizes microtubules → blocks depolymerization, arrests cells in metaphase	Peripheral neuropathy; myelosuppression	β-tubulin point mutations/isoform switching; over-expression of the efflux pump P-gp/ABCB1	[26]
Conventional microtubule poisons—depolymerizers	Vinblastine, Vincristine (vinca alkaloids)	Bind β-tubulin at the Vinca site and destabilize microtubules → prevent spindle formation	Neurotoxicity; haematological toxicity	P-gp–mediated drug efflux; β-tubulin mutations	[32]
Spindle-specific: Kinesin-5 (Eg5) inhibitors	Ispinesib, Filanesib	Allosterically block the loop L5 pocket of Eg5/KIF11 → centrosome separation fails → monopolar spindle	Neutropenia / leukopenia	Compensatory up-regulation of KIF15; point mutations in Eg5 loop L5	[33]
Spindle-specific:Aurora-A kinase inhibitors	Alisertib	Inhibits Aurora A catalytic activity → defective spindle checkpoint, abnormal chromosome alignment	Grade 3–4 neutropenia; mucositis	AURKA amplification or gatekeeper mutations; PLK1 up-regulation (reported in pre-clinical models)	[34]
Spindle-specific: PLK1 inhibitors	Volasertib	Blocks polo-like kinase-1 phosphorylation events required for G2/M transition	Haematological toxicity (thrombocytopenia, neutropenia)	PLK1 over-expression/mutation; activation of compensatory mitotic kinases	[35]

Among these spindle-specific targets, kinesin motor proteins have garnered considerable interest because of their direct role in the mechanochemistry of spindle assembly and function. Kinesins are ATP-dependent molecular motors that translocate along microtubules, facilitating intracellular transport and contributing to critical mitotic events such as spindle pole separation [36]. Notably, overexpression of the kinesin Eg5 (also known as KIF11 or KSP) has been reported in a wide array of human malignancies, including but not limited to glioma and colon, lung, and liver cancers, and its elevated expression levels often correlate with poor patient prognosis and

aggressive tumor behavior [37]. Given its indispensable role in establishing bipolar spindle assembly, Eg5 has emerged as an attractive target for cancer therapies. The therapeutic rationale and progress in targeting Eg5 will be elaborated in subsequent sections [38].

2.1.4 Kinesin family overview and mitotic kinesins

Motor proteins are molecular machines that convert chemical energy, typically derived from ATP hydrolysis, into mechanical energy to drive various forms of cellular movement and intracellular transport. These proteins operate along cytoskeletal filaments, primarily microtubules and actin. Myosins function on actin filaments, and kinesins and dyneins are the principal motor proteins that translocate along the microtubule tracks. Collectively, these motor proteins are fundamental to a multitude of cellular processes including chromosome segregation during mitosis and meiosis, transport of organelles and vesicles, cell migration, and dynamic remodeling of the cytoskeleton [39].

The kinesin superfamily comprises a large and diverse group of motor proteins crucial in eukaryotic cells for the directed transport of various cargo, including organelles, vesicles, and macromolecular complexes, along microtubule tracks. The first kinesin was identified in 1985 through studies of axonal transport in giant squid neurons, named after the Greek word *kinein* ("to move"). Kinesins are ubiquitous in eukaryotes and perform conserved functions across diverse cell types and species [40, 41].

Structurally, conventional kinesins (e.g., Kinesin-1) typically exist as dimers of two identical heavy chains and are often associated with light chains that primarily mediate cargo interactions. The N-termini of the heavy chains form globular motor (head) domains, which harbor ATP and microtubule-binding sites responsible for force generation and processive movement. Adjacent to the motor domain, the neck linker is a flexible region that is crucial for coordinating the hand-over-hand stepping mechanism and determining directionality along the microtubule [42]. Heavy chains dimerize via an extended coiled-coil stalk region, and their C-terminal tail domains, often in conjunction with the associated light chains, are responsible for specific cargo recognition and binding (Figure 2.7) [43].

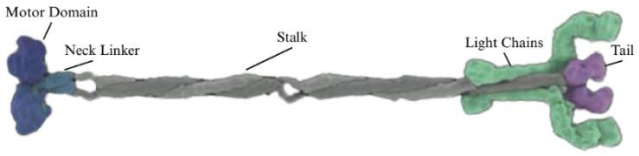

Figure 2.7 : Structural organization of conventional kinesin.

Kinesins are broadly classified based on the position of their motor domain within the polypeptide chain into N-type kinesins (motor domain at the N-terminus), M-type kinesins (central motor domain), and C-type kinesins (C-terminal motor domain). The majority of kinesins, including the Kinesin-5 family to which Eg5 belongs, are N-type and typically translocate towards the plus-end of microtubules (Figure 2.8) [44].

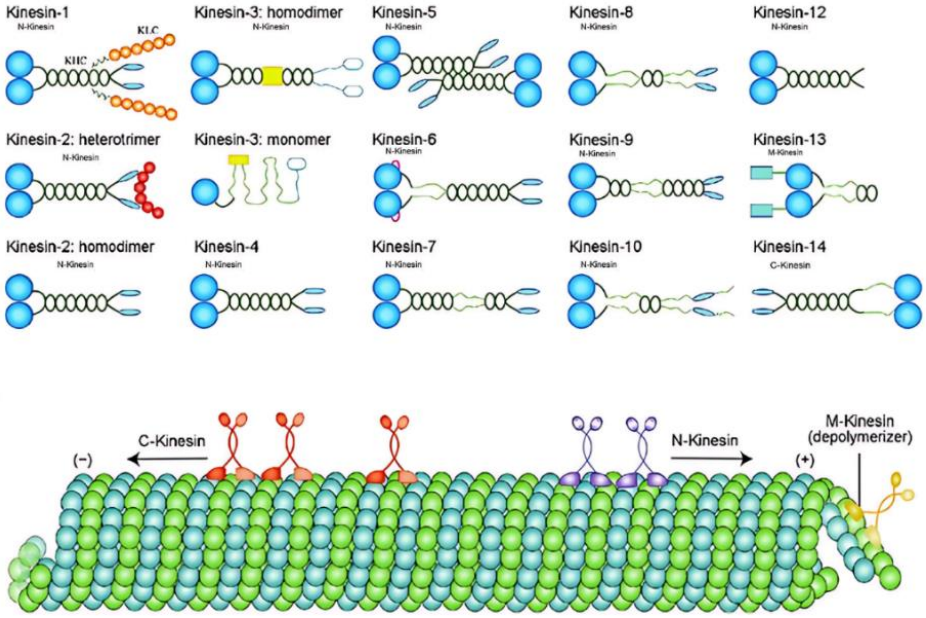

Figure 2.8 : Classification of kinesin superfamily members based on motor domain location and directionality.

Based on the sequence homology within their conserved ~340 amino acid motor domains, kinesins are further categorized into at least 14 distinct subfamilies (Kinesin-1 to Kinesin-14). In humans, approximately 45 different genes encode kinesin superfamily members (KIFs) (Table 2.2). This structural classification largely reflects the functional diversification. While most kinesins are plus-end directed motors, some subfamilies, notably Kinesin-14, exhibit minus-end directed motility, enabling a wide array of specialized cellular functions [40].

Table 2.2 : Overview of human kinesin subfamilies: structure, function, and disease associations.

Subfamily	Motor Domain Position	Directionality	Representative Genes	Structure	Primary Function	Major disease / clinical link
Kinesin-1	N-terminal (N-type)	Plus-end	KIF5A, KIF5B, KIF5C	Homodimer	Organelle and vesicle transport	Hereditary spastic paraplegia 10; ALS [45, 46]
Kinesin-2	N-terminal	Plus-end	KIF3A, KIF3B, KIF3C	Heterotrimer	Intraflagellar transport (IFT)	Primary ciliary dyskinesia, situs inversus [45]
Kinesin-3	N-terminal	Plus-end	KIF1A, KIF1B, KIF13A	Monomer/Dimer	Long-distance vesicle transport	KIF1A-associated neurological disorder (KAND) [47]
Kinesin-4	N-terminal	Plus-end	KIF4A, KIF4B	Dimer	Chromosome positioning and condensation	X-linked intellectual disability, epilepsy [48]
Kinesin-5	N-terminal	Plus-end	KIF11 (Eg5)	Homotetramer	Bipolar spindle formation	Over-expressed in various cancers and solid tumors [49]
Kinesin-6	N-terminal	Plus-end	KIF20A, KIF20B	Dimer	Cytokinesis and spindle organization	Autosomal-recessive microcephaly (KIF20B/23 variants) [50]
Kinesin-7	N-terminal	Plus-end	CENP-E (KIF10)	Monomer	Kinetochore-microtubule attachment	Primary microcephaly 13 (MCPH-13) [51]
Kinesin-8	N-terminal	Plus-end	KIF18A, KIF18B	Dimer	Microtubule length regulation and chromosome alignment	Dependency of CIN-high cancers; KIF18A inhibitors in oncology [52]
Kinesin-9	N-terminal	Plus-end	KIF9	Dimer	Ciliary function and motility	Male infertility (asthenozoospermia) [53]
Kinesin-10	N-terminal	Plus-end	KIF22 (Kid)	Monomer	Chromosome congression	Spondylo-epimetaphyseal dysplasia (SEMD-JL2) [54]
Kinesin-11	N-terminal	Plus-end	KIF26A, KIF26B	Monomer	Cell signaling and adhesion	Fetal akinesia & spina bifida spectrum [55]
Kinesin-12	N-terminal	Plus-end	KIF15	Dimer	Spindle assembly and maintenance	Over-expressed in breast, pancreas, other cancers [56] (prognostic marker)
Kinesin-13	Central (M-type)	Non-motile	KIF2A, KIF2B, KIF2C	Monomer	Microtubule depolymerization	Complex cortical malformations / lissencephaly [57]
Kinesin-14	C-terminal (C-type)	Minus-end	KIFC1, KIFC2, KIFC3	Dimer	Spindle pole focusing and retrograde transport	Globozoospermia & other male-infertility phenotypes [58]

Functionally, kinesins exhibit remarkable diversity and span a wide array of cellular roles. For instance, Kinesin-2 is indispensable for intraflagellar transport (IFT), which is crucial for the assembly and maintenance of cilia and flagella, whereas members of the Kinesin-13 subfamily act as potent microtubule depolymerases, profoundly influencing cytoskeletal dynamics throughout the cell cycle, including mitosis [59,

60]. In particular, the process of mitosis relies on a specialized cohort of these motor proteins to ensure precise choreography of chromosome segregation and cell division. Among these are Kinesin-4 (such as KIF4A), which aids in chromosome congression to the metaphase plate and contributes to spindle organization [61]; Kinesin-6 (like MKLP1/KIF23), which is essential for forming the central spindle during anaphase and driving cytokinesis [62]; Kinesin-7 (for example, CENP-E), which plays a vital role in attaching chromosomes to spindle microtubules and facilitating their movement, as well as participating in the spindle assembly checkpoint [63]; Kinesin-8 (such as KIF18A), which meticulously regulates microtubule length at kinetochores for accurate chromosome alignment [64]; and Kinesin-14 (including HSET/KIFC1), typically minus-end directed motors that assist in focusing spindle poles and maintaining overall spindle integrity [65].

Among these crucial mitotic kinesins, the Kinesin-5 family, particularly Eg5, the primary member of this family in humans, stands out for its specific characteristics and is of particular interest in cancer therapy [66]. The homotetrameric structure of kinesin Eg5 and its capacity to cross-link antiparallel microtubules and function as a plus-end motor driving them in opposing directions define it most fundamentally from others [67]. This repulsive force is essential for the separation of spindle poles (centrosomes) and, thus, the formation of the bipolar structure of the mitotic spindle, a process schematically illustrated where active Eg5 maintains a normal bipolar spindle by crosslinking antiparallel microtubules (Figure 2.9a) [68].

Although many other mitotic kinesins may result in different spindle abnormalities, the suppression of Eg5 activity is unique in that it typically results in mitotic arrest by producing monopolar spindles, thereby highlighting the special and necessary role of Eg5 in preserving bipolarity [4, 5]. This outcome is clearly visualized in both schematic representations, which show the prevention of centrosome separation upon inhibition (Figure 2.9a), and in fluorescence microscopy images. These images starkly contrast a normal mitotic cell with a bipolar spindle with a cell treated with an Eg5 inhibitor, which exhibits characteristic monopolar spindle formation and mitotic arrest (Figure 2.9b).

Moreover, Eg5 activity is mostly limited to mitosis, which makes it a desirable therapeutic target in oncology, as its suppression primarily affects rapidly developing cancer cells and potentially causes less harm to non-dividing normal cells [69]. Eg5

also has a well-defined allosteric inhibitor-binding site (L5 allosteric binding site), which highlights its potential pharmaceutical importance. This important and particular function of kinesin Eg5 in spindle bipolarity makes it a focus point for the development of novel anticancer therapies, which will be discussed in detail in the following sections [70].

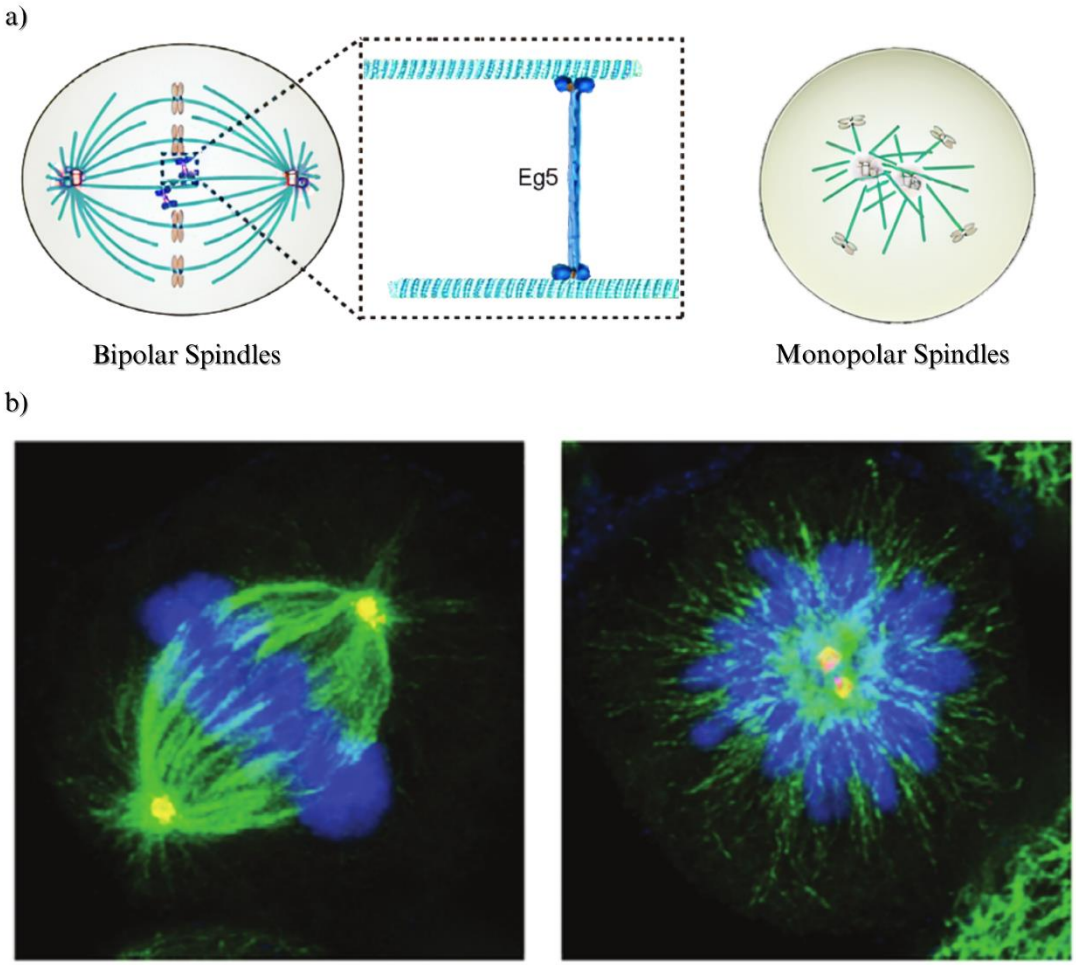

Figure 2.9 : Role of Eg5 in mitotic spindle assembly and phenotypic outcome upon inhibition [71].

2.2 Kinesin Eg5 (KIF11): Discovery, Evolutionary Conservation, and Therapeutic Significance

Given the significant limitations and side effects associated with current antimitotic therapies, there is an urgent need for more specific and tolerable novel molecular targets for cancer treatment [23]. In this context, as detailed in the preceding section, human Kinesin Eg5, an essential constituent of the Kinesin-5 (BimC) family, comes to the forefront because of its pivotal function in mitotic spindle dynamics. This motor

protein was discovered during cDNA library screening in Xenopus laevis, which sought to identify transcripts that swiftly break their poly(A) tails after fertilization [72, 73]. The motor domain of Eg5 exhibits significant sequence similarity (approximately 30%) to BimC from Aspergillus nidulans, a kinesin-related protein that is essential for nuclear division [74]. Furthermore, the motor domain of BimC shares up to 42% similarity with the kinesin heavy chain (DmKHC) of Drosophila melanogaster [75].

These findings indicate a high degree of conservation across kinesins in the motor domain despite the possibility of species-specific changes in the stalk and tail areas [76]. Disruption of BimC has been shown to impair spindle pole body (SPB) segregation without compromising the operation of other organelles [73, 77].

This discovery prompted additional investigations into related Kinesin-5 proteins, resulting in the identification of the Cut7 protein from Schizosaccharomyces pombe, which is crucial for the separation of spindle pole bodies (SPB) [78]. Eg5 was later recognized as a member of the Kinesin-5 family, exhibiting high expression levels during the early divisions of embryonic development in Xenopus, a subject of considerable research [79].

In humans, Eg5 primarily facilitates the formation of bipolar spindle assembly and advancement of mitosis. Blangy and colleague's research suggested that phosphorylation regulates Eg5's binding to spindle structures [80], but comparative analysis by Sawin and Mitchison demonstrated its motor activity and evolutionary conservation across species [69]. The preservation of Kinesin-5 family proteins across several eukaryotes underscores their pivotal functions in mitotic spindle assembly [81]. This evolutionary continuity highlights the significance of their biological functions and positions Eg5 as a widely relevant and conserved target protein for cancer treatment. The inhibition of Eg5, leading to monopolar spindle formation, subsequent mitotic arrest, and ultimately apoptosis, particularly in rapidly dividing cancer cells, is the cornerstone of its therapeutic potential [82].

The evolutionary conservation of Eg5 and pioneering research demonstrates its critical role in mitotic spindle dynamics clearly establish the biological significance of this motor protein [69]. The distinctive homotetrameric configuration of Eg5 and its capacity to generate the necessary force for spindle pole separation via the sliding of

antiparallel microtubules are vital for the mitotic process [83]. The preferential high expression of Eg5 in cancer cells compared with its low expression in normal, non-dividing cells (e.g., neurons) offers a significant therapeutic window for targeted treatments [84].

These specified traits provide unique prospects for targeted treatment approaches, particularly in conditions characterized by increased cell proliferation rates, such as cancer [33]. Consequently, a comprehensive understanding of these attributes and the mechanism of action of Eg5 are essential for the development of novel targeted anticancer therapeutics.

This section provides a thorough analysis of Eg5's structural biology, investigating its mechanism of action on microtubules, various inhibition strategies, the binding sites of inhibitors as revealed by crystal structures, and a detailed review of the currently available small-compound inhibitors.

2.2.1 Structural biology of Eg5

Kinesin Eg5 is an essential motor protein involved in the assembly of bipolar mitotic spindles [5]. The human Eg5 protein consists of ~1057 amino acids and three main structural domains: the motor domain (N-terminal ~1–367 aa), stalk domain (~368-761 aa) (rod-shaped middle part), and tail domain (762-1056 aa) (C-terminal tail part) [85].

Eg5 initially resides in the cytosol as an auto-inhibited, single-chain monomer (Figure 2.10a) [86]. The parallel alignment of the first coiled-coil segment in its stalk brings two monomers together, producing a dimer (Figure 2.10b) [83]. The central BASS (Bipolar Assembly) module then zippers two parallel dimers in an antiparallel fashion, creating a bipolar homotetramer (Figure 2.10c) approximately 80 nm in length [87]. Motor pairs positioned at opposite ends of the tetramer bind to antiparallel microtubules (Figure 2.10d), allowing each Eg5 complex to function as a cross-linking microtubule-sliding bridge that drives spindle pole separation [67]. The precise mechanisms governing this stepwise oligomerization and the ensuring functional capacity of the Eg5 tetramer as a microtubule-sliding bridge facilitating spindle pole separation will be elucidated in greater detail in subsequent subsections, particularly in the dedicated Mechanism and Movement section.

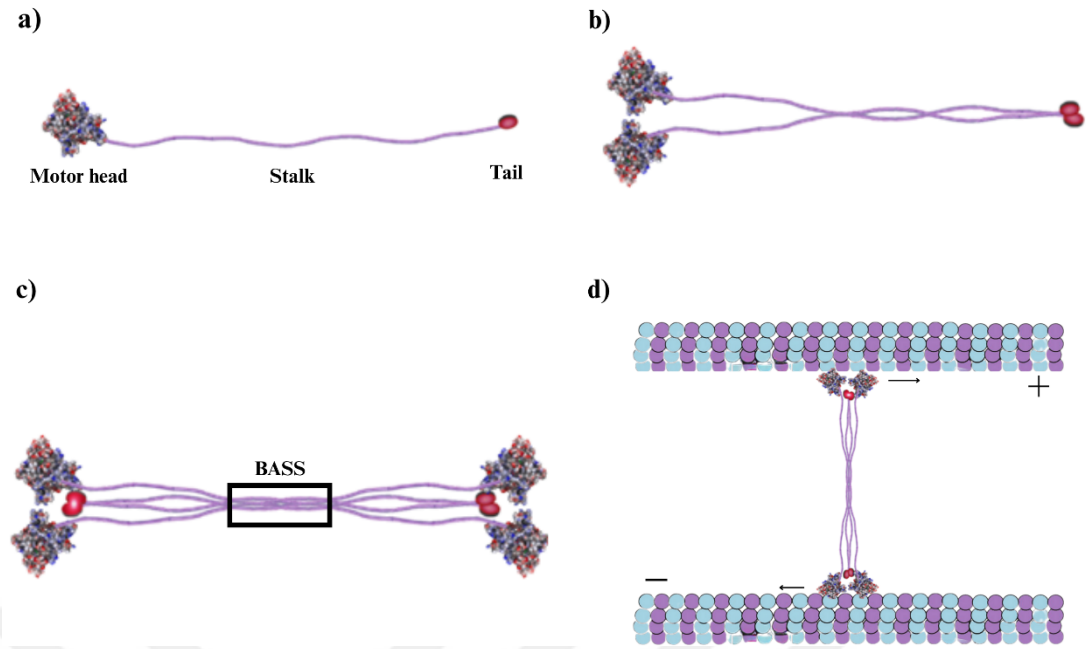

Figure 2.10 : Stepwise oligomerization of Eg5 and its role as a microtubule cross-linking motor [71].

2.2.1.1 Eg5 motor domain architecture and subdomains

The N-terminal motor domain of human Eg5 encompasses the first 360 amino acids. It adopts the conserved kinesin α/β motor fold, which consists of a central eight-stranded β-sheet flanked (Figure 2.11a) by α-helices on either side (Figure 2.11b). Additionally, a small, three-stranded β-sheet was located near the N-terminus [88]. Similar to other P-loop NTPases (e.g., myosins and G-proteins) , the motor domain of Eg5 contains three fundamental nucleotide-binding motifs: the P-loop (Walker A) (Figure 2.11c), which binds the β-phosphate of the nucleotide; Switch I, which senses γ-phosphate; and Switch II, which triggers conformational changes following phosphate release [60, 89].

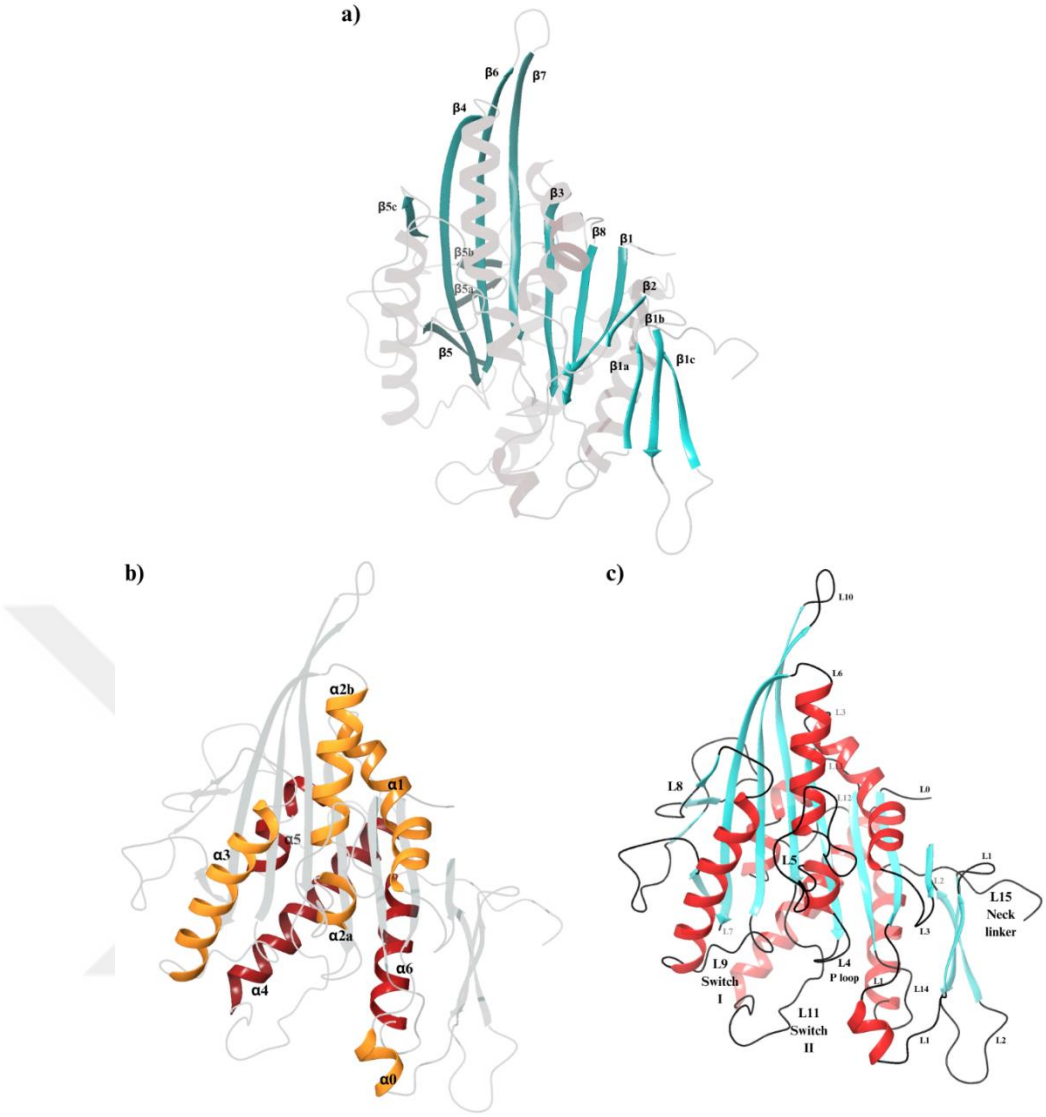

Figure 2.11 : Structural architecture of the human Eg5 motor domain.

These motifs are situated within the nucleotide-binding pocket of the domain core. Surrounding them are structural regions specialized for function: a "tubulin-binding" subdomain that interfaces with the microtubule, and a "switch" subdomain containing Switch I/II, which undergoes substantial movements during the ATPase cycle [90]. The motor neck linker (a flexible segment, ~aa 345–365) extends from the C-terminus of the domain and, by docking onto the catalytic core in different nucleotide states, couples the chemical cycle to mechanical stepping [91].

Structurally, the motor domain of Eg5 can be divided into three subdomains that move relative to each other during the catalytic cycle: (1) the P-loop subdomain (comprising the P-loop and adjacent structural elements), (2) the Switch I/II subdomain (harboring

switch motifs that sense and trigger nucleotide state changes), and (3) the tubulin-binding subdomain (carrying the surface that binds microtubules) [92].

The tubulin-binding interface, which is predominantly positively charged, includes critical elements such as helix-α4 (part of the Switch II region), helix-α5, loop L2 (often containing the conserved KVD finger), loop L8, and loop L12 [90]. This architecture facilitates nucleotide-dependent interactions with the microtubule track, enabling a tightly bound state with ATP or ADP·Pi and a weakly bound state with ADP [93].

Several conserved motifs within the motor domain are crucial for ATP binding and hydrolysis:

P-loop (Walker A): This glycine-rich loop (residues 100-108, sequence GPTGVGKT), also known as loop L4, is located between the β3 strand and helix-α2 [89]. It forms the base of the nucleotide pocket. The invariant Lys111 within the GxxxxGKT motif coordinates the β-phosphate of ATP, and the loop's flexibility allows it to close over the nucleotide. Mutations in this region, such as T107N, abrogate ATP hydrolysis [94, 95].

Switch I: Located C-terminal to helix-α3, this loop contains the conserved NXXSSR motif (in Eg5: NKCDSS, residues 215–221), where Asn216 senses the γ-phosphate of ATP [88]. Switch I transitions between an "open" conformation in the ADP state and a "closed" conformation in the ATP state, a movement essential for stabilizing the γ-phosphate and positioning the catalytic water. Its proximity to helix-α3 and loop L5 makes it sensitive to their conformational changes [96].

Switch II: This region contains the conserved DLAGSE sequence (residues 268-273), which is characteristic of kinesins. Asp268 coordinates the catalytic magnesium and interacts with Switch I, while the subsequent helix-α4 (residues 274–285) is a core component of the microtubule-binding interface. The release of the γ-phosphate triggers a significant conformational change in Switch II, moving helix-α4 and transmitting the signal from the nucleotide pocket to the microtubule-binding surface. This coupling mechanism is critical for motor function, toggling the motor between a "closed" (ATP, ordered α4) and an "open" (ADP, flexible α4) state [96].

Unique Loop L5 Insertion: a 17-residue segment (117-133) that interrupts helix-α2, splitting it into α2a and α2b and creating a prominent surface arm adjacent to the

nucleotide pocket. Rich in polar and aromatic residues with two conserved prolines (Pro121, Pro131), L5 acts as a flexible conformational latch, modulating nucleotide affinity and exchange; deletion of L5 accelerates ADP release, whereas the intact loop stabilises a closed state [97]. Its distinctive size and sequence form an Eg5-specific allosteric pocket that is exploited by selective small-molecule inhibitors and absent from kinesins with shorter L5 loops [98].

Neck Linker: Immediately C-terminal to the motor domain, the flexible neck linker (~15 amino acids, residues 340–360) is essential for processive movement. Upon ATP binding, the neck linker "docks" onto the motor head, initiating the power stroke that drives the forward step of the partner head [91].

In summary, the Eg5 motor domain's function is governed by the coordinated movement of its subdomains and conserved motifs. The unique Loop L5, in conjunction with helices α2 and α3, forms a specific allosteric module (Figure 2.12) that not only controls the nucleotide state but also serves as a regulatory hotspot for both endogenous and synthetic inhibitors, making it a key target for therapeutic intervention [99].

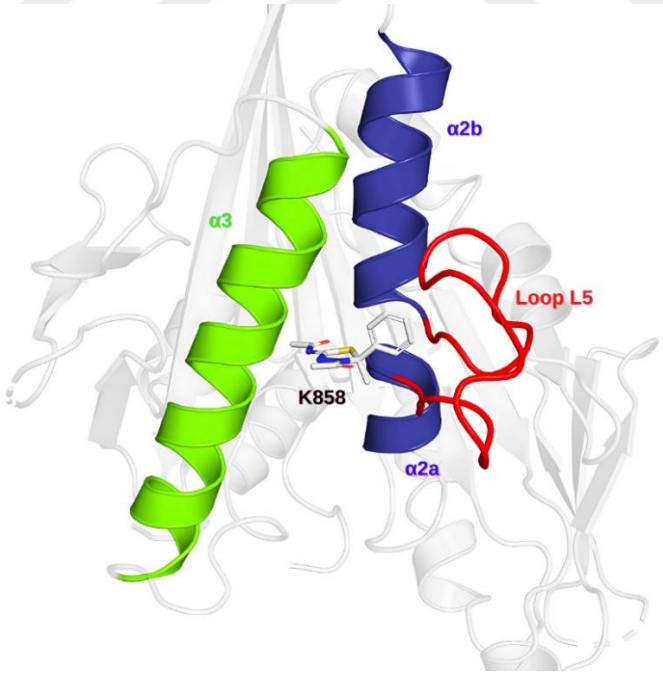

Figure 2.12 : Structural view of the Eg5 loop L5 binding pocket with K858.

In our thesis, the detailed structural and functional analysis of conserved motifs, such as the P-loop, Switch I, Switch II, and particularly the Eg5-specific Loop L5 insertion, has formed the basis for understanding the binding mechanisms of potential inhibitor

candidates and evaluating their selectivity. The position and chemical properties of these motifs were considered when filtering virtual screening results and rationalizing the possible interactions of hit compounds with Eg5. Specifically, the unique structure of Loop L5 has been one of our main focal points in the design and selection of new inhibitors that target this region. Although the literature reveals that Loop L5 forms an Eg5-specific allosteric pocket and is a target for many inhibitors, the interaction potential of this pocket with different chemical groups and how these interactions modulate motor function have not been fully elucidated. This study aims to contribute to the body of knowledge in this field by discovering new and diverse chemical scaffolds that can target the Loop L5 pocket and offer alternatives that could overcome the limitations of existing inhibitors.

Conformational States: ATP-bound, ADP-bound, Apo forms and microtubule interactions

ADP-bound state (Nucleotide-reduced affinity state): In the presence of ADP, Eg5 adopts a conformation conducive to low microtubule affinity. The crystal structures of Eg5-ADP (e.g., PDB 1II6 at 2.1 Å) show an "open" pocket geometry where the Switch I loop is canted away from the nucleotide, expanding the active site. In this state, helix-α3 is relatively straight and extended, while loop L5 adopts a conformation that points outwards from the active site [100, 101]. The Switch II region, including helix-α4, is partially disordered in some ADP structures, reflecting a more "relaxed" state of the motor, a factor which may influence the binding of inhibitors targeting the L5 pocket [102]. A stabilizing salt-bridge cluster between Arg234 and Leu239 at the β6/α3 junction further opens the cleft (Figure 2.13) [103].

An interesting feature of Eg5 compared to conventional kinesin-1 is that its neck-linker is not fully disengaged in the ADP state. In the first solved structure, the neck-linker adopted a novel, ordered conformation angled away from the catalytic core, unlike in kinesin-1 where it is fully peeled back [76]. Cryo-electron microscopy (Cryo-EM) studies of microtubules have shown that neck-linker residues (350-358) become completely undocked, underscoring context-dependent flexibility [77]. Overall, ADP-bound Eg5 possesses an open nucleotide cleft poised for ADP release and exhibits low microtubule affinity [104].

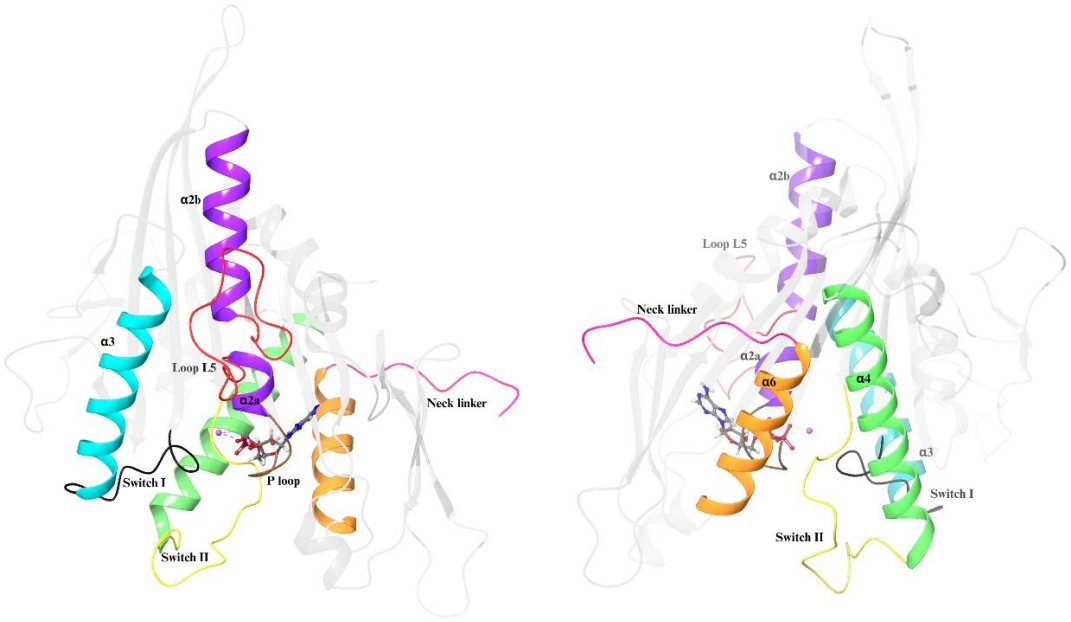

Figure 2.13 : Eg5 ADP-bound state (PDB:1II6).

ATP (or ATP-analog) bound state (pre-powerstroke, tight-binding state): Upon binding a non-hydrolyzable ATP analog, such as AMPPNP or ADP·AlF_x, Eg5 transitions into an ATP-like conformation that confers strong microtubule affinity and corresponds to the pre-power-stroke state. In the Eg5–AMPPNP crystal structure (PDB 3HQD), the Switch I loop folds over the γ-phosphate, while the adjacent Switch I helix partially unwinds to accommodate this closure [96, 105]. The P-loop Lys111 pivots ~2 Å toward the γ-phosphate, completing the catalytic triad. Similarly, Switch II adopts a compact, ordered arrangement as the DLAGSE loop and helix-α4 snap shut, recreating the tight-binding configuration observed on the microtubule lattice [106]. A catalytic water pair (W1/W2) that coordinates Mg^{2+} in 3HQD, absent in ADP and apo structures, marks a hallmark of the transition state (Figure 2.14). Concomitantly, helix-α3 shortens and kinks, while loop L5 swings inward toward the nucleotide pocket, establishing direct hydrophobic contact with α3. This reorientation fashions a cavity bordered by L5, α2b, and α3, stabilizing the motor in its ATP-like geometry [106–108]. ATP-analog-bound structures reveal the motor domain clamped in a transition-state-like posture primed for force generation: the microtubule-binding subdomain is rotated, and the neck-linker begins to "zipper" forward along the filament [109]. In effect, nucleotide binding locks Eg5 into a closed and strained configuration—with Switch I and II sealed, L5 folded inward, and the neck linker docked to the catalytic

core—representing the high-affinity state in which the motor tightly grips the microtubule while awaiting the power stroke [109, 110].

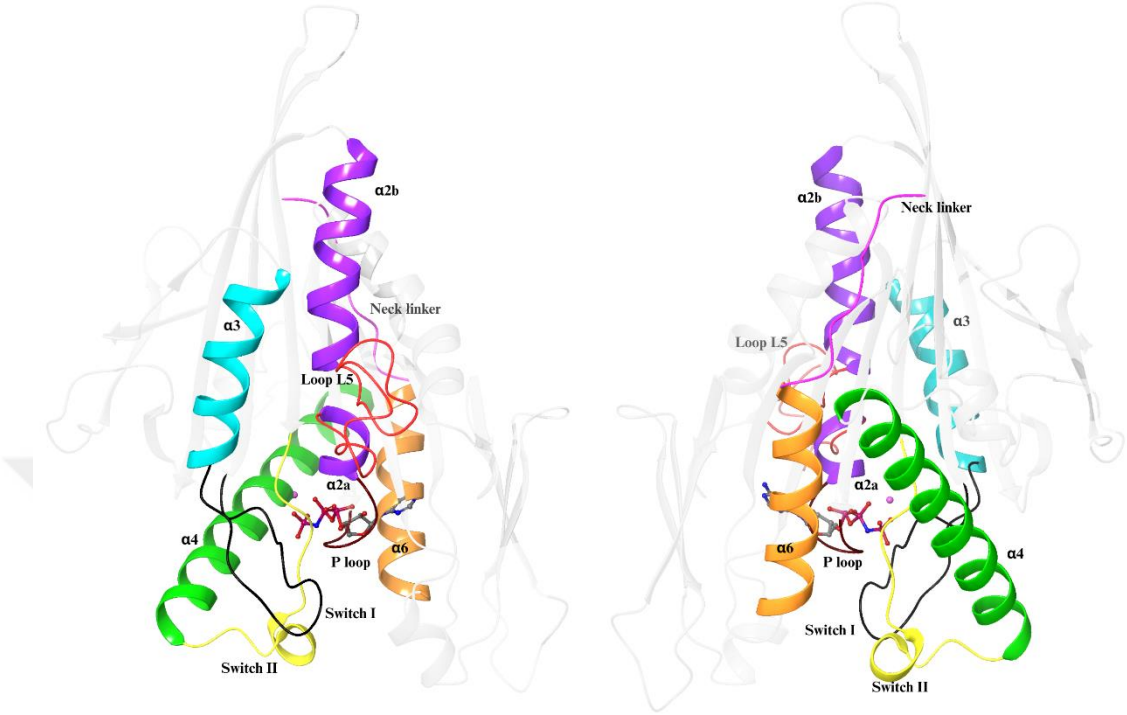

Figure 2.14 : Eg5-AMPPNP(ATP-analog) bound state (PDB:3HQD).

Nucleotide-free (apo) state: The apo state, defined as the moment after ADP release but before a new ATP molecule has docked, is intrinsically transient [90]. Although capturing a truly nucleotide-free Eg5 structure is experimentally challenging, a "rigor-like" conformation can be inferred from biochemical and structural surrogates . The crystal structure 3ZCW, obtained with the benzimidazole BI8 bound in the α4/α6 pocket, shows an empty nucleotide site and serves as a soluble apo proxy (Figure 2.15) [111]. In most kinesins, complete nucleotide loss clamps the motor onto the microtubule with an even higher affinity than the ATP state; Switch I typically remains closed, preserving an ATP-like geometry while the catalytic pocket is empty [91]. For Eg5, this behavior can be investigated using the P-loop mutant T107N, which binds microtubules without release. Cryo-EM reconstructions of this locked apo analog (e.g., PDB 6TA2) reveal that in the absence of nucleotide, the P-loop and both Switch loops retain an ATP-like fold but tilt slightly inward, thereby enhancing the tubulin interface. These data indicate that nucleotide-free Eg5 conserves most ATP-like contacts (closed switches, forward-docked neck-linker) but exhibits subtle distortions that compensate for the lack of the γ-phosphate's stabilizing charge [10].

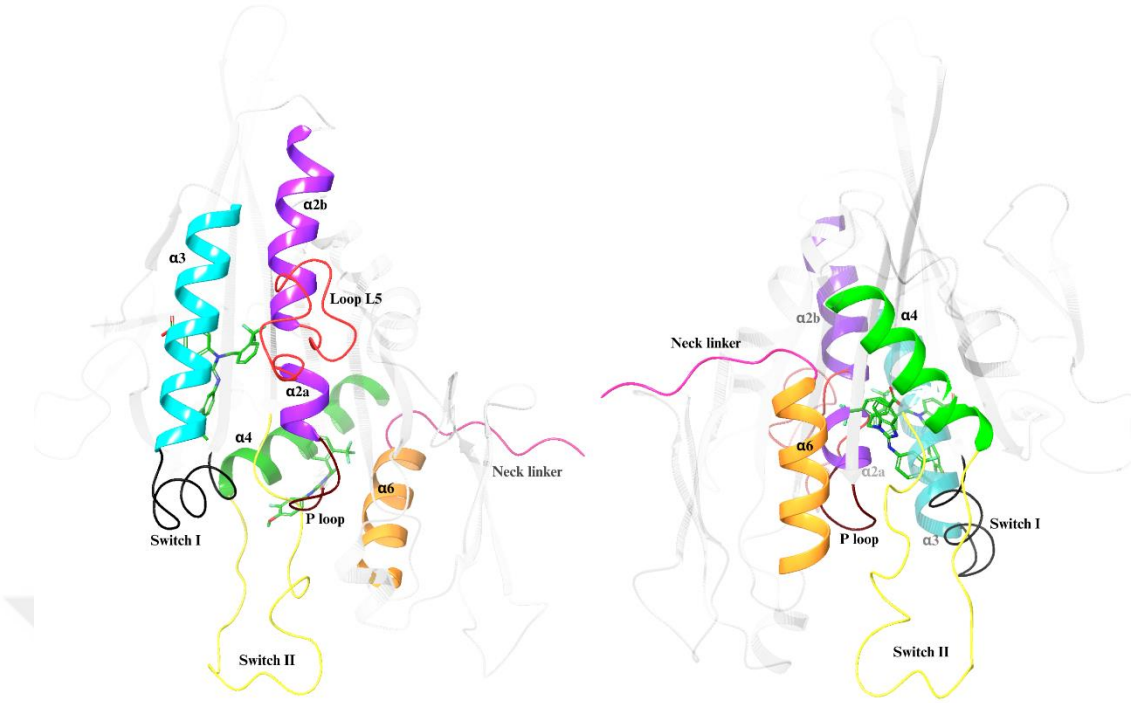

Figure 2.15 : Eg5 nucleotide free state (PDB: 3ZCW).

Eg5 on microtubules and AlphaFold2 perspective: The states described above become more pronounced in the presence of microtubules (MT), as MT binding further biases the conformations. Eg5 bound to a microtubule and captured with an ATP analog shows a strong binding configuration, with the Switch II region interacting with tubulin and the neck-linker docked [101]. Recent Cryo-EM of human Eg5 bound to microtubules with AMPPNP at ~4 Å resolution (PDB 6TA4) shows the Eg5 subdomains in an ATP-like, docked state on tubulin (Figure 2.16) [98]. Conversely, Eg5–ADP tends to bind weakly to microtubules; structural data suggest that upon initial binding, it rapidly releases ADP to enter the apo state, as microtubule binding greatly accelerates ADP release in kinesins. Thus, on microtubules, Eg5 most likely exists predominantly in a closed ATP-like state or an apo state, with the ADP state being short-lived [103, 112].

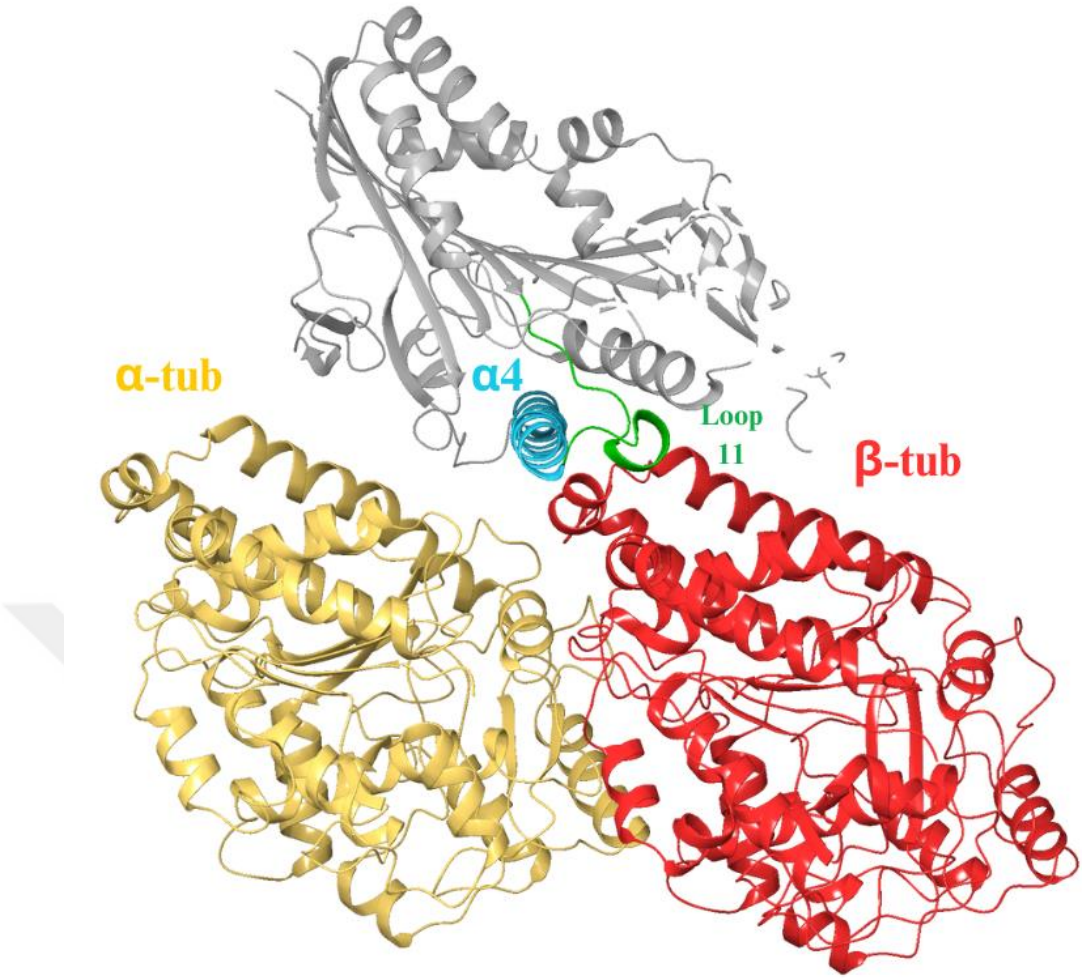

Figure 2.16 : Eg5 on microtubule (PDB: 6TA4).

The AlphaFold2 predicted model for Eg5 aligns well with known crystal structures, predicting a closed, ATP-like conformation for the motor domain with a well-formed core and a partially ordered Loop L5 [113]. However, AlphaFold tends to model Loop L5 in an averaged position, and experimental structures show that L5 can be repositioned significantly between states; a single model, therefore, does not capture its full flexibility [114, 115]. Confidence in the AlphaFold prediction drops for L5 and the neck-linker, indicating that these regions are intrinsically flexible, consistent with experimental findings [116].

In summary, the structural plasticity of the Eg5 motor is crucial for its function. Elements like L5 and the switch loops act as molecular switches, transitioning between an open, relaxed state (ADP) and a closed, strained state (ATP), with microtubule binding shifting the equilibrium toward the strained, force-generating conformations.

2.2.1.2 Crystal structures of Eg5

More than 35 structures of the Eg5 motor domain, capturing various liganded states, have been solved using X-ray crystallography or cryo-EM (Table 2.3). The table below summarizes the key representative structures, their states, and conformations:

Table 2.3: Representative X-Ray and Cryo-EM structures of selected critical human Eg5 complexes.

PDB ID	Method	Resolution (Å)	Bound Ligand(s), Co-factors	Conformation / Notes
1II6	X-ray	2.1	ADP, Mg^{2+}	ADP-bound state (no inhibitors). Open nucleotide pocket, L5 extended, α3 long (ADP-like). Neck linker undocked. First Eg5 structure solved [88].
3HQD	X-ray	2.3	AMPPNP (ATP analog), Mg^{2+}	ATP-like state. Switch I/II closed, L5 folded inward contacting α3, neck linker docked. Pre-powerstroke conformation [96].
1X88	X-ray	1.8	ADP, Mg^{2+}; Monastrol	Allosteric inhibitor (L5 pocket) bound. L5 folded over (ATP-like), but α3 and Switch I remain in ADP-like position. "Hybrid" conformation; monastrol locks motor in an ADP-bound, inhibited state [117].
2IEH	X-ray	2.0	ADP, Mg^{2+}; Mon-97 (monastrol analog)	Similar to 1X88: inhibitor in L5/α2/α3 pocket. High affinity state with L5 closed over pocket. Validates monastrol binding mode with improved ligand [118].
4AP0	X-ray	2.6	ADP, Mg^{2+}; Ispinesib (clinical inhibitor)	Eg5–ADP–ispinesib ternary complex. L5/α2/α3 allosteric pocket occupied by ispinesib, locking Eg5 in ATP-like conformation with ADP trapped. Switch II is perturbed to a "closed" state. Represents how ispinesib (Kis <10 nM) stalls the motor in an ADP-bound, tight MT-binding form [119].
5ZO7	X-ray	2.8	ADP, Mg^{2+}; STLC analog (chlorinated)	Eg5 is bound to a S-trityl-L-cysteine type inhibitor. Inhibitor in L5 pocket; induces pocket closure. Shows enhanced inhibitor interactions with Eg5. Switch I/II remain in ADP state (inhibited) [120].
6HKY	X-ray	2.7	ADP, Mg^{2+}; Arry-520 (Filanesib)	Eg5 with filanesib (thiadiazole allosteric inhibitor). L5/α3 pocket occupied; conformation like ispinesib-bound (ADP trapped). Filanesib was a Phase II compound. Shows Eg5's L5 can accommodate diverse chemotypes [121].
3WPN	X-ray	3.5	ADP, Mg^{2+}; PVZB1194 (biaryl inhibitor)	Eg5 with a helix-α4/α6 pocket inhibitor (MT-independent inhibitor). Some regions (α3, α4, Switch loops, L5) are disordered in this MT-free structure, reflecting flexibility. Helped identify the alternate allosteric site between helices α4 and α6 [122].
6TA3/6TIW (pair)	Cryo-EM	3.8	*No nucleotide* (apo); GSK-1; Microtubule	MT-bound Eg5 trapped with GSK-1 (a biaryl ATP-competitive inhibitor). Reveals allosteric site at α4/α6 on the MT interface. Two confirmations reported for motor domain. GSK-1-bound Eg5 shows tight MT attachment (tubulin-binding subdomain locked) and an empty, distorted nucleotide pocket (P-loop perturbed, nucleotide cannot bind). Demonstrates how GSK-1 "freezes" Eg5 on MT in a rigor-like state [98].

Table 2.3 (continue): Representative X-Ray and Cryo-EM structures of selected critical human Eg5 complexes.

PDB ID	Method	Resolution (Å)	Bound Ligand(s), Co-factors	Conformation / Notes
6TA4	Cryo-EM	4.2	AMPPNP, Mg^{2+}; Microtubule	Eg5 motor domain bound to microtubule in AMPPNP state. ATP-like, strongly bound state on tubulin – Switch I/II closed, and neck linker docked. Confirms Eg5-tubulin interface and subdomain rotations in the pre-powerstroke state. Used as reference for inhibited complexes [98].
6G6Y	X-ray	1.8	ADP, Mg^{2+}; K858	Eg5–ADP–K858 complex. K858 binds to the L5/α2/α3 allosteric pocket, locking Eg5 in an ADP-bound state by preventing ADP release. This structure provides insights into the design of thiadiazole-containing inhibitors targeting Eg5 [123].
1Q0B	X-ray	1.9	ADP, Mg^{2+}; Monastrol	First inhibitor-bound Eg5 structure. Monastrol binds to an induced-fit pocket between loop L5 and α3, stabilizing an inhibited conformation with the neck linker undocked [124].
2G1Q	X-ray	2.4	ADP, Mg^{2+}; Inhibitor 9h	Eg5 bound to a dihydropyrazole inhibitor at the L5/α2/α3 pocket, revealing interactions critical for high-affinity binding [125].
2FL6	X-ray	2.5	ADP, Mg^{2+}; Inhibitor 6	Eg5 complexed with a dihydropyrrole inhibitor, demonstrating enhanced potency and solubility through modifications at the L5/α2/α3 site [126].
2GM1	X-ray	2.3	ADP, Mg^{2+}; Pyrrolotriazine-based inhibitor	Eg5 in complex with a pyrrolotriazine inhibitor, illustrating a unique binding mode at the L5/α2/α3 pocket [127].

The extensive collection of solved Eg5 structures provides a detailed atomic-level understanding of its function and inhibition. These structures capture the motor domain in its key nucleotide states (ADP- and ATP-bound) and in complex with a diverse range of inhibitors targeting its distinct allosteric pockets. A key insight from these structures is the confirmation of two primary allosteric inhibition mechanisms. Inhibitors targeting the L5/α2/α3 pocket, such as monastrol and ispinesib, consistently trap the motor with ADP locked in the active site [121]. In contrast, inhibitors of the α4/α6 site, such as GSK-1, bind in a different conformation: GSK-1 binding causes Eg5 to eject its nucleotide (apo) and bind strongly to the microtubule [128]. Thus, these structures illustrate two distinct allosteric mechanisms of inhibition (discussed later).

For structure-based drug discovery targeting the loop L5 pocket of Eg5, it is advantageous to use receptor structures that already contain an inhibitor and a nucleotide in the pocket. Such "co-crystal" conformations accurately define the shape and interactions of the allosteric site. In Eg5, the binding of an inhibitor establishes a distinct pocket geometry and stabilizes loop L5 against helix α3 [107]. For instance, the binding of ispinesib or monastrol analogs locks Eg5 in the ADP-bound conformation, which is the same state captured by K858 in the PDB entry 6G6Y. In the 6G6Y structure, ADP is present, and loop L5 is closed over the ligand. Virtual

screening against this structure thus targets the "final inhibitor-bound state" which aligns with the intended mechanism (blocking ADP release). Moreover, high-resolution ligand-bound structures minimize the ambiguity in side-chain positions and pocket waters [119, 123]. In practice, many Eg5 inhibitor discovery campaigns have utilized Eg5-ADP-inhibitor complexes as docking templates (discussed later). This approach follows precedent: Eg5 inhibitors (such as ispinesib) have been optimized using ADP-bound structures, which consistently exhibit similar L5 pocket arrangements. In summary, using ligand/ADP-bound Eg5 structures (e.g., 6G6Y) for virtual screening ensures that the loop L5 allosteric pocket is in its appropriate "closed" conformation, thereby increasing the probability of finding compounds that will bind and lock Eg5 in its inhibited state.

The Eg5/ADP/K858 crystal structure

The 6G6Y structure, representing the Eg5 motor domain in complex with an inhibitor, is the crystal structure of human Eg5 bound to ADP and the small compound inhibitor K858. This structure was solved at 1.8 Å resolution and revealed the ternary complex of the Eg5 motor protein: in the 6G6Y structure, Eg5 was crystallized as a monomeric motor domain (367 amino acids, chain A) and had bound ADP, a Mg^{2+} ion, and the K858 compound (PDB ligand code: EOK). This complex represents the "locked" state of the Eg5 motor induced by the inhibitor [123]. The complex structure illustrates a state in which ADP release from the Eg5 motor is blocked, which is a critical step for Eg5 activity because ADP dissociation is necessary for new ATP binding and for the cycle to continue [111]. In our thesis, the various Eg5 crystal structures summarized in Table 2.3 and discussed in the literature served as fundamental reference points, especially for defining inhibitor binding pockets and developing virtual screening protocols. Specifically, for the detailed analysis of the Loop L5 allosteric pocket, which is the focal point of our study, and for virtual screenings targeting this pocket, the high-resolution 6G6Y structure, complexed with the K858 inhibitor and ADP, was selected as the primary receptor structure. The rationale for this choice was based on its accurate representation of the 'locked' inhibited state and successes in similar studies in literature. Although numerous Eg5 structures are available, how new potential inhibitors with different chemical scaffolds might bind to these structures and what advantages they might offer compared to existing inhibitors remain active areas of research. Our thesis focuses on discovering new and diverse chemical structures,

not previously identified for Eg5, by specifically targeting the Loop L5 allosteric pocket defined around the 6G6Y structure. In this way, it is aimed to identify next-generation ligands that could offer alternative solutions to problems such as resistance and toxicity encountered with current Eg5 inhibitors.

2.2.1.3 Binding pockets of Eg5

Nucleotide binding pocket of Eg5

The ATP-binding site in the Eg5 motor protein has a mostly hydrophilic character. ATP's triphosphate tail is highly negatively charged and polar in aqueous solution, so the pocket is lined with polar side-chains—principally Lys111 of the P-loop, Ser219 of Switch I, and Asp268/Glu273 of Switch II—that coordinate the phosphate groups together with a catalytic Mg^{2+} ion [129]. The pocket is not entirely hydrophilic, however: the adenine ring of ATP packs against hydrophobic aromatics such as Phe87 and Tyr104, creating π–π stacking that anchors the adenosine moiety deep inside the cleft [130]. Function-critical sequence blocks (P-loop GxxxxGKT, Switch I NXXSSR, Switch II DLAGSE) are > 90 % conserved across the kinesin super-family—including kinesin-1, kinesin-3 and kinesin-5 (Eg5)—highlighting the evolutionary pressure to maintain an invariant ATP-handling mechanism (Figure 2.17) [112].

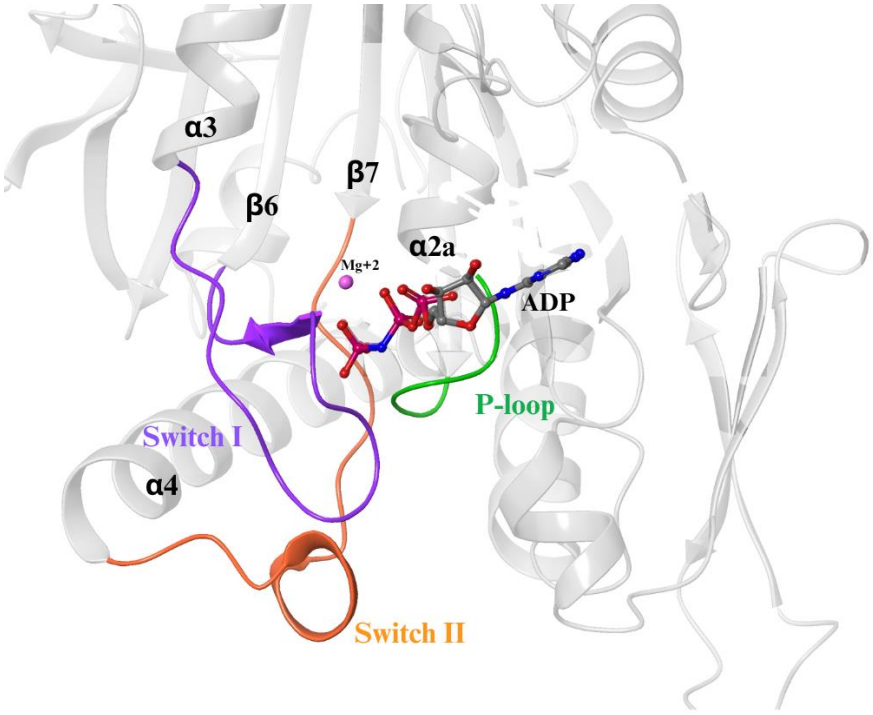

Figure 2.17 : Nucleotide binding pocket.

Allosteric inhibitor binding pockets of Eg5

Eg5 has proven to be a rich target for allosteric inhibitors that bind to specific pockets within the motor domain that are distinct from the ATP-binding site []. These compounds interfere with the force-generating cycle by trapping Eg5 in an inactive conformation [131]. Two main allosteric pockets have been characterized in the motor domain of Eg5 (Figure 2.18).

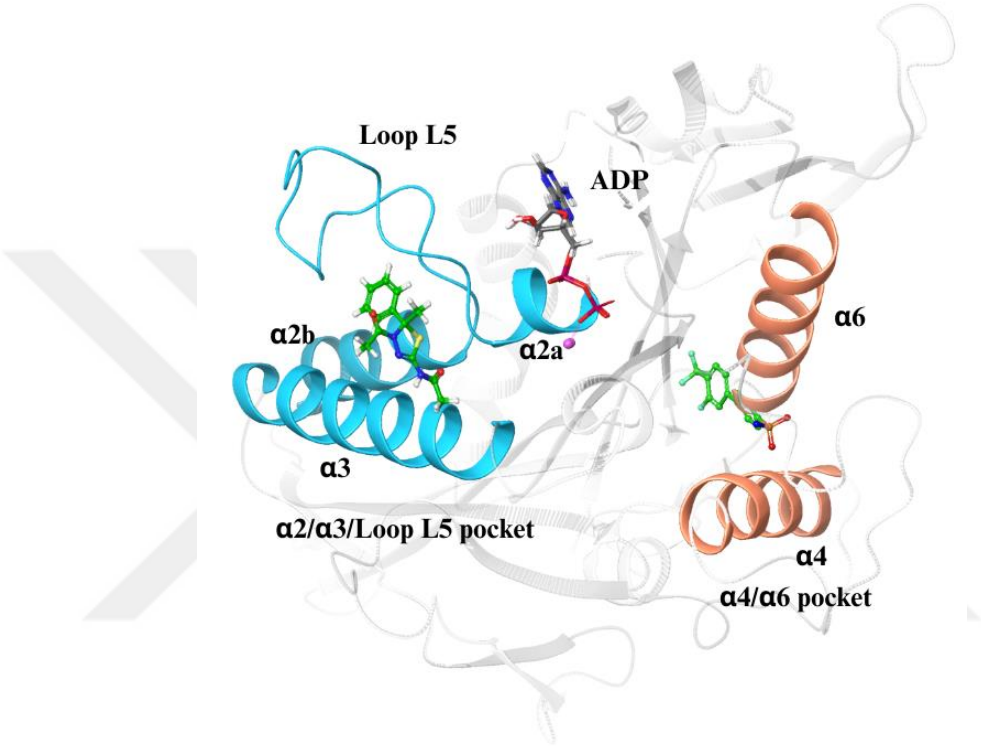

Figure 2.18 : Allosteric inhibitor binding pockets.

The Loop L5 Pocket (α2/L5/α3 interface) (Allosteric Site – I)

This allosteric pocket sits on the upper surface of the motor head, bordered laterally by helix α2b (Ile134–Leu148) and helix α3 (Val210–Leu214), and capped dorsally by the long loop L5 (Gly117–Gly134) [117]. The cleft lies ~10 Å from the nucleotide site; when vacant, it is shallow, but ligand binding triggers an "induced-fit" swing of L5 (hinging at Pro121 and Pro131) that deepens the pocket [132, 133]. The floor and walls are built principally from Glu116-Glu118-Arg119 (α2a/L5 junction), Trp127-Asp130-Asp133 (central L5), and Tyr211-Leu214 (α3) (Figure2.19) [134]. These residues create a composite surface of hydrogen bond donors/acceptors and aromatic platforms. Clinically advanced inhibitors nestle beneath the folded L5 arm. Hydrogen bonds to Glu116 or Glu118 anchor the ligand's heteroatom-rich core, π–π or CH–π

stacking against Trp127 and Tyr211 stabilizes the aromatic rings, and a signature salt bridge or H-bond to Asp130 secures the dihydropyrimidine or trityl moieties. Leu214 provides a hydrophobic buttress that defines the ligand depth [135]. Crystal and stopped-flow studies have shown that the occupancy of this pocket traps ADP in the catalytic site: Switch I/II remains in an ATP-like configuration, whereas the exit channel is occluded by the folded loop [132]. Mutations D130V or D133A disrupt the key polar lock, abolishing inhibitor binding with >400-fold shifts in IC_{50}, while leaving basal motility largely intact, underscoring the central role of these acidic side chains in ligand recognition and signal transmission [136]. Because kinesins 1, 2, and others possess only a short L5, most of the above residues (particularly Trp127, Asp130, and Leu214) are absent or repositioned in their motor cores; therefore, the pocket is unique to kinesin-5, explaining the exquisite selectivity of L5-site drugs [119]. In cells, such inhibitors "freeze" Eg5 on microtubules before force generation, producing monopolar spindles (monoasters) within a single mitosis—a phenotype that can be reversed upon wash-out, confirming that inhibition hinges on this ligand-induced, residue-specific pocket closure [5, 88].

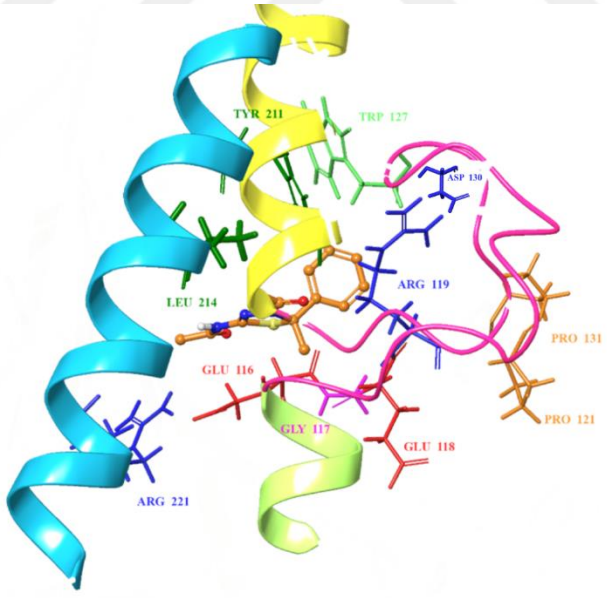

Figure 2.19 : α2–L5–α3 inhibitor binding pocket – close-up view, K858 within the pocket, key residues were highlighted.

Helix-α4/α6 pocket (Switch II allosteric region) (Allosteric Site – II)

Subsequently, a second allosteric drug-binding site was identified at the interface between helix-α4 and helix-α6 of the motor domain [111]. This region is located closer to the side of the motor that directly contacts the microtubule. The pocket is partially

blocked when Eg5 is microtubule-bound and becomes fully accessible only in a specific post-power-stroke conformation, typically at the MT plus-end. Inhibitors targeting this pocket are typically ATP-competitive; they displace or prevent ATP but bind ≈12–15 Å from the nucleotide triphosphate itself, and exert maximal potency when the motor is attached to microtubules (hence their designation as microtubule-dependent inhibitors) [98]. Prototypical compounds include the biaryl inhibitors GSK-1 and GSK-2 developed by GlaxoSmithKline. High-resolution cryo-EM at 3.8 Å (Eg5–MT–GSK-1, PDB 6TA3) shows GSK-1 wedged between the C-terminal half of α4 (Glu255–Gly268), loop L11 (Asn271–Ile288) and the N-terminus of α6 (Leu292–Arg355) [98, 137]. This binding pocket abuts the Switch II cluster and overlaps with part of the microtubule-binding interface. GSK-1 occupies an allosteric cleft at the α4/α6 interface of Eg5 and, when bound to microtubule-attached motors, distorts the P-loop, expels ADP, and prevents re-entry of ATP, thereby locking the kinesin in a rigor-like state that immobilizes spindle microtubules [98]. This mechanism differs from that of loop-L5 inhibitors, which trap ADP within the catalytic site; α4/α6 ligands instead block the γ-phosphate pocket after ADP ejection, producing a “hyper-locked” motor [128]. The pocket comprises residues Glu255, Glu258, Glu270 (α4), Asn287-Asn289 (loop L11) and Leu292-Arg355 (α6), with Tyr104 on the β-sheet rim acting as a key recognition element—its Y104C mutation confers >100-fold resistance to BRD9647 [103, 138, 139]. Because this surface lies outside loop L5, α4/α6 inhibitors such as GSK-1/GSK-2 retain efficacy against L5-site–resistant mutants (e.g., D130V/A133D). Both the L5 and α4/α6 pockets ultimately halt Eg5’s mechanochemical cycle and provoke mitotic arrest, but through distinct allosteric routes.

Contemporary drug-discovery programs prioritise these two unique allosteric cavities as the most tractable sites for selective Eg5 inhibition.

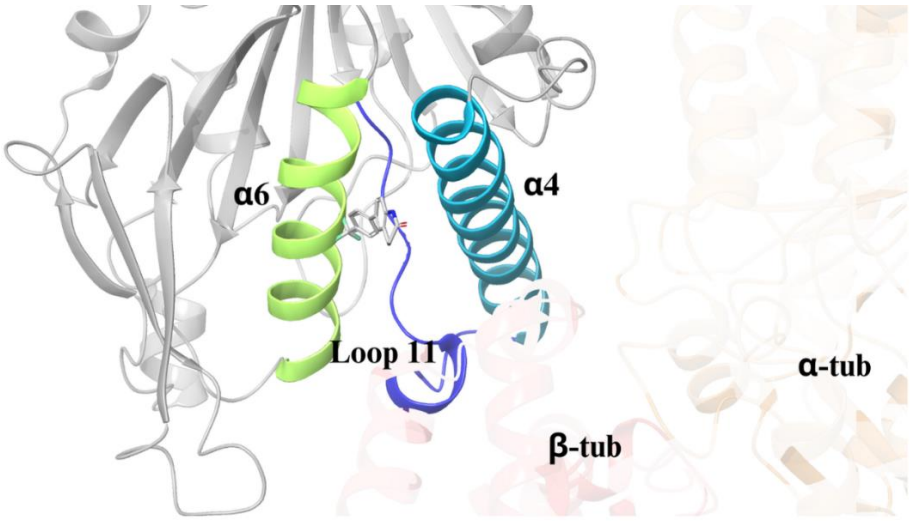

Figure 2.20 : α4-α6 binding pocket (6TA3).

Emerging new potential sites

The allosteric landscape of Eg5 extends beyond the canonical L5 and α4/α6 sites, offering new avenues to overcome drug resistance. Evidence for this includes the natural product Terpendole E, which inhibits Eg5 mutants resistant to ligands of both known pockets, suggesting a novel, uncharacterized binding site or mechanism [140]. Additionally, a distinct allosteric pocket adjacent to the neck-linker junction, formed by helix-α5 and surrounding sheets, has been identified and validated []. Ligands targeting this site, such as the pyrimidine-indazole SRI35566, allosterically inhibit the ATPase cycle and block the power stroke, providing a strategic alternative to bypass classical monastrol-resistance mutations [141].

In summary, the human Eg5 motor domain comprises a well-understood structural core containing conserved nucleotide-processing motifs (P-loop, Switch I/II) and unique regulatory elements (notably, loop L5) that differentiate it. Different nucleotide states induce major conformational changes in regions such as helix-α3, L5, and Switch I/II, directly coupling ATP chemistry to changes in microtubule affinity. Decades of structural work (X-ray, cryo-EM, and computational) have painted a detailed picture of how Eg5 works and how it can be inhibited by small molecules. The α2/L5/α3 allosteric pocket is exploited by numerous Eg5-specific inhibitors to lock the motor in an ADP-bound, inactive state, whereas newer α4/α6 pocket inhibitors lock the motor in a nucleotide-vacant rigor state on microtubules. This knowledge not only informs drug design for anti-cancer therapeutics but also advances our fundamental understanding of kinesin motor mechanics and allosteric regulation.

A detailed understanding of these binding pockets, particularly the well-characterized Loop L5 pocket and the emerging α4/α6 pocket, has been instrumental in guiding the virtual screening strategies employed in this thesis. Our primary focus has been the Loop L5 pocket, leveraging the extensive structural information available, including the 6G6Y structure, to design specific pharmacophore models and develop docking protocols. The characterization of key interacting residues within these pockets (e.g., Glu116, Glu118, Arg119, and Trp127 for L5; Tyr104 for α4/α6) allowed for the rational filtering of binding poses and prioritization of candidates likely to form meaningful interactions. Although the L5 pocket is a validated target, the emergence of resistance to L5 inhibitors highlights the need to explore alternative binding sites or novel chemotypes for the L5 pocket itself. This thesis aims to contribute to the identification of novel scaffolds for the L5 pocket, providing a basis for future strategies to overcome resistance or achieve synergistic inhibition.

2.2.1.4 Eg5 stalk domain and tail domain architecture and subdomains

Beyond the motor domain, the stalk and tail regions of Eg5 are essential for its tetrameric assembly, regulation, and overall function in the mitotic spindle.

The stalk domain and BASS-mediated tetramerization

The Eg5 stalk is composed of four sequential coiled-coil segments (CC1-CC4) that mediate dimerization through hydrophobic heptad repeats. Flexibility is conferred by minor linker regions between these segments. A key feature within the stalk is the Bipolar Assembly (BASS) domain, a structural module unique to the kinesin-5 family responsible for tetramerization [87, 142]. Spanning approximately 200 amino acids in the central part of the stalk, the BASS domain is essential for kinesin-5 activity and cell viability. Crystal structure studies have revealed that its core consists of two antiparallel coiled-coils stabilized by alternating hydrophobic and ionic four-helix interfaces [143]. From this central region, the helices splay outwards, swap partners, and form parallel coiled-coils shifted by 90°. This unique architecture is proposed to play a crucial role in transmitting forces between the motors located at opposite ends of the molecule. It creates a rigid four-helix bundle that provides an axis for the symmetric distribution of forces and absorbs mechanical strain, which is crucial for its function as a microtubule crosslinker (Figure 2.21) [144].

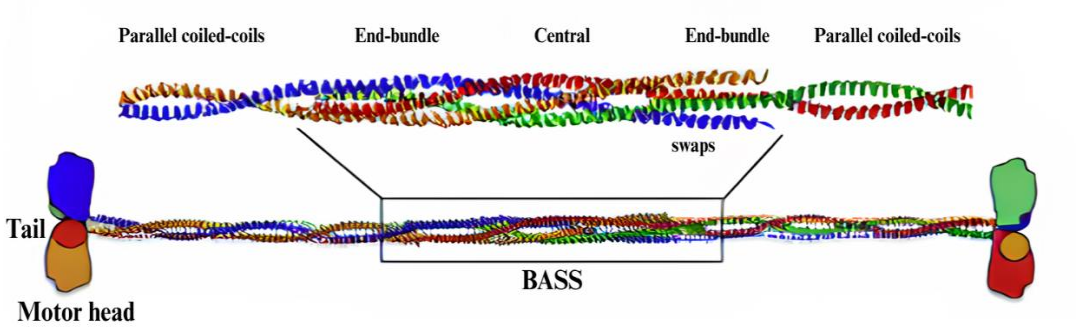

Figure 2.21 : Structural organization of the Eg5 bipolar tetramer, illustrating the stalk domain and central BASS assembly.

The C-terminal tail domain and auto-regulation

The C-terminal tail domain (approx. aa 913–1056) is a crucial regulatory hub. In the nucleotide-free or ADP-bound state, the tail domain can fold back and directly engage with the α0-helix of the motor domain. This interaction functions as an auto-inhibitory mechanism, slowing ATP binding and ATPase activity, which in turn enhances the motor's force output and promotes robust, coordinated sliding of microtubules [145].

The tail domain contains several structurally important motifs, including the "BimC box" (approx. residues 927-948), which is a known phosphorylation site critical for regulating Eg5's localization within the mitotic spindle. The most distal part of the tail (approx. residues 1020-1056) constitutes the primary motor domain interaction site, mediating the fold-back interaction [75, 145]. The structural integrity of the tail and its ability to engage with the motor domain are essential, as confirmed by cryo-EM studies showing that its deletion abrogates key regulatory functions [145].

While our primary investigation in this thesis is centered on the allosteric Loop L5 pocket (Allosteric Site I) within the motor domain, understanding the stalk and tail domain architecture is crucial for a holistic view of Eg5's function and regulation. The role of the stalk domain in dimerization and tetramerization directly influences the spatial arrangement and cooperative action of the motor domains. Furthermore, the auto-inhibitory and regulatory functions mediated by the tail domain, including its interactions with the motor domain, can allosterically impact the accessibility or dynamics of drug-binding sites, such as the L5 pocket. Although this study does not directly target the stalk or tail domains for ligand binding, knowledge of their structure and interactions provides a more complete context for interpreting the behavior of the motor domain and for considering broader future strategies in Eg5-targeted drug

design that may involve modulating these inter-domain communications or oligomerization interfaces. A comprehensive understanding of these distal domains also informs the interpretation of cellular assay results in which the full-length protein is active.

2.2.2 Mechanism and movement of Eg5

Whereas the preceding section, 2.2.1, detailed the static structural biology of Eg5, this section focuses on the dynamic functions of the motor. Although a member of the kinesin superfamily, Eg5 has evolved a unique mechanochemical profile that diverges significantly from canonical transport kinesins. This profile is precisely adapted for its specialized role: to operate in large ensembles within the complex and crowded environment of the mitotic spindle, generating robust, outward-pushing forces. The following subsections will dissect the layers of this mechanism, from the fundamental ATP-driven cycle of a single motor domain to the collective behavior of motor ensembles and the complex regulatory networks that control its function in time and space.

2.2.2.1 Overview of the mechanochemical cycle of Eg5

Eg5 is a plus-end-directed motor protein that converts the chemical energy from ATP hydrolysis into mechanical work to move along microtubule (MT) tracks [71, 146]. While its movement follows the generally accepted "hand-over-hand" model for kinesins, Eg5 incorporates significant adaptations to this basic template that are specific to its biological function[96]. Its most fundamental structural feature is its operation as a homotetramer with two motor domains at each end; this bipolar architecture allows it to cross-link antiparallel microtubules and slide them apart. This process is a precision mechanism in which each 8 nm step is powered by the hydrolysis of one ATP molecule [146].

One of the most fundamental properties distinguishing Eg5 from highly processive transport motors like Kinesin-1 (i.e., those that can travel long distances before detaching) is its "minimally processive" nature. Studies have shown that an Eg5 dimer can, on average, take only a few steps (~8–10 steps) before detaching from a microtubule. While this might initially appear to be an inefficiency, it is a result of Eg5's evolutionary adaptation. This low processivity is compensated by a high "duty ratio," meaning the motor spends a large fraction of its ATPase cycle strongly bound

to the microtubule. This characteristic makes Eg5 a highly effective force generator and load-bearer when working within an ensemble [147–149].

In this context, the attributes of Eg5 as a single motor that might be perceived as "weaknesses" (low processivity and slow speed) are, in fact, "strengths" evolved for its biological role. Eg5 is not a cargo transporter designed for long-distance travel but rather a "scaffolding motor" optimized for sustained force production and load-bearing within a crowded and constantly stressed structure—the mitotic spindle. The mitotic spindle is under continuous tension created by opposing forces from other motors (e.g., dynein). In this environment, the ability to grip the microtubule, maintain position, and apply force is prioritized over speed. The high duty ratio ensures that even if one motor in an ensemble detaches, many others remain firmly bound, which translates to uninterrupted force transmission and load sharing. Therefore, low processivity is not a flaw but a feature. It allows for the dynamic reorganization of individual motors, permitting them to detach and re-attach where needed, thereby distributing the force across the entire ensemble. This creates a more robust system than relying on a few highly processive motors that might stall or break under high load. This perspective positions Eg5 not as a "solo sprinter" but as a collective "team player" [90, 97, 150].

2.2.2.2 Nucleotide processing and microtubule affinity

The intricate sequence of events that constitute the mechanochemical cycle, from initial microtubule binding to the force-generating power stroke, is visually summarized in (Figure 2.22) The initial and critical interaction of an Eg5 dimer with a microtubule involves unique nucleotide-processing steps that prime the motor for its mechanochemical cycle. Free Eg5 in solution exists predominantly in an ADP-bound state, which exhibits a low affinity for microtubules. Upon encountering a microtubule, a defining event that distinguishes Eg5 from other kinesins occurs: both motor domains in the dimer release their bound ADP molecules [147, 151].

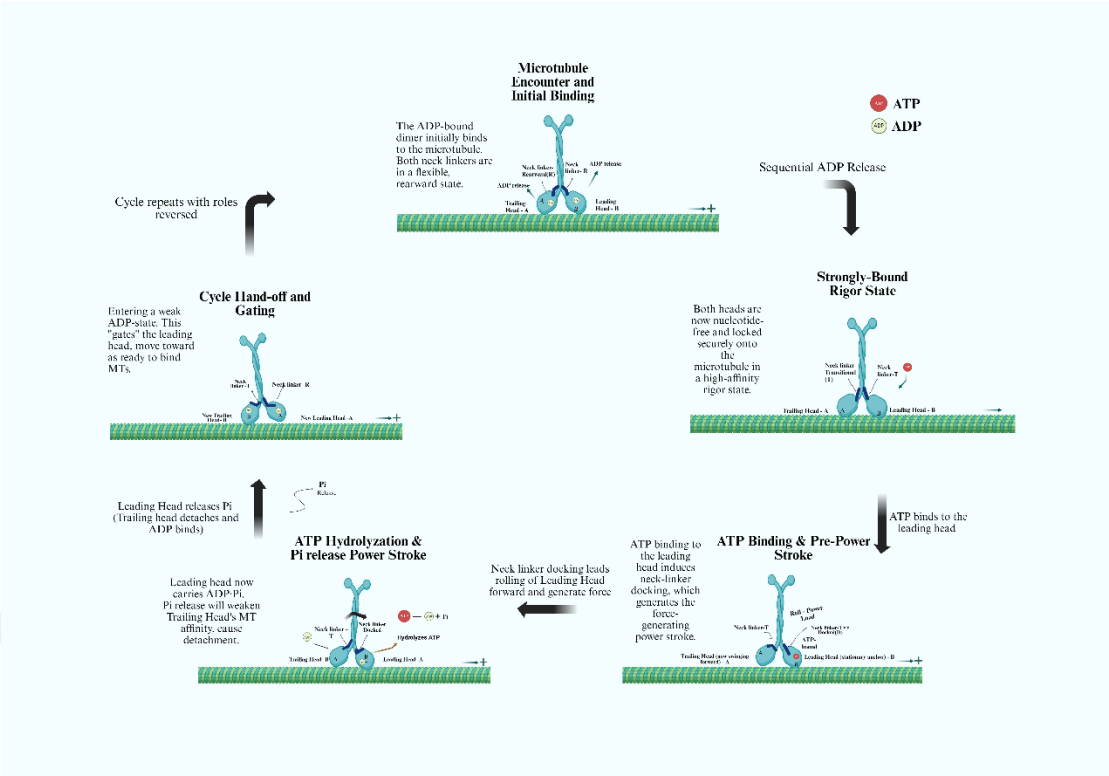

Figure 2.22 : A schematic representation of the five-step mechanochemical cycle of the Eg5, illustrating the coordinated hand-over-hand mechanism driven by ATP hydrolysis.

This ADP release process is not symmetrical. While the first motor head releases its ADP rapidly, the second head undergo a slow conformational change (isomerization) at a rate of approximately 1 s−1 before releasing its ADP. This kinetically constrained step is a fundamental element of Eg5's gating mechanism, allowing the motor to enter a nucleotide-free "rigor" state in which both heads are strongly bound to the microtubule. This state serves as the starting point for the motor's processive run [147, 152].

This slow, ~1 s−1 isomerization step functions as a "kinetic gate" or a "binding checkpoint." The existence of this mechanism is a reflection of the motor's evolutionary adaptation. The primary function of Eg5 is to cross-link and push antiparallel microtubules, which requires the formation of a stable, force-transducing bridge between the two filaments. If both heads were to release ADP and bind ATP too quickly, the motor might initiate a power stroke before a stable two-headed attachment is secured, leading to inefficient force production or premature detachment. This pause of approximately one second acts as an effective timer. It provides the motor with a window of opportunity to establish a robust, two-headed rigor bond,

ensuring it is properly locked onto the microtubule track. This "lock-before-you-load" mechanism maximizes the efficiency of subsequent force production, which is vital for its role in separating the spindle poles. This feature distinguishes Eg5 from transport kinesins, which are optimized for rapid initiation of movement [100, 153].

2.2.2.3 Structural transitions: focus on loop L5–switch II–neck linker

The mechanical movement of Eg5 is driven by the coordinated conformational changes of specific structural elements within the motor domain. These elements include the N-terminal neck linker (NL), the Switch I and Switch II loops, and the kinesin-5-specific L5 loop. The movements of these structures during the ATPase cycle translate chemical signals, such as nucleotide binding and hydrolysis, into mechanical motion [90, 152, 154]. The binding of ATP to the nucleotide-free head triggers a cascade of conformational changes. At the center of this cascade is the L5 loop, which is situated on the motor's surface adjacent to the active site. Upon ATP binding, the L5 loop retracts from the active site, while the Switch I and Switch II loops close over the nucleotide—a movement essential for catalysis. This closure, via an allosteric communication network, triggers the forward "docking" of the flexible neck linker (NL) onto the motor's catalytic core. This docking, unlike the "disorder-to-order" transition in Kinesin-1, is an "order-to-order" transition in Eg5, as the NL is already partially structured in the ADP-bound state. This docking motion of the neck linker constitutes the primary force-producing "power stroke," which propels the partner head forward by approximately 16 nm to the next binding site [90, 149, 154]. The role of the L5 loop in this process is that of a critical regulatory "latch". Its unique length and conformation in Eg5 modulate the rates of nucleotide binding and release. The fact that mutations in L5 significantly alter ATPase kinetics underscores this loop as a key control element. The binding of allosteric inhibitors like monastrol and STLC to the L5 pocket stalls the cycle by locking the motor in a closed, ATP-like conformation even when ADP is bound [66, 116, 155].

These findings reveal that the L5 loop is not merely an inhibitor binding site but the central processing unit of an allosteric computer that integrates the nucleotide state of the active site with the mechanical state of the neck linker. The binding of an inhibitor like monastrol mimics the conformational state of L5 in the ATP-bound form. This means the inhibitor essentially "tricks" the motor: it locks one part of the system (L5)

in a "go" signal conformation while jamming another part (the nucleotide pocket) in a "stop" signal state (ADP-bound). This creates a conformational conflict that paralyzes the motor. The neck linker receives the signal to dock, but the catalytic site cannot proceed with hydrolysis and release. Thus, L5 is the essential link that couples the chemical event (nucleotide binding) to the mechanical event (NL docking). Its unique size and sequence in the Kinesin-5 family create a specific allosteric vulnerability not present in other kinesins. This explains both the high selectivity of L5-pocket inhibitors and the distinctive kinetics of Eg5 [66].

2.2.2.4 Head-head coordination and gating

The processive movement of an Eg5 dimer requires "hand-over-hand" coordination between its two motor domains. This coordination is ensured by a mechanism known as "gating," which prevents both heads from detaching from the microtubule simultaneously. This mechanism varies among members of the kinesin family and is critical for the motor's biological function [156].

Kinesin-1 exhibits a "tight" front-head gating mechanism, in which strain on the leading head prevents it from detaching until the trailing head has completed its own cycle. This tight control is the basis for Kinesin-1's high processivity. In contrast, Eg5 has a more flexible or "lax" gating mechanism. The communication between the heads is not as tightly controlled as in Kinesin-1. As a consequence, it is a common event for both heads of Eg5 to simultaneously enter the weakly-binding ADP state, leading to frequent detachment of the motor and resulting in low processivity. This lax coordination also explains why the stall force of a single Eg5 dimer (~2 pN) is much lower than that of Kinesin-1 (~6–7 pN), as the motor tends to detach easily under load [65, 87].

This "lax gating" mechanism of Eg5 is a functional adaptation for working in an ensemble in a high-load environment. This mechanism prioritizes force distribution and dynamic rearrangement over the long-distance fidelity of a single motor. The mitotic spindle is an environment where hundreds of Eg5 motors cross-link microtubules and work against opposing forces from motors like dynein and Kinesin-14. In this "tug-of-war", if a motor were too tightly processive like Kinesin-1, it could become permanently stalled or even break under excessive load, turning into a static obstacle. Eg5's lax gating mechanism, however, acts as a "mechanical fuse." It allows

an individual motor to detach easily when the local force on it exceeds its low stall force. This releases the tension, which can then be absorbed by the dozens of other motors in the ensemble. The detached motor is then free to re-bind elsewhere and contribute to force production. In effect, lax gating transforms Eg5 from a rigid machine into a dynamic component of a resilient, force-distributing network. It is a strategy of "letting go to hold on stronger" at the ensemble level [100, 157].

2.2.2.5 Tetrameric organization and antiparallel shift geometry

One of the most distinguishing features of Eg5 is its tetrameric structure, formed by the antiparallel interaction of two dimeric units, which is fundamental to the motor's function. This higher-order organization is mediated by the Bipolar Assembly (BASS) domain in the center of the stalk. The BASS domain is a novel, torsionally stiff, four-helix bundle structure approximately 26 nm in length. This structure positions the two motor dimers with a rotational offset of approximately 90–100 degrees relative to each other. This specific geometry is perfectly designed for Eg5 to simultaneously cross-link two antiparallel microtubules and "walk" toward the plus-ends of both concurrently. This movement generates a pushing force that drives the microtubules apart [100, 146, 158]. The full-length tetramer, including the tail domain, has been shown to be necessary for efficient and orderly sliding motion. Constructs lacking the tail domain have been observed to lead to inefficient "scissoring" motions rather than effective sliding. This highlights the critical role of the full tetrameric structure in producing effective sliding forces at the molecular level [158].

The BASS domain functions as a "molecular jig" that precisely orients the two motor pairs in three-dimensional space. This pre-organization solves the geometric problem of binding to two separate and antiparallel filaments and converts the simple linear motion of the motor domains into a powerful outward sliding force. Without this "jig," the motor dimers would be floppy and would struggle to form a stable, force-transducing cross-link. The torsional stiffness of the BASS domain is also critical. It ensures that the force generated by one motor dimer is efficiently transmitted along the stalk to the other microtubule, rather than being lost to bending or twisting. Thus, the BASS domain is the key innovation that elevates Eg5 from a simple dimeric motor to a specialized machine for generating separation forces [67].

2.2.2.6 Ensemble kinetics and spindle forces

Within the mitotic spindle, hundreds of Eg5 motors work collectively. The ensemble kinetics of these motors and their concerted action, which differ from the dynamics of a single motor, give rise to the macroscopic forces required for mitosis. Although an individual Eg5 dimer is not highly processive, a multitude of Eg5 molecules can provide a continuous sliding motion by successively binding to the microtubule tracks. Studies in Xenopus egg extracts have shown that the poleward flux of microtubules in the spindle is driven by Eg5 ensembles, and pharmacological inhibition of Eg5 reduces or halts this flux [159]. While an individual Eg5 motor produces a modest force of approximately 1.5–2 pN, the total force increases when multiple motors act on the same microtubule pair, and it has been suggested that this force is nearly a linear sum of the individual forces produced by the motors. This collective force is responsible for maintaining the distance between the spindle poles and for the alignment of chromosomes. Load sharing among the motor ensembles is a critical feature; if one motor detaches, the others take on the load, ensuring continuous force production and the robustness of the system. This "relay race" model, facilitated by Eg5's flexible gating mechanism, is vital for the stability of collective motor actions in the cell [90, 160]. Within the spindle, a "push-pull equilibrium" exists between plus-end-directed motors like Eg5 and minus-end-directed motors such as dynein and Kinesin-14. Eg5 provides an outward-pushing force for spindle elongation, which is balanced by inward-pulling forces to maintain a steady-state spindle length during metaphase. Eg5 ensembles can also exhibit a "braking" behavior. If microtubules are moved too rapidly by other forces, Eg5 motors, due to their slow stepping rate and tendency to remain strongly bound, generate a viscous drag that resists and dampens the movement, preventing excessively rapid spindle dynamics. This property allows the Eg5 ensemble to act as a dynamic regulator of spindle length and stability [71, 161, 162].

In this context, the Eg5 ensemble functions not just as a simple motor, but as a "dynamic rheostat" for the mitotic spindle. It can generate pushing force when needed, but it can also passively resist and dampen motion to prevent the spindle from over-extending or collapsing. This dual capability makes the Eg5 ensemble a sophisticated regulator. It creates a self-regulating system that maintains the mechanical homeostasis of the spindle, much like a rheostat controls electrical current.

2.2.2.7 Microtubule end interactions

Recent research has revealed that Eg5 may have unexpected interactions with microtubule plus-ends (+ ends), expanding its functional spectrum beyond its classic role as a midzone cross-linker. Studies have shown that Eg5 motor domains can act as a "polymerase" by promoting polymerization at microtubule plus-ends, independent of motor activity. This effect is thought to occur because Eg5 preferentially recognizes and straightens the curved tubulin protofilaments at the growing microtubule tip, thereby accelerating filament elongation. This function contributes to the control of spindle length by coupling microtubule sliding and elongation [151, 159, 161, 162].

Furthermore, the observed localization of Eg5 at free microtubule + ends, likely mediated by the adaptor protein TPX2, suggests potential interactions with + TIPs (+End Tracking Proteins) such as EB1. This would allow Eg5 to be loaded onto newly polymerized microtubule regions, enabling it to immediately perform its functions of promoting sliding and polymerization [159]. This dual function of Eg5 as both a "slider" and a "builder" represents a powerful positive feedback loop for spindle assembly. By simultaneously extending the tracks on which it slides, Eg5 ensures the creation of robust, overlapping microtubule arrays, which are the substrate for its own force-generating activity. In the spindle midzone, the plus-ends of antiparallel microtubules interdigitate. If Eg5 motors in this region not only slide the existing overlaps but also promote further growth of the plus-ends, they actively increase the length of the overlap zone. A longer overlap zone allows more Eg5 motors to bind and generate force. This synergistic mechanism makes spindle assembly much more efficient and robust. Consequently, Eg5 is a multifaceted actor that not only crosslinks microtubules but also influences +end dynamics, thereby regulating spindle size, flux, and overall stability.

2.2.2.8 Multilayered Regulation

The activity of the Eg5 motor protein is tightly regulated by its intrinsic properties, interdomain interactions, post-translational modifications, and auxiliary proteins to ensure it functions at the correct places and times during mitosis. This multilayered regulatory network harnesses the powerful capabilities of Eg5 for precise mitotic functions. These regulatory mechanisms are summarized in Table 2.4.

Table 2.4: Summary of Eg5 regulatory factors and their effects.

Regulatory Factor	Site of Action/Interaction on Eg5	Key Effect on Eg5 Activity/Localization	Relevant Phosphorylation Sites (if any)
Tail Domain	Motor domain (specifically α0 helix)	Auto-regulation/auto-inhibition: slows ATP binding rate, prolongs microtubule attachment time (resulting in slower but more forceful steps), influences spindle localization and tetrameric function; acts as an "activity switch" [158].	The BimC box within the tail is phosphorylated by Cdk1 (e.g., Thr-926), which can relieve auto-inhibition [163].
Cdk1 (Cyclin-dependent kinase 1)	Tail domain (BimC box)	General activation of Eg5 at the onset of mitosis, likely by alleviating tail-mediated auto-inhibition; facilitates recruitment to the spindle and localization to the midzone [73].	Thr-926 (Human Eg5) [73].
Nek6 Kinase	A site located near the C-terminus of the motor domain	Phosphorylates a small fraction of the Eg5 pool, predominantly localized at the spindle poles; this phosphorylation is critical for Eg5's centrosome separation function [163].	Ser-1033 (Human Eg5) [163].
TPX2	Eg5 motor domain and Microtubules	Can directly suppress Eg5 motor activity ("braking" effect) and tether Eg5 on microtubules, preventing excessive spindle elongation; facilitates recruitment and proper localization of Eg5 to the spindle (e.g., polar regions) [159].	The Eg5–TPX2 interaction can be modulated by Nek9 kinase phosphorylation of TPX2 [159].

The C-terminal tail domain contains a conserved BimC box and plays a significant role in the auto-regulation of Eg5 activity. Cryo-electron microscopy (Cryo-EM) studies have shown that in the ADP or nucleotide-free state, the tail domain can fold back and bind to the motor domain's α0-helix. This interaction stabilizes the motor in an "open" conformation that slows ATP binding, thereby reducing motor speed but increasing the duration of strong, dual-headed binding to microtubules. This auto-inhibitory mechanism enables the full-length Eg5 to generate greater force and ensure robust, unidirectional microtubule sliding, a function that is lost in tailless constructs [158].

This auto-inhibition is dynamically controlled by post-translational modifications (PTMs). At the onset of mitosis, the Cdk1/cyclin B complex phosphorylates a conserved threonine residue (Thr926 in human Eg5, Thr937 in *Xenopus*) located within the BimC box of the tail domain. This phosphorylation does not affect the motor's velocity but significantly increases its binding affinity for microtubules. This likely activates the Eg5 motor pool by relieving tail-mediated auto-inhibition and ensures that a sufficient amount of motor is loaded onto the spindle. Additionally, the Nek6 kinase phosphorylates a small and specific pool of Eg5 at Ser1033, which is

predominantly found at the spindle poles. This localized phosphorylation is critical for the separation of centrosomes during early spindle formation [73, 80, 158, 163].

Finally, auxiliary proteins such as the spindle assembly factor TPX2 provide local fine-tuning by directing Eg5 to spindle microtubules and simultaneously suppressing its motor activity. This interaction acts as a "brake" that prevents excessive spindle elongation by tethering Eg5 and reducing its velocity. This multilayered regulatory network demonstrates that Eg5 operates not like an "on/off" switch, but like a highly tunable "rheostat" with multiple "gears." The tail domain provides an intrinsic "low gear" (slow, high force), while Cdk1 phosphorylation acts as a "clutch" that engages the motor with the track (the spindle) at the onset of mitosis. This is a sophisticated system that provides precise control over a powerful motor [73, 159].

2.2.2.9 Mechanistic Comparison of Eg5 with Other Kinesins

The kinesin superfamily comprises at least 14 subfamilies adapted for diverse cellular functions, leading to significant differences in motor characteristics. Eg5 (kinesin-5), specialized for mitosis, exhibits a unique profile when compared to other kinesins. This comparison also highlights why Eg5 is a specific and attractive target for cancer therapy, as inhibiting its specialized mitotic function is distinct from disrupting the vital long-range transport roles of other kinesins, which may lead to reduced neurological side effects in patients. Table 2.5 compares the motility characteristics of selected kinesin motors.

Table 2.5: Comparison of motility characteristics of selected kinesin motors.

Kinesin Motor (Family)	Average Speed (µm/s)	Stall Force (pN)	Processivity (run length)
Kinesin-1 (Conventional) – e.g., KIF5B	~0.5–0.8 µm/s (fast)	~6 pN (high)	High processivity: ~1–2 µm (some up to several µm) [152]
Kinesin-3 – e.g., KIF1A (neuronal vesicle motor)	~1–2 µm/s (very fast)	~1 pN (low) (estimated)	Super-processive: ~10 µm or more (thousands of steps) [149]
Kinesin-5 (Eg5)	~0.02–0.1 µm/s (very slow) *	~2 pN (medium-low)	Low processivity: ~0.1–0.3 µm (only a few steps) [148]
Kinesin-14 – e.g., Ncd/HSET (mitotic, minus-end motor)	~0.1–0.3 µm/s (slow) (estimated)	<1 pN (very low)	Non-processive: <0.1 µm (detaches after a few steps) [149]

Note: Although the speed of Eg5 at the single-compound level (as a free dimer) is measured to be approximately ~0.1 µm/s, full-length Eg5 motors under antiparallel microtubule sliding conditions translocate at extremely low average speeds, such as

0.005–0.01 μm/s. This suggests a strategy in which Eg5 sacrifices its speed to augment its propulsive force [145].

The evolved kinetic differences among kinesin families are contingent on their specialized functions. Cargo-transporting motors, such as Kinesin-1 and Kinesin-3, prioritize high speed and long-distance displacement, exhibiting more "athletic" behavior. In contrast, spindle organizers like Eg5 and the opposing minus-end-directed Kinesin-14 prioritize high force generation and multi-motor synergy, exhibiting more of a "collective team play." Therefore, a direct comparison of Eg5 with a motor like Kinesin-1 is akin to comparing a freight truck with a racing motorcycle; one is optimized for power and torque, the other for speed and maneuverability. Structural differences also reflect these functional adaptations; for instance, Kinesin-3 motors possess features like a "K-loop" for tight microtubule adhesion, whereas Eg5 is specialized with its unique L5 loop and tetramerization domains to maximize force production in an ensemble [82, 152].

In conclusion, Eg5 occupies a unique position with respect to its motility parameters, consistent with its cellular role as a slow but robust force generator within multi-motor ensembles. These comparisons show that the kinesin superfamily is a "molecular toolkit" that has been evolutionarily optimized for different physical tasks. This reveals a design principle where Kinesin-1 is optimized for speed and endurance, while Eg5 is optimized for collective force and stability.

2.2.2.10 The mechanism of Eg5 inhibition and its cellular consequences

A detailed understanding of Eg5's structure and mechanochemical cycle provides a clear rationale for why its inhibition leads to potent antimitotic effects. The primary strategy for selective Eg5 inhibition has focused on allosteric sites, particularly the Loop L5 pocket, which offers a specificity that cannot be attained by targeting the highly conserved ATP-binding pocket.

Inhibitors that bind to the L5 pocket, such as the archetypal compound monastrol, function by "jamming" the motor's mechanical cycle. As previously detailed, the release of ADP from the nucleotide-binding pocket is a rate-limiting step that is tightly coupled to conformational changes in the L5 Loop and Switch I/II regions. The binding of an allosteric inhibitor to the pocket formed by helices [121].

α2/α3 and the L5 Loop stabilizes the motor domain in a specific ADP-bound conformation. This binding event prevents the necessary conformational changes, effectively trapping ADP in the active site and blocking the binding of a new ATP molecule. Without the ability to bind and hydrolyze ATP, the Eg5 motor is rendered powerless and cannot generate the force required to "walk" along microtubules [66, 71].

This molecular paralysis has profound consequences at the cellular level. An Eg5 tetramer with its motor domains locked in an inactive, ADP-bound state can still cross-link antiparallel microtubules but cannot produce the outward sliding force necessary to push the spindle poles apart. Consequently, in proliferating cells entering mitosis, the duplicated centrosomes fail to separate. Instead, the microtubules emanating from both centrosomes are pulled together into a single aster-like structure, with the chromosomes arranged radially around them. This characteristic phenotype is known as a "monopolar spindle" or "monoaster" [71, 162].

The formation of a monopolar spindle triggers a powerful cellular surveillance mechanism known as the Spindle Assembly Checkpoint (SAC). The SAC ensures the fidelity of cell division by monitoring the proper attachment of chromosomes to bipolar spindles. In the presence of a monopolar spindle, chromosomes cannot achieve bipolar attachment and tension, leading to sustained SAC activation. This activation arrests the cell during mitosis, preventing it from proceeding to anaphase. Prolonged mitotic arrest is often an unsustainable state for many cancer cells, which are highly dependent on rapid proliferation. This condition ultimately triggers the intrinsic apoptotic pathway, leading to programmed cell death [71, 162].

This cascade of events—from the allosteric locking of the motor domain to monopolar spindle formation, SAC activation, and subsequent apoptosis—forms the mechanistic basis for the therapeutic potential of Eg5 inhibitors. Eg5 inhibition does not simply block an enzyme; it converts a dynamic motor into a static and "poisonous" clamp. The inhibited Eg5 tetramer becomes a dominant-negative saboteur of the spindle structure. Instead of becoming inert and diffusing away, the inhibited motors remain cross-linked to antiparallel microtubules. These paralyzed motors now act as static, non-functional cross-links, physically preventing any remaining active Eg5 motors from sliding the microtubules apart. This creates a dominant-negative effect where a relatively small number of inhibited motors can halt the function of the entire

ensemble. The outward pushing force collapses, and the inward pulling forces from other motors (like dynein) prevail, causing the spindle to collapse into a monoaster. This is a more potent mechanism than simple competitive inhibition of an enzyme and leads to an easily detectable cellular phenotype (the monoaster) that reliably triggers cell death pathways. This explains why, despite clinical setbacks, Eg5 remains such an attractive therapeutic target.

While this general pathway is well-established, significant gaps in the literature persist. The precise molecular determinants that make certain cancer cells more susceptible to this apoptotic fate, while allowing others to escape via mitotic slippage, are not fully understood. Furthermore, despite the discovery of many L5 pocket inhibitors, challenges related to potency, specificity, and the development of resistance continue. This thesis addresses this need by identifying novel chemical scaffolds for the L5 pocket. In the subsequent sections, a general assessment of the inhibitors developed to date will be presented, providing a context for the novel compounds investigated in this study.

2.2.3 Eg5 inhibitors

Considering the comprehensive analysis presented in earlier sections regarding the mechanism of mitotic division, which is fundamental to the unchecked growth of cancer cells (Section 2.1), along with the critical and distinctive function of Eg5, especially in the establishment of the bipolar mitotic spindle, it is clear that Eg5 represents a targeted and promising avenue for cancer treatment. Stopping the cell cycle during mitosis and disrupting its typical progression are considered one of the most effective strategies in the fight against cancer currently in use. The essential role of Eg5, particularly in the context of the survival of rapidly proliferating cancer cells as opposed to normal cells, offers a compelling scientific basis for the development of therapeutic agents aimed at targeting this protein.

Indeed, conventional antimitotics that target microtubule dynamics, such as taxanes and vinca alkaloids, which are widely used in clinical practice, exhibit significant limitations despite their efficacy [2, 164]. These classical drugs can lead to severe side effects, most notably severe peripheral neuropathies and myelosuppression, which are dose-limiting toxicities because microtubules are essential not only in the mitotic spindle but also in non-dividing normal cells (e.g., neurons) [2, 165]. Additionally, the

emergence of resistance, toxicity, and mitotic slippage significantly limit the efficacy of these microtubule-targeting agents [165]. Therefore, in the pursuit of addressing certain limitations associated with MTAs, components involved in mitosis and the Spindle Assembly Checkpoint (SAC) have attracted significant attention as potential microtubule-independent targets. This interest has led to a novel category of pharmaceuticals known as second-generation antimitotics (SGAs) [166]. Alternatively, Eg5 inhibitors have attracted intensive research interest as less toxic mitotic inhibitors. The development of inhibitors directed against molecular targets such as Eg5, which have more specific roles in the mitotic process, has emerged as a significant need, offering the prospect of greater selectivity, reduced toxicity, and potentially overcoming resistance mechanisms in cancer treatment [123, 167].

In line with these expectations, the discovery and development of Eg5 inhibitors gained substantial momentum towards the end of the twentieth century (particularly as scientists recognized the importance of selective KSP inhibitors in the early 2000s). The first groundbreaking development in this field was the discovery of Monastrol in 1999 by Mayer and colleagues [4]. Identified through a phenotype-based high-throughput screening (HTS) method, Monastrol was found to selectively inhibit the ATPase activity of the Eg5 motor protein via an allosteric mechanism (specifically by blocking the helix α2/loop L5/helix α3 region) and induced the formation of characteristic monopolar spindles (monoasters) in cells [168]. The discovery of Monastrol proved that Eg5 is a 'druggable' cancer target and sparked intense interest in both academia and industry for the identification of novel small compounds targeting Eg5 [4]. Following Monastrol, S-trityl-L-cysteine (STLC) [169, 170], which belongs to a different chemical class and also binds to the L5 allosteric pocket to inhibit Eg5 with higher potency, further reinforcing the potential of this target.

Encouraging early stage preclinical findings has led to the engagement of important pharmaceutical companies. By the mid-2000s, first-generation potent Eg5 inhibitors, including Ispinesib (SB-715992), had been developed and progressed to Phase I/II clinical trial stages [171]. Among the approximately ten Eg5 inhibitors that underwent clinical evaluation, Ispinesib and Filanesib (ARRY-520) emerged as the most promising small compounds against the KSP protein in clinical trials [171–173]. These initial clinical studies revealed that Eg5 inhibition is, in principle, a viable strategy in humans and can exhibit some signs of antitumor activity. Nonetheless, the clinical

effectiveness of most of these inhibitors has been found to be inadequate, and they typically do not advance beyond Phase II trials [172]. The primary challenges involved the common occurrence of neutropenia, which serves as a dose-limiting toxicity, along with the inability to reach the anticipated objective response rates in numerous solid tumors [174].

These initial clinical experiences did not dampen interest; instead, they motivated efforts to discover compounds with improved PK/PD profiles, alternative binding modes and resistance-bypass potential, and to explore combination regimens (e.g., ispinesib + docetaxel; filanesib + bortezomib/dexamethasone) as well as newer modalities such as antibody–drug conjugates and gene-silencing approaches [172–175]. The shortcomings of first-generation agents therefore define a clear, unmet need for next-generation Eg5 inhibitors. The present thesis addresses this gap by applying advanced in-silico techniques to mine large chemical spaces for fresh scaffolds that target loop L5, with the aim of delivering candidates possessing superior potency, selectivity and pharmacological properties.

The following sections first detail the specific chemical classes of inhibitors developed to date, providing a comprehensive context against which the novelty and potential of the compounds discovered in our work can be measured. These topics will be examined in detail in the subsequent sections.

2.2.3.1 Main Eg5 inhibitor classes: Their structures, binding mechanisms, and key SAR findings

Small-compound inhibitors targeting the Eg5 motor protein, whose historical development process is summarized in Section 2.2.3, exhibit considerable diversity in terms of their chemical structures and mechanisms of action on Eg5. In this section, the inhibitors are first classified according to their primary binding sites on the Eg5 motor domain (allosteric pockets or orthosteric ATP pocket). Subsequently, the predominant chemical classes that targeted these regions were examined in detail. For each class, the structural characteristics of prototype compounds, their molecular interactions with Eg5, key structure-activity relationship (SAR) findings derived from these interactions, and their potential therapeutic significance will be addressed in light of the current literature, particularly the recently published studies provided for this work. It was observed that the vast majority of the most potent and selective Eg5

inhibitors developed to date target specific allosteric sites distinct from the ATP-binding pocket of the protein; therefore, these allosteric pockets and the ligands that target them will constitute the main focus of this section. Notably, the comprehensive molecular dynamics simulations and binding energy analyses presented by Jia et al. (2024) offer a valuable perspective for understanding the binding patterns of different inhibitors to the L5 and α4/α6 pockets of Eg5 [135].

Loop L5 allosteric pocket-targeted inhibitors

The L5 pocket, which is the most extensively studied allosteric site on the Eg5 motor domain, is located at the interface formed by helix α2b, loop L5, and helix α3, as previously detailed in the "Structural Biology" section. Owing to its unique structural characteristics specific to Eg5 and its essential role in regulating motor function, this pocket has emerged as the main focus for the development of selective inhibitors. Inhibitors that interact with this region typically influence Eg5's nucleotide cycle by preventing its normal function, thereby stabilizing the motor in an inactive conformation that is bound to ADP. The following discussion focuses on the main chemical classes of inhibitors that target the L5 allosteric pocket [132].

Dihydropyrimidine derivatives (DHPMs)

Investigation and Initial Compound Development (Monastrol): Monastrol, recognized as the prototype of the dihydropyrimidinone class, stands out as the first selective, cell-permeable Eg5 inhibitor identified in 1999 by Mayer and colleagues. This discovery emerged from a phenotype-based high-throughput screen (HTS) designed to identify small compounds that interfere with mitotic spindle bipolarity [4]. Its chemical structure incorporated a 4-aryl-3,4-dihydropyrimidine-2(1H)-thione core (Figure 2.23a). It suppresses Eg5 ATPase activity at micromolar concentrations: the average IC_{50} against the basal enzyme is reported as ≈ 30 μM, and between 25 and 40 μM in the microtubule-stimulated state, leading to the characteristic monopolar spindle (monoaster) phenotype by preventing centrosome separation in cells [5]. This discovery established Eg5 as a "druggable" cancer target by revealing that it possesses an allosteric site suitable for selective small-compound inhibitors, distinct from the ATP-binding pocket, thereby igniting intense interest in both academic and industrial research [166].

Binding Mode and Mechanism: Monastrol and its derivatives interact with the allosteric pocket, which were referred to as the L5 pocket, created by helix α2, loop L5, and helix α3 within the motor domain of Eg5. Crystal structure studies, particularly of Aspergillus nidulans BimC (Eg5 homolog) complexed with Monastrol (PDB:1X88) and human Eg5 complexed with various Monastrol, have shown that these compounds settle into a hydrophobic niche immediately beneath loop L5 [7, 118, 176]. The binding of the inhibitor induces a movement in loop L5 towards the inhibitor, resulting in the adoption of a "closed" conformation. In this binding scenario, the dihydropyrimidine thione ring of monastrol typically forms hydrogen bonds with Glu116. Meanwhile, the substituents present on the phenyl ring, including the hydroxyl group, interact with residues such as Asp130 located in loop L5. Furthermore, the remaining portions of the compound engaged in van der Waals interactions with hydrophobic residues located within the pocket, such as Arg119, Ala133, Pro137, Leu214, and Ala218 (Figure 2.24) [117, 135]. The allosteric effect notably impedes the release of ADP from the catalytic pocket of Eg5, consequently stabilizing the motor in a state bound to ADP that is unable to participate in ATP hydrolysis. This inhibition disrupts the ATP hydrolysis cycle, thereby diminishing the motor's capacity to exert force on the microtubules [7].

Structure-Activity Relationship (SAR) and Analog Development: Following Monastrol's discovery, extensive SAR studies have been conducted to improve its weak micromolar activity and "non-drug-like" properties. These studies revealed that modifications to the dihydropyrimidinone core, particularly the phenyl ring, had significant effects on activity. Enastron (Figure 2.23b) and Dimethylenastron (Figure 2.23c) were developed through modifications such as cyclization of Monastrol's ester side chain (formation of a cyclic ketone in dimethylenastron) and/or replacement of the thione group with oxygen. These modifications improved the rigidity of the compound, allowing for a better fit within the L5 pocket and encouraging further interactions with residues such as Tyr211. Consequently, Enastron (IC_{50} ~2 μM) and especially dimethylenastron (IC_{50} ~110-200 nM) demonstrated approximately 15-fold and as much as ~150-fold increased potency in comparison to Monastrol [4, 7]. Halogenated analogs, including (R)-fluorastrol, which were created by incorporating fluorine atoms into the phenyl ring, demonstrated significant activity (Figure 2.23e) . The phenyl extension of (R)-fluorastrol has been shown to engage with a hydrophobic

region of the L5 pocket, establishing further pi-interactions with residues such as Glu116, Leu214, Gly217, and Arg221 [33]. Research on structure-activity relationships has demonstrated that the S-enantiomers within this category of inhibitors generally display enhanced activity relative to the R-enantiomers, especially in the instances of Monastrol and Class I DHPMs. Nevertheless, certain cases demonstrate a preference for the R-configuration, as evidenced by Class II DHPMs, such as Mon-97 (Figure 2.23d) and Fluorastrol [7, 33]. The observed stereoselectivity can be explained by chiral interactions occurring within the L5 pocket. The results reported by Jia et al. (2024) revealed that fluorastrol (5) exhibited the lowest binding free energy when compared to the other compounds examined in this study [135].

Clinical Development: Although the dihydropyrimidinone class provided proof-of-concept for targeting Eg5 and demonstrated the "druggability" of the L5 pocket, no compound from this class has directly progressed to clinical trials. Monastrol and its optimized analogs remain valuable chemical tools for studying Eg5 function and inhibition. Suboptimal pharmacokinetic properties (e.g., metabolic stability and solubility) may have played a role in this outcome [167, 177].

The journey of the Monastrol class serves as a crucial benchmark for this thesis. This establishes the validity of targeting the L5 pocket and defines the key chemical interactions required for inhibition. However, the limitations of this class of drugs, particularly their modest potency and poor drug-like properties, underscore the need for novel chemical scaffolds. Our computational screening efforts aim to identify compounds that retain the successful binding paradigm of monastrol (e.g., key hydrogen bonds) but are built upon entirely new, potentially more optimizable molecular frameworks, thereby addressing the deficiencies that prevented this foundational class from reaching the clinic.

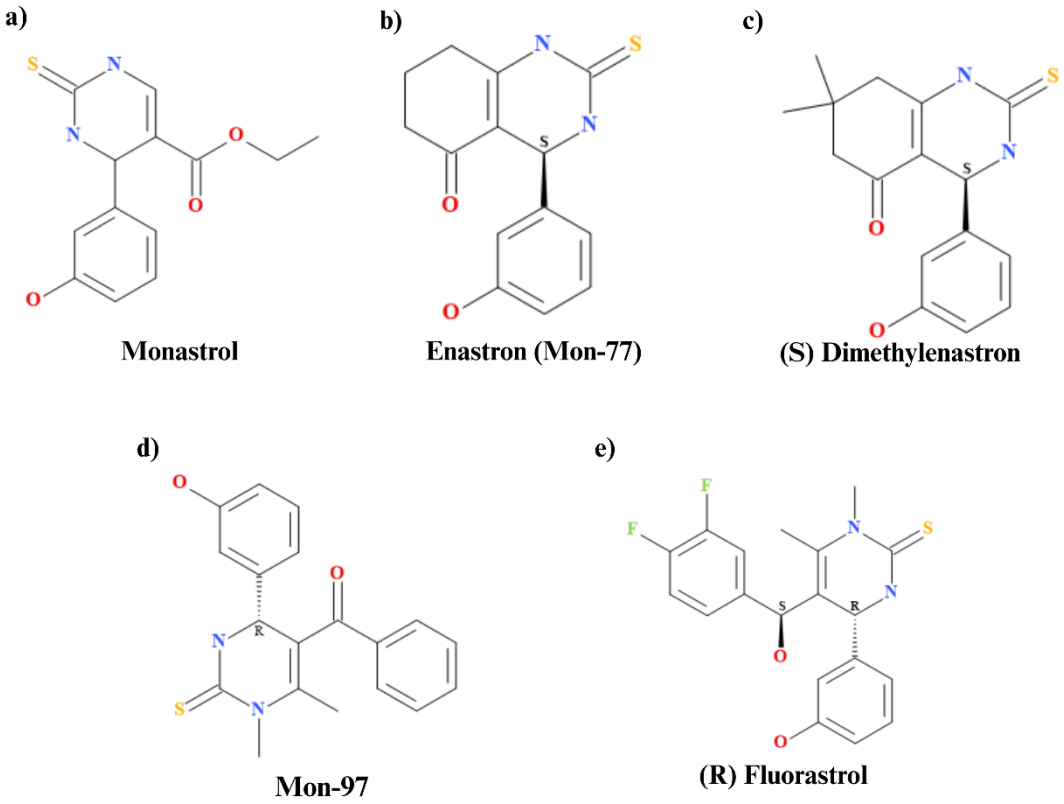

Figure 2.23 : 2D chemical structures of dihydropyrimidine derivatives.

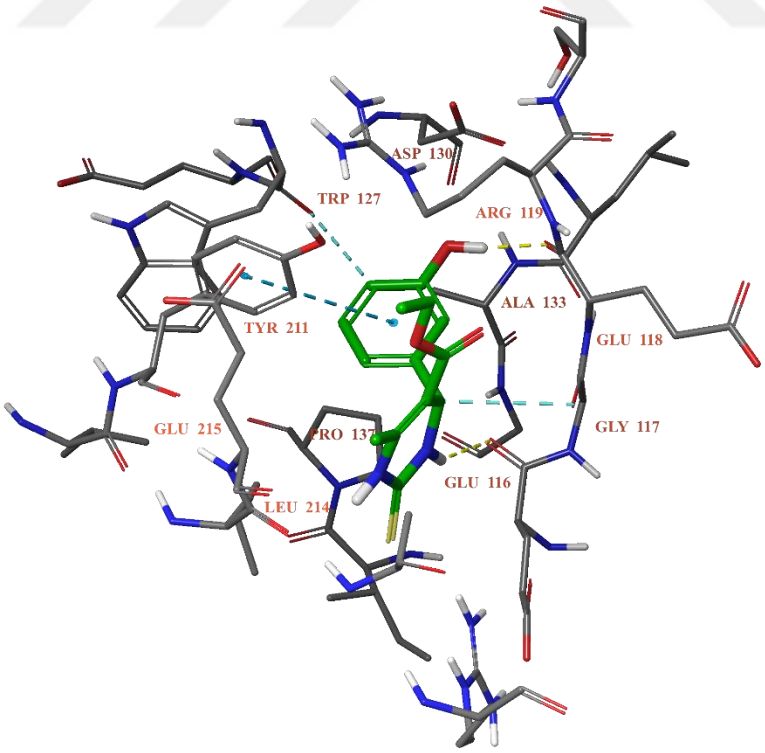

Figure 2.24 : Binding pose of monastrol in the allosteric Loop L5 pocket of Eg5 (PDB: 1X88).

Trityl Amino Acid Derivatives

Discovery and Prototype Compound (S-Trityl-L-cysteine - STLC): An important development that followed the discovery of Monastrol, during the investigation of different chemical scaffolds targeting Eg5, was the recognition of S-trityl-L-cysteine (STLC) as an antimitotic agent within the NCI-60 panel, as reported by Paull et al. in 1992, with a GI_{50} of approximately 1.3 μM [178]. In 2004, Brier et al. demonstrated that STLC serves as a specific and high-affinity inhibitor of Eg5, indicating that the L5 allosteric pocket is receptive to chemical diversity. STLC (Figure 2.25a) inhibits Eg5 ATPase activity with an IC_{50} of approximately 40 nM under basal conditions and approximately 140 nM when stimulated by microtubules, resulting in the distinct monoaster phenotype observed in cells [169]. According to Shahin and Aljamal (2022), STLC demonstrates a greater capacity to inhibit mitosis than monastrol, leading to a more significant reduction in the viability of MCF-7 and SK-BR-3 breast cancer cells, as well as several glioblastoma cell lines [33].

Binding Mode and Mechanism: Like Monastrol, STLC interacts with the L5 allosteric pocket of Eg5. Investigations into crystal structures, particularly those analyzing human Eg5 in complex with STLC (PDB: 2WOG, 3KEN, 5X70) [7], have demonstrated that STLC is situated within the L5 pocket. The prominent trityl group engages in crucial van der Waals and π–π interactions with the hydrophobic sections of the pocket, including residues such as Trp127 in loop L5 and Tyr211 in helix α3 [7, 179]. At the N-terminus of loop L5, interacting with residues such as Glu116 and Gly117, the cysteine component establishes hydrogen bonds with more polar sections of the pocket. The interaction of STLC enables loop L5 to wrap around the inhibitor, hence generating a "closed" conformer. Similar to Monastrol, this compound immobilizes the motor in a state bound to ADP by preventing ADP release from the catalytic pocket of Eg5. This action disrupts the mechanochemical cycle by preventing the motor from progressing via ATP hydrolysis. Studies have shown that the interaction between STLC and Eg5 shows notable binding affinity and is reversible [170]. Garcia-Saez and Skoufias (2021) find that the binding of L5 pocket inhibitors, including STLC, causes substantial conformational changes in the Switch II cluster and the neck-linker region, thereby stabilizing the neck-linker in a docked position on the motor domain [121].

Structure-Activity Relationship (SAR) and Analog Development: Following its identification, STLC became the subject of extensive structure-activity relationship (SAR) studies. The primary goal of these investigations was to improve upon the compound's initial profile, which, despite its potency, was hampered by suboptimal physicochemical properties and poor cellular permeability. These medicinal chemistry efforts have established several foundational principles for this scaffold. The trityl group is considered indispensable for activity, and potency can often be augmented by substitutions at the para-position of its phenyl rings. For example, (S)-methoxy-trityl-L-cysteine (MeO-STLC) (Figure 2.25b), which incorporates a para-methoxy group, exhibits a nearly fivefold increase in inhibitory activity ($IC_{50} \approx 28.5$ nM) compared to the parent molecule [180]. Conversely, the core structure demonstrates low tolerance for change; isosteric replacement of the central sulfur atom is not permitted, and while the L-cysteine moiety is preferred, the free amino acid group is essential for activity. To circumvent pharmacological limitations and enhance tumor selectivity, a GGT-activated prodrug strategy has been explored. This innovative approach involves modifying the amino acid portion of STLC to mimic glutathione, allowing for targeted release of the active compound in the tumor microenvironment, where GGT is frequently overexpressed [181].

Clinical Development: STLC and its optimized analogs, similar to the Monastrol class, have shown considerable effectiveness as chemical tools for elucidating the mechanism of Eg5 inhibition and investigating the properties of the L5 pocket. Although their significant effectiveness has been highlighted in reviews [22], none of the STLC derivatives have progressed to clinical trials thus far. This could be associated with potential pharmacokinetic issues, such as inadequate physicochemical properties, limited cellular permeability, and metabolic stability, or may involve off-target interactions [33].

The success of the STLC class in achieving nanomolar potency, a significant improvement over the monastrol class, further validated the L5 pocket as a prime target. However, the failure of this class to enter clinical trials, likely due to ADMET issues associated with the bulky and lipophilic trityl group, highlights a critical challenge: achieving high potency without sacrificing the drug-like properties. This thesis aimed to bridge this gap by searching for novel scaffolds that are structurally distinct from both dihydropyrimidinones and trityl amino acids. The goal was to

identify compounds that can replicate the high-affinity interactions of the STLC class, but with more favorable physicochemical profiles, making them better candidates for future therapeutic development.

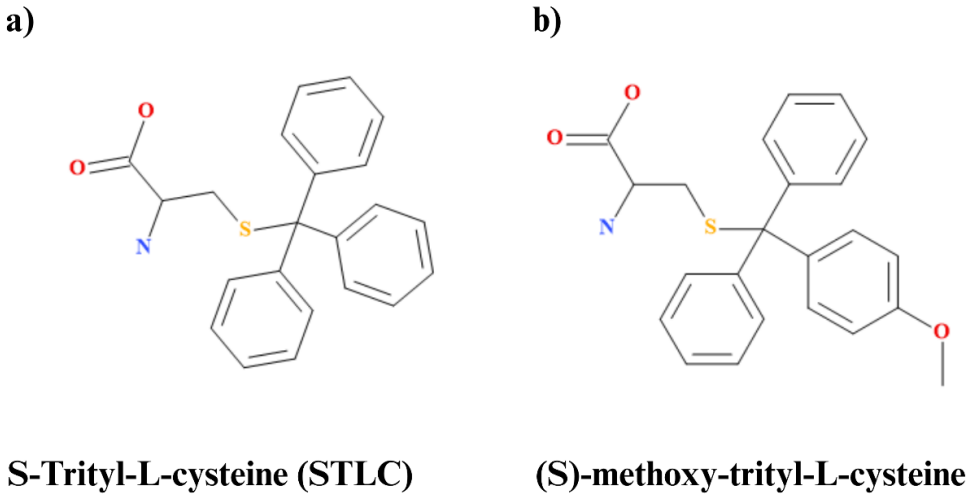

Figure 2.25 : Chemical structures of trityl amino-acid–based Eg5 inhibitors.

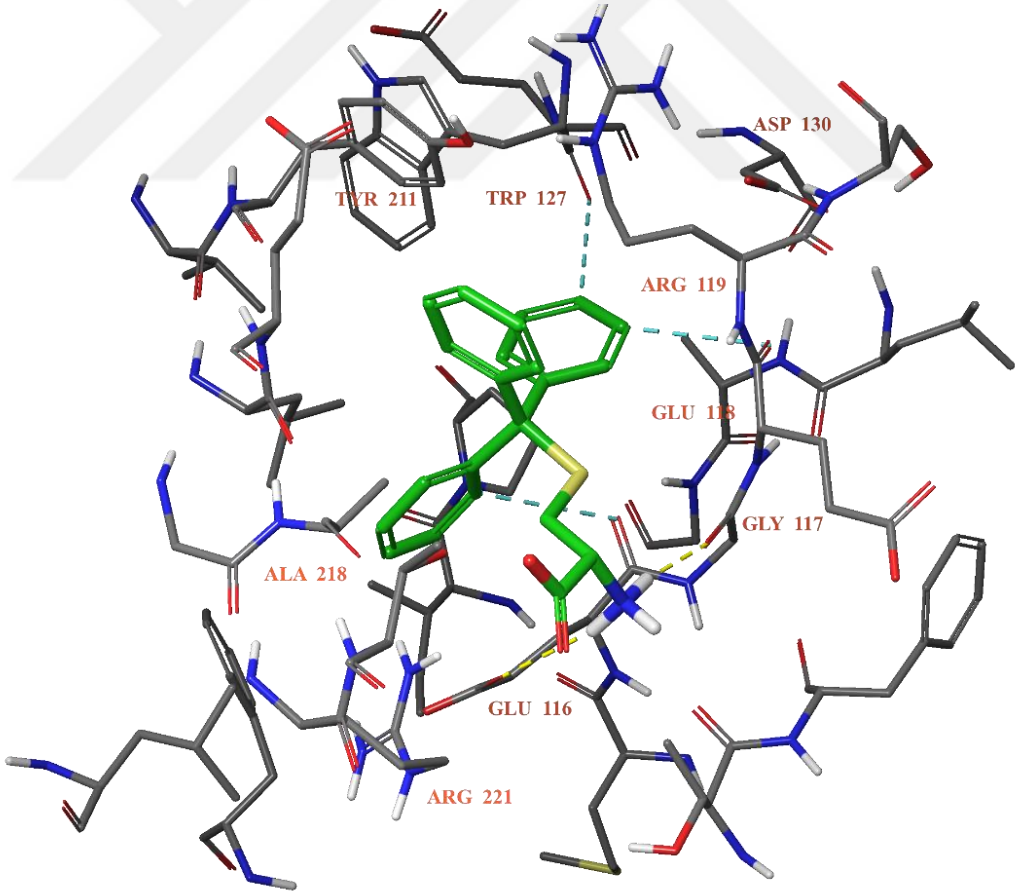

Figure 2.26 : Binding pose of S-trityl-L-cysteine (STLC) in the allosteric Loop L5 pocket of Eg5 (PDB: 2WOG).

Thiadiazoles and Related Heterocycles (including K858)

Overview and Prototype Compound (K858): Compounds derived from sulfur- and nitrogen-containing heterocyclic ring systems, particularly thiazoles and 1,3,4-thiadiazoles, represent another major class of potent L5 pocket inhibitors. A prominent example is K858, first reported by Nakai et al. in 2009 [182]. K858 (Figure 2.27a) features a 2,4,5-trisubstituted 1,3,4-thiadiazole core and inhibits Eg5 ATPase activity in the low micromolar range ($IC_{50} \approx 1.3$ μM), inducing the characteristic monoaster phenotype in cancer cells [182]. Its potent antiproliferative activity has been demonstrated in various preclinical models, including breast, glioblastoma, gastric, and head and neck squamous carcinoma cell lines [183–185]. The discovery of K858 validated the thiadiazole scaffold as a viable pharmacophore for targeting the L5 pocket and served as a crucial template for further drug design.

Binding mode and mechanism: The binding mechanism of K858 was elucidated by the high-resolution (1.8 Å) crystal structure of the Eg5-ADP-K858 complex (PDB: 6G6Y). This structure revealed that the (R)-enantiomer of K858 binds selectively to the L5 allosteric pocket, where it is stabilized by a network of hydrophobic and polar contacts. The binding of K858 induces a "closed" conformation of the L5 loop, which "caps" the inhibitor. This allosterically prevents the release of ADP from the catalytic site, thereby locking the motor in an inactive state and inhibiting its function [123].

SAR and Analog Development: Subsequent structure-activity relationship (SAR) studies have provided insights into the chemical features required for potency. These studies, which analyzed a large library of K858 analogues, confirmed that the 5-phenyl group is a crucial substituent for optimal steric and electronic interactions within the pocket. While modifications to this phenyl group or the 5-methyl group significantly affect inhibitory activity, the core 1,3,4-thiadiazole scaffold has proven to be a robust platform for developing even more potent inhibitors.

Clinical Development and Other Related Heterocycles: Although K858 itself, despite its potent preclinical activity and lack of neurotoxicity, did not proceed to clinical trials, the structural and SAR insights gained from its study were instrumental in the development of next-generation clinical candidates like ARRY-520 (Filanesib) and (R)-Litronesib, which share the same core scaffold [123, 186, 187].

Litronesib (LY2523355): A 1,3,4-thiadiazole derivative developed by Kyowa Kirin/Eli Lilly; in vitro KSP ATPase IC_{50} = 26 nM (Figure 2.27b). Multiple Phase I/II studies managed neutropenia with G-CSF support; limited responses in solid tumors and partial responses in multiple myeloma were observed, and the program was discontinued in the early 2020s [188, 189].

Filanesib (ARRY-520) (Array BioPharma / Pfizer / Karyopharm):

Discovery and Development: Filanesib (ARRY-520) (Figure 2.27c) was developed by Array BioPharma, which was subsequently acquired by Pfizer, and its rights were later transferred to Karyopharm Therapeutics. This compound is recognized as a highly potent inhibitor of Eg5, exhibiting an ATPase IC_{50} of approximately 6 nM and an in vivo potency ranging from 0.4 3.1 nM [190]. Its chemical structure possesses a complex benzamide backbone and is sometimes classified as a phenylalanine benzamide derivative or, owing to the thiadiazole ring in its structure, as a thiadiazole derivative [191].

Binding Mode and Mechanism: Filanesib binds to the L5 allosteric pocket of Eg5 and inhibits the motor by preventing ADP release. Crystal structure studies (e.g., PDB: 4AP8, 4AP9, 6HKY) have detailed Filanesib's specific interactions within the L5 pocket. The 2-phenyl group orients toward the protein's interior, engaging in hydrophobic interactions with the side chains of Trp127, Pro137, and Ala133. It also establishes a π-cation interaction with the guanidinium group of Arg119 and a hydrophobic interaction with the propyl region of Arg119's side chain. The 2,5-difluorophenyl substituent is wedged between the Glu116-Arg221 ion pair and Leu214, surrounded by various hydrophobic side chains (Ile136, Phe239, Leu160). This binding mode is similar to how the quinazolinone ring of ispinesib binds in the wild-type Eg5 structure [121, 135].

SAR Findings: The structure of Filanesib was designed to establish optimal interactions with the L5 pocket. Jia et al. (2024) noted that Arry-520 structurally matches the L5 pocket well (ΔE_{VDW} = −49.29 kcal/mol) and has a considerably low binding free energy [135].

Thiazolopyrimidine

AZD4877: AstraZeneca's thiazolopyrimidine core inhibitor; KSP ATPase $IC_{50} \approx 2$ nM (Figure 2.27d). Six Phase I/II trials (solid tumors, AML, lymphoma) achieved

pharmacodynamic target engagement, but the clinical benefit was limited, and development was terminated in 2013 [33].

In conclusion, thiadiazoles and related heterocycles have been able to target the Eg5 L5 pocket with high chemical diversity, and have provided a platform to achieve nanomolar potency. However, the fact that K858 remained preclinical and more potent candidates such as Litronesib/AZD4877 faced challenges in the efficacy-toxicity balance indicates the need to broaden the therapeutic window for this chemical class as well.

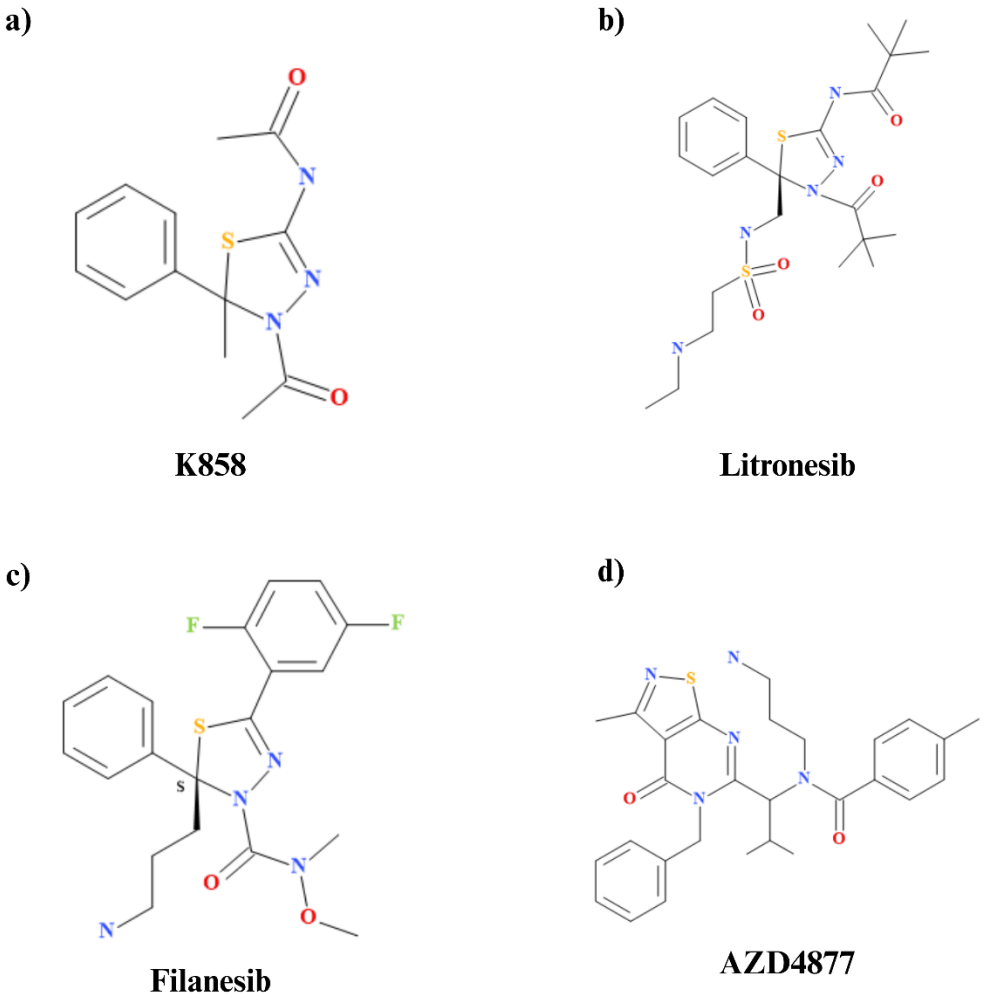

Figure 2.27 : Chemical structures of representative thiazole-derived Eg5 inhibitors.

Quinolone and Quinazolinone Derivatives (Including Chromen Replacement)

Among the chemical structures targeting the L5 allosteric pocket of Eg5, compounds containing heterocyclic ring systems such as quinolone and quinazolinone have also

shown significant potential. These classes offer scaffolds that allow for diverse substitutions and their pharmacological properties can be optimized.

ARQ 621 (ArQule)

Discovery and Development: Discovered by ArQule in 2008, ARQ 621 (Figure 2.28a) was identified as a quinazolinone derivative and joined the Merck & Co. portfolio in 2019 [192].

Chemical Structure and Potency: It contains various amine and halogen substituents attached to a quinazolinone core; Eg5 ATPase $IC_{50} \approx 1.8$ nM, and it demonstrated antiproliferative effects across a broad panel of cancer cells [].

Binding Mode and Mechanism: Its crystal structure has not been published; due to SAR similarity, it presumably binds to the L5 allosteric pocket, blocking ADP release and locking Eg5 in an inactive conformation [33].

SAR findings: Preclinical studies reported relatively lower toxicity to bone marrow-derived cells; aryl/heteroaryl groups added to various positions of the quinazolinone core enhanced the binding affinity and pharmacokinetics [66].

EMD 534085

Discovery and Development: The hexahydro-pyrano[3,2-c]quinolines (HHPQ) scaffold was identified by Merck Serono during high-throughput screening (HTS), leading to the development of highly potent Eg5 inhibitors [193].

Chemical Structure and Potency: The core is a hexahydro-pyranoquinoline with a quinoline ring fused to a pyran ring, and the representative compound EMD 534085 (Figure 2.28b) has an Eg5 IC_{50} of 8 nM, with some analogs reaching the picomolar range [193].

Binding Mode and Mechanism: Crystal and docking studies, such as those represented by PDB 3L9H, have demonstrated that the HHPQ core occupies the L5 pocket. This interaction involves both hydrophobic (Tyr211 and Leu214) and polar residues (Glu116 and Arg119), effectively inhibiting the release of ADP and maintaining Eg5 in an inactive conformation [135].

SAR Findings: Substitutions on the quinoline and pyran rings were determinants for potency and pharmacokinetics; in particular, aryl/heteroaryl groups compatible with

the hydrophobic pocket increased activity, while hydrophilic modifications improved solubility [66].

In general, quinoline- and quinazoline-based Eg5 inhibitors have exemplified the chemical diversity with which the L5 allosteric pocket can be targeted and have provided scaffolds capable of achieving high potency. However, the emergence of clinically successful drugs from these classes has encountered various challenges, such as many other Eg5 inhibitor classes. The clinical failure of promising scaffolds, such as quinazolines and HHPQs, despite achieving high potency, further reinforces the central challenge in Eg5 inhibitor development: the difficulty of decoupling potent on-target inhibition from dose-limiting toxicities. This suggests that achieving low IC_{50} is insufficient for clinical success. This thesis addresses this challenge by not only searching for potent compounds but also by employing a dual-track virtual screening approach. The aim was to identify a wide diversity of novel chemical scaffolds, with the hope that some of these new structural classes will possess inherently different ADMET profiles, potentially offering a wider therapeutic window than scaffolds that have been clinically tested and discontinued.

Ispinesib (SB-715992) and SB-743921 (GSK):

Discovery and Development: Developed through a collaboration between GlaxoSmithKline and Cytokinetics, ispinesib (Figure 2.28c) was the first Eg5 inhibitor to enter clinical trials [132]. Chemically, it features a quinazolinone core attached to a phenylurea side chain and inhibits Eg5 ATPase activity with high potency (IC_{50} in the range of 1.2–9.5 nM) [8]. A subsequent analog, SB-743921 (Figure 2.28d), was developed through the isosteric replacement of the quinazolinone ring with a chromen-4-one core (Figure 2.28b). This modification resulted in an approximately fivefold increase in potency ($IC_{50} \approx 0.1$ nM), demonstrating the scaffold's tunability [8, 132].

Binding Mode and Mechanism: Crystal structure studies of both ispinesib (PDB: 4AP0) and SB-743921 (PDB: 4BXN) confirm that they bind to the L5 allosteric pocket. The binding of these compounds induces the closure of loop L5, which allosterically prevents ADP release and locks the motor in an ATP-hydrolysis-incompetent state. Structure-activity relationship (SAR) studies have shown that the aryl and heterocyclic groups linked to the central core have been highly refined to optimize hydrophobic and polar interactions within the L5 pocket, with the chromen-

4-one core of SB-743921 providing an improved geometric fit and enhanced potency [119, 121, 132, 135].

Pyrrole derivative

MK-0731 (Merck):

Discovery and development: MK-0731 (Figure 2.28e), developed by Merck, represents another effective Eg5 inhibitor that specifically targets the L5 allosteric pocket (KSP ATPase IC_{50}=2.2 nM). The chemical structure is occasionally characterized as 2,4-diaryl-2,5-dihydropyrrole, or more broadly, as a sophisticated heterocyclic system [180, 187, 194].

Binding Mode and Mechanism: It has been shown to inhibit Eg5 activity by binding to the L5 pocket. Jia et al. (2024) stated that MK-0731 has a low binding free energy to the L5 pocket and structurally matches the pocket well (ΔE~VDW~ = −48.69 kcal/mol) [135].

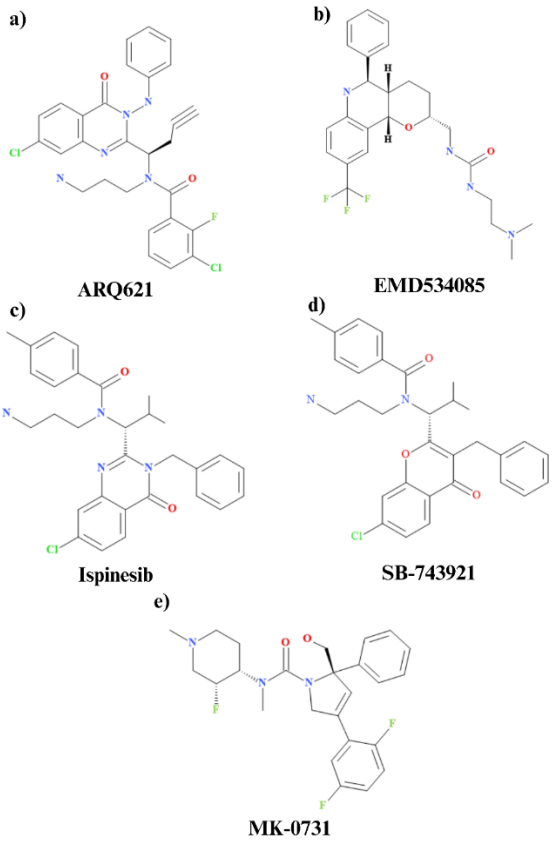

Figure 2.28 : Chemical structures of representative quinolone/quinazolinone/pyrrole core based Eg5 inhibitors targeting the L5 allosteric pocket in clinical development.

.

Natural product-derived L5 pocket inhibitors

Beyond the primary synthetic classes, natural products serve as valuable reservoirs of intricate and bioactive compounds that can target Eg5. These compounds illustrate a wider range of chemical diversity and may provide alternative interaction modes or pharmacological profiles compared to traditional synthetic compounds.

Biflavonoids ((+)-Morelloflavone and Kolaflavanone): Several biflavonoids, a class of plant-derived polyphenols, have been identified as Eg5 inhibitors. (+)-Morelloflavone (Figure 2.29a), for example, inhibits Eg5 ATPase activity in the micromolar range ($IC_{50} \approx 83–98$ μM) and leads to characteristic monopolar spindle formation in cells [134]. Similarly, Kolaflavanone (KLF) also suppresses Eg5 activity at a comparable concentration range [195]. Computational studies for both compounds suggested a binding mode within the L5 allosteric pocket, similar to known synthetic inhibitors.

Terpendole-E: This indole-diterpenoid natural product (Figure 2.29b) was particularly noteworthy. The IC_{50} value for the inhibition of microtubule-activated ATPase activity for Terpendole E has been determined to be 23 μM, retains its activity against Eg5 mutants that are resistant to both L5-pocket inhibitors (like STLC) and α4/α6-pocket inhibitors (like GSK-1) [196]. This suggested that Terpendole-E may bind to a novel yet uncharacterized third allosteric site or inhibit Eg5 through a completely different mechanism, such as destabilizing the tetramer. However, it has been experimentally demonstrated that the binding site of Terpendole E to Eg5 is located in the α2/L5/α3 allosteric pocket, similar to many other Eg5 inhibitors [197].

The identification of these natural products is important for two main reasons. First, they proved that the Eg5 motor domain could be effectively targeted by complex and diverse chemical scaffolds beyond those found through synthetic chemistry. Second, given that natural products often possess distinct Absorption, Distribution, Metabolism, Excretion, Toxicity (ADMET) profiles, their scaffolds could provide a starting point for the development of semi-synthetic derivatives that might overcome the challenges of toxicity and resistance encountered by previous inhibitor classes. The discovery of these natural products, especially Terpendole-E, reinforces a key motivation for this thesis: that the landscape of Eg5 inhibition is not limited to two well-characterized allosteric sites. While our primary focus remains the validated L5

pocket, the existence of other potentially druggable sites supports the broader search for novel chemicals. It highlights the importance of exploring a wide chemical space, as we have done, which may lead to compounds with unexpected binding modes or superior pharmacological properties.

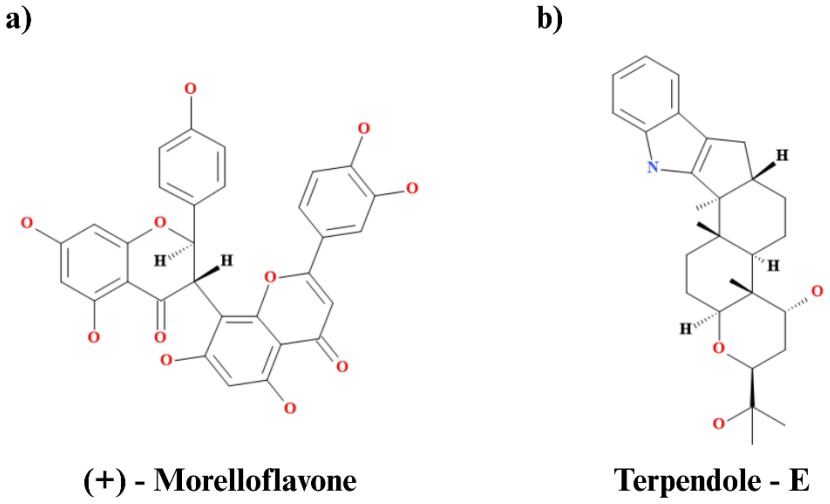

Figure 2.29 : Chemical structures of natural product-derived Eg5 inhibitors.

Helix α4/α6 allosteric pocket-targeted inhibitors

A second allosteric binding site, distinct from the well-known L5 pocket, was identified at the interface of helices α4 and α6, closer to the microtubule-interacting surface. Inhibitors targeting this site offer a different mechanism of action and a potential strategy for overcoming resistance to L5 pocket ligands [10].

Prototype compounds and mechanism: Prototypical compounds for this site include biaryl derivatives such as GSK-1 and GSK-2 and benzimidazoles such as BI8 and PVZB1194. Unlike L5 inhibitors, their efficacy is microtubule-dependent and they act as functional ATP-competitive inhibitors despite binding outside the ATP pocket [10, 198]. Their mechanism involves binding to Eg5 on the microtubule, which distorts the P-loop and prevents ATP re-binding. This locks the motor in a tight, "rigor-like" state with an empty nucleotide pocket, effectively stalling it on the microtubule track [98, 128]. This mode of action contrasts sharply with that of L5 inhibitors, which trap ADP within the motor.

SAR and clinical outlook: The α4/α6 pocket is formed by residues from helices α4, α6, and loop L11, with Tyr104 forming the outer rim []. A major advantage of this site is that inhibitors targeting it can remain effective against Eg5 variants with L5 pocket

resistance mutations (e.g., D130V) . Despite this promise, compounds targeting the α4/α6 pocket, such as GSK-1 and PVZB1194, have not yet succeeded in clinical trials, suggesting that they may face distinct pharmacokinetic or toxicity challenges. Nonetheless, this site remains a viable alternative target for future drug discovery [10].

Clinical development of Eg5 inhibitors: Trials, results, and lessons learned

The critical role of Eg5 in mitotic spindle formation and the discovery of its allosteric pockets led to the perception of this protein as a promising target in cancer therapy, prompting the development of many potent inhibitors in clinical trials since the early 2000s. Nonetheless, despite robust preclinical findings and well-founded drug design initiatives, the clinical development pathway for Eg5 inhibitors has encountered considerable obstacles and has largely not met anticipated outcomes. This section aims to provide a summary of the trial processes involving notable clinical candidates, highlighting the key results achieved and insights gained from these experiences.

Principal clinical candidates and their general outcomes

Several Eg5 inhibitors have entered clinical development, each with its unique journey and outcomes:

Ispinesib (SB-715992/CK0238273):

Early Trials: Ispinesib was one of the first Eg5 inhibitors to enter clinical trials. It demonstrated potent and selective Eg5 inhibitory activity (approximately 40,000 times more selective for Eg5 over other kinesins) and potent anticancer activity against various cell lines in vitro (IC50 = 1.2–9.5 nM) as well as in xenograft tumor models [199].

Clinical Outcomes: A total of 16 clinical trials were conducted on ispinesib, with 14 completed and two terminated, but none of these trials conclusively demonstrated clinical benefit. Phase I/II trials investigated ispinesib as monotherapy or in combination with other agents (e.g., docetaxel, capecitabine, carboplatin) in various cancers including recurrent renal cell cancer, breast cancer, head and neck cancer, liver and colorectal cancer, and prostate cancer. While some studies reported stable disease (SD), complete or partial responses were rare [200]. For instance, in prostate cancer, ispinesib was found ineffective, possibly due to the low mitotic index and Eg5 expression in these tumors [132].

Resistance: Ispinesib resistance, sometimes observed in clinical trials, may be attributed to multidrug resistance. or mutations in the Eg5 inhibitor binding pocket, such as D130V . Despite this, it was still fully capable of inhibiting growth in Eg5(L214A) mutant cells, unlike STLC and ARRY-520 [121, 200].

Filanesib (ARRY-520)

Early trials: Filanesib (ARRY-520) is a thiadiazole derivative that has shown promising antitumor activity, particularly in hematological and taxane-resistant tumor models [201]. Its structure is similar to MK-0731, and it binds to the α2/L5/α3 pocket without forming direct hydrogen bonds with Eg5 in MD simulations [135].

Clinical Outcomes: While monotherapy efficacy in solid tumors was limited94125, filanesib has proved to be a promising agent in phase I and II clinical trials for patients with relapsed or refractory multiple myeloma. It has progressed to Phase III clinical trials [132, 172, 202]. In combination with low-dose dexamethasone, it showed an overall response rate of 15-16% in heavily pretreated multiple myeloma patients125. Combining filanesib with proteasome inhibitors like bortezomib or carfilzomib further improved response rates (up to 42% and 37% respectively) [132].

Biomarkers and Resistance: The efficacy of filanesib in clinical trials has been correlated with expression levels of α-1 acidic glycoprotein (AAG), suggesting AAG as a potential predictive biomarker. Mutations like D130A and L214A in the Eg5 allosteric binding site can confer resistance to filanesib, similar to STLC. Its better pharmacokinetic profile and reduced dosing frequency compared to other Eg5 inhibitors may contribute to its wider therapeutic index, especially in highly proliferative cancers like multiple myeloma.

AZD4877:

Trials and Results: AZD4877, from AstraZeneca, inhibits Eg5 with an IC50 of 2 nM in vitro. It entered six phase I/II trials against solid and hematological tumors [203, 204].

Outcomes: Clinical results were largely disappointing. A Phase I study in patients with relapsed/refractory solid tumors and lymphoma showed neutropenia as the only dose-limiting toxicity (DLT) at 15 mg, but no complete or partial responses were observed, indicating no clear clinical benefit. Similar modest efficacy was reported in

trials for AML and urothelial cancer. No further development for AZD4877 is planned [187].

MK-0731:

Trials and Outcomes: MK-0731 underwent Phase I clinical trials for advanced solid tumors. The maximum tolerated dose (MTD) was 17 mg/m², with neutropenia as the major DLT. Only a small percentage of patients (4.3%) achieved stable disease, and no objective responses were reported [187].

Termination: Development of MK-0731 has been terminated.

4SC-205 (AEGIS):

Clinical Progression: 4SC-205 is an orally available Eg5 inhibitor. A first-in-human Phase I study (NCT01065025) in patients with advanced malignancies (solid tumors and malignant lymphomas) has been completed [205, 206].

Dosing and Results: Patients received 4SC-205 once or twice weekly. The MTD was determined to be 150 mg/m² for once-weekly and 75 mg/m² for twice-weekly administration. Neutropenia was the most common DLT, consistent with other Eg5 inhibitors. No complete or partial responses were reported in the Phase I study, with 28% of patients achieving stable disease as the best response. Neoplasm was stabilized in 67% of patients for 100 days, and median survival time was prolonged (details not directly in provided sources for this statement, but based on prior conversation output from a comprehensive response that drew on external knowledge) [187].

LY2523355 (Litronesib):

Preclinical Activity: In preclinical xenograft models, it showed broad-spectrum anticancer activity, causing rapid apoptosis following spindle assembly checkpoint (SAC)-mediated mitotic arrest. Histone H3 phosphorylation was identified as a promising pharmacodynamic marker for its activity [207].

Clinical Outcomes: Litronesib entered at least seven phase I/II trials, mainly for solid tumors, as monotherapy or with G-CSF. Efficacy results were generally disappointing. DLTs included neutropenia and leukopenia. Only a small percentage of patients (3.7%) showed partial responses, and stable disease was achieved in 26-36.5% of patients [188, 189].

Discontinuation: Despite its promising preclinical profile, further development of Litronesib has been discontinued [187].

EMD534085:

Development and Trials: It showed an Eg5 ATPase IC50 of 8 nM and inhibited human tumor xenograft models without toxic side effects at certain doses [180].

Clinical Outcomes: A phase I trial in patients with solid tumors or lymphoma as a single agent defined the MTD as 108 mg/m². Neutropenia was the most common DLT. No objective responses were reported, with 52% of patients achieving stable disease. No further development is planned for this inhibitor [187].

ARQ 621:

Trials and Outcomes: ARQ 621 displayed anticancer activity in preclinical settings. It was clinically inspected in phase I trials against hematological malignancies and metastatic solid tumors. No clear responses were registered, and no further development was reported [33].

ALN-VSP02:

Unique Mechanism: ALN-VSP02 is distinct as an siRNA lipid nanoparticle formulation that targets the expression of two genes critical for cancer growth: Eg5 and vascular endothelial growth factor (VEGF) [187].

Clinical Outcomes: Two phase I trials were completed in patients with solid tumors. It demonstrated a favorable safety profile. Notably, a complete response was observed in one patient with endometrial cancer with multiple hepatic metastases, and prolonged disease stabilization for 1-1.5 years in others. Despite these promising results, further clinical studies on ALN-VSP were put on hold, likely due to legal disputes regarding intellectual property for lipid nanoparticle delivery technology [208, 209].

Lessons Learned from Clinical Development

The clinical experiences with Eg5 inhibitors have provided several crucial insights:

Limited Monotherapy Efficacy in Solid Tumors: A recurring theme across most Eg5 inhibitors is their limited efficacy as monotherapies in solid tumors. This may be due to the slower proliferation rates of cells in solid tumors compared to those in preclinical xenograft models, limiting the number of mitotic cells available to respond

to the Eg5 inhibition during chemotherapy. Additionally, reduced drug penetration into solid tumors may play a role [180, 187].

Dose-Limiting Toxicities (DLTs): While Eg5 inhibitors largely avoided the neurotoxicity associated with tubulin-targeting agents [66], neutropenia was consistently identified as the most common DLT across various compounds (e.g., 4SC-205, AZD4877, EMD534085, Ispinesib, Litronesib, MK-0731) [187]. Other reported side effects included nausea, fatigue, anemia, leukopenia, thrombocytopenia, and diarrhea [180]. This highlights that while Eg5 is preferentially expressed in dividing cells, it still plays an essential role in the proliferation of normal cells, leading to anticipated target-mediated toxicities in highly proliferative normal tissues like bone marrow .

Importance of Pharmacokinetic/Pharmacodynamic (PK/PD) Optimization: The clinical results underscored the critical need for careful dose and schedule selection based on a deep understanding of PK/PD relationships [151, 180, 187, 200]. Eg5 is expressed only briefly during mitosis, and cancer cell-cycle progression is asynchronous. Therefore, Eg5 inhibitors must be available at or above a threshold concentration for an adequate duration to ensure sustained mitotic arrest of most dividing cancer cells as they asynchronously progress into mitosis. Early Eg5 inhibitors, which followed dosing paradigms established for antitubulins, showed little clinical activity, possibly due to insufficient half-lives for optimal dosing [207]. The better pharmacokinetic profile (e.g., longer half-life) of successful compounds like filanesib may explain their improved therapeutic index [187].

Resistance Mechanisms: Many mutations conferring resistance to Eg5 inhibitors are located in the α2/L5/α3 allosteric binding pocket. While these mutations do not always directly prevent inhibitor binding, they can disrupt the allosteric communication network through which the inhibitor's effect is transmitted to the ATP-binding site [132].

D130A/V and A133D/I/V/M/P mutations: These mutations, located in loop L5, have been identified in ispinesib-resistant tumor cells. Although these mutations may not significantly alter the binding of monastrol or STLC to the motor domain, they can overcome inhibition by blocking the allosteric communication network to the ATP-binding site. For instance, the A133D mutation, while still allowing inhibitor binding,

can block the perturbation pathway to the nucleotide site, effectively uncoupling inhibitor binding from its inhibitory effect. Molecular dynamics simulations have shown that in the Eg5(A133D)-SB743921 complex, local rearrangements of hydrogen bonding and salt bridging transmit structural changes throughout the complex, blocking the inhibitor-induced allosteric transmission to the nucleotide site and rendering the motor resistant to inhibition. This suggests that these mutations can "close" an energy gate, preventing the transmission of the inhibitory signal [121, 136, 210–212].

L214A mutation: This mutation, located in helix α3, has also been shown to confer resistance to inhibitors like STLC, monastrol, and ARRY-52073. Molecular dynamics simulations indicate that for ARRY-520, the L214A mutation leads to an increased volume in the inhibitor-binding pocket, allowing for greater ligand flexibility and "wobbling" of the difluorophenyl group, which compromises its effectiveness as a tight-binding inhibitor and results in resistance. In contrast, ispinesib was highly sensitive to the L214A mutation, suggesting that its binding dynamics were less affected, enabling it to retain inhibitory activity [121].

These findings highlight that resistance to Eg5 inhibitors is not always scaffold-independent, meaning that different chemical scaffolds can exhibit differential resistance patterns to the same mutations. This opens the door for designing new inhibitors that can retain their activity against specific resistant mutants. For example, biphenyl-type inhibitors like GSK-1 and PVZB1194, which bind to the α4/α6 pocket, have shown activity against ispinesib-resistant tumor cells carrying D130V and A133D mutation.

Mutations at other sites, like Y104C (between α4 and α6 helices), can also lead to resistance [132]. Cells can also adapt by upregulating alternative pathways, such as Kif15 [213].

Combination Therapy as a Promising Strategy: Despite the limited success of Eg5 inhibitors as monotherapies in many solid tumors, several sources suggest that combination therapies hold significant promise. Filanesib's success in combination with proteasome inhibitors for multiple myeloma exemplifies this . Preclinical studies with 4SC-205 have also shown additive effects when combined with standard chemotherapies or ALK/MEK inhibitors, providing a strong rationale for combination

strategies in future clinical trials. This approach may overcome resistance and achieve better therapeutic outcomes by synergistically targeting multiple pathways or by addressing the issue of tumor heterogeneity and asynchronous cell division [180, 187, 200].

In conclusion, while the initial excitement surrounding Eg5 inhibitors was tempered by disappointing monotherapy results in clinical trials, particularly for solid tumors, the lessons learned are invaluable . The field has moved towards understanding the critical importance of optimal PK/PD, identifying predictive biomarkers, and exploring combination strategies, especially for highly proliferative cancers like hematological malignancies, where Eg5 inhibitors have shown more promising results. Continued research into novel binding sites and strategies to overcome resistance remains a vital part of developing these agents into effective anticancer therapeutics [132].

Upregulation of drug efflux pumps: Overexpression of ATP-binding cassette (ABC) transporter proteins, most notably P-glycoprotein (MDR1/ABCB1), is a classic mechanism of multi-drug resistance. These pumps can actively remove Eg5 inhibitors from the cell, lowering their intracellular concentrations to below the effective threshold. Several Eg5 inhibitors have been identified as substrates for these pumps, making it a clinically relevant resistance mechanism [180, 214, 215].

Activation of compensatory (bypass) pathways: Cells can develop resistance by activating alternative mechanisms to compensate for the loss of Eg5 function. The most significant of these involves KIF15 protein, a member of the Kinesin-12 family. Preclinical studies have shown that KIF15 plays an important Eg5's role in establishing and maintaining spindle bipolarity [216]. Consequently, cancer cells with high intrinsic levels of KIF15 or those that upregulate KIF15 upon Eg5 inhibition can display significant resistance to Eg5 inhibitors .

Resistance to the STLC Eg5 inhibitor, for example, has been associated with the expression of a rigor mutant of Eg5, G268V. This G268V mutation, located in the switch II nucleotide binding area, traps the Eg5 motor in a microtubule-binding state, leading to strong microtubule crosslinking. In such resistant cells, KIF15 changes its distribution, becoming enriched throughout the entire spindle instead of being restricted solely to kinetochore microtubules, thereby contributing to the formation of

a normal bipolar spindle. This adaptation is not necessarily due to KIF15 overexpression but rather to changes in its localization and activity [132, 139].

The observation of KIF15's compensatory role suggests that a combination of Eg5 inhibitors and KIF15 inhibitors could be an effective anti-proliferative chemotherapeutic regimen for rapidly dividing cancer types. Lead compounds for KIF15 inhibitors, such as GW108X, KIF15-IN-1, and benserazide, have been identified for further development. Additionally, in an SB743921-resistant cell model that relies on KIF15 activity, inhibiting the centrosome-associated AURKA kinase has been shown to be an effective strategy to overcome resistance to Eg5 inhibitors [213, 217, 218].

Evasion of mitotic arrest via mitotic slippage: Eg5 inhibition activates the Spindle Assembly Checkpoint (SAC), leading to prolonged arrest in mitosis. However, rather than undergoing apoptosis as a result of this arrest, some cancer cells can adapt and exit mitosis without completing cell division, a phenomenon known as "mitotic slippage." This process results in viable but genetically unstable tetraploid (4N DNA content) cells, which can contribute to tumor survival and the emergence of a drug-resistant population [187].

The quest for predictive biomarkers

Several lines of evidence indicate that reliable biomarkers will be essential for translating Eg5 inhibitors into precision oncology. Early hypotheses suggested that highly proliferative tumours—those with strong Ki-67 staining—would be the most susceptible, because Eg5 is required only during mitosis; however, pre-clinical and phase-I studies with ispinesib failed to show a consistent correlation between Ki-67 levels and drug response [219].

By contrast, the tumour-suppressor p53 appears to influence cell fate after Eg5 blockade: cells that retain wild-type p53 undergo apoptosis following prolonged mitotic arrest, whereas p53-mutant cells are more likely to slip out of mitosis and survive, a behaviour documented for kinesin-5 inhibitors and other anti-mitotics [220].

Perhaps the strongest predictive signal identified so far is the Eg5 / KIF15 expression balance. Because the plus-end motor KIF15 can functionally compensate for Eg5, tumours with high Eg5 and low KIF15 expression are markedly more sensitive,

whereas up-regulation of KIF15 drives both intrinsic and acquired resistance; thus the Eg5:KIF15 ratio is emerging as a rational stratification marker [139].

For on-treatment response, several clinical trials of the potent inhibitor filanesib have shown that low baseline serum α1-acid glycoprotein (AAG) correlates with higher objective-response rates and longer progression-free survival, making AAG a promising pharmacodynamic/selection biomarker for this agent [172].

A comprehensive understanding of these resistance mechanisms and an ongoing search for biomarkers are central to the rationale of this thesis. The prevalence of on-target resistance through mutations in the L5 pocket underscores the need to discover inhibitors that either bind differently to this pocket or that have a high affinity to overcome common mutations. Furthermore, the existence of bypass pathways such as KIF15 highlights the importance of developing highly potent inhibitors that can induce rapid and irreversible mitotic arrest before the cell has time to adapt. Our computational approach was designed to tackle this challenge by screening for novel chemical scaffolds that form highly stable and optimal interactions within the L5 pocket, with the ultimate goal of identifying compounds that are less susceptible to these known resistance mechanisms.

2.2.3.3 Computational and AI-assisted approaches in the discovery of Eg5 inhibitors

The clinical journey of first-generation Eg5 inhibitors has been met with the challenge of a narrow therapeutic window, primarily because of dose-limiting neutropenia and rapidly developing resistance mechanisms. These hurdles have propelled the field away from traditional high-throughput screening (HTS) methods towards data-driven discovery strategies grounded in physics-based simulations, machine learning (ML), and ultimately, generative artificial intelligence (AI). The therapeutic need for this target is underscored by the association of high Eg5 expression levels with aggressive tumor phenotypes and poor prognosis. This section details the methodological evolution and provides a conceptual framework for the advanced computational work undertaken in this thesis.

The foundational era: from early models to large-scale virtual screening

The initial wave of in silico techniques focused on deciphering the rules of activity from the structural features of known active compounds, and using these rules to filter large chemical libraries.

Pharmacophore and QSAR modeling: The first pharmacophore model mapped the essential three-dimensional chemical features of molecules that bind to the L5 pocket. A four-feature model derived from 25 inhibitors by Liu et al. was used as a reference template for the field. Shortly thereafter, structure-based 3D-QSAR studies have added quantitative predictive power to these models. For example, the CoMFA/CoMSIA models developed by Luo et al. (2012) for a dihydropyrazole series achieved high statistical quality ($q^2 = 0.74$, $r^2 = 0.98$), visualizing the critical hydrophobic and hydrogen bond interactions required for activity [221].

The rise of virtual screening: in silico scaffold hopping: In 2007, Jiang et al. published the first docking study of Eg5 inhibitors in the allosteric binding pocket [222]. The increasing availability of high-resolution crystal and cryo-electron microscopy structures has enabled structure-based virtual screening (SBVS) campaigns. A pioneering 2012 study by Nagarajan et al. demonstrated the power of in silico screening by uncovering entirely new quinazoline and thiooxazolidine cores through large-scale docking [223]. This strategy was expanded to target alternative pockets to overcome resistance mechanisms; a key example is the discovery of SRI-35566 by Zhang et al [141]. Ligand-based virtual screening (LBVS), using similarity to known actives, led to the discovery of high-affinity tetrazoles like YL001 [205].

Hybrid and integrated workflows: A multi-pronged strategy for hit discovery

Recognizing that no single computational method is infallible, contemporary research has converged on a more robust philosophy: the construction of multistage, integrated workflows. This paradigm operates similar to a computational funnel designed to manage the immense scale of modern compound libraries. Computationally inexpensive, faster methods are employed at the top of the funnel to triage millions or billions of molecules, while the more physically rigorous, and thus computationally expensive, methods like molecular dynamics and free energy calculations are reserved for a much smaller, highly enriched set of promising candidates at the bottom. This

strategic layering of filters maximizes the efficiency and significantly increases the probability of identifying true high-quality hits.

A compelling illustration of the application of this hierarchical filtering approach to natural products is the 2023 study by Bodun et al.. Confronted with the chemical diversity of a Traditional Chinese Medicine (TCM) library, they first deployed a QSAR model trained to recognize general inhibitor-like features as a coarse, rapid primary filter. This initial step dramatically reduced the chemical space. Subsequently, a structure-derived pharmacophore model provided a more stringent and geometrically aware second filter. Only compounds that successfully navigated these initial gates were advanced to the resource-intensive final stages of molecular docking and MM-GBSA rescoring. This cascaded workflow efficiently distilled a vast library down to five promising candidates with predicted binding energies superior to that of the reference inhibitor, monastrol [224].

The versatility of these integrated workflows was further demonstrated by their application to alternative allosteric sites, which is a key strategy for circumventing resistance. A recent work by Elseginy et al. (2024) exemplifies this. To discover novel inhibitors of the challenging α4/α6 pocket, they constructed a sophisticated pipeline beginning with a 3D-pharmacophore model specific to that site. This was followed not by a single docking run but by parallel campaigns using two distinct algorithms (MOE and AutoDock). A hit was considered a high-confidence candidate only if it was prioritized by both programs, a technique known as consensus scoring. This algorithmic cross-validation adds a crucial layer of stringency, mitigating the risk of artifacts or biases inherent in any single docking program and boosting confidence in the predicted binding poses. Their comprehensive and rigorous approach culminated in the experimental validation of a potent inhibitor with a novel benzo[b][1,4]oxazine-6-carboxamide scaffold, which exhibited an IC_{50} of 1.37 ± 0.15 μM [191].

This multi-step philosophy extends beyond the initial hit identification to encompass rigorous post-screening validation, which is critical for de-risking candidates before committing to chemical synthesis. The study by Shahin and Aljamal, which targeted the canonical L5 pocket, provides a clear example. Their initial screening cascade, involving pharmacophore and docking methods, was critically augmented by a subsequent in-depth validation phase. Each promising hit was subjected to long-term MD simulations to assess the dynamic stability of the protein-ligand complex over

time. This was followed by end-point free energy calculations (MM/PBSA) to provide a more physically grounded estimate of binding affinity than docking scores alone. This rigorous end-to-end process, from broad screening to deep dynamic validation, successfully yielded novel hits built on a 4-aminoquinoline scaffold, showcasing a mature and robust computational discovery engine [33].

The role of molecular dynamics

While static docking and pharmacophore models are powerful, they do not capture the dynamic nature of protein-ligand interactions. Molecular Dynamics (MD) simulations have become a critical tool in modern Eg5 inhibitor discovery, allowing researchers to examine the stability of binding poses over time and accounting for protein flexibility. This is particularly important for the L5 allosteric pocket, in which flexible-loop regions can adopt different conformations.

MD simulations are now routinely used to validate the results of virtual-screening studies. For example, the aforementioned study by Elseginy et al. subjected their lead compounds to 100 ns MD simulations, confirming that the compounds remained stably bound in the α4/α6 allosteric pocket [191]. Furthermore, advanced techniques such as binding free energy calculations (e.g., MM/PBSA and MM/GBSA) are performed on MD trajectories to provide a more quantitative ranking of inhibitor candidates than docking scores alone, as was also performed in both the Elseginy and Bodun studies to confirm the favorable binding energy of their hits [191, 224].

A comprehensive 2024 study by Jia et al. illustrated the power of MD simulations for comparative analysis. They employed extensive MD simulations and MM/PBSA calculations to systematically compare the binding mechanisms of seven known inhibitors targeting the L5 pocket (Monastrol, STLC, Ispinesib, and Filanesib) and one inhibitor targeting the α4/α6 pocket. Their analysis revealed that van der Waals forces were the primary drivers for L5 pocket inhibitors, while electrostatic interactions also played a significant role in the α4/α6 pocket ligand. This study confirmed the stability of all inhibitors in their respective pockets and provided a quantitative ranking of binding free energies that correlated well with experimental potencies, offering a molecular-level explanation for their varied activities [135].

Another notable application of MD is in a 2022 drug repurposing study by Sebastian et al.. They used MD simulations of Eg5 with known inhibitors to generate stable

"closed" pocket conformations, which were then used to build a pharmacophore model for screening approved drugs. This led to the identification of tolvaptan, a vasopressin receptor antagonist, as a novel Eg5 inhibitor . Subsequent experimental validation strongly supported the computational predictions: tolvaptan was determined to bind directly to the Eg5 motor domain with a K~D~ of 27 ± 3.8 μM and was observed to cause mitotic blockage in HeLa cells, inducing the characteristic unipolar spindle phenotype. Intriguingly, the study also found that tolvaptan is a dual inhibitor, exhibiting weak affinity for the colchicine binding site on tubulin (K~D~ ~100 μM). Furthermore, the combination of tolvaptan with other antimitotics, such as STLC or vinblastine, resulted in synergistic enhancement of mitotic defects. This study demonstrated how MD-assisted strategies can uncover new uses of existing drugs and reveal complex pharmacological profiles [225].

The AI revolution: from intelligent screening to autonomous discovery

Machine learning for smarter screening: As computational power and data availability have grown, machine learning has been increasingly integrated into VS workflows, often as a highly efficient pre-filtering step. A study on the discovery of Eg5 inhibitors from natural products illustrates this modern approach. Researchers began with a virtual library of 1.830 secondary metabolites isolated from fungi. Instead of immediately proceeding to computationally expensive docking, they first employed a machine learning model. A Random Forest (RF) classifier was trained on a dataset of known Eg5 inhibitors and non-inhibitors to learn the chemical features that distinguish active from inactive compounds. This predictive model demonstrated excellent performance, with a Receiver Operating Characteristic (ROC) area of 0.99, indicating a near-perfect ability to classify compounds correctly [226].

De novo design with generative AI: The core of this work was a generative model based on a Long Short-Term Memory (LSTM) neural network. This model was first pre-trained on a massive database of molecules from ChEMBL, allowing it to learn the fundamental rules of chemical structure and syntax (represented as SMILES strings). This general chemistry model was then fine-tuned on a small, specific dataset of known Eg5 inhibitors. This transfer learning step effectively taught the model the specific chemical features associated with Eg5 activity.

The fine-tuned generative model was then used to design a library of entirely new molecules tailored to the Eg5 chemical space. This de novo generated library was then subjected to a validation pipeline:

- A separate ML classifier was used to predict the bioactivity of the generated molecules.
- The molecules were filtered for drug-like properties (e.g., using Lipinski's Rule of Five).
- The most promising candidates were validated using molecular docking.
- The top hits from docking were subjected to extensive 200-nanosecond MD simulations to assess their binding stability and interactions in detail.

The results were remarkable. The workflow produced five novel compounds that were predicted to have better binding affinities to Eg5 than the reference ligand that was co-crystallized in the protein structure. The top-performing hit, which was a bioisostere (a structural analog with similar physical or chemical properties) of the known ligand, showed a significantly improved calculated binding free energy. This study exemplifies an end-to-end in silico design process. It demonstrates the immense potential of generative AI to navigate and explore the almost infinite chemical space to create bespoke, optimized inhibitors from scratch, promising to accelerate the discovery of novel therapeutics for challenging targets like Eg5 [227].

Emerging therapeutic modalities and novel discovery strategies for Eg5

Although the development of potent allosteric small-molecule inhibitors targeting the L5 pocket of Eg5 has been a primary focus of the field, the clinical translation of these agents has been consistently hampered by a narrow therapeutic index, with neutropenia being the dose-limiting toxicity. This on-target toxicity highlights a fundamental challenge; merely increasing inhibitor potency may not resolve the issue, as it could simply lower the dose at which toxicity occurs without widening the therapeutic window. The core difficulty lies in decoupling efficacy in tumor cells from toxicity in healthy, proliferating cells.

To overcome these limitations, the field is evolving beyond traditional screening, exploring novel discovery platforms, and pioneering new therapeutic modalities to enhance tumor selectivity and efficacy.

Novel discovery platforms

Fragment-based drug discovery (FBDD): FBDD builds potent leads by identifying small, low-molecular-weight "fragments" that bind efficiently to the target, which are then synthetically "grown" to create higher-affinity compounds [228]. Additionally, Fragment-based discovery with CNS focus. A 2025 fragment-based drug-discovery (FBDD) campaign by Naseer et al. identified a 3-phenyl-1H-pyrazol-5-carboxylate scaffold (lead fragment Mol-121026) that binds the L5 pocket and, based on MD permeability metrics and low polar surface area, is predicted to evade P-gp–mediated efflux and penetrate the blood–brain barrier—an important step toward Eg5 inhibition in glioblastoma and other CNS tumors [229].

Advanced therapeutic modalities

Gene silencing (RNAi): Strategies have been developed to target Eg5 at the genetic level. The clinical candidate ALN-VSP02, for instance, uses siRNA delivered in a lipid nanoparticle to suppress Eg5 protein production, demonstrating that genetic targeting is a viable, though challenging, strategy [208, 209].

Antibody-drug conjugates (ADCs) and degrader-antibody conjugates (DACs): This approach links a highly potent payload, for which an Eg5 inhibitor is an excellent candidate monoclonal antibody that targets a tumor-specific surface antigen. This allows for selective delivery of therapeutic agents to cancer cells, directly addressing the therapeutic index problem by minimizing systemic exposure. Preclinical Eg5-targeting ADCs and DACs such as dKIF976 have shown promise [230].

Targeted protein degradation (PROTACs): Targeted protein degradation has emerged as a powerful therapeutic modality, with Proteolysis-Targeting Chimeras (PROTACs) at its forefront. PROTACs are heterobifunctional molecules engineered to eliminate specific proteins of interest from the cell. Structurally, they consist of two distinct ligands joined by a chemical linker: one ligand binds to the target protein, while the other recruits an E3 ubiquitin ligase, such as Cereblon (CRBN) or Von Hippel-Lindau (VHL) [231, 232].

The mechanism of action is fundamentally different from that of traditional small-molecule inhibitors, which typically function through occupancy-driven inhibition of a protein's active site. Instead, PROTACs act catalytically by inducing proximity between the target protein and the E3 ligase. This proximity facilitates the tagging of

the target protein with ubiquitin chains, marking it for recognition and subsequent degradation by the 26S proteasome. Because a single PROTAC molecule can induce the degradation of multiple target proteins before it is eventually metabolized, this event-driven, catalytic process can elicit a more profound and durable biological response[231].

This strategy has been effectively applied to Eg5. As demonstrated in the provided research, a Eg5 inhibitor can be chemically linked to an E3 ligase ligand to create a potent degrader. For example, the optimal compound from the study, compound 21, induced dose- and time-dependent degradation of Eg5, leading to G2/M cell cycle arrest and apoptosis. Crucially, this Eg5-targeting PROTAC exhibited significantly greater tumor growth inhibition in a xenograft model compared to its parent inhibitor compound alone. This underscores the potential of PROTAC technology to not only degrade previously "undruggable" targets but also to achieve superior efficacy and a more sustained therapeutic effect than conventional inhibitors [231].

The evolution of these discovery strategies recontextualizes the contributions of this study. By identifying new, high-affinity scaffolds for the L5 pocket, this thesis provides essential starting points that could be optimized into next-generation inhibitors or, perhaps more significantly, serve as ideal "warheads" for advanced modalities such as ADCs or PROTACs. The exceptional binding energies and novel scaffolds identified herein could therefore represent a critical enabling step in solving the "payload problem," providing the potent targeting component necessary to unlock the full therapeutic potential of these more advanced and tumor-selective strategies.

2.2.4 Pathophysiological implications of Eg5

The preceding sections established the indispensable role of Eg5 in the mechanics of mitotic spindle assembly. However, the importance of Eg5 extends beyond its canonical mitotic function to nonmitotic processes, particularly neuronal development. Consequently, any dysregulation in its expression or function has profound pathological consequences, positioning Eg5 at the crossroads between cancer progression and developmental disorders. This section reviews the key pathophysiological implications of Eg5 dysregulation, focusing on its multifaceted role in specific diseases.

2.2.4.1 Eg5 in cancer: Overexpression, prognosis, and molecular effects

A recurring theme in oncology is the significant overexpression of the KIF11 gene, leading to elevated levels of Eg5 protein in a wide array of human cancers compared to normal tissues [180, 233, 234]. This overexpression is not merely a bystander effect of proliferation but is considered an active contributor to tumorigenesis. Due to their high proliferation rates and inherent genomic instability, cancer cells exist under constant "mitotic stress." This state makes them exceptionally dependent on the flawless execution of mitotic machinery for survival [235, 236]. Eg5, as a key component of this machinery, thus becomes an "Achilles' heel," representing a non-oncogene addiction where cancer cells are far more vulnerable to its inhibition than their normal counterparts. The importance of Eg5 is supported by studies conducted in almost all major cancer types, in which its expression level and function are intricately linked to tumor biology [49, 184, 233, 237].

Lung and head-neck cancers

Eg5 is frequently upregulated in lung cancers, including non-small cell lung cancer (NSCLC), and its overexpression contributes to genomic instability and tumor progression [238]. In lung adenocarcinoma (LUAD), this is driven by specific oncogenic pathways, such as inactivation of the tumor-suppressive miRNA miR-30a by the long non-coding RNA lncRNA VPS9D1-AS1 [239]. Similarly, high Eg5 expression has been reported in laryngeal and head and neck squamous cell carcinomas (HNSCC), where it is critical for tumor cell viability [185].

Gastrointestinal (GI) cancers

The GI system is another major area where Eg5 is frequently overexpressed and is functionally important. Eg5 is a powerful independent predictor of poor overall and disease-free survival in hepatocellular carcinoma (HCC) [240].

KIF11 is one of the most strongly overexpressed motor protein genes in colon cancer. High Eg5 levels correlate with invasive tumor size (T3/T4) and distant metastasis (M1) [241, 242]. Mechanistically, Eg5 silencing not only reduces tumor growth but also sensitizes colon cancer cells to oxaliplatin by affecting the p53 and GSK3β signaling pathways [243]. High KIF11 expression is associated with poor patient survival in pancreatic cancer (PDAC). Beyond its mitotic role, recent studies have revealed a link between Eg5 and tumor metabolism. Eg5 interacts with and stabilizes SREBP2, the

main regulator of cholesterol biosynthesis. The resulting activation of the mevalonate (MVA) pathway enhances pancreatic cancer cell growth and migration, indicating that Eg5 can support tumor progression through metabolic reprogramming [37, 244, 245].

In gastric cancer, Eg5 inhibition has a dual effect, halting cell division and inhibiting angiogenesis by reducing the secretion of factors such as VEGF via the PI3K/Akt and ERK signaling pathways [246].

Breast and gynecological cancers

Eg5 is extensively implicated in breast and gynecological cancers, and its overexpression is linked to advanced clinical stages, higher recurrence rates, and poorer overall survival [247–250].

Molecular pathways in breast cancer: The stability of the Eg5 protein is actively maintained through specific molecular pathways. For instance, the E3 ligase adapter TRAF4 protects Eg5 from degradation by inhibiting the ubiquitin ligase Smurf2. The resulting TRAF4-Eg5 axis leads to stably high levels of Eg5, which in turn accelerates proliferation and suppresses apoptosis [251].

Furthermore, Eg5 plays a crucial role in the biology of cancer stem cells (CSCs), particularly in aggressive subtypes like triple-negative breast cancer (TNBC). Eg5 inhibition has been shown to reduce CSC markers (e.g., ALDH1), while KIF11 overexpression enhances stem-like properties [252, 253]. This positive correlation between Eg5 and the breast cancer stem cell phenotype establishes Eg5 as an attractive therapeutic target for eliminating not only the bulk tumor but also the resilient, recurrence-driving CSC population.

Ovarian Cancer: The prognostic significance of Eg5 is also pronounced in ovarian cancer. In serous ovarian carcinomas, KIF11 overexpression correlates strongly with high histological grade, advanced stage, and lymph node metastasis. Consequently, it serves as a significant independent indicator of shorter overall and progression-free survival for patients [249, 250].

Urogenital, CNS, and Hematological Malignancies

Urogenital cancers: In prostate cancer, KIF11 overexpression is associated with a higher T stage and bone metastases, whereas nuclear accumulation of the protein is a particularly poor prognostic sign [254, 255]. In clear cell renal cell carcinoma (RCC),

high Eg5 expression correlates with a high nuclear grade and advanced TNM stage [256].

Central nervous system tumors: In glioblastoma (GBM), the most aggressive primary brain tumor, Eg5 levels are significantly elevated. Its inhibition has been shown to arrest cell division, reduce cell migration, and reverse EMT characteristics [257]. Furthermore, Eg5 is critical for the self-renewal of glioblastoma stem-like cells, making it a multifaceted target for this difficult-to-treat cancer [258].

Hematological malignancies: Among hematological malignancies, the link between Eg5 function and disease progression is most clinically established in Multiple Myeloma (MM). In this context, Eg5 has been validated as a compelling therapeutic target, with the inhibitor filanesib (ARRY-520) demonstrating notable clinical response rates, particularly when used in combination therapies [259].

2.2.4.2 Eg5 in non-cancer pathological conditions

In contrast to the consequences of its overexpression in cancer, the loss of Eg5 function highlights its critical role in normal development and cellular functions.

Genetic diseases related to KIF11 mutations

Inherited heterozygous loss-of-function mutations in KIF11 cause rare genetic diseases in humans. Among these is microcephaly with or without Chorioretinopathy, Lymphedema, or Intellectual Disability (MCLMR), an autosomal dominant inherited syndrome [260]. This condition, linked to KIF11 haploinsufficiency, presents with a variable spectrum of symptoms, including microcephaly (86% of patients), ocular problems, such as chorioretinopathy (78%), and lymphedema (46%) . Microcephaly is thought to stem from apoptosis of neural progenitor cells due to mitotic failure during brain development [82, 261]. Ocular defects are linked to Eg5's role in maintaining the structure of retinal photoreceptors [262]. Another related genetic disease is Familial Exudative Vitreoretinopathy (FEVR), which can be caused by KIF11 mutations [263]. The complete absence of Eg5 (homozygous knockout) is embryonically lethal in mouse models, confirming its fundamental importance in life [82].

Eg5 in nervous system and neurological diseases

Beyond its mitotic functions, Eg5 plays important roles in the maintenance of non-dividing nerve cells, contributing to the organization of axons and dendrites [264]. Consequently, dysregulation of Eg5 function has been directly linked to the pathology

of Alzheimer's disease (AD). In AD, the accumulation of pathological proteins—specifically beta-amyloid (Aβ) and hyperphosphorylated tau—has been shown to inhibit Eg5 motor activity [265, 266]. This inhibition disrupts critical cellular processes, contributing to synaptic dysfunction and eventual neuronal cell death. This presents a "loss-of-function" pathology, contrasting with its "gain-of-function" role in cancer. Underscoring this connection, preclinical studies have demonstrated that increasing Eg5 levels in AD mouse models can protect against cognitive deficits, suggesting that restoring or supporting Eg5 function could represent a novel therapeutic strategy for neurodegeneration [267].

Eg5 in immunological and infectious conditions

The Eg5 motor is also critical for immune system functions. The HIV-Tat protein has been shown to directly bind to and allosterically inhibit Eg5, leading to mitotic arrest and death of CD4+ T lymphocytes, which helps explain T cell depletion in AIDS [129]. Furthermore, Eg5 is essential for the effector function of Natural Killer (NK) cells; it is required for the transport of cytotoxic granules to kill targets, such as the fungus *Cryptococcus*, explaining another facet of immunosuppression in HIV-infected patients [268]. Preliminary links have also been made between altered KIF11 expression and other conditions, such as pulmonary arterial hypertension (PAH) and the systemic response to COVID-19 [198, 269].

2.2.4.3 Conclusion: the "two-faced" nature of Eg5 and the therapeutic imperative

The Eg5 motor protein holds a position in the scientific world similar to the two faces of Janus. On one hand, it is an "undesirable overactive" target in cancer, whose deep involvement in multiple hallmarks of the disease validates it as a high-value therapeutic target. On the other hand, it is a "protective deficient" element in developmental and neurological conditions, where its absence or dysfunction leads to severe pathology.

The dual nature of Eg5's pathophysiology provides the ultimate and most powerful rationale for this thesis. The devastating effects of its dysfunction underscore the necessity of highly selective inhibitors. The opposing roles of Eg5 in health and disease define a critical 'therapeutic window': while cancer cells become "addicted" to high

levels of Eg5 activity for their survival, most healthy, non-dividing tissues, such as mature neurons, have a low requirement for it [167]. This difference in dependency forms the fundamental basis for selective targeting of cancer cells. An ideal therapeutic agent must potently inhibit the overactive Eg5 in cancer cells without significantly affecting its basal function elsewhere, a goal that can best be achieved by targeting a non-conserved, specific allosteric site [132]. Our focus on discovering novel inhibitors for the L5 allosteric pocket is directly aligned with this critical therapeutic imperative: to design compounds that can precisely and selectively shut down the Eg5 engine in cancer, leveraging the very dependency that makes the disease so aggressive while sparing the essential physiological roles of this "motor of life."

2.3 Computational Approaches in Drug Discovery

Computer-Aided Drug Design (CADD) is the cornerstone of modern pharmaceutical research, providing a powerful framework for identifying and optimizing novel drug candidates through in silico methodologies [270]. By leveraging computational power to screen vast chemical libraries and predict molecular interactions, CADD significantly accelerates the discovery pipeline, reduces the financial burden associated with experimental high-throughput screening (HTS), and increases the probability of identifying effective hit compounds [271]. Over the last two decades, the exponential growth in computational capacity coupled with broader access to macromolecular structures and large-scale chemical data has solidified the role of CADD in drug discovery [272]. The methodologies within CADD are broadly categorized into two primary strategies based on the availability of the three-dimensional (3D) structure of the biological target: structure-based drug design (SBDD) and ligand-based drug design (LBDD) [273].

2.3.1 Ligand-based drug design

Ligand-Based Drug Design (LBDD) methods leverage the collective knowledge of known active compounds to infer the necessary features for biological activity. This approach is indispensable when the 3D structure of a target macromolecule is unknown, and serves as a powerful complementary strategy even when it is available [274]. The core of LBDD is the molecular similarity principle: the assumption that structurally similar compounds are likely to exhibit similar biological functions [1].

While traditional two-dimensional (2D) similarity methods, such as fingerprint-based comparisons, are valued for their speed, the field has increasingly shifted toward 3D approaches. These methods, which compare 3D molecular shapes and pharmacophoric features, are particularly effective for "scaffold hopping"—the discovery of topologically novel compounds that retain the essential interaction features of a known active, thereby accessing new and potentially more advantageous chemical space [275].

2.3.1.1 Shape similarity screening

Shape-similarity screening is a robust LBDD technique that identifies novel candidate compounds by comparing their 3D volumes and shapes against those of known active ligands [276]. his method is a powerful engine for lead discovery and scaffold hopping, as it can identify compounds that are structurally distinct from the reference ligand but possess the requisite steric compatibility for the target's binding site [275]. The technique is implemented through various algorithms, including descriptor-based methods, surface-based comparisons using spherical harmonics, and overlay approaches that utilize atom-centered Gaussian functions [277]. Because molecular flexibility is a critical determinant of shape, modern applications consider an ensemble of low-energy conformers for each compound. The advent of GPU-accelerated algorithms, such as those implemented in OpenEye FastROCS [278] and the Schrödinger Suite [198] , has made it feasible to screen millions of conformers in a time-efficient manner, thereby improving the throughput of shape-based virtual screening [279]. In this thesis, GPU-accelerated shape screening was employed as the initial, broad filter in Track 1 of our virtual screening campaign, enabling the rapid prioritization of compounds from the ZINC20 Anodyne database that shared the general 3D shape of known Eg5 inhibitors.

2.3.1.2 Phase pharmacophore modeling and screening

A pharmacophore is an abstract representation of key steric and electronic features (e.g., hydrogen bond donors/acceptors, aromatic rings, and hydrophobic centers) and the spatial arrangement necessary for a compound to interact with a specific biological target. Pharmacophore models serve as 3D search queries for virtual screening, allowing for rapid identification of compounds that possess the required interaction motifs [280, 281]. These models can be generated using two approaches.

Ligand-based pharmacophore modeling: In the absence of a target structure, a pharmacophore hypothesis is inferred from a set of known active ligands by identifying their common chemical features. This is a classical CADD approach [282].

Structure-based pharmacophore modeling: When a protein-ligand complex structure is available, a pharmacophore can be derived directly from the key interactions observed in the binding site. This approach, exemplified by the "e-Pharmacophore" methodology, often yields more precise and target-specific models by generating energetically optimized hypotheses from scoring functions like Glide XP [283].

A critical step in developing a reliable pharmacophore is its rigorous validation to ensure the model can effectively distinguish known active compounds from inactive "decoy" compounds. [284]. This is typically performed by screening a benchmark dataset, such as those provided by the DEKOIS2.0 database, which is specifically curated for validating virtual screening protocols [285–287]. The performance of a pharmacophore model is then quantified using metrics, such as the Receiver Operating Characteristic (ROC) curve and its associated Area Under the Curve (AUC) value [288], which measures the model's overall ability to separate actives from decoys. In this thesis, a structure-based pharmacophore model was developed from the Eg5-K858 crystal structure and rigorously validated against the DEKOIS2.0 KIF11 benchmark set. The high AUC value obtained confirmed the model's predictive power, justifying its use as a primary filter to enrich for potential Eg5 inhibitors in our screening workflow (see Section 4.1.1.1).

2.3.2 Structure-based drug design

When the three-dimensional structure of a biological target is elucidated at atomic or near-atomic resolution, typically through X-ray crystallography, cryo-electron microscopy (cryo-EM), or NMR spectroscopy, Structure-Based Drug Design becomes a preeminent strategy for rational inhibitor discovery [289, 290]. In contrast to LBDD, which infers requirements from active ligands, SBDD directly leverages the architectural details of the binding site of the target. This "receptor-centric" approach enables the design or identification of compounds with high shape and chemical complementarity, akin to designing a key for a known lock. The ultimate goal is to computationally predict and prioritize compounds that form stable high-affinity

complexes with the target. The most prominent and widely used technique in the SBDD toolbox is molecular docking [291].

2.3.2.1 The ZINC molecular database

Modern virtual screening campaigns are made possible by the existence of massive, publicly accessible chemical databases. Among the most widely used databases is the ZINC database, a free, curated collection of commercially available compounds prepared specifically for virtual screening [11]. Now containing billions of unique compound structures in ready-to-use 3D formats with precalculated properties and annotations, a key feature of ZINC is its organization into various subsets tailored for different discovery strategies.

For this thesis, two specific ZINC20 subsets were strategically utilized:

The ZINC Anodyne Subset: This collection of approximately 11 million "lead-like" compounds was used for the focused screening in Track 1. The "Anodyne" designation signifies that this subset has been computationally filtered to remove molecules with reactive or promiscuous functional groups (e.g., PAINS - Pan-Assay Interference Compounds), which are known to cause artifacts in biological assays. This pre-curation provides a higher-quality, lower-risk chemical space ideal for a traditional, structure-based virtual screening campaign aimed at identifying high-confidence hits.

The ZINC ML-Ready Subset: For the large-scale active learning campaign in Track 2, the ~1 billion compound "ML" (machine learning) subset was employed. This ultra-large library is specifically prepared for AI-driven applications and is provided with pre-calculated 2D Morgan fingerprints for each molecule. This pre-processing is critical, as it eliminates the computationally expensive step of generating these descriptors on-the-fly, making it feasible to train and run machine learning models on a library of this immense scale.

The sheer scale and pre-processed nature of these two distinct ZINC subsets provided the essential chemical space for our comprehensive, dual-track inhibitor discovery efforts.

2.3.2.2 Molecular docking: Predicting and ranking molecular interactions

Molecular docking is a computational method that aims to predict the preferred binding orientation(s) and conformation(s) of a ligand—referred to as its "pose"—

within a target's binding site. This process involves two fundamental components: a search algorithm to explore the vast conformational space of the ligand and a scoring function to evaluate and rank the generated poses [292]. In a typical virtual screening campaign, libraries containing thousands to billions of compounds are computationally docked and the resulting scores are used to prioritize a manageable number of candidates for subsequent experimental testing [293].

The field offers a wide array of docking programs, including academic staples such as AutoDock Vina [294, 295], and industry standards such as Glide [296], GOLD [297], and FRED [298]. These programs differ in their underlying methodologies. Search algorithms can range from systematic, grid-based methods to more complex stochastic approaches, such as genetic algorithms (which mimic natural selection to evolve better poses) or Monte Carlo simulations (which randomly sample conformational space). Scoring functions are equally diverse and can be broadly classified into physics-based (calculating terms such as van der Waals and electrostatic energies), empirical (using regression-fitted terms to reproduce experimental binding data), and knowledge-based (deriving potentials from statistical analysis of known protein-ligand structures) formalisms. Ultimately, all docking programs must negotiate a critical trade-off between computational speed and predictive accuracy [299, 300].

Glide Docking: Developed by Schrödinger, Glide is an industry-leading program renowned for its accuracy and its use of a hierarchical screening approach. It employs a series of scoring functions with increasing computational rigor—from High-Throughput Virtual Screening (HTVS) for rapid initial placement, to Standard Precision (SP) for refinement, and finally to Extra Precision (XP) for highly accurate scoring and pose prediction. This tiered approach allows for an efficient balance of speed and precision, making it ideal for focused screening campaigns. In this thesis, the full Glide HTVS-SP-XP cascade was employed in Track 1 to meticulously filter and prioritize candidates from the enriched Anodyne library.

Qvina2 Docking: Qvina2 is a fork of the widely used academic program AutoDock Vina, specifically optimized for speed and parallel processing, making it exceptionally well-suited for large-scale virtual screening. While retaining the core search algorithm and scoring function of its predecessor, Qvina2's enhanced performance is critical for its application in iterative, AI-driven workflows [301]. In this thesis, Qvina2 served as the essential high-throughput docking engine for the DeepDocking active learning

campaign in Track 2, where its ability to rapidly dock millions of compounds per iteration was indispensable for training the deep learning models.

A significant and well-documented challenge in virtual screening is the propensity of scoring functions to produce a high rate of false positives, thereby ranking inactive compounds as promising candidates. This underscores the necessity of validating the docking protocol before its application in a large-scale screening campaign. Validation is a multi-step process. First, the protocol's ability to reproduce the experimentally observed binding mode is tested via "redocking," where a co-crystallized ligand is removed and docked back into its own receptor; a low root-mean-square deviation (RMSD) between the docked and crystal poses indicates good pose prediction accuracy [302, 303]. Second, and more critically, the ability of the protocol to distinguish a set of known active compounds from a much larger set of presumed inactive "decoy" compounds must be assessed using a benchmark dataset, such as those provided by DEKOIS2.0.

While the standard ROC-AUC is a useful global metric of a model's ability to distinguish actives from decoys, its utility for virtual screening is limited because it gives equal weight to the entire dataset. In practice, researchers will only experimentally test a very small fraction of top-ranked compounds (e.g., the top 1%). The standard ROC-AUC does not specifically reflect the model's performance in this critical top-tier ranking [304].

To provide a more rigorous and practically relevant assessment, a predictive Receiver-Operating Characteristic (pROC) curve and its corresponding pROC-AUC metric were developed. The pROC curve is a semi-logarithmic plot that magnifies the initial segment of the ROC curve using a logarithmic scale for the False Positive Rate (FPR) on the x-axis. This focuses on the evaluation of the model's performance at very low FPRs, directly corresponding to top-ranked virtual hits. Consequently, the pROC-AUC value is a measure of the “early enrichment” capability of a protocol. A high pROC-AUC value provides much stronger confidence that the docking protocol is not just globally accurate but is specifically proficient at placing true active compounds at the top of the ranked list, which is the primary goal of any virtual screening campaign [305].

In this study, both the Glide (for Track 1) and Qvina2 (for Track 2) docking protocols were rigorously validated using the DEKOIS2.0 benchmark set. The calculated pROC-AUC values confirmed that both protocols exhibited strong early enrichment performance, thereby providing confidence in their ability to successfully prioritize potential Eg5 inhibitors from the large ZINC libraries (see Sections 4.1.1.2 and 4.1.1.3).

2.3.2.3 Post-docking refinement: Molecular dynamics simulations and free energy calculations

While molecular docking provides an invaluable snapshot of a potential protein-ligand interaction, its reliance on a largely rigid protein structure and simplified scoring functions are inherent limitations. Static binding poses provide a valuable but incomplete picture, often failing to account for the dynamic nature of biological systems, including subtle protein conformational changes upon ligand binding (induced fit), and the explicit role of water compounds. To address these limitations and provide a more rigorous assessment of binding stability and affinity, molecular dynamics (MD) simulations have become a critical post-docking refinement step [306].

MD simulations offer a dynamic atomistic view of a biomolecular system by solving Newton's equations of motion for every atom over a period of time, typically in the range of nanoseconds to microseconds. These simulations are performed using powerful, specialized software packages such as GROMACS [307], AMBER [308], NAMD [309], and CHARMM [310]. For this thesis, Desmond, a high-performance MD engine developed by D. E. Shaw Research and integrated into the Schrödinger Suite, was utilized. Desmond is particularly well-suited for drug discovery applications due to its optimization for running long-timescale simulations on Graphics Processing Units (GPUs), leveraging scalable algorithms to efficiently model complex biomolecular interactions [311].

The standard process for a conventional (unbiased) MD simulation involves placing the docked protein-ligand complex into a simulation box, solvating it with explicit water molecules, and adding ions to mimic physiological conditions. After an initial energy minimization and an equilibration period to stabilize the system's temperature and pressure, a "production" simulation is run [312]. While the unbiased MD

simulations employed in this thesis are ideal for assessinGbinding pose stability, it is worth noting that more specialized techniques such as steered MD (SMD) or umbrella sampling can be used to probe ligand dissociation pathways or calculate rigorous absolute free energies of binding, respectively. The resulting trajectory provided a wealth of information [313]. Key analyses include:

Root-Mean-Square Deviation (RMSD): The RMSD of the protein backbone and heavy atoms of the ligand were monitored over time. A low, stable ligand RMSD relative to the protein indicated that the compound had a stable binding pose and did not drift out of the binding pocket [314].

Interaction Analysis: The persistence of specific interactions (e.g., hydrogen bonds, hydrophobic contacts, and salt bridges) between the ligand and protein residues can be tracked throughout the simulation. This dynamic analysis is far more informative than the single snapshot provided by docking, confirming the interactions that are truly stable and significant [314].

The persistence of specific interactions (e.g., hydrogen bonds, hydrophobic contacts, and salt bridges) between the ligand and protein residues can be tracked throughout the simulation. This dynamic analysis is far more informative than the single snapshot provided by docking, confirming the interactions that are truly stable and significant.

In this thesis, MD simulations served as a critical final filter used to confirm that the top-ranked candidates from both screening tracks formed stable complexes with Eg5 and maintained key binding interactions over time, thereby filtering out promising-looking but dynamically unstable binding poses.

Furthermore, the trajectories generated by MD simulations can be used to calculate binding free energies with greater accuracy than docking scores. End-point methods, such as MM/PBSA and MM/GBSA (molecular mechanics/Poisson-Boltzmann or generalized Born surface area), calculate the free energy of binding by averaging energetic contributions over hundreds or thousands of snapshots from the equilibrated trajectory [315]. The binding free energy (ΔGbind) was calculated as the difference between the energy of the complex and the energies of the individual proteins and ligands, including contributions from the molecular mechanics energy (ΔEMM), solvation free energy (ΔGsolv), and conformational entropy (−TΔS). While computationally intensive and often neglecting the entropic term for throughput, these

methods generally provide a better correlation with experimental binding affinities than docking scores and are invaluable for differentiating between high-quality candidates [316]. This approach was central to the final prioritization of compounds for experimental testing in our study.

2.3.3 ADMET prediction

For a compound to become a successful drug, its potent activity against its target is not sufficient; it must also possess a favorable safety and pharmacokinetic profile. This profile is often summarized by the acronym ADMET: Absorption, Distribution, Metabolism, Excretion, and Toxicity. Poor ADMET properties are a major cause of late-stage failures in drug development [317, 318]. Therefore, the computational prediction of these properties at an early stage is crucial for prioritizing candidates with a higher likelihood of clinical success. One of the most influential guidelines for predicting good oral bioavailability is Lipinski's Rule of Five, which sets thresholds for molecular weight, lipophilicity (LogP), and hydrogen bonding capacity [319]. Throughout the filtering cascade of this thesis, particularly after the initial large-scale screenings, ADMET-based filters, including Lipinski's rules and other criteria, such as polar surface area, were applied to eliminate compounds with predicted unfavorable properties and enrich for more "drug-like" candidates.

2.4 Artificial Intelligence in Drug Discovery

The traditional drug discovery pipeline, fraught with high costs, long timelines, and staggering attrition rates, is undergoing a paradigm shift driven by the integration of Artificial Intelligence (AI) [320, 321]. AI, a broad field of computer science encompassing Machine Learning (ML) and its subfield Deep Learning (DL), leverages complex algorithms to learn from vast datasets. This enables predictive modeling and automated decision-making capabilities that can accelerate and de-risk nearly every stage of pharmaceutical R&D, from initial target identification to preclinical development [322, 323].

The trajectory of AI in drug discovery has evolved from early expert systems and classical QSAR models to the sophisticated deep learning architectures of today [321, 324]. This rapid evolution has been fueled by a confluence of three critical factors: the explosion of "big data" from high-throughput technologies (e.g., genomics,

transcriptomics, proteomics, HTS); significant advances in computational hardware, particularly the parallel processing power of Graphics Processing Units (GPUs); and the concurrent development of novel AI algorithms and model architectures [320, 322, 323]. The impact of this convergence is profound. Landmark achievements, such as DeepMind's AlphaFold, which can now predict protein structures with experimental-level accuracy, have fundamentally altered the landscape of structural biology and target-based design [325]. Furthermore, the progression of AI-designed drugs into clinical trials by companies like Insilico Medicine and BenevolentAI has provided tangible proof that AI has moved from a conceptual novelty to a practical and indispensable tool in modern drug discovery [320].

AI-driven strategies are now being applied with increasing sophistication across the discovery pipeline:

Target Identification and Validation: This initial stage, which determines the entire direction of a drug discovery program, involves identifying biological compounds whose modulation could lead to a therapeutic effect. AI excels at integrating and mining heterogeneous, large-scale data to uncover novel disease-target relationships that may be missed by human researchers. For example, Natural Language Processing (NLP) models, such as those based on transformer architecture (e.g., BioBERT), can scan millions of scientific articles, patents, and clinical trial records to perform named-entity recognition and relationship extraction, thereby identifying latent connections between genes, pathways, and diseases [326]. Furthermore, Knowledge Graphs (KGs) structure this extracted information into a vast network of interconnected biological entities. By applying graph-based algorithms to these KGs, researchers can infer non-obvious relationships, such as predicting that a protein involved in one pathway might also play a role in a disease previously thought to be unrelated, thus generating novel testable hypotheses [327]. Perhaps most transformatively, the advent of highly accurate protein structure prediction models such as AlphaFold allows for the immediate computational assessment of a novel target's "druggability"—the presence of viable binding pockets—on a proteome-wide scale, a task that previously required years of experimental work [328].

Hit Discovery and De Novo Design: Once a target is validated, the search for "hit" compounds begins. AI accelerates this process in two ways. First, predictive models are used to screen ultra-large virtual libraries. Machine learning models, such as

Random Forests, Gradient Boosting Machines, or Deep Neural Networks, are trained on large datasets of known ligand-target interactions (e.g., from databases such as ChEMBL) to learn a complex function that maps a compound's chemical features to its biological activity. These trained models can then be deployed to rapidly provide a 'hit-likeness' score for billions of compounds, prioritizing a small, enriched subset for more resource-intensive analysis like docking or experimental testing [329, 330]. Second, a more advanced strategy is generative de novo design. Instead of merely finding compounds in existing libraries, these models create entirely new compounds. Architectures such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and transformers learn the underlying "grammar" of chemical structures and stability from vast datasets. They can then be conditioned to generate novel compounds that are not only chemically valid but also optimized a priori for specific properties, such as a high predicted affinity for the target protein and favorable ADMET characteristics [219, 331].

Lead Optimization and ADMET Prediction: A "hit" compound rarely possesses all the properties required for a drug. The lead optimization phase involves iteratively modifying its chemical structure to improve potency, selectivity, and ADMET profile. AI provides crucial guidance for this multi-parameter optimization (MPO) challenge. Instead of the traditional, often intuitive, 'one-compound-at-a-time' synthesis cycle, ML models can predict the impact of a proposed chemical modification on a whole suite of properties simultaneously [332]. For instance, specialized Graph Neural Network (GNN) models can predict a compound's potential to block the hERG potassium channel (a key indicator of cardiotoxicity), identify the most likely sites of metabolism by cytochrome P450 enzymes, or flag it for potential hepatotoxicity based on its substructures [333, 334]. By integrating these predictions early in the design cycle, chemists can steer optimization efforts away from compounds destined for failure, thereby significantly improving the efficiency and success rate of preclinical development.

2.4.1 Navigating ultra-large chemica spaces: From hierarchical screening to AI-acceleration

The advent of "make-on-demand" chemical libraries from vendors such as Enamine has revolutionized hit discovery, expanding the accessible chemical space from

millions to tens of billions of compounds [335]. While this presents an unprecedented opportunity to find novel chemical matters, it also creates a significant computational challenge: the brute-force virtual screening of such ultra-large libraries with conventional docking methods is computationally intractable, as it would require years of continuous computation even on large clusters [336].

Several strategies have been developed to address these challenges. Early approaches relied on hierarchical filtering, where a massive library was sequentially passed through a cascade of increasingly complex and computationally expensive filters. This often begins with simple 2D property filters (e.g., ADMET rules and substructure alerts), followed by faster 3D methods such as pharmacophore or shape-similarity screening, to progressively reduce the library size before committing to full molecular docking on a much smaller, enriched subset [337, 338]. While effective, this approach risks discarding novel scaffolds that may not conform to predefined filters based on known activities.

The true breakthrough in navigating these vast chemical spaces has resulted from the application of Artificial Intelligence. Machine learning models trained on smaller, representative datasets can learn to predict the outcome of a more expensive calculation (such as a docking score) for the rest of the library. This allows for intelligent pre-selection of candidates, drastically reducing the number of compounds that require full docking [12, 339]. This principle has given rise to a new generation of AI accelerated screening methods.

One of the most powerful and widely adopted methods is active learning, which is a strategy in which the machine learning model is iteratively improved during the screening process. The DeepDocking (DD) protocol, developed by Gentile et al., is a prime example of an active learning framework designed specifically for ultralarge library virtual screening [12, 340]. It fully combines the physical realism of the established docking programs with the predictive speed of deep learning. The protocol operates in a cyclical iterative manner.

Initial docking and training: The process begins by docking a small, randomly sampled subset (e.g., 0.1%) of the massive library using a conventional docking program (e.g., Glide and Qvina2). The resulting docking scores and their corresponding molecular fingerprints were used to train an initial Deep Neural

Network (DNN). The DNN learns to approximate the complex, nonlinear relationship between a compound's 2D structure (represented by its fingerprint) and its predicted 3D binding score from docking [341].

Inference and prioritization: The trained DNN, which can make predictions thousands of times faster than the docking program, was then used to rapidly screen the entire remaining portion of the library. It assigns a predicted "hit-likeness" score to every compound, prioritizing a new smaller subset of compounds that are most likely to achieve high docking scores [341].

Iterative refinement: This newly prioritized subset is docked using a conventional program. New, high-quality data were added to the training set, and the DNN was retrained. This new, more intelligent model was used to screen the library again [341].

The active learning cycle was repeated several times. With each iteration, the training set became progressively enriched with high-scoring compounds, allowing the DNN to become increasingly accurate in identifying the most promising areas of the vast chemical space. This intelligently guided exploration allows researchers to focus their limited computational resources on the compounds that matter the most.

The DD protocol has been demonstrated to accelerate the screening of billion-scale libraries by 50- to 100-fold compared with conventional brute-force docking, while successfully retrieving the vast majority of top-scoring hits [340]. It has been successfully applied to discover novel inhibitors for a range of targets, including the SARS-CoV-2 main protease, by screening libraries of as many as 40 billion compounds [342].

The application of the DeepDocking protocol constitutes the second and most expansive screening track in this thesis. By employing this state-of-the-art active learning strategy, we were able to efficiently explore the 1 billion-compound ZINC20-ML library for novel Eg5 inhibitors, a scale that would be intractable with conventional methods alone. This approach allowed for a much broader search for chemical novelty than the more focused classical screening of Track 1, maximizing our chances of discovering structurally unique and potent hit compounds.

2.5 Hypothesis

The extensive review of the literature presented in the preceding sections culminates in a clear yet paradoxical conclusion: the kinesin motor protein Eg5 is an exceptionally well-validated and high-value target for anticancer therapy, yet no inhibitor targeting it has successfully reached the market. The literature unequivocally establishes that: (i) Eg5 is indispensable for mitotic spindle bipolarity and is thus essential for the proliferation of cancer cells; (ii) its overexpression is a hallmark of numerous human malignancies and is strongly correlated with poor clinical prognosis; and (iii) selective allosteric inhibition of its non-conserved Loop L5 pocket is a proven strategy for inducing mitotic arrest and apoptosis. The persistent clinical failure, therefore, does not invalidate the target itself but rather points to a critical deficiency in the chemical matter of the inhibitors developed to date, which have been consistently hampered by a narrow therapeutic window, dose-limiting toxicities such as neutropenia, and the emergence of resistance.

Therefore, the central challenge is the discovery of novel chemical scaffolds that can retain potent and selective L5-pocket-mediated inhibition while possessing superior pharmacological and safety profiles. This dissertation is founded on the premise that this challenge can be effectively addressed by employing a comprehensive dual-track computational strategy that combines established and state-of-the-art AI-driven methodologies to explore the chemical space on an unprecedented scale.

Based on this comprehensive understanding, this dissertation is founded on the following central hypothesis and its subordinate computationally testable components.

General hypothesis:

The central hypothesis of this dissertation is that a rationally designed, multi-step CADD workflow integrating advanced computational techniques can successfully identify novel and structurally diverse small-compound inhibitors of human kinesin, Eg5. By combining focused, structure-based screening (Track 1) with a more exploratory, AI-accelerated active learning campaign (Track 2), this study posits that it will uncover distinct chemical scaffolds targeting the validated allosteric Loop L5 pocket. It is hypothesized that these novel compounds will represent superior starting points for the development of next-generation therapeutics capable of overcoming the well-documented clinical limitations of previous candidates.

Specific hypotheses:

To address this overarching hypothesis, the following specific computationally testable hypotheses were formulated:

Hypothesis of dual-track discovery and scaffold diversity: Focused, structure-based virtual screening of the ZINC Anodyne library (Track 1) will effectively "scaffold hop" beyond the known chemical series to identify novel, active chemotypes. In parallel, the AI-driven DeepDocking campaign against the billion-scale ZINC-ML library (Track 2) will explore a much larger chemical space, leading to the discovery of entirely new and structurally diverse chemical classes that are inaccessible through conventional screening methods.

Hypothesis of stable binding and mechanistic fidelity: The top-ranked novel compounds emerging from both distinct screening tracks, when subjected to rigorous, all-atom molecular dynamics (MD) simulations, demonstrate dynamically stable and persistent binding within the L5 allosteric pocket. It is predicted that these stable binding modes will not only be energetically favorable, but also recapitulate the key intermolecular interactions critical for potent inhibition, such as hydrogen bonding with residues such as Glu116 and Arg119, and hydrophobic packing against Tyr211 and Leu214, providing a biophysical basis for a durable inhibitory effect beyond a static docking score.

H**ypothesis of superior binding energetics from AI-driven discovery:** Post-simulation analyses using binding free energy calculation methods (e.g., MM/GBSA) will provide a quantitative rationale for the high predicted affinity of prioritized candidates. It is hypothesized that the candidates from the AI-accelerated Track 2 will exhibit profoundly more favorable calculated binding free energies not only compared to known reference inhibitors, but also compared to the candidates identified via the more conservative Track 1, providing a strong energetic rationale for their superior potential potency.

Hypothesis of improved therapeutic potential through chemical novelty: The portfolio of computationally and experimentally validated lead compounds possessing novel scaffolds identified from both tracks will represent a rich set of starting points for future medicinal chemistry optimization. It is postulated that the structural diversity of these frameworks will increase the probability of identifying candidates that

circumvent the liabilities of previous inhibitor classes, offering the potential for improved ADMET profiles and reduced susceptibility to class-specific resistance mechanisms, thus directly addressing the primary obstacles that have hindered the clinical translation of Eg5 inhibitors to date.

The successful validation of these hypotheses through the computational and experimental methodologies detailed in subsequent chapters will provide a set of novel, high-potential lead compounds, thereby making a significant contribution to the ongoing quest for a clinically successful Eg5-targeted anticancer therapeutic and potentially revitalizing this promising field of research.

In summary, the literature highlights the need for novel Eg5 inhibitors and the promise of advanced computational techniques. Based on this, we hypothesize that combining structure-based and AI-driven virtual screening will yield new Eg5 inhibitor scaffolds. In the next chapter, we detail the materials and methods used to implement this strategy.

3. MATERIALS AND METHODS

3.1 In silico Calculations

3.1.1 General workflow organization

The virtual screening campaign proceeded along two parallel tracks, as illustrated in Figure 3.1.

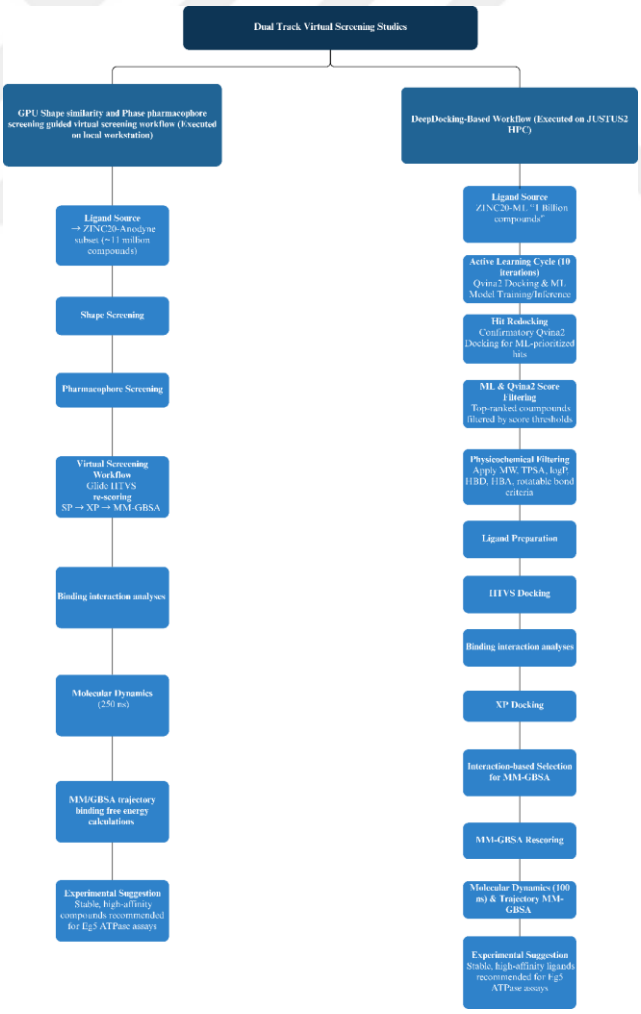

Figure 3.1 : General overview of the dual track of the virtual screening studies.

GPU Shape similarity and Phase pharmacophore screening guided virtual screening workflow

An initial library of ~11 million lead-like compounds was retrieved from the ZINC20-Anodyne subset [11]. The compounds were filtered sequentially by GPU Shape similarity screening [343] and Phase [344] pharmacophore screening queries that captured the key steric and hydrogen-bond features of the Eg5 AS-I pocket. Obtained compounds were docked with Glide in HTVS mode and re-scored using SP and XP modes and Prime MM/GBSA, respectively [283, 345]. Poses that satisfied the predefined interaction pattern were subjected to 250 ns explicit-solvent molecular dynamics simulations, followed by MM/GBSA trajectory binding free energy calculations. Compounds that remained stably bound and displayed favorable binding energies were selected for subsequent ATPase assays.

DeepDocking – Qvina2 workflow

A second virtual screening campaign was executed on the JUSTUS2 high-performance cluster between January and July 2024, starting from the 1 billion-compound ZINC20-ML molecular database [13]. Active learning procedure were carried out for 10 iterations: in each iteration a random sample was docked with Qvina2 [301], a DNN model was trained on the resulting scores, and the model was used to prioritize the next sample compounds identified through the DeepDocking pipeline [340] and subsequent filtering stages (including Qvina2 docking of ML-prioritized hits, ML score, Qvina2 score, and ADME property evaluation) were re-prepared using LigPrep (preserving existing stereoisomers and charge states) [346], redocked with Glide HTVS and XP, rescored by MM/GBSA, and evaluated in 100 ns MD simulations and MM/GBSA trajectory energy calculations. Complexes that remained stable and displayed favorable energies entered the same experimental validation queue as candidates from the first track.

The dual approach was adopted to preserve chemotype diversity while ensuring adequate coverage of ultra-large chemical space. Detailed parameters, software versions and validation metrics are provided in the following subsections.

3.1.2 Computational infrastructures and software versions

Computational studies were conducted utilizing a combination of local high-performance workstations and the JUSTUS2 high-performance computing (HPC) cluster. Job management and scheduling on the JUSTUS2 cluster were handled using

Slurm [347]. A dedicated Conda environment was configured for the DeepDocking workflow to ensure reproducibility and manage dependencies.

The Schrödinger Suite [348]was extensively used for various stages of the "GPU Shape similarity and Phase pharmacophore screening" track and for the post DeepDocking refinement of hits from the second track. Work was performed using versions ranging from 2021-4 to 2024-4 (Schrödinger, LLC, New York, NY), with specific modules and their corresponding versions within this range applied as appropriate for each stage of the research.

Key modules included:

Maestro: Used for overall project management, visualization, protein preparation (via the Protein Preparation Wizard), ligand preparation (LigPrep), grid generation, binding interaction analysis, and general molecular modeling tasks.

Epik: Employed within the Protein Preparation Wizard for predicting protonation and tautomeric states [346].

Create Shape Data File module & GPU Shape Screening module: Used for 3D shape-based virtual screening, leveraging GPU acceleration on graphics processing units.

Phase: Employed for pharmacophore model development (E-Pharmacophore), generation of conformer databases, and pharmacophore screening.

Glide: Used for all docking calculations (HTVS, SP, and XP modes).

Prime: Utilized for MM/GBSA energy calculations and minimizations.

Desmond: Employed for performing explicit-solvent molecular dynamics simulations [311].

Canvas: Used for similarity and clustering analyses [349].

For the "DeepDocking workflow", subsequent analysis, and visualization, the following software and tools were central:

DeepDocking protocol: Following the general procedures outlined by Gentile et al., with specific adaptations as described herein.

Python: Utilized for custom scripting, data parsing, workflow automation, and analysis of simulation results. Key libraries included pandas for data manipulation, and matplotlib/seaborn for generating plots [350–353].

Bash Scripting: Used for automating the multi-step workflows and managing job submissions on the HPC cluster [354].

Qvina2 v2.1: The primary docking engine for the high-throughput docking stages within the DeepDocking active learning iterations.

ADFRsuite v1.1dev: Used for receptor preparation, including hydrogen addition and PDBQT file generation [355].

PyMOL Schrödinger, LLC: Utilized for molecular visualization, selection of specific protein chains, removal of unwanted compounds, and determination of grid box coordinates [356].

Meeko v0.5.0: Utilized for ligand preparation via scripts using for converting molecular files into desired format [357, 358].

RDKit v2024.03.1: Utilized for the calculation of molecular properties (e.g., LogP, TPSA, molecular weight) for drug-likeness filtering, 3D conformer generation from SMILES strings via the script using the MMFF94 force field [359], custom Python scripts employing RDKit [360] for reading and batching SDF files, and for processing PDBQT files via the Meeko API.

MayaChemTools: Utilized to generate the scripts for 3D conformer generation from SMILES strings [361].

Conda: Utilized for managing software packages and creating isolated, reproducible environments to ensure all dependencies for the DeepDocking workflow were correctly installed and version controlled.

Docker: Used to create a containerized environment for the automated DeepDocking workflow, including protein and ligand preparation steps, ensuring standardized execution and reproducibility [362].

GraphPad Prism 9: Employed for the statistical analysis of in vitro assay data, calculation of IC_{50} values, and generation of dose-response curves [363].

The operating system for computational tasks was predominantly Linux-based. Data storage and management relied on dedicated high-speed storage solutions, including the use of temporary directories on the HPC cluster for intermediate docking files. Access to molecular databases such as ZINC20 (Anodyne subset dated March 2020, and ML subset dated March 2021; both accessed and utilized in early to mid-2024) and benchmark sets like DEKOIS 2.0 [285–287] were as described in their respective methodological sections.

3.1.3 Protein preparation

The crystal structure of Eg5 in complex with ADP and the inhibitor K858 (PDB Entry: 6G6Y; 1.80 Å) was chosen for our virtual screening studies and obtained from the RCSB Protein Data Bank [364]. The protein, ADP, and K858 atoms were retained, while all other atoms were deleted. Hydrogen atoms (pH 7.4) and any missing side chain atoms were incorporated using the Protein Preparation Wizard module [365]. During the Epik protonation step, alternative protomer/tautomer states within ±1 pK_a were sampled, and the lowest-energy states for ADP and K858 were selected. The N- and C-terminals of the protein were capped with N-Acetyl and N-Methyl Amide, respectively. Finally, the system was minimized using the OPLS4 force field [366].

For the DeepDocking workflow utilizing Qvina2, the Eg5 structure (PDB: 6g6y) underwent a separate preparation pipeline. Initially, hydrogen atoms were added using the reduced program from ADFRsuite. The resulting structure was loaded into PyMOL, where Chain A was isolated by removing Chain B, solvent compounds, organic cofactors (excluding the ligand of interest if needed for grid definition but not part of the receptor PDBQT), and metal ions. The cleaned structure containing only Chain A was saved as a PDB file. This prepared PDB file was then converted into the PDBQT format required by Qvina2 using the receptor preparation utility from ADFRsuite. This script automatically added polar hydrogens, merged non-polar hydrogens, assigned Gasteiger charges, and set AutoDock4 [357] atom types. The final receptor PDBQT file was stored in a designated directory. These protein preparation steps were incorporated into automated scripts managed within the Docker environment for the DeepDocking workflow.

3.1.4 Ligand preparation

3.1.4.1 ZINC20 - anodyne molecular database

The ZINC20 - Anodyne database, consisting of ~11 million compounds, was selected for the first path of the study. The structures were downloaded in SMILES format and converted into shape data files for shape screening using the Create Shape Data File module. The ligands were prepared using the LigPrep tool. To this end, all protomeric states at pH 7.0 ± 2.0 were created. The chiral centers of compounds with known stereochemistry were maintained, and all possible stereoisomers were created for all other chiral centers. Subsequently, all generated ligand structures were energy-minimized using the OPLS4 force field.

3.1.4.2 ZINC20 - ML ready-to-use molecular database

The "1 billion compounds" subset of the ZINC20-ML directory (100 × 10 million SMILES, about 1 billion compounds) was obtained as open-access, ready-to-use data for the DeepDocking workflow. The source archive, dated March 2021, maintains detailed stereochemistry and provides a unique ZINCID_index identification for each compound, distinguishing specific stereoisomers. The ZINC team preselected the dominant tautomeric and protonation states at pH 7.4, as necessary.

The downloaded archive contained 100 files (for example, smiles_all_*. txt, each with 10 million compounds) was extracted, and its contents were stored in a conveniently named directory. The associated archive, containing precomputed Morgan fingerprints (radius 2, 1,024 bits), was extracted to another directory (e.g., within the main DeepDocking project folder). The integrity and continuity of these files were confirmed using the DeepDocking protocol. This setup provided a foundational chemical space for active learning campaigns.

3.1.5 Development and validation of pharmacophore models.

Pharmacophore hypotheses were developed based on the ligand-protein binding interactions observed in the Eg5-K858 co-crystal structure using the E-Pharmacophore model of the Phase module, with default settings applied. The binding pocket was determined by selecting all amino acids located within 5 Å of the co-crystallized ligand, and these amino acids were used to define an excluded volume. Essential

pharmacophoric characteristics have been recognized to develop models, particularly hydrogen bond acceptors (A), hydrophobic groups (H), and aromatic groups (R).

The pharmacophore models that were created were validated using the DEKOIS 2.0 benchmark set, which includes 40 active compounds and 1,200 decoys. For this purpose, 50 conformers were generated for each compound in the benchmark set using the Create Phase Database module [344]. Subsequently, the generated database was evaluated using pharmacophore models through a Phase pharmacophore screening module. The performance of the generated pharmacophore models was evaluated against the DEKOIS 2.0 benchmark set, and the model demonstrating the best ability to distinguish known actives from decoys (based on standard metrics such as enrichment factor and ROC AUC) was selected for screening the Anodyne set.

3.1.6 Development and Validation of Docking Protocols

Glide protocol: A grid box was created around the center of the co-crystallized ligand within the protein crystal structure to delineate the binding pocket, utilizing the Grid Generation module. All side chain hydroxyl groups within 5 Å of the cocrystallized ligand, which were not part of the protein hydrogen-bonding network, were permitted to undergo free rotation. Docking studies were conducted using the Glide module with default settings. A retrospective docking study was initially performed using HTVS settings, during which the RMSD of the docked and co-crystallized K858 poses was assessed. In the second phase, the DEKOIS2.0 benchmark set was subjected to docking using the same protocol, and the pose with the highest score for each benchmark ligand was preserved. Subsequently, the poses were re-scored using the SP, XP, and MM/GBSA functions. The predictive performance of the docking protocol was assessed by generating pROC curves and calculating the area under the curve (AUC) values , which indicated its suitability for subsequent screening.

Qvina2 protocol: For protocol validation and docking with Qvina2, the grid box was defined based on the coordinates of an atom within the co-crystallized ligand K858 in chain A of the PDB entry 6G6Y. These coordinates were obtained using PyMOL, yielding the center of the grid (e.g., center_x = 10.77, center_y = 17.62, and center_z = 1.32). Standard dimensions of 20 × 20 × 20 Å (size_x, size_y, size_z) were used for the grid box. These parameters were specified in the Qvina2 configuration file. The receptor PDBQT file was prepared as detailed in Section 3.1.3. Ligands for validation

(DEKOIS 2.0 KIF11 actives and decoys) were prepared as batched SDF files (Section 3.1.10.1) and subsequently converted to the PDBQT format using the procedure described in Section 3.1.10.2. The docking runs employed default settings. After docking, heavy-atom RMSD values for the re-docked ligand and early enrichment statistics for the benchmark set were calculated, as described in the Results and Discussion section.

3.1.7 GPU shape screening

The GPU shape-screening methodology is a ligand-centric virtual screening approach aimed at identifying compounds that exhibit three-dimensional structural similarity to reference inhibitors. A reference collection of effective and structurally varied Eg5 inhibitors was established, including 32 co-crystallized ligands and eight compounds sourced from the CHEMBL database (Figure 3.2) [367]. This dataset was categorized using the Canvas Similarity and Clustering tool [293] with the following parameters: Tanimoto similarity metric; atom classification system classified by functional type: {H}, {C}, {F, Cl}, {Br, I}, {N, O}, {S}, {other}; bond hybridization; dendritic fingerprint type; and centroid linkage approach. A question was formulated using the 13 structurally distinct substances shown within the red circles in Figure 3.2. The shape data files of the Anodyne ligands were later filtered using this query with default settings without applying any similarity score threshold.

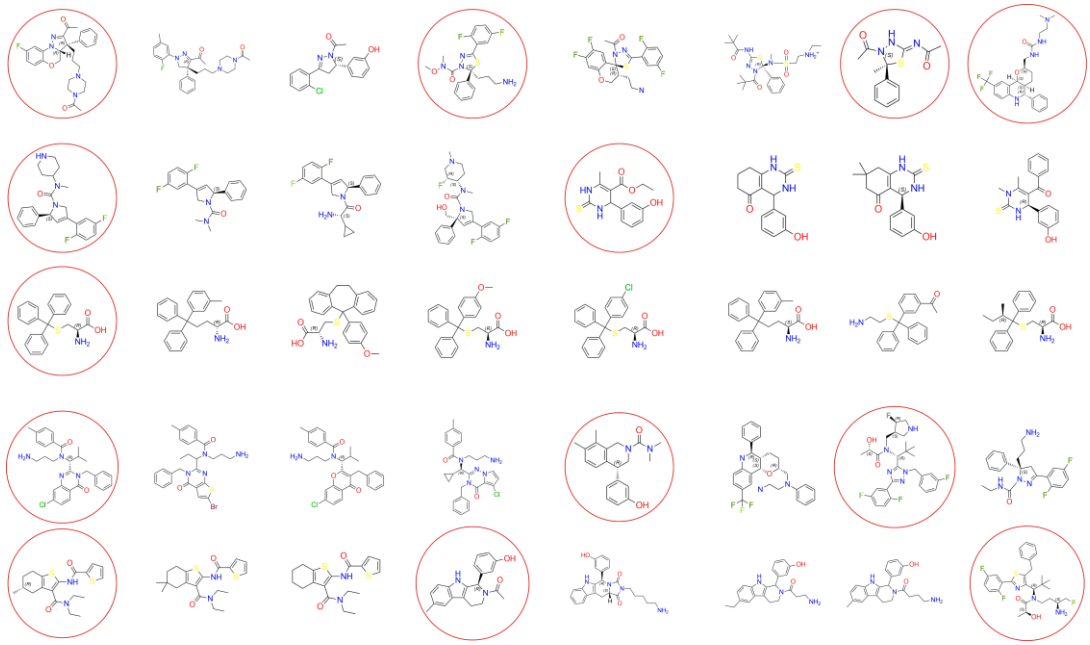

Figure 3.2 : Representative Eg5 inhibitors retrieved from the RCSB Protein Data Bank and the CHEMBL database.

3.1.8 Phase pharmacophore screening

All the compounds that survived the shape-based filter were subjected to Phase pharmacophore screening. The workflow followed the default module settings**.** Using Create Database (Phase GUI), each ligand was expanded into up to 50 low-energy conformers with ConfGen (OPLS4; energy window = 10 kcal mol^{-1}; RMSD cut-off = 1.0 Å; ring sampling = enhanced; amide trans enforcement = on). The protonation states, tautomers, and stereoisomers previously generated in LigPrep were kept unchanged, and no additional enumeration was performed at this stage. Screening employed the validated PM1 e-pharmacophore (features: H-bond acceptor, aromatic ring, and hydrophobic group) derived from the 6G6Y-K858 complex. Default Phase tolerances were retained**:** feature distance tolerance = 2.0 Å, vector/angle tolerance = 30°, and volume shell sampling = 1.0 Å. The tree-based flexible-ligand algorithm explored all conformers; at least two of the three pharmacophore features were required, aromatic ring equivalence was allowed, and duplicate alignments within 0.5 Å RMSD were discarded. Hits were exported in SDF format for subsequent docking. This procedure produced a pharmacophore-filtered library, which was immediately passed to the Glide docking stage.

3.1.9 Docking, re-scoring, and candidate selection

The compounds that had successfully passed the Phase pharmacophore filter (as detailed in Section 3.1.8) were subsequently advanced to a meticulous, hierarchical docking and rescoring protocol using the Schrödinger Suite. All computational efforts were specifically targeted at the allosteric Loop L5 pocket of Eg5. For this, a predefined grid was utilized, based on the co-crystallized K858 ligand found in the 6G6Y PDB structure (the setup of which is described in Section 3.1.6). The screening and selection strategy involved the following carefully sequenced stages.

High-Throughput Virtual Screening (HTVS) Docking: Initially, each compound was positioned within the L5 pocket using the HTVS mode of Glide. At this stage, the main objective is to produce a varied assortment of possible binding poses. In this phase, standard parameters were utilized, which involved adjusting the non-polar van der Waals radii of the ligand atoms to 0.80, and implementing a partial atomic charge threshold of 0.15 e. Among the various poses generated for each ligand, the top 10%

were chosen for further evaluation based on their HTVS docking scores, indicating their superior performance in the initial screening.

Standard Precision (SP) re-scoring: The selected promising poses from the HTVS stage underwent a rescoring process. The Glide SP scoring function was used for this objective. This phase aimed to enhance initial scores through a more thorough evaluation process. Following the rescoring process, the set of poses was subjected to further filtration, retaining only those that were classified within the top 10%, according to their SP docking score.

Extra Precision (XP) re-scoring: The poses that successfully passed the SP rescoring filter were subjected to an even more stringent evaluation using the Glide XP scoring function. Consistent with the tiered approach, only the poses ranking in the highest 10% based on their XP docking cores were included in the final rescoring step.

C

Final Candidate Selection: Compounds, represented by their top-scoring and thoroughly rescored poses, that passed the MM/GBSA score threshold were subjected to further in-depth examination. This focused on their predicted binding interactions and overall compatibility within the L5 allosteric site, involving two distinct sub-steps:

Automated Interaction Analysis: The retained poses were automatically analyzed using Maestro's interaction fingerprint utility. This served to identify and confirm the presence of key interaction motifs that are known to be characteristic of L5 site binders. The specific interactions monitored included hydrogen bonds with the side chains of Glu116 and Glu118, π–π stacking interactions with Trp127, and significant hydrophobic contacts within the pocket defined by residues such as Leu119 and Pro137 [135]. Only compounds whose poses exhibited at least two of the predefined key interactions were retained for the final visual check.

Manual visual inspection: The remaining candidates were subjected to rigorous manual visual inspection using Maestro software. The aim of this study was to confirm that the ligand pose demonstrated adequate shape complementarity within the binding pocket and to ensure the absence of significant steric clashes with receptor residues. It was also verified that the ligand conformations and torsional angles were energetically reasonable and geometrically plausible.

Compounds that successfully navigated all stages of this hierarchical pose generation, rescoring, and multifaceted filtering protocol were considered the most promising candidates from this screening track. These selected compounds were subsequently advanced to detailed molecular dynamics simulations (as described in Section 3.1.12) for a more thorough evaluation of their binding stability and dynamic behavior within the target site.

3.1.10 DeepDocking workflow

An active learning strategy employing the DeepDocking framework was implemented between January and June 2024 to identify potential hit compounds iteratively. This process involved multiple cycles of molecular docking, machine learning model training, and inference, as detailed below. The entire workflow, including the Qvina2 docking pipeline and the associated protein and ligand preparation steps, was automated using custom scripts and managed within a Docker container to ensure a consistent and reproducible computational environment. The workflow was executed using Slurm for job submission to a high-performance computing cluster. The computations were performed in a dedicated conda environment. A text file was created in the project folder to specify the project paths, library locations, and other parameters, as required by the DeepDocking scripts.

3.1.10.1 Initial compound library sampling and conformer generation

At each of the 10 DeepDocking iterations, an initial sampling or selection step was performed. For the first iteration, the relevant bash script named “phase_1” was used to randomly sample 1,000,000 compounds from the related directory. This script generated text files (containing ZINC IDs of the sampled compounds), corresponding SMILES files within a smile subdirectory, and importantly, corresponding Morgan fingerprint files within a morgan subdirectory of the current iteration folder.

The SMILES files were then processed to generate the 3D conformers. Using the script generated by MayaChemTools, SDF files were created for each set (training, validation, and testing). This script employed the ETKDG v2 [368] embedding protocol (--enforceChirality yes), generated a single conformer per compound, used the MMFF94 force field, and had a maximum of 1,000 optimization steps. The resulting SDF files are saved in a subdirectory within the current iteration folder.

For subsequent Qvina2 docking, these generated SDF files were moved to a a directory within Qvina2 working path. Here, the custom Python script (utilizing RDKit's Chem.MultithreadedSDMolSupplier) was employed to divide the large SDF files into smaller, more manageable batches to facilitate parallel processing during docking.

3.1.10.2 Ligand preparation for Qvina2 docking

Each batched SDF file served as an input for the PDBQT conversion. This was performed using the mk_prepare_ligand.py script (attributed to Meeko v0.5.0), executed with the command:

mk_prepare_ligand.py --rigid_macrocycles -i {input_sdf_batch} --multimol_outdir {temp_pdbqt_dir}.

The output PDBQT files for each ligand were written to a temporary directory. Molecular docking was then performed using Qvina2 against the prepared Eg5 receptor PDBQT (Section 3.1.3) and a defined configuration file (Section 3.1.6). Docking for each ligand PDBQT file was executed in parallel using multi-processing module of Python (typically utilizing 46 processes). Each Qvina2 instance was run with the command:

./qvina --receptor {receptor.pdbqt} --config {config.txt} --ligand {ligand.pdbqt} --cpu 8 --out {output_ligand.pdbqt} --num_modes 1.

Binding poses were saved in PDBQT format in a subdirectory within the temporary directory.

3.1.10.3 Post-docking processing and score extraction

Following Qvina2 docking, the output PDBQT files, containinGbinding poses and REMARK lines with docking scores (e.g., "REMARK VINA RESULT: -9.1 ..."), were processed. A custom Python script utilizing the Meeko library and RDKit converts each docked PDBQT file back into an SDF format. During this conversion, docking scores and other relevant information from the PDBQT REMARK lines were parsed and embedded as properties in the output SDF files. These individual SDF files are saved in a temporary output directory. Subsequently, all SDF files corresponding to a given set (e.g., training set for iteration 1) were merged into a single SDF file and stored in a related directory.

To prepare data for machine learning model training, the generated and modified scripts processed these merged SDF files (or the intermediate PDBQTs/SDFs) to extract the ZINC ID and primary docking score for each compound. This information was compiled into comma-separated value files named text files, which were stored in the project directory respectected to iteration. These label files served as direct inputs for the subsequent machine learning model training phase.

3.1.10.4 Machine learning model training

Following docking and score extraction, deep learning models were trained using the phase 4 script, utilizing the generated labels containing text files (from Section 3.1.10.3) and the corresponding Morgan fingerprints [369] of the sampled compounds (from the morgan subdirectory mentioned in Section 3.1.10.1). In each of the 10 iterations, 24 machine learning models were trained. The script was configured as summarized below, running from its location within the DeepDocking Protocol directory.

The phase_4.sh script was executed for each iteration with parameters defining the current iteration number (1 to 10), number of label/fingerprint sets (3, for training, validation, and testing), paths to project (~/DeepDocking/projects) and project name (DD_Eg5), GPU partition (gpu-partition), total number of iterations (10), progressively decreasing hit percentage thresholds for the current iteration (from 1.0 in iteration 1 down to 0.1 in iteration 10), a fixed hit percentage for the overall last iteration (0.01), a recall value of 0.90, a maximum training time (00-20:00 days-hours:minutes), and the Conda environment name (dd-env). Upon completion of training in each iteration, the trained models were saved using the DeepDocking framework within a subdirectory of the respective iteration folder. The framework also generates a text file within the iteration directory, which documents the performance metrics (e.g., recall, precision on the test set) of all trained models and identifies the best-performing model based on these metrics (typically highest recall on the test set).

General command structure for phase_4.sh:

sbatch phase_4.sh <current_iteration> <num_label_and_fingerprint_sets> <project_path> <project_name> <gpu_partition> <total_iterations_setting> <hit_percent_current_iter> <hit_percent_last_iter_overall> <recall> <max_time> <conda_env>

3.1.10.5 Inference and selection for next iteration

After training, the script called “phase_5” was run in each iteration with arguments specifying the current iteration number, project path, project name, the recall value (0.90, consistent with training), GPU partition, and Conda environment name as per the DeepDocking protocol. Crucially, this script utilizes the best-performing model identified in the previously mentioned text file from the training phase of the current iteration (Phase 4), loading it from its saved location within the iteration's model directory. The selected model was then used to predict virtual hits from the larger unexplored chemical space (using fingerprints from the main library fingerprint directory). These predictions (specifically, ZINC IDs of compounds listed in the Morgan _1024_predictions text file within the respective iteration directory, e.g., projects/DD_Eg5/iteration_N/morgan_1024_predictions/) formed a pool of candidates for the sampling/selection stage of the subsequent iteration.

3.1.10.6 Final hit extraction and filtering

Upon completion of the 10 active learning iterations, the final_extraction.sh script from the DeepDocking utilities was employed:

sbatch --cpus-per-task 48 final_extraction.sh ~/DeepDocking/library_prepared /DeepDocking/projects/protein_test_automated/iteration_10/morgan_1024_predictions 48 ‘all_mol’ dd-env

This script generated two primary output files: smiles.csv, containing the SMILES strings of the compounds predicted as prospective virtual hits in the final iteration's morgan_1024_predictions file, and id_score.csv, containing the corresponding ZINC IDs and their ML-predicted virtual hit-likeness scores (probabilities, ranging from 0.0 to 1.0). The compounds listed in smiles.csv (corresponding to all entries in morgan_1024_predictions) were subjected to a confirmatory docking round using Qvina2, employing the same docking protocol and parameters as described in Section 3.1.10.2. This step aims at obtaining reliable docking scores for all ML-prioritized candidates.

Subsequently, a multi-step filtering process was applied:

ML score filtering: Compounds with an ML-predicted score (from id_score.csv) of ≥ 0.8 were retained. This threshold was chosen to select compounds with a high

predicted probability of being active, balancing the enrichment of true positives with a manageable number of compounds for subsequent validation.

Qvina2 docking score filtering: From the subset above, only compounds with a Qvina2 docking score of ≤ -10.0 kcal/mol were kept. This ensured that selected compounds exhibited strong predicted binding affinity in a physics-based scoring model.

ADME properties filtering: The remaining compounds were evaluated for drug-likeness based on ADME properties, which were calculated using the RDKit cheminformatics toolkit. This step aimed to prioritize compounds with favorable pharmacokinetic profiles, increasing the likelihood of downstream success. The criteria applied were:

- Molecular Weight (MW): ≤ 500 Da
- LogP (octanol-water partition coefficient): ≤ 5
- Hydrogen Bond Donors (HBD): ≤ 5
- Hydrogen Bond Acceptors (HBA): ≤ 10
- Topological Polar Surface Area (TPSA): ≤ 140 $Å^2$
- Number of Rotatable Bonds (rotB): ≤ 10

These ranges are based on widely accepted guidelines for oral bioavailability, such as Lipinski's Rule of Five, with additional common filters for TPSA and rotatable bonds to improve developability prospects.

Compounds that survived all of these filtering stages were then re-prepared using LigPrep. During this re-preparation, the existing stereoisomers and charge states were preserved, and the compounds were energy-minimized using the OPLS4 force field. This ensured that they were suitably formatted for a subsequent hierarchical Glide docking cascade starting with Glide HTVS. Compounds with a Glide HTVS docking score ≤ -5.0 kcal/mol progressed to Glide XP docking. The -5.0 kcal/mol threshold for HTVS served as an initial, simpler filter to pinpoint compounds with at least moderate predicted binding affinity prior to implementing more computationally demanding and precise XP scoring. After the Glide XP docking process, compounds were filtered according to a docking score of ≤ -5.0 kcal/mol. The threshold for the XP score was established to identify compounds that exhibited a more advantageous and precise

forecasted binding energy. Complexes that met the criteria of this XP score filter and demonstrated interactions with the essential residues of the Loop L5 binding site (as detailed below) were subsequently analyzed through MM/GBSA rescoring using Prime (Schrödinger Suite) with the VSGB solvent model and OPLS4 force field. A stringent MM/GBSA binding free energy threshold of ≤ -60 kcal/mol was applied to select compounds with highly favorable predicted binding affinities.

The final set of compounds satisfying the Glide XP score, MM/GBSA score, and interaction criteria was then subjected to a detailed visual inspection of their binding poses and interactions within the Loop L5 allosteric site of Eg5, using Maestro's interaction-fingerprint utility. The focus was directed towards verifying the formation of essential interactions with pivotal residues identified for this pocket, including hydrogen bonds (e.g., with Glu116 and Glu118), π-π stacking (e.g., with Trp127), and notable hydrophobic/van der Waals contacts (e.g., with Arg119, Pro137, Tyr211, Leu214, and Ala218), as detailed for recognized L5 binders and confirmed by existing literature. Furthermore, good shape complementarity and absence of steric clashes or unrealistic torsions were ensured. A selected number of compounds demonstrating optimal binding modes, robust interactions with these critical residues, and favorable MM/GBSA scores were then advanced to molecular dynamics simulations.

3.1.11 Ligand drug-likeness and structural similarity profiling

A dedicated in-silico profiling step was carried out immediately before the molecular-dynamics stage to quantify (i) the oral drug-likeness of the ten candidate ligands identified in the two parallel virtual-screening tracks and (ii) their structural relatedness to the 1.278 experimentally confirmed Eg5 actives curated from PubChem [370].

All calculations were performed in Python 3.11 (RDKit 2024.03.1, pandas 2.2). The SMILES strings for the active set were stripped of salts and duplicates during an earlier curation step; the same canonicalization procedure was applied to the ten candidates.

For each compound, the physicochemical descriptors used in the previous section were extracted using RDKit. Lipinski compliance was evaluated again with conventional thresholds of MW ≤ 500 Da, LogP ≤ 5, TPSA ≤ 140 Å², HBD ≤ 5, HBA ≤ 10, and rotB ≤ 10.

Structural similarity was assessed using extended-connectivity fingerprints (Morgan, radius 2, 1 024 bits). A Tanimoto similarity matrix between each candidate and the active ligand was generated. For ease of interpretation, four statistics were retained for each candidate: maximum, mean, and median Tanimoto coefficients and the percentage of pairwise values exceeding 0.40, a boundary frequently regarded as the onset of moderate similarity.

3.1.12 Molecular dynamic simulations

Selected candidate compound-Eg5 complexes, identified from both the GPU Shape & Phase Screening track (after visual inspection, Section 3.1.9) and the DeepDocking–Qvina2 workflow (after final filtering, Glide XP docking, MM/GBSA re-scoring, and interaction analysis, Section 3.1.10.6), were subjected to explicit-solvent molecular dynamics (MD) simulations using the Desmond tool [26] (Schrödinger Suite). The same approach for system setup was used across all complexes: each complex was solubilized in an orthorhombic box filled with TIP5P water compounds, ensuring a minimum distance of 10 Å from the box boundaries. The system was neutralized by adding NaCl ions to a concentration of 0.15 M. Prior to the production run, the system experienced a standard relaxation protocol, often including initial minimization with limited solute heavy atoms, followed by short simulations under NVT and NPT ensembles to achieve temperature and pressure equilibrium.

Production MD simulations were performed under the NPT ensemble at 300 K (Nose-Hoover chain thermostat, 1.0 ps relaxation time) and 1.0 bar (Martyna-Tobias-Klein barostat, 2.0 ps relaxation time, isotropic coupling) using the OPLS4 force field. No restraints were applied to the solute during the production phase. Trajectories were saved at regular intervals (e.g., 100 ps) for subsequent analysis. The simulation duration differed based on the origin of the complex: compounds from the GPU Shape & Phase Screening track were simulated for 250 ns, while those from the DeepDocking–Qvina2 workflow were simulated for 100 ns.

Analyses of the MD trajectories included the calculation of root-mean-square deviation (RMSD) for the ligand and protein Cα atoms to assess system stability, and monitoring of ligand-protein interactions over time. For complexes exhibiting stable binding and maintaining promising interactions throughout the MD simulation, binding free energies were calculated from the simulation trajectories using the

thermal_mmgbsa.py script of Schrodinger suite. This script executes the MM/GBSA approach, average energies across a designated segment of the equilibrated trajectory to evaluate the binding affinity. Appropriate graphical representations of the MD findings (e.g., RMSD plots, interaction diagrams) and MM/GBSA calculations were produced for further analysis.

3.2 In vitro calculations

3.2.1 Eg5 inhibition assay

To estimate the in vitro Eg5 inhibitory activity of compounds ZINC299785885, ZINC102735853, ZINC243914928 and ZINC1857672049 a kinesin ATPase end-point assay (Biochem Kit, Cytoskeleton Inc, USA) [371] was performed. The overmentioned kit was optimized to evaluate the amount of Pi generated after kinesin Eg5 inhibition [150]. K858 was used as a reference compound to verify the adaptation of the kit to our purpose. All compounds were first dissolved in DMSO, and then different concentrations diluted in kinesin reaction buffer tested: 0, 0.1, 1, 5, 10, 50 and 100 μM. Briefly, the reaction was conducted in a 96-well plate in a final volume of 300 μL for each well, containing microtubules, human recombinant Eg5 motor domain protein (Cytoskeleton Inc, USA) and Eg5 inhibitors. The reaction started by adding 0.3 mM of ATP (Merck Life Science, Milan, Italy). The amount of Pi released (which is a measure of the Eg5 ATPase activity) from each reaction was spectrophotometrically read by a microplate reader at 650 nm wavelength (Multiskan GO, Thermo Scientific, Waltham, MA, USA). The IC50 value of each compound was calculated by Prism 9 (GraphPad) [363].

4. RESULTS AND DISCUSSION

4.1 In Silico Calculations

To discover novel inhibitors for the Eg5 kinesin, this study employed a complementary dual-track virtual screening strategy, as detailed in the Materials and Methods section (3.1.1). The first, more conservative track, performed targeted scaffold hopping through multi-stage, structure-based filtering of the ~11 million compound ZINC20 Anodyne library. In parallel, the second, more exploratory track utilized a DeepDocking active learning campaign to navigate the vast chemical space of the ~1 billion compound ZINC20-ML database to search for novel chemotypes.

Following these initial large-scale screenings, the most promising candidates emerging from both workflows were subjected to rigorous post-screening validations. This crucial phase included extensive all-atom molecular dynamics (MD) simulations and binding free energy calculations to assess the dynamic stability and predict the binding affinity of the compounds, thereby filtering the compounds with the highest potential for durable Eg5 inhibition. This chapter presents the detailed outcomes of this comprehensive approach, beginning with the validation of the screening protocols and culminating in the identification and in-depth analysis of potential lead candidates from each track.

4.1.1 Validation of Computational Protocols

Prior to large-scale screening, core computational methods, including pharmacophore modeling and docking protocols (Glide and Qvina2), were rigorously validated to ensure their reliability and predictive power against the Eg5 target.

4.1.1.1 Pharmacophore model development and validation

Pharmacophore models were generated by a detailed analysis of the co-crystallized inhibitor K858 within the α2/L5/α3 allosteric pocket of Eg5. The K858-bound structure was chosen as the template because its high resolution and representation of a stably inhibited state provided an ideal blueprint for the key chemical interactions required for potent allosteric modulation. The analysis focused on delineating these critical features, namely hydrogen bonding and hydrophobic interactions (Figure 4.1a, b). The selected model, PM1, was constructed to encompass three essential features: a hydrogen bond acceptor (A3), a hydrophobic group (H7), and an aromatic ring (R8). Crucially, the model incorporates excluded volume spheres to represent the steric boundaries of the binding pocket. This feature is vital for proactively eliminating compounds that, while matching the pharmacophoric points, would be sterically incapable of fitting into the pocket, thereby improving screening efficiency.

To rigorously assess the discriminatory power and practical effectiveness of the model, it was validated against DEKOIS 2.0. The primary goal of this validation was to determine the success of the model in identifying known active compounds. The PM1 model correctly identified 39 of the 40 known actives as 'hits' (97.5% hit rate), demonstrating high sensitivity and confirming that it effectively captures the essential chemical features for Eg5 recognition. This high hit rate was particularly significant because this stage of the virtual screening workflow was designed as a binary (match/no match) filter rather than a score-based ranking system. As all compounds that matched the pharmacophore model were advanced to the next stage without a scoring threshold, the primary task of the model was not to rank the best candidates but to enrich the pool with potentially active c ompounds and eliminate irrelevant ones. Therefore, the high recognition rate confirmed that the PM1 model was a robust and reliable tool for this "gatekeeper, effectively enhancing the efficiency of the screening funnel.

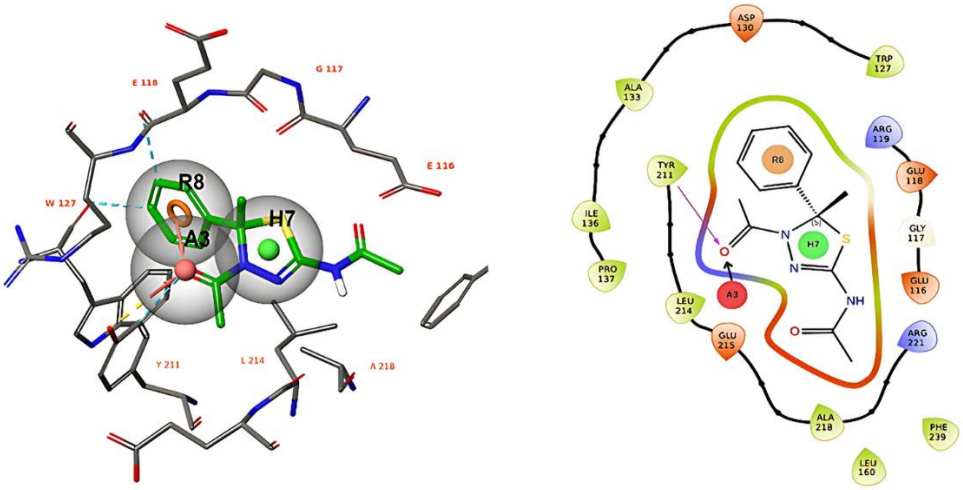

Figure 4.1 : Representations of pharmacophore model PM1.

4.1.1.2 Validation of the Glide docking and re-scoring protocol

The hierarchical Glide docking and rescoring protocol was first validated for its ability to accurately reproduce the experimental binding pose of a known inhibitor, which is a critical measure of pose-prediction accuracy. This was confirmed by retrospective docking of the co-crystallized inhibitor K858. As illustrated in Figure 4.2, which superimposes the binding pose onto the experimental structure, the protocol achieved a root-mean-square deviation (RMSD) of 0.2819 Å. This low RMSD value provides high confidence that the Glide docking algorithm and OPLS4 force field are well calibrated for the Eg5 L5 pocket and accurately capture the geometric and energetic factors that govern ligand binding. This strongly suggests that the binding poses generated for the novel compounds during the screening phase are likely to be physically meaningful and reliable representations of potential interactions.

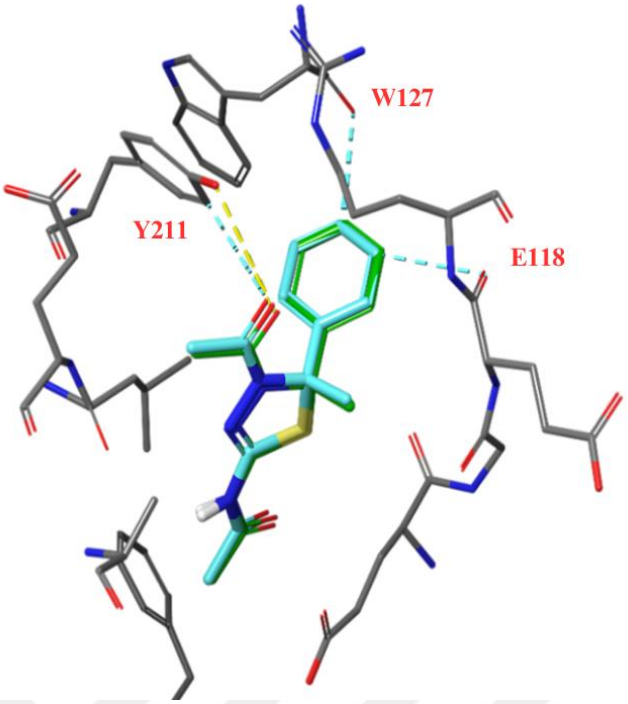

Figure 4.2 : Glide HTVS retrospective binding pose (turquoise) for Eg5 in complex with K858 (green)

Second, and more critically for virtual screening, the protocol's ability to enrich true positives was assessed using DEKOIS 2.0. The results demonstrated a clear and progressive improvement in enrichment when more rigorous scoring functions were applied (Figure 4.3). Although the initial HTVS and SP docking modes yielded a respectable predictive receiver operating characteristic area under the curve (pROC-AUC) of 1.79, this value increased to 1.90 with a more demanding XP scoring function and peaked at 1.96 after MM/GBSA rescoring. This trend is highly significant because it provides quantitative justification for our hierarchical workflow. This confirms that the increased computational investment at each successive stage directly translates into a more accurate ranking of potential hits.

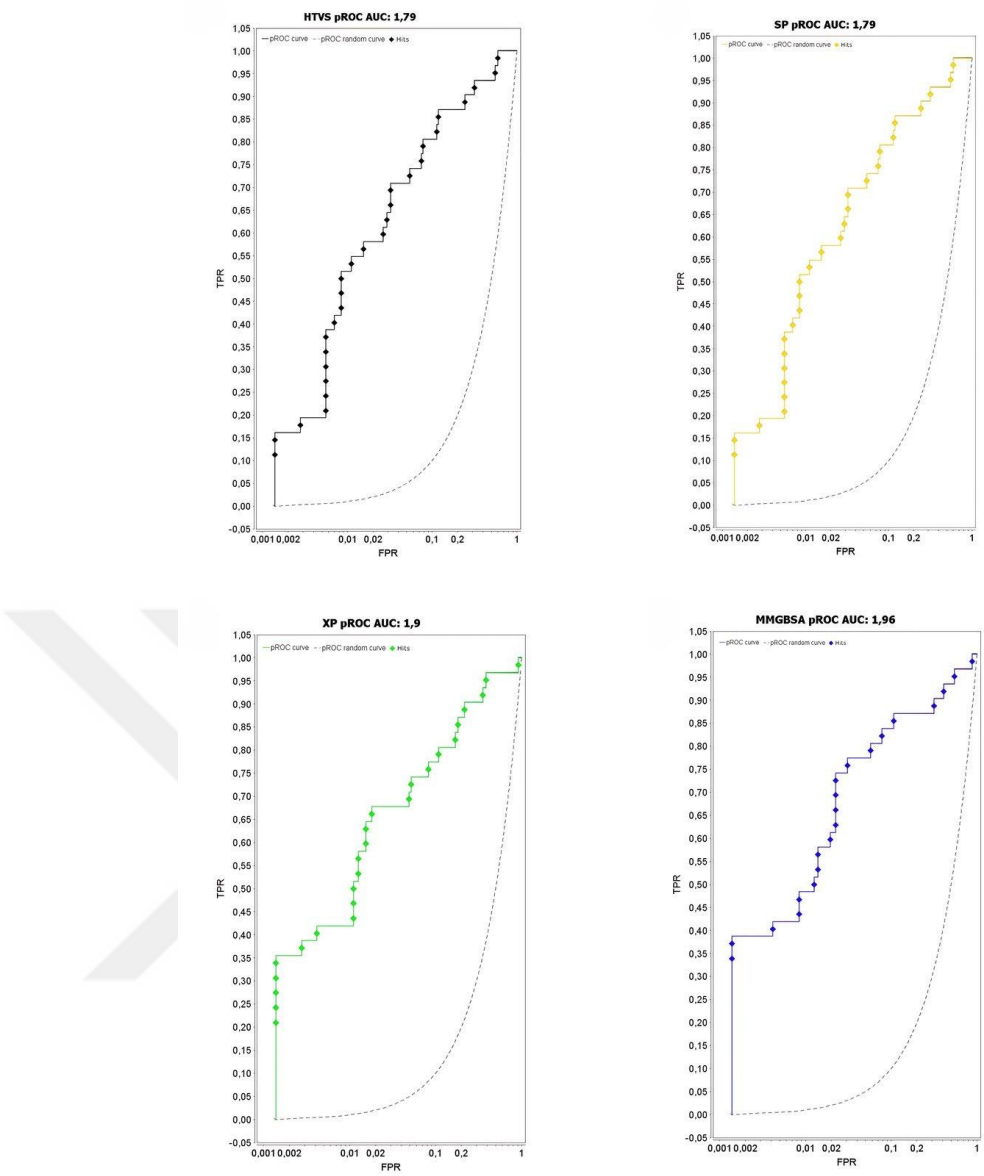

Figure 4.3: pROC curves and AUC values for the Glide docking protocol with different scoring functions (HTVS, SP, XP, and MM/GBSA) on the DEKOIS 2.0 benchmark set.

The practical value of this enhanced accuracy is most evident in the analysis of early enrichment data. At a strict False Positive Rate (FPR) of 0.1%, representing the top-ranked candidates most likely to be selected for experimental testing, the True Positive Rate (TPR) increased dramatically from 0.15 (SP) to 0.40 (MM/GBSA). This near three-fold improvement indicates that the full rescoring protocol is exceptionally effective in placing genuine active compounds at the top of the hit list. This robust early enrichment capability is the cornerstone of an efficient virtual screening campaign, as it maximizes the probability of success while minimizing the resources spent on testing false positives. Taken together, these validation results underscore the

reliability and efficacy of the selected Glide docking and rescoring cascade, confirming its suitability for discovering and prioritizing prospective Eg5 inhibitors from a large chemical library of compounds.

4.1.1.3 Validation of the Qvina2 docking protocol

The DeepDocking workflow requires a docking engine capable of processing millions of compounds iteratively, demanding a balance between computational speed and predictive accuracy. For this purpose, the Qvina2 docking protocol was selected and validated. Retrospective docking of K858 yielded an RMSD of 0.3840 Å compared to the co-crystallized pose (Figure 4.4). Although this value is slightly higher than that achieved with Glide, it remains well within the acceptable threshold (< 2.0 Å) for accurate pose prediction. This result confirmed that Qvina2, despite its computational efficiency, could reliably identify the correct binding mode for ligands within the L5 pocket, which is a prerequisite for generating meaningful data to train the active learning model.

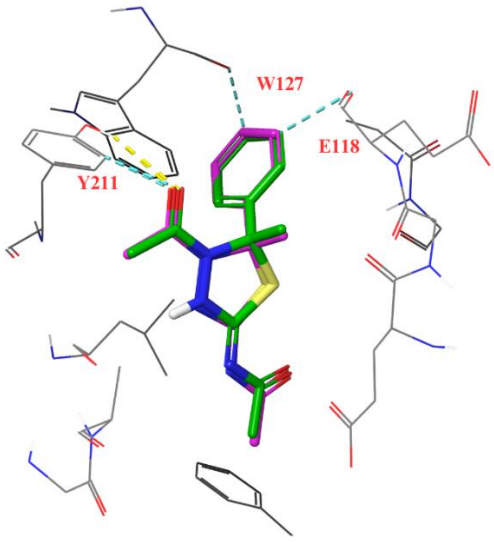

Figure 4.4: Qvina2 retrospective binding pose (pink) of K858 (green)

When evaluated against the DEKOIS 2.0 benchmark set, the Qvina2 protocol achieved a pROC-AUC of 1.71 (Figure 4.5). This value is crucial because it demonstrates that Qvina2 possesses substantial enrichment capability and performs significantly better than random chance. The shape of the pROC curve, with a notable initial rise, indicates a good rate of early enrichment, which is sufficient to provide the active learning algorithm with a high-quality signal-rich dataset in each iteration. For the DeepDocking strategy, the primary role of the docking engine is not to achieve a

perfect ranking in a single pass but to generate reliable data that can effectively train the neural network to navigate the chemical space. The performance metrics of Qvina2 confirmed its fitness for this role.

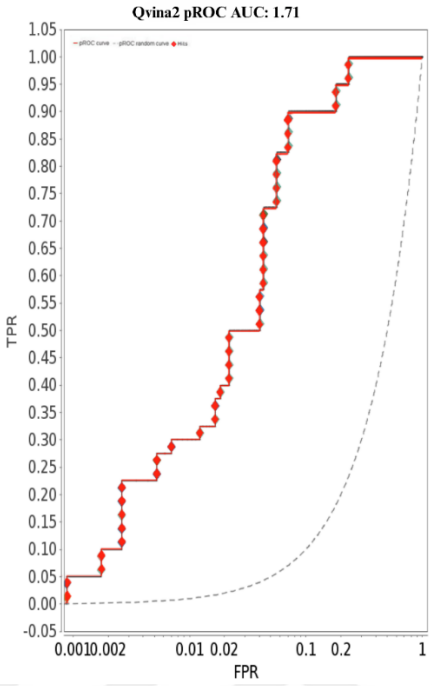

Figure 4.5: pROC curves and AUC values for the Qvina2 docking protocol on the DEKOIS 2.0 benchmark Set.

Ultimately, the choice of Qvina2 represents a strategic compromise in this study. Although the multi-stage Glide protocol demonstrated superior enrichment, its computational cost renders it prohibitive for the iterative, large-scale screening required by DeepDocking. Qvina2, with its validated accuracy and significantly higher throughput, provides the optimal balance of speed and predictive power, making it an effective and appropriate engine to drive the active learning campaign.

4.1.2 Virtual screening studies track 1: Structure-based filtering of ZINC20 anodyne database

4.1.2.1 GPU Shape similarity and Phase pharmacophore screening

The first stage of this screening track served as a broad, computationally efficient filter to rapidly reduce the vast ZINC20 Anodyne library (~11 million compounds) to a more manageable and enriched subset. The process began with GPU-accelerated shape screening. The decision to use a query set of 13 structurally diverse inhibitors rather than a single ligand was a key strategic choice designed to broaden the scope of the

search and increase the probability of discovering novel scaffolds that might otherwise be missed. This initial pass, which processed approximately 140 million conformers, yielded approximately 3 million unique compounds with a general 3D shape compatible with the L5 pocket.

This shape-filtered set was then subjected to the validated PM1 pharmacophore model. This step adds a crucial layer of chemical intelligence to the purely steric filter, ensuring that the shape-matched compounds also possess the necessary electronic features (hydrogen bond acceptors, hydrophobic groups, etc.) required for productive interactions. The synergy between shape and pharmacophore screening is a powerful combination, as it filters both the volumetric fit and the specific chemical grammar of binding. This combined approach proved to be highly effective, reducing the library to 1.335.248 unique compounds. In essence, these initial filters successfully eliminated approximately 90% of the starting library, creating a high-quality, enriched set for the subsequent, more computationally intensive docking.

4.1.2.2 Hierarchical Glide docking, re-scoring, and candidate selection

The 1.335.248 compounds emerging from the initial filters were then processed through a "computational funnel" of increasing rigor using the validated Glide protocol. This hierarchical approach is designed to balance throughput with accuracy by applying faster, less precise methods first and reserving the most computationally expensive calculations for a smaller, more promising set of candidates.

Initial docking with Glide HTVS provided a rapid geometric assessment, followed by sequential rescoring using the more accurate SP and XP scoring functions. The pragmatic application of a top 10% cutoff at each rescoring stage effectively winnowed the candidate pool while retaining chemical diversity. The most significant refinement was obtained using Prime MM/GBSA rescoring, a physics-based method that provides a more robust estimate of the binding affinity by incorporating implicit solvent effects. Applying a reasonably stringent MM/GBSA energy cutoff of < -40 kcal/mol was a critical decision point, reducing the set of potential ligands to 5.620 and filtering out compounds with poor predicted binding thermodynamic properties.

To further enhance the quality of the hit list, a filter based on specific biochemical knowledge was applied to the results. This moved the selection process beyond simple energy scores to require that candidates exhibit a chemically plausible binding mode.

By requiring the formation of key interactions known to be critical for potent L5 pocket inhibition (e.g., hydrogen bonds with Glu116/Glu118 and π-π stacking with Trp127), we could effectively eliminate false positives that may have achieved a good score through fortuitous but non-productive contacts. This interaction filter yielded 410 compounds.

The final and indispensable step was a rigorous visual inspection. This expert-driven evaluation is crucial for identifying subtle structural liabilities, such as unfavorable bond torsions or imperfect shape complementarity, which may be overlooked by automated algorithms. This meticulous process ultimately led to the selection of 30 high-confidence compounds for molecular dynamics simulations.

4.1.3 Virtual screening track 2: DeepDocking-driven exploration of ZINC20-ML database

4.1.3.1 Performance of the DeepDocking active learning campaign

To explore a chemical space far exceeding that of the Anodyne subset, the DeepDocking active learning protocol was executed for ten iterations against the 1 billion-compound ZINC20-ML database (Figure 4.6). This approach was essential because a brute-force docking campaign of this magnitude would be computationally intractable. Active learning provides an intelligent solution that enables the efficient and targeted exploration of this vast library.

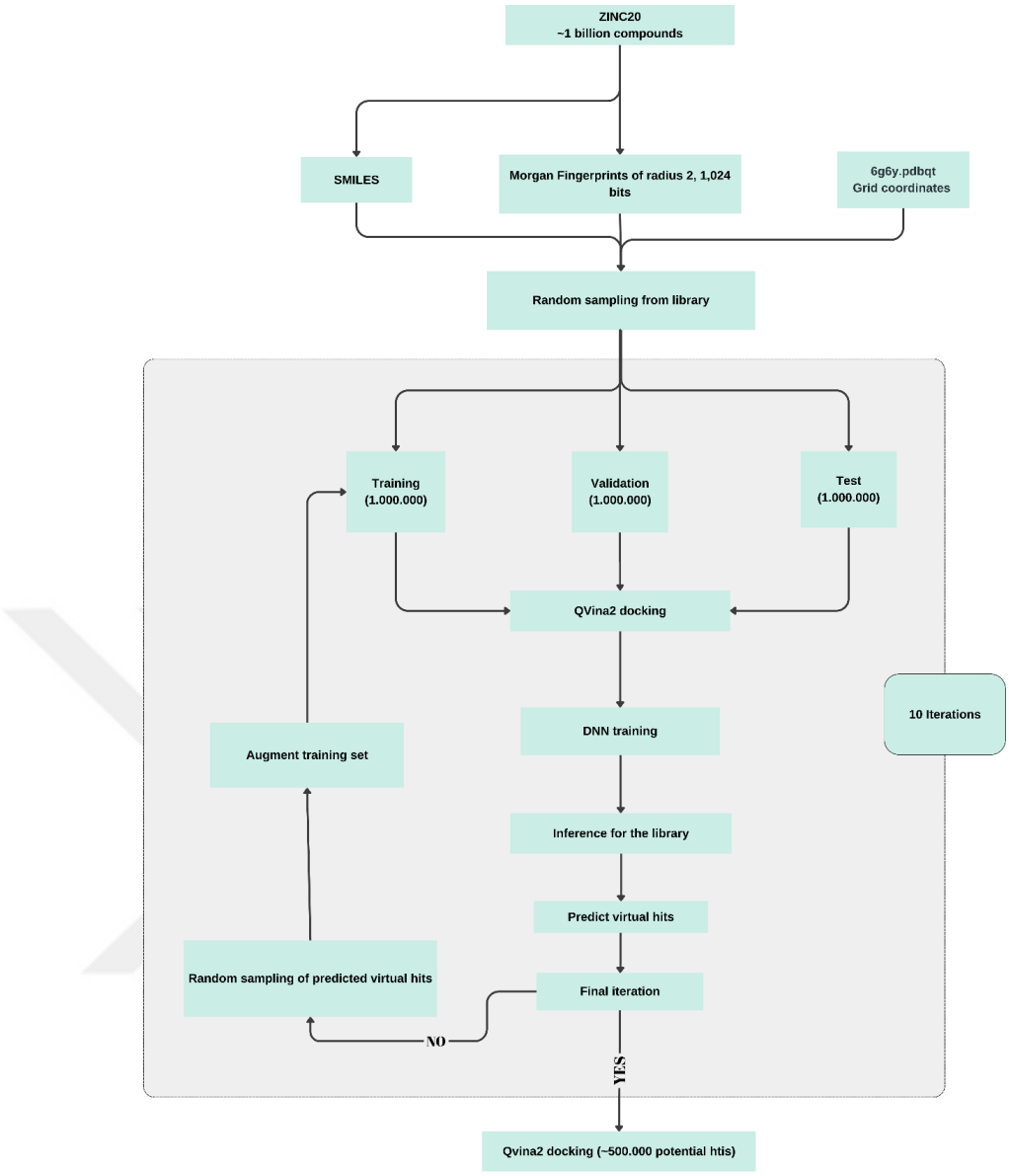

Figure 4.6 : General overview of DeepDocking campaign workflow.

The process began by docking a random sample of 1 million compounds with Qvina2, providing the initial "lesson" for the Deep Neural Network (DNN) models. In each subsequent iteration, the trained DNN predicted the docking scores for the remainder of the library, and a new set of promising compounds was selected for docking. This new data were then used to retrain and improve the model.

The iterative enhancement of the predictive power of the DNN models was demonstrated qualitatively and quantitatively. Visually, the performance improvement was evident in the held-out test sets, where the area under the ROC curve (AUC) for predicting Qvina2 scores increased progressively from an initial value of 0.968 in iteration 1 to a near-perfect 0.998 by iteration 10 (Figure 4.7). This trend is the

hallmark of a successful, active learning campaign. This indicates that with each cycle, the AI model became substantially more proficient at identifying compounds likely to achieve high docking scores, effectively acting as a fast and accurate surrogate for the much slower physics-based docking calculation.

This increasing proficiency is further substantiated by the detailed performance metrics of the best-performing model from each iteration (Table 4.1). The model cutoff value, representing the Qvina2 docking score threshold learned by the model to classify a compound as a "hit", became increasingly stringent with each cycle, moving from -9.4 kcal/mol in the first iteration to -10.7 kcal/mol in the final iteration. This demonstrates that the model successfully learned to focus on compounds with progressively more favorable (i.e., lower) binding energies, effectively raising the quality bar for the predicted hits.

The model's ability to recover the majority of true positives in the test sets was consistently high, as indicated by the model recall values, which remained around a robust 0.87-0.90 for the first nine iterations. This high recall was a critical operational parameter that ensured that the model did not prematurely discard significant portions of the promising chemical space. The slight decrease in recall to 0.78 in the final iteration is a typical and acceptable trade-off for a significant gain in precision as the model becomes more specialized.

This gain in model precision, which measures the fraction of predicted hits that are true positives, showed a notable trend in the results. Although fluctuating in the initial stages, it more than doubled from the first to the tenth iteration, reaching a peak of 0.21. This indicates that as the selectivity and confidence of the model increased, its predictions became significantly more accurate and less noisy. Concurrently, the number of compounds deemed promising by the model decreased with each iteration (Figure 4.8), reflecting the increasing selectivity.

The combination of steadily increasing AUC values, a progressively more stringent cutoff, high recall, and improving precision provides powerful quantitative evidence for a highly effective and efficient AI-driven exploration of the chemical space. The protocol effectively "zoomed in" on the most fertile regions of the library, focusing on the limited computational docking resources, where they were most likely to yield high-quality hits. This iterative refinement was key to making the exploration of a

billion-compound library feasible, ultimately allowing for a much broader and more diverse search for novel Eg5 inhibitors.

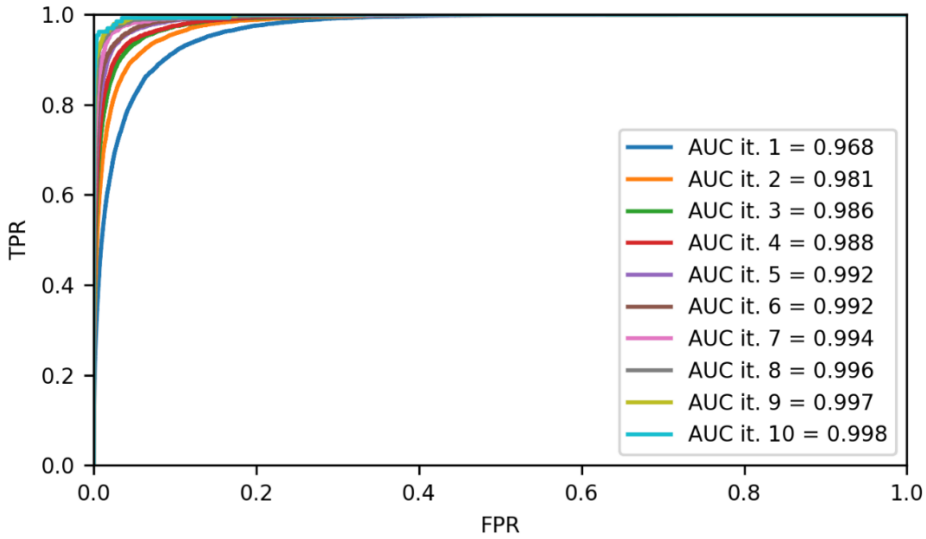

Figure 4.7: ROC Curves for DeepDocking model performance across iterations.

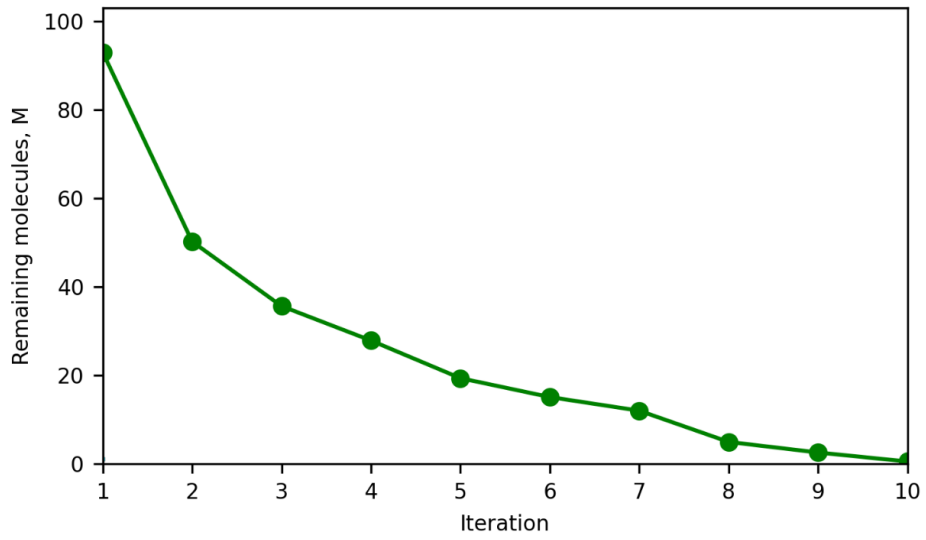

Figure 4.8: Reduction in the number of candidate compounds across DeepDocking iterations.

Table 4.1 : Performance metrics of the best-performing DNN model in each DeepDocking iteration.

Iteration	Model Cutoff (kcal/mol)	Model Precision	Model Recall	Model AUC
1	–9.40	0.097	0.896	0.968
2	–9.50	0.133	0.890	0.981
3	–9.60	0.140	0.895	0.986
4	–9.70	0.133	0.891	0.988
5	–9.80	0.137	0.893	0.992
6	–9.90	0.128	0.892	0.992
7	–10.00	0.118	0.902	0.994
8	–10.20	0.141	0.884	0.996
9	–10.40	0.134	0.875	0.997
10	–10.70	0.211	0.785	0.998

4.1.3.2 Final hit generation and multi-stage filtering from the DeepDocking track

Upon completion of the 10 active learning iterations, the top 531,202 candidate compounds prioritized by the final DNN model were subjected to a comprehensive filtering cascade. This multistage process was critical for systematically refining this large set into a small number of high-confidence hits. Although the AI model is adept at predicting high-scoring compounds, it does not inherently account for physicochemical properties, drug-likeness, or the nuanced accuracy of more rigorous physics-based models. Therefore, the subsequent cascade was designed to integrate these essential considerations, creating a robust pipeline for identifying genuinely promising candidates, as shown in Figure 4.9.

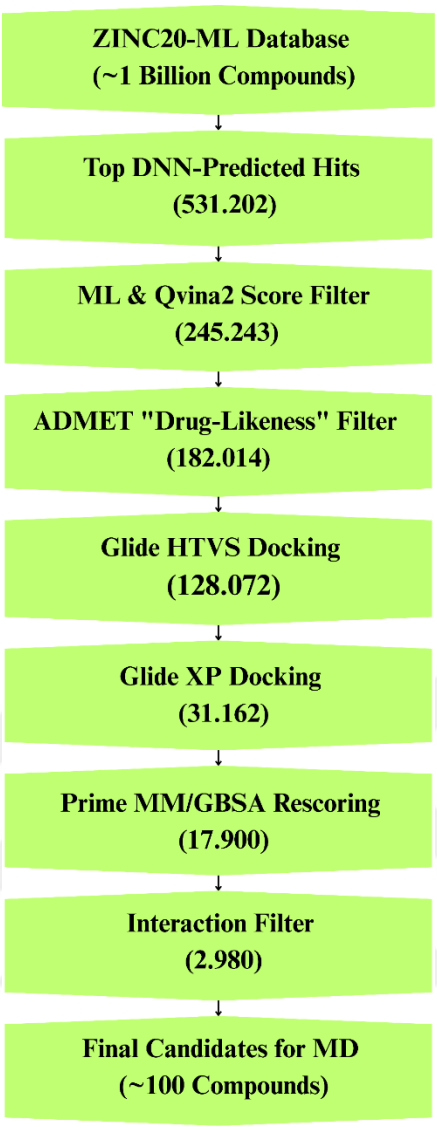

Figure 4.9: Schematic representation of the multi-stage filtering cascade for candidates from the DeepDocking track.

The process began with a crucial validation step: a confirmatory docking round using Qvina2 for all 531.202 compounds, which served to ground the subsequent filtering using a more reliable physics-based metric. A dual filter based on both the ML-predicted score (≥ 0.8) and the confirmatory Qvina2 docking score (≤ -10.0 kcal/mol) was applied, reducing the set to 245.243 candidates. Next, a strategically placed ADME property filter eliminated compounds with poor predicted drug-like properties, yielding 182.014 candidates. The workflow then transitioned to the more accurate but computationally intensive Glide docking protocol. The 182.014 compounds were first docked using Glide HTVS, and a score cutoff of ≤ -5.0 kcal/mol reduced the set to 128.072 compounds. These were then subjected to Glide XP docking. The application of a more stringent XP scoring function, with a cutoff of ≤ -5.0 kcal/mol, served as a

key refinement step, yielding 31.162 compounds. For these top candidates, Prime MM/GBSA rescoring was performed with a stringent threshold of ≤ -50 kcal/mol, identifying 17.900 compounds with excellent predicted binding thermodynamics. Finally, an automated interaction filter requiring favorable contact with critical L5 pocket residues (e.g., Glu116 and Tyr211) reduced the set to 2.980 compounds. A final, meticulous visual inspection of these top-ranked poses led to the selection of approximately 100 high-confidence compounds for MD simulations.

The results of this rigorous funneling strategy highlight several key points. The initial dual-score filter successfully validated the AI's predictions against a physics-based model, immediately focusing resources on a quarter-million candidates with a consensus of high potential for use as catalysts. The early application of the ADMET filter was a critical, pragmatic decision that effectively eliminated nearly 28% of the remaining compounds based on potential downstream liabilities, highlighting the importance of integrating drug-like properties early in the discovery pipeline. The subsequent hierarchical Glide docking and MM/GBSA rescoring stages demonstrated their power in refining the hit list based on increasingly accurate energy estimations. This comprehensive and logical filtering strategy, which successfully distilled an initial list of over half a million AI-prioritized compounds to a manageable set of approximately 100 promising candidates, powerfully validated the efficiency and precision of our integrated workflow. The final set of compounds, notable for their structural diversity and dissimilarity to known Eg5 inhibitor classes, not only represents a significant enrichment of chemical novelty but also provides a robust foundation for subsequent resource-intensive dynamic analyses required to identify the most stable and potent inhibitors.

4.1.3.3 Ligand property and similarity analysis

All ten candidates fulfilled the adapted Lipinski criteria, with no violations recorded for any compound. The molecular weights spanned 242–498 Da (mean 396 Da), LogP values 0.38–4.91 (mean 2.80), and TPSA values 67–132 Å² (mean 90 Å²), indicating that the hit set occupies the sweet spot for passive oral absorption yet remains sufficiently polar to avoid excessive lipophilicity-driven toxicity. Together with hydrogen-bond capacities (HBD ≤ 4, HBA ≤ 8) and modest flexibility (rotB ≤ 8 for

every compound), the profile suggests favorable developability and supports the advancement of the series to ADME-focused experimental assays (Table 4.2).

Table 4.2 : Key drug-likeness descriptors for the ten suggested compounds

Compound ID	MW	LogP	TPSA	HBD	HBA	rotB	violations
ZINC102735853	298.30	2.84	103.44	4	7	1	0
ZINC243914928	306.32	0.43	125.71	3	7	7	0
ZINC299785885	369.49	1.15	67.69	3	3	6	0
ZINC1857672049	242.24	0.38	87.97	2	5	3	0
ZINC28804061	491.63	4.91	79.32	3	3	8	0
ZINC580729109	424.50	3.49	103.44	4	7	6	0
ZINC410422515	497.63	3.40	101.56	3	6	9	0
ZINC484655776	441.57	3.76	105.38	4	7	6	0
ZINC1159968549	410.52	3.87	96.33	4	6	8	0
ZINC1817238515	481.56	3.44	101.56	4	7	8	0

Similarity analysis revealed pronounced scaffold diversity relative to the Eg5 reference chemotype space (Table 4.3). The mean Tanimoto coefficients ranged narrowly between 0.124 and 0.150 (grand mean ≈ 0.13), well below the 0.30 threshold often associated with analog derivation. Even the highest single pairwise similarity observed (0.41 for compound ZINC102735853) barely exceeded the 0.40 moderate-similarity demarcation, and only 8 % of its 1 278 comparisons crossed that line. The remaining nine candidates showed no individual similarity above 0.29 and no comparisons above 0.40, underscoring their novelty.

Table 4.3 : Tanimoto 2-D similarity (Tanimoto coefficient) statistics

Compound ID	Tan_max	Tan_mean	Tan_median	Tan_%>0.4
ZINC102735853	0.407	0.127	0.113	0.083
ZINC243914928	0.250	0.133	0.133	0
ZINC299785885	0.224	0.124	0.128	0
ZINC1857672049	0.233	0.124	0.125	0
ZINC28804061	0.286	0.150	0.149	0
ZINC580729109	0.321	0.146	0.145	0
ZINC410422515	0.265	0.124	0.123	0
ZINC484655776	0.276	0.143	0.140	0
ZINC1159968549	0.214	0.120	0.119	0
ZINC1817238515	0.237	0.124	0.124	0

Taken together, these data validate the dual screening strategy: the physicochemical attributes demonstrate drug-likeness, while the consistently low Tanimoto values confirm that the workflow achieved genuine scaffold hopping rather than mere optimization of known chemotypes. Therefore, the ensuing molecular dynamics simulations probe fresh chemical matter in the Eg5 binding landscape, a prerequisite for overcoming historical resistance and patent crowding around classical monastrol-derived inhibitors.

4.1.4 Molecular dynamic simulations and MM/GBSA calculations

Selected candidates from both screening tracks were subjected to explicit-solvent MD simulations to provide the most rigorous computational assessment of binding mode stability of the ligands. This step is indispensable for moving beyond static binding poses, as it allows for the evaluation of the dynamic behavior of the protein-ligand complex over time, accounting for both protein flexibility and the influence of the solvent environment. The stability of the key interactions and the overall binding free energy, calculated from the simulation trajectories, served as the ultimate filters for

identifying the most promising candidates for experimental validation of the docking results.

4.1.4.1 MD simulation results of Eg5 - K858 and comparative discussion

To establish a robust benchmark for our computational findings, the co-crystallized reference inhibitor K858 was subjected to an identical 250 ns MD simulation protocol. This analysis serves as a standard against which the performance of our newly discovered candidates is measured. The results of this simulation, as detailed in Figure 4.10, confirmed the stable binding of K858 within the L5 allosteric pocket.

The initial binding pose, as illustrated in the 3D and 2D diagrams (Figure 4.10a, b), shows K858 fitting into the allosteric pocket, anchored by a hydrogen bond with Tyr211 and aromatic interactions with Glu118 and Trp127.

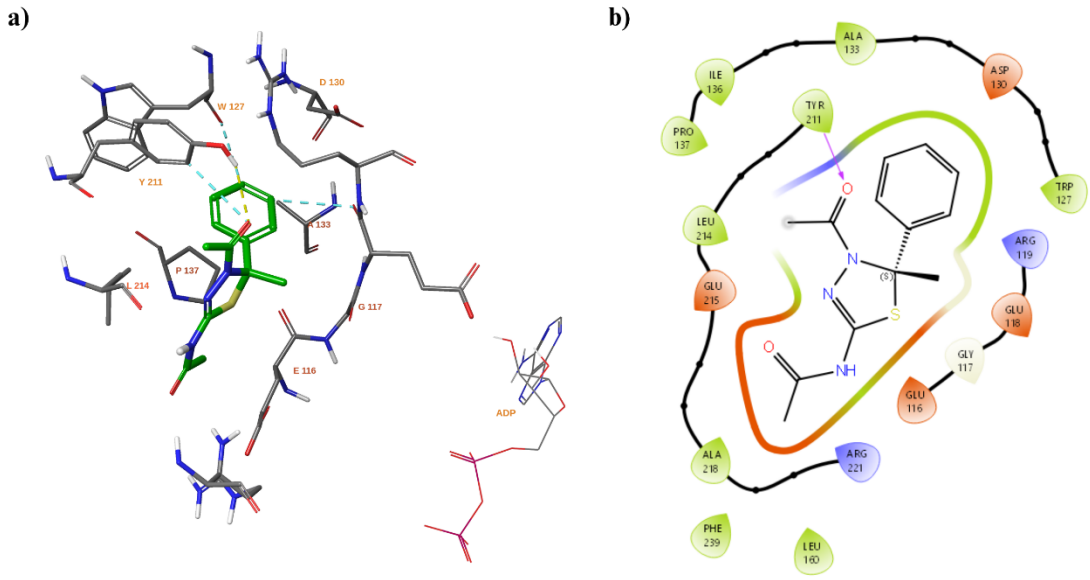

Figure 4.10: Binding pose representations of K858.

MD simulations were then used to test the dynamic stability of this pose. The stability of the complex was quantitatively demonstrated by the root-mean-square deviation (RMSD) plots (Figure 4.11a), where both the protein Cα-atom and ligand heavy atoms remained low throughout the simulation (generally < 2.7 Å). The interaction analysis over the trajectory (Figure 4.11b) revealed that while the initial direct hydrogen bonds fluctuated, a highly persistent water-bridged hydrogen bond was formed with Leu214 (89% occupancy), underscoring the critical role of solvent compounds in stabilizing the ligand. The average MM/GBSA binding free energy, which represents the overall binding affinity throughout the dynamic simulation (Figure 4.11c), was calculated to

be -53.66 ± 3.08 kcal/mol. This value represents an energetic reference point for a known potent inhibitor and provides a fundamental benchmark for evaluating the potential of newly discovered candidates.

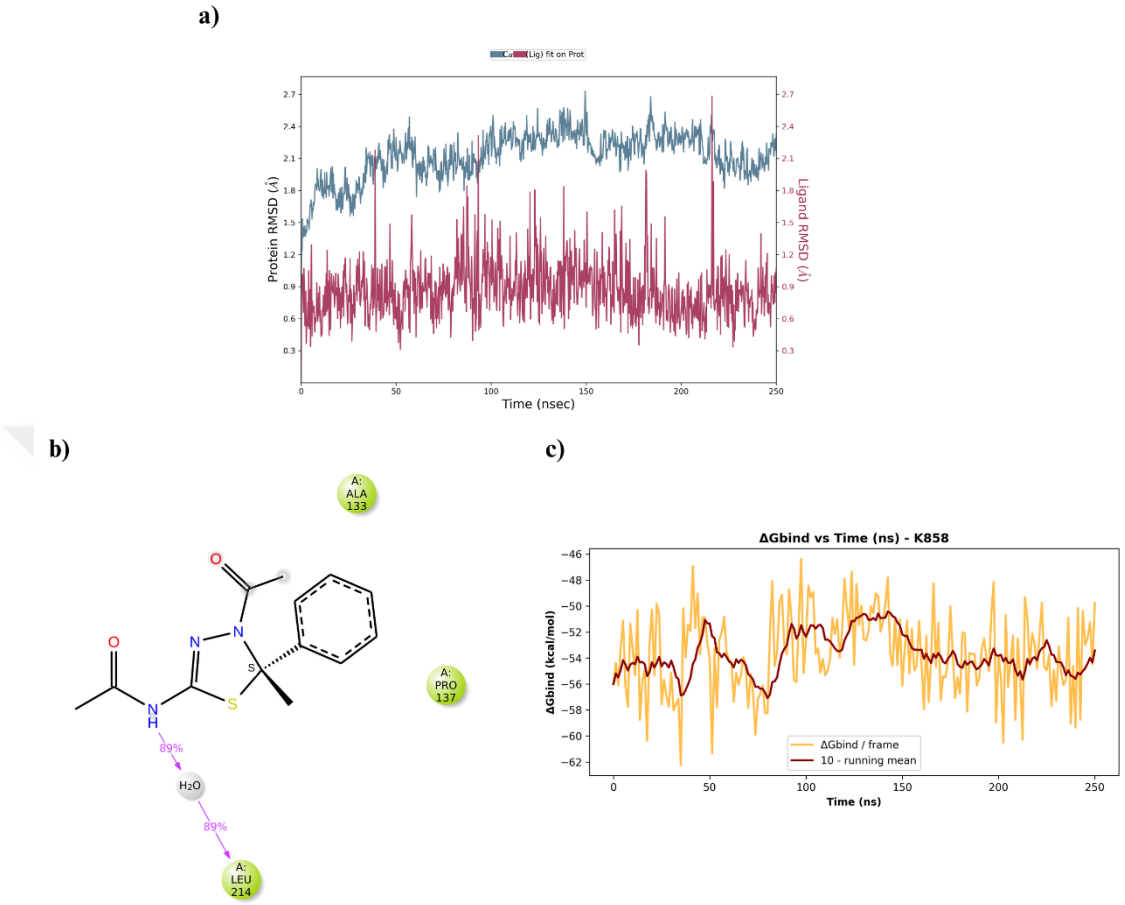

Figure 4.11: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - K858.

4.1.4.2 Stability and interaction analysis of candidates from track 1

From the 30 compounds advanced from the Anodyne screening track, 250 ns MD simulations were performed on each compound. This crucial step filtered out 26 candidates that displayed significant conformational instability or dissociated from their binding pockets. However, four compounds were observed to bind stably. Notably, for each of these chiral candidates, other stereoisomers were also generated and evaluated during the initial screening phases; they were subsequently eliminated because of inferior docking scores or less favorable interaction profiles, underscoring the stereospecificity of the binding and validating the selection of the specific isomers that advanced to this final stage of analysis.

The 2D chemical structures of these four promising candidates are shown in Figure 4.12. A detailed compound-by-compound analysis of each of these successful candidates is provided below, with each analysis supported by a comprehensive five-panel figure.

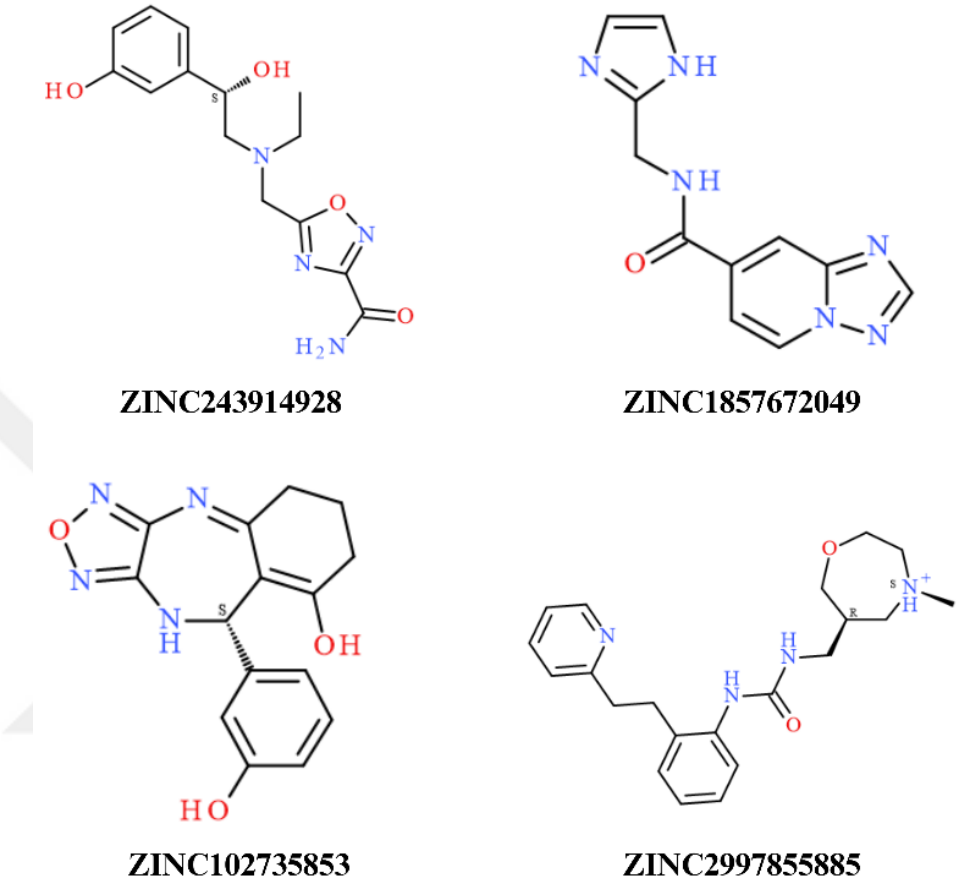

Figure 4.12: 2D chemical structures of the four lead compounds identified from track 1.

The key findings for the four candidates are summarized in Table 4.4, allowing for direct comparison with the reference inhibitor. A detailed analysis of each compound, supported by its corresponding comprehensive figure, is provided below.

Table 4.4 : Comparison of key values and key interactions for track 1 candidates.

Compound ID	Glide XP Score	MM/GBSA binding free energy (kcal/mol)	Key Interactions in the Binding pose	Average Trajectory MM/GBSA binding free energy (kcal/mol)
ZINC102735853	-7.452	-63.78	H-bond - Glu116, H-bond - Glu118, H-bond - Leu214 Aromatic H-bond - Trp127	-63.98 ± 5.10
ZINC299785885	-7.680	-66.88	H-bond - Glu116, H-bond - Gly117, Aromatic H-bond - Glu118, Aromatic H-bond - Leu214, Salt bridge - Glu118	-63.04 ± 6.03
ZINC1857672049	-7.714	-48.60	H-bond - Glu116, H-bond - Glu118, H-bond - Arg221	-49.25 ± 3.99
ZINC243914928	-6.967	-56.33	H-bond - Glu116, H-bond - Gly117, H-bond - Glu118	-49.34 ± 6.04

Analysis of ZINC102735853: As detailed in Figure 4.13, ZINC102735853 emerged as one of the most promising candidates. The initial binding pose, visualized in three dimensions (Figure 4.13a), revealed how the complex, multi-ring structure of the compound achieved excellent shape complementarity within the L5 allosteric pocket. The 2D interaction diagram (Figure 4.13b) further elucidates the specific molecular contacts that anchor the ligand; the key interactions include hydrogen bonds with the backbones of Glu116, Glu118, and Leu214, supplemented by an aromatic hydrogen bond with the side chain of Trp127.

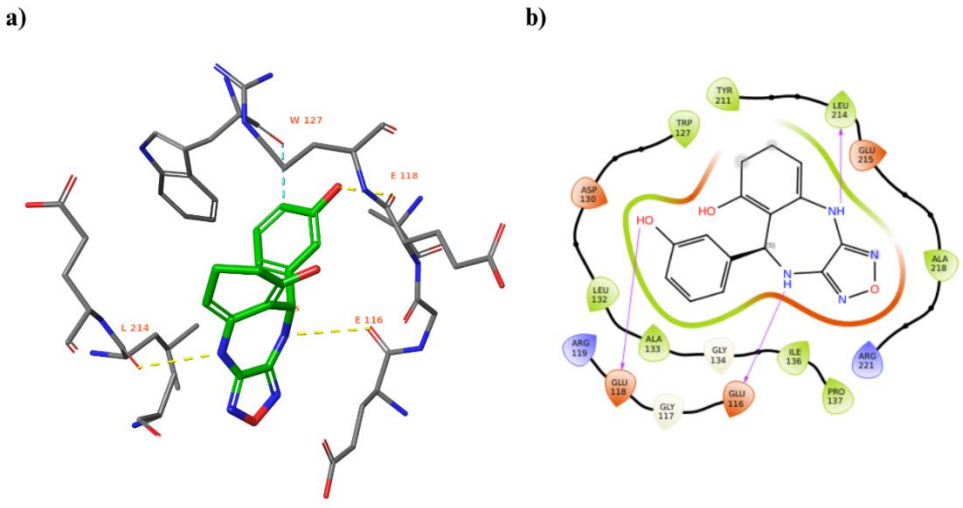

Figure 4.13: Binding pose representations of ZINC102735853.

MD simulations were then used to assess the durability of this binding mode. The protein Cα atoms remained stable (RMSD < 2.7 Å), and although the ligand RMSD plot (Figure 4.14a) indicated some flexibility within the pocket (remaining below 4.0 Å), the compound maintained its core position. The interaction timeline (Figure 4.14b) shows that the binding is dynamically anchored by highly persistent hydrogen bonds with the side chains of Glu116 (82% occupancy) and Asp130 (83% occupancy), which is further stabilized by a water-bridged interaction with Arg119 (39% occupancy). on profile is reflected in its exceptional MM/GBSA binding free energy of -63.98 ± 5.10 kcal/mol (Figure 4.14c). This value is significantly more favorable than that of the reference inhibitor K858 (− 53.66 kcal/mol), suggesting a predicted intrinsic binding affinity that could be substantially higher.

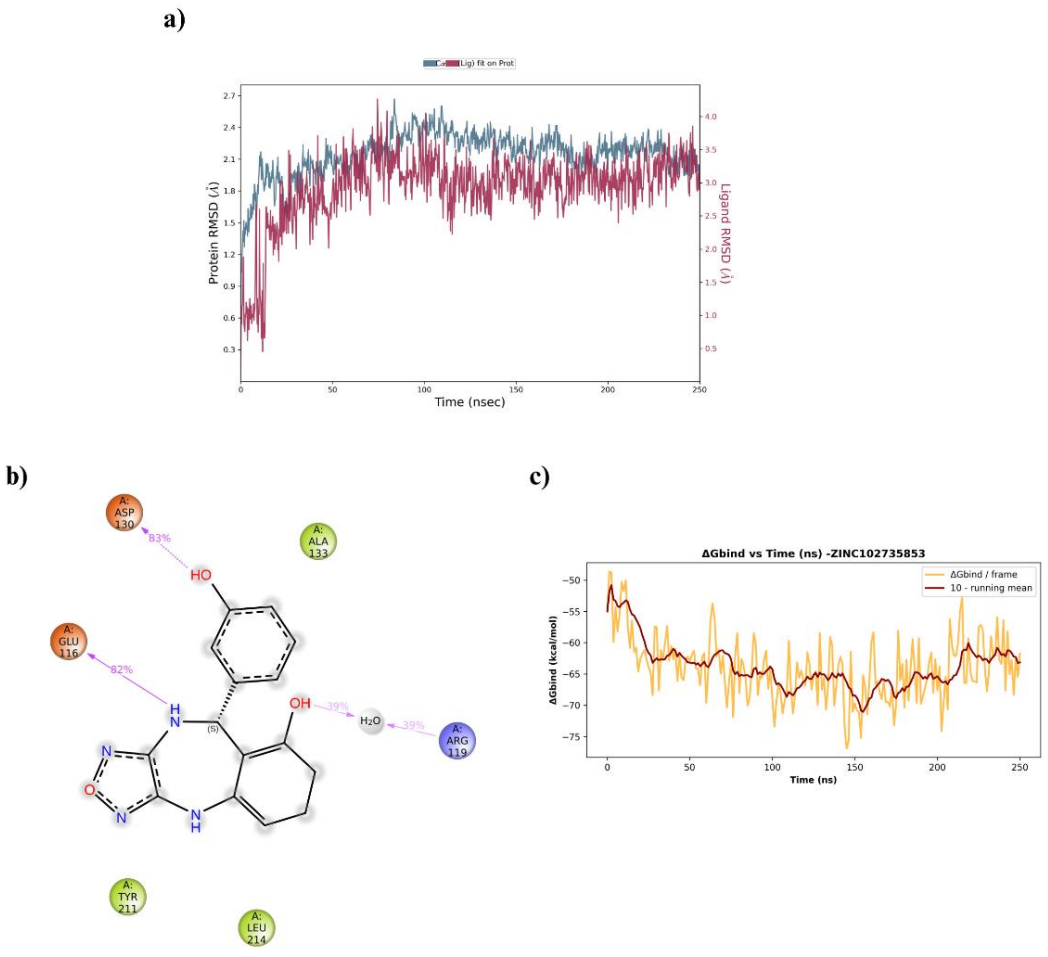

Figure 4.14: Molecular dynamics simulations and MM/GBSSA biNding free energy analysis of Eg5 - ZINC102735853.

Analysis of ZINC299785885: Compound ZINC299785885 also presented a highly compelling profile, as shown in Figure 4.15. The initial binding pose (Figure 4.15a, b) was stabilized by a rich network of interactions, including hydrogen bonds with the backbones of Glu116 and Gly117, aromatic hydrogen bonds with Glu118 and Leu214, and a salt bridge with the side chain of Glu118.

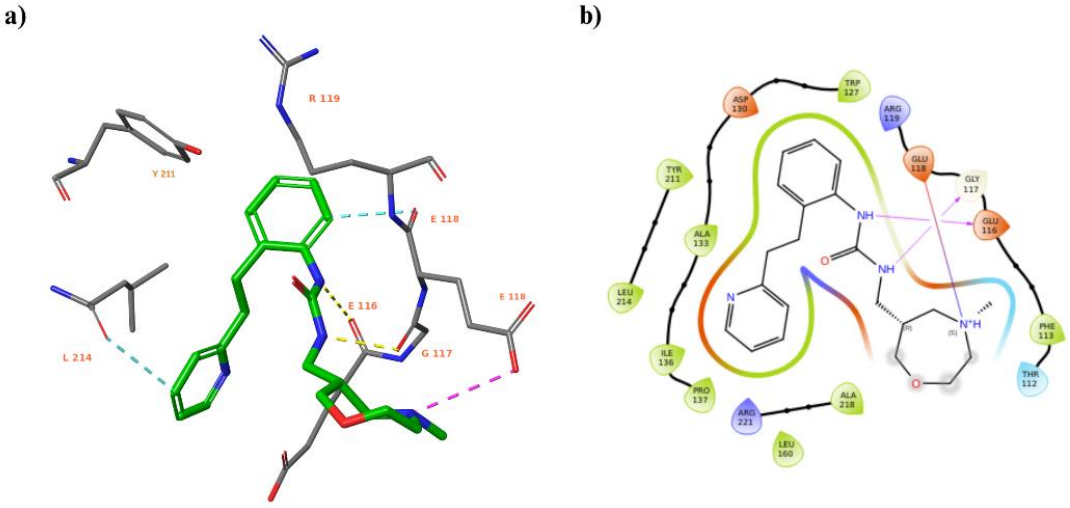

Figure 4.15: Binding pose representations of ZINC299785885.

The subsequent 250 ns MD simulation confirmed the exceptional stability of this complex, with both protein and ligand RMSD values remaining low and stable (< 2.7 Å) throughout the simulation (Figure 4.16a). The interaction timeline (Figure 4.16b) demonstrates that the binding is dynamically maintained by persistent hydrogen bonds, primarily with the backbone of Glu116 and a strong, enduring electrostatic interaction with the side chain of Glu118. This robust and stable binding is corroborated by the calculated MM/GBSA binding free energy of -63.04 ± 6.03 kcal/mol (Figure 4.16c), which is markedly superior to that of the reference inhibitor, further solidifying its potential as a high-affinity binder:

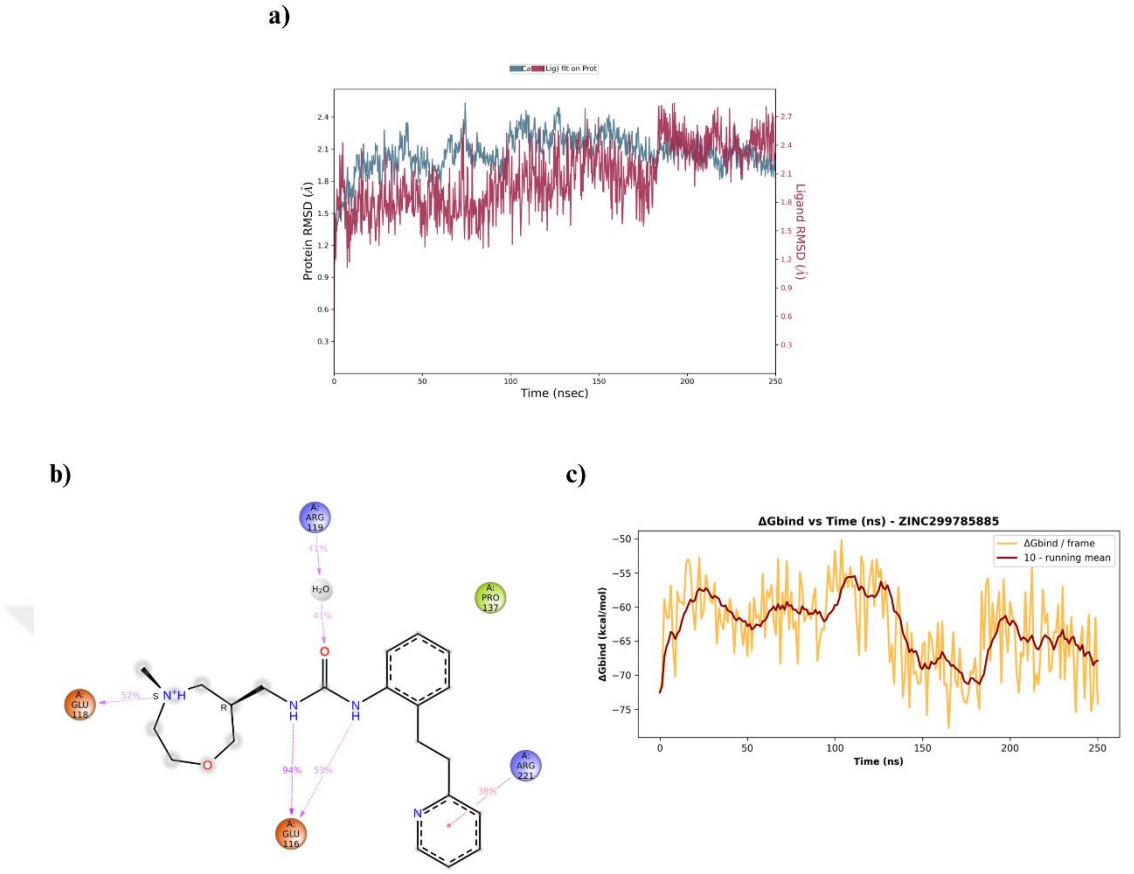

Figure 4.16: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC102735853.

Analysis of ZINC1857672049: The dynamic profile of ZINC1857672049 is shown in Figure 4.17. The initial binding pose (Figure 4.17a, b) revealed that the compound was anchored by hydrogen bonds with the key residues Glu116, Glu118, and Arg221.

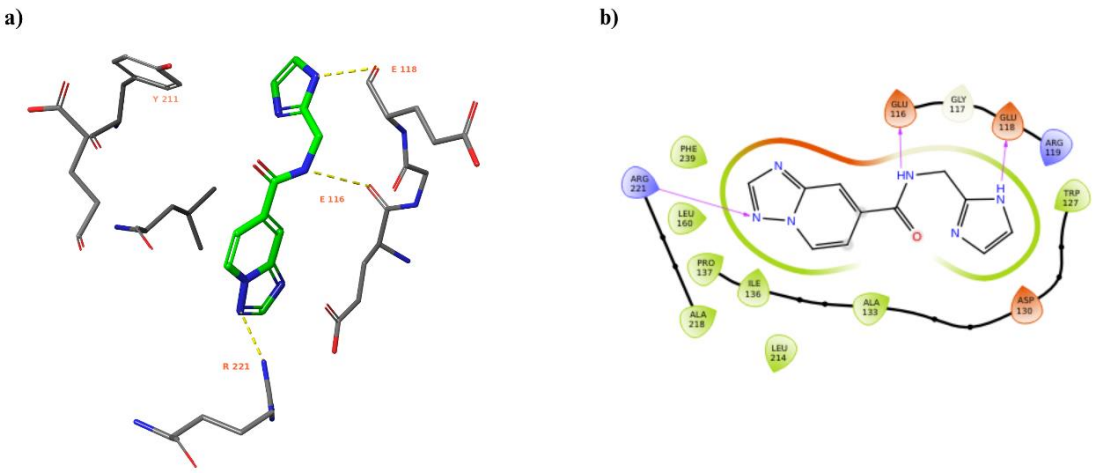

Figure 4.17: Binding pose representations of ZINC1857672049.

MD simulation confirmed the robustness of this binding mode, as the compound maintained excellent stability with low RMSD values (< 2.7 Å, Figure 4.18a). The interaction timeline (Figure 4.18b) further demonstrates the consistent maintenance of the critical hydrogen bonds with Glu116 and Glu118, which are essential for potent inhibition. The calculated MM/GBSA binding energy of -49.25 ± 3.99 kcal/mol (Figure 4.18c) was found to be slightly less favorable than that of the reference inhibitor. However, its confirmed dynamic stability and the persistence of these crucial interactions strongly validate it as an important compound. This result exemplifies the principle that, while the binding energy is a key metric, proven dynamic stability is an equally critical criterion for advancing a candidate, as it confirms the viability of the binding pose beyond a static model.

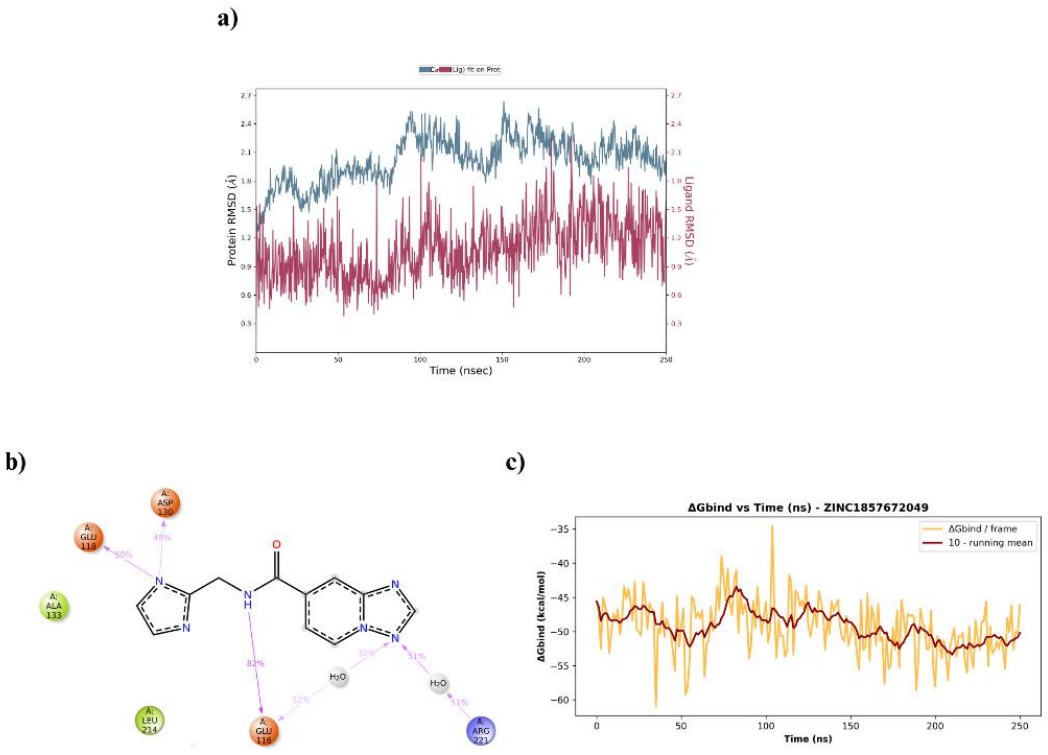

Figure 4.18: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC1857672049.

Analysis of ZINC243914928: The analysis of ZINC243914928 is shown in Figure 4.15. Its initial binding pose was anchored by a network of hydrogen bonds with Glu116, Gly117, and Glu118 (Figure 4.19a, b).

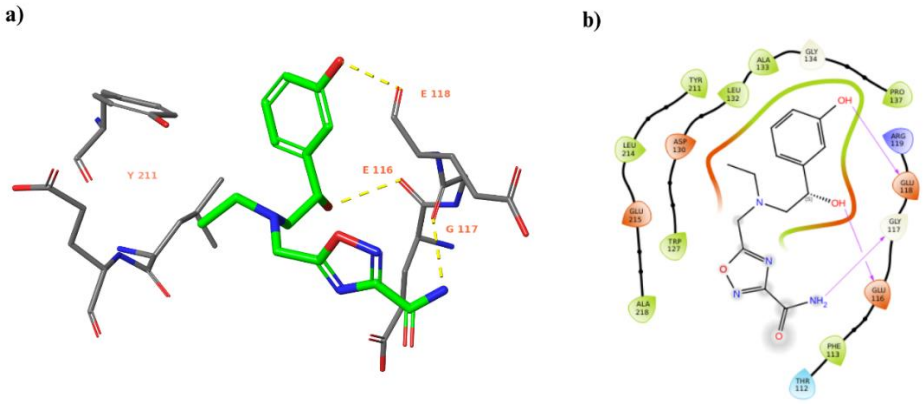

Figure 4.19: Binding pose representations of ZINC243914928.

MD simulation confirmed that the compound maintained a stable presence in the binding pocket, although it exhibited higher flexibility, as indicated by its ligand RMSD (< 5.4 Å) (Figure 4.20a). Despite this flexibility, the interaction timeline (Figure 4.20b) shows that it preserved a key hydrogen bond with Glu118 with high occupancy (83%), which is crucial for anchoring the compound. The calculated MM/GBSA binding energy of -49.34 ± 6.04 kcal/mol (Figure 4.20c) was comparable to that of ZINC1857672049 and slightly less favorable than that of the reference. Nevertheless, like ZINC1857672049, its proven stability in the binding pocket throughout the simulation makes it a reliable starting point for future medicinal chemistry optimizations.

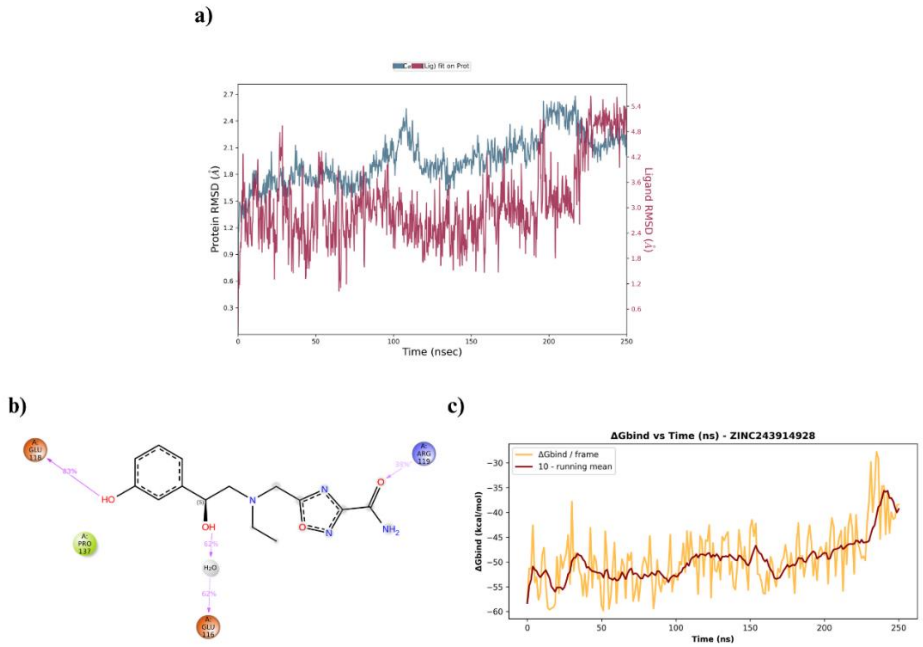

Figure 4.20: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC243914928.

The comprehensive analysis of these four MD trajectories provides crucial and nuanced insights into the outcomes of structure-based screening. The results revealed a clear stratification among the hits, with two compounds (ZINC102735853 and ZINC299785885) exhibiting calculated binding free energies that were significantly more favorable than those of the reference inhibitor, suggesting the potential for superior potency.

The other two compounds (ZINC1857672049 and ZINC243914928), although possessing slightly less favorable binding energies, were nevertheless validated as high-quality hits because of their confirmed dynamic stability and maintenance of essential interactions within the binding pocket. This differentiation, which would be difficult to achieve with static docking scores alone, highlights the power of trajectory-based energy calculations in distinguishing promising but energetically diverse candidates. Consequently, a robust rationale was established for the in vitro biological evaluation of all four compounds. The decision to progress all four was strategic, valuing not only the compounds with the most promising energetic profiles but also those that, despite more modest scores, demonstrated stable binding and presented structurally distinct scaffolds. This approach provides a broader and more diverse set of validated starting points for experimental testing and future medicinal chemistry studies.

4.1.4.3 Stability and interaction analysis of candidates from track 2

The exploratory DeepDocking campaign, designed to navigate a vast chemical space, culminated in the identification of a set of promising compounds. Following a comprehensive evaluation of the final candidates from the filtering cascade, a focused group of six lead candidates was selected for in-depth analysis based on their superior in silico metrics, structural novelty, and strong interaction profiles observed in the initial assessments. The analysis of these candidates reveals some of the most striking results of this thesis, powerfully demonstrating the success of a high-risk, high-reward strategy.

As detailed in our research summary, these compounds belong to at least four distinct chemical classes, including indole/indazole derivatives, triphenylmethane-like structures, and polycyclic amines, confirming that the primary goal of achieving significant scaffold diversity was successfully achieved. The 2D chemical structures

of these six candidates, which form the basis of the following analyses, are shown in Figure 4.21.

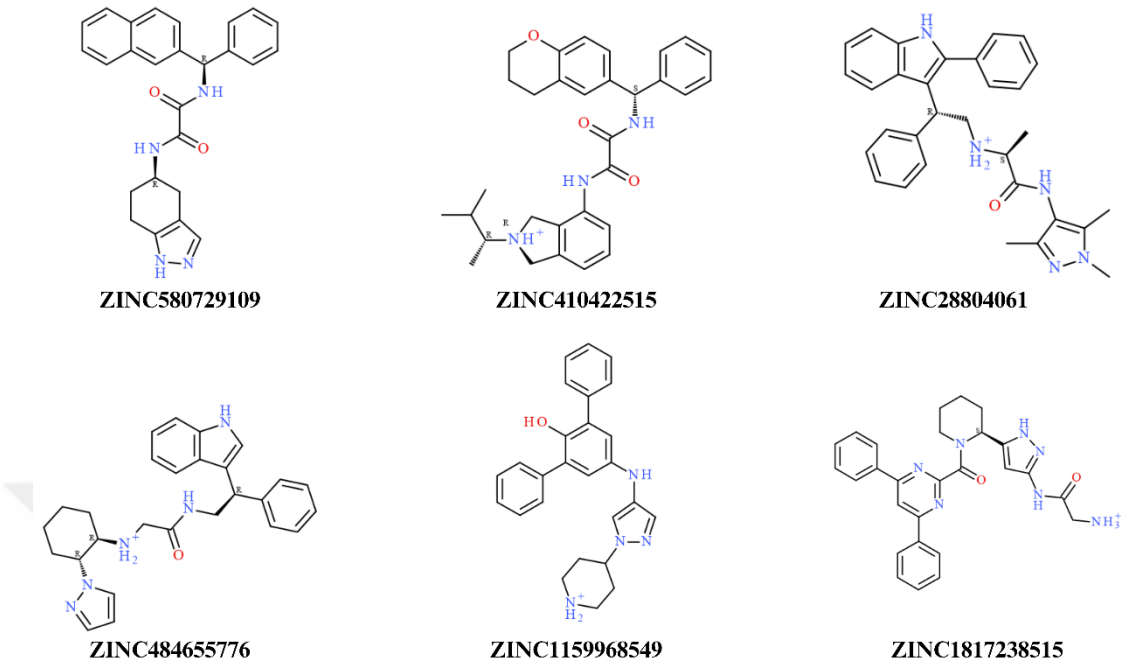

Figure 4.21: 2D chemical structures of the four lead compounds identified from track 1.

The comprehensive scoring metrics from the filtering cascade for the six compounds are summarized in Table 4.5. A direct comparison of their MM/GBSA rescoring energies with the trajectory-averaged binding free energy of the reference inhibitor K858 (–53.66 kcal/mol, as established in Section 4.1.4.1) is particularly enlightening. As the data indicate, nearly all candidates from this track showed markedly more favorable binding energies. This is not merely a statistical improvement but a testament to the power of the active learning track to discover entirely new chemical entities that may possess a potentially much higher affinity than known inhibitors.

Table 4.5 : Comparison of key values and key interactions for track 2 candidates.

Compound ID	DNN score (id_score)	Qvina2 Score	Glide XP Score	MM/GBSA value (kcal/mol)	Key Interactions in the Binding pose	Average Trajectory MM/GBSA binding free energy (kcal/mol)
ZINC580729109	0.9974	-11.400	-9.221	-79.46	H-bond - Glu116, Gly117, Arg119, Arg221, Tyr211 Aromatic H-bond - Tyr211, Glu116, Trp127, Leu214	-79.18 ± 5.38
ZINC410422515	0.9955	-10.800	-9.374	-95.44	H-bond - Glu118, Gly117, Arg119 Salt bridge - Glu118 Aromatic H-bond - Glu118, Trp127	-73.79 ± 4.78
ZINC28804061	0.8590	-10.500	-9.441	-99.12	H-bond - Glu116, Gly117, Leu214 Aromatic H-bond - Trp127 Pi-pi stacking - Tyr211	-71.65 ± 4.89
ZINC484655776	0.9913	-10.100	-9.397	-79.81	H-bond - Glu116, Gly117, Leu214 Aromatic H-bond - Trp127, Glu118	-69.26 ± 4.41
ZINC1159968549	0.9529	-10.600	-9.520	-77.35	H-bond - Glu118, Arg119 Aromatic H-bond - Trp127, Leu214, Arg119	-68.87 ± 4.58
ZINC1817238515	0.9840	-11.300	-10.603	-82.65	H-bond - Glu118, Arg119 Aromatic H-bond - Trp127, Arg119 Pi-pi stacking - Tyr211	-67.60 ± 5.78

The quality and potential of these novel hits are illustrated through detailed compound-by-compound analyses, integrating the findings from the MD simulations with the known pharmacophore requirements of the L5 pocket. This will include an in-depth examination of their dynamic stability and interaction profiles, providing a comprehensive evaluation of their potential as next-generation inhibitors.

Analysis of ZINC580729109 (Indole/Indazole Derivative)

ZINC580729109, an indole/indazole derivative featuring a urea-like scaffold, emerged as one of the most compelling candidates from the DeepDocking track. A detailed analysis integrating both static and dynamic data is presented in Figure 4.22.

The initial binding pose, visualized in three dimensions (Figure 4.22a), revealed how the compound achieved excellent shape complementarity within the L5 allosteric pocket. The 2D interaction diagram (Figure 4.22b) further elucidates the specific and extensive network of interactions anchoring the compound. The compound establishes a multifaceted connection with the receptor, forming direct hydrogen bonds with the key residues Glu116, Gly117, Arg119, and Arg221. Its interaction with Tyr211 is particularly noteworthy because it engages in both a direct hydrogen bond and an

aromatic interaction. The hydrophobic character of the binding is reinforced by additional aromatic interactions with residues Glu116, Trp127, and Leu214. This rich network of contacts in the initial pose suggests a strong potential for high affinity binding.

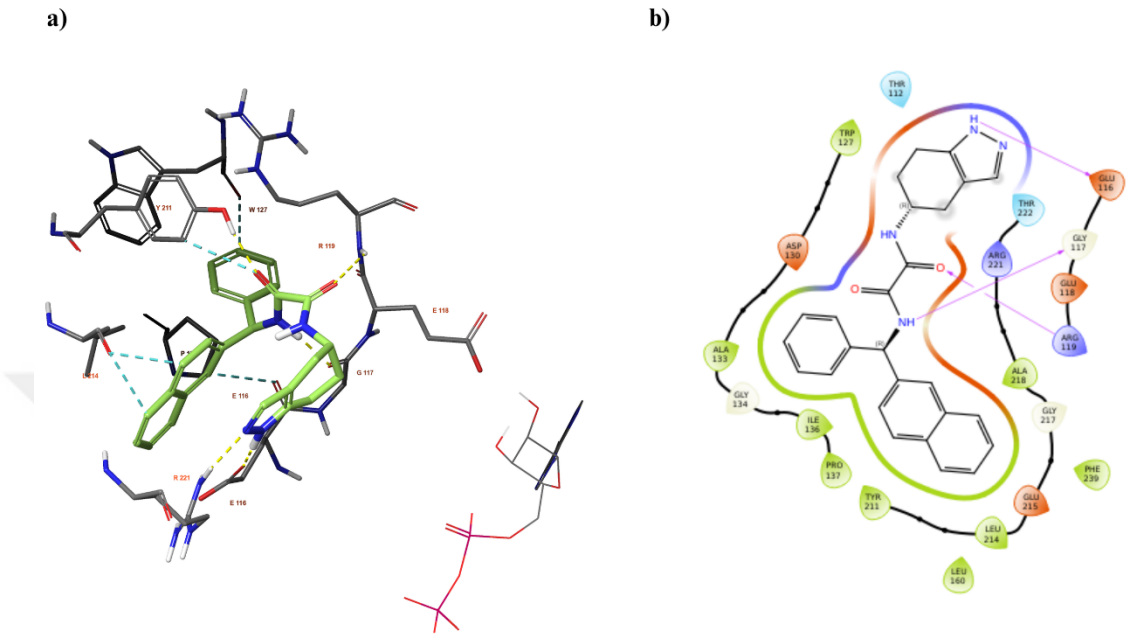

Figure 4.22: Binding pose representations of ZINC580729109.

A 100 ns MD simulation confirmed the exceptional stability of this complex. The ligand RMSD remained consistently low, generally below 2.5 Å after an initial equilibration period (Figure 4.23a), indicating a tight and well-anchored fit with minimal conformational fluctuations.

Interaction analysis over the trajectory provides a dynamic validation of the initial pose, demonstrating that a core subset of these interactions is maintained with high persistence, providing a durable anchor for the ligand. As illustrated in the interaction summary diagram (Figure 4.23b), the ligand established a remarkably robust polar interaction network. A highly persistent hydrogen bond with the side chain of Glu116 was maintained for 83% of the simulation time, serving as a primary electrostatic anchor. This is exceptionally supplemented by a dual hydrogen bonding interaction with the backbone of Gly117, with occupancies of 48% and 87%, respectively, effectively locking the central urea-like scaffold in place. Furthermore, the complex was stabilized by two distinct water-bridged hydrogen bonds with Arg119, each maintained for 63% of the simulation time. This dense and multipoint polar network, a feature not available to many neutral reference inhibitors, explains the exceptional

stability of the compound. This dynamically robust interaction network is reflected in the calculated MM/GBSA binding free energy of -79.18 ± 5.38 kcal/mol (Figure 4.23c). This value is approximately 25 kcal/mol more favorable than that of the reference inhibitor K858 (–53.66 kcal/mol) and significantly better than those of the best candidates from Track 1, highlighting the remarkable potential of this novel scaffold.

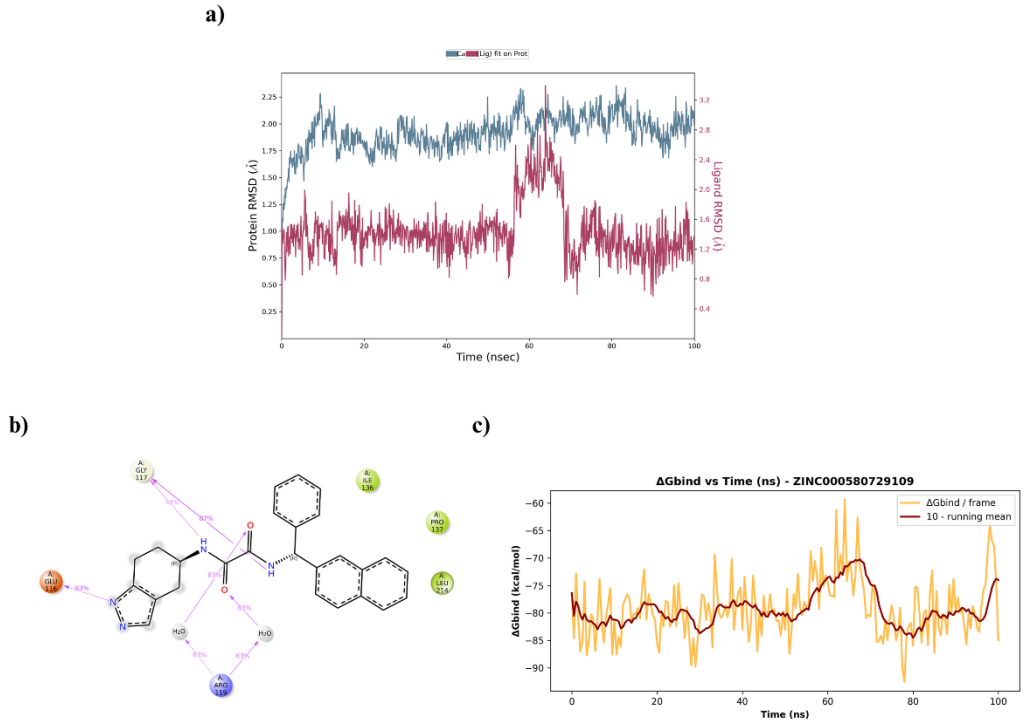

Figure 4.23: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC580729109.

Analysis of ZINC410422515 (Polycyclic Amine Derivative)

The second noteworthy candidate, ZINC410422515, features a novel polycyclic amine scaffold linked to a urea moiety, representing another distinct chemical class identified in this study. A comprehensive analysis is presented in Figure 4.24. The initial binding pose (Figure 4.24a, b) illustrated a remarkably dense network of interactions, where the compound was anchored by a multifaceted connection with Glu118, engaging in a hydrogen bond, salt bridge, and aromatic interaction simultaneously. This is further supported by hydrogen bonds to the backbones of Gly117 and Arg119 and an aromatic contact with Trp127, highlighting the scaffold's potential for high-affinity binding.

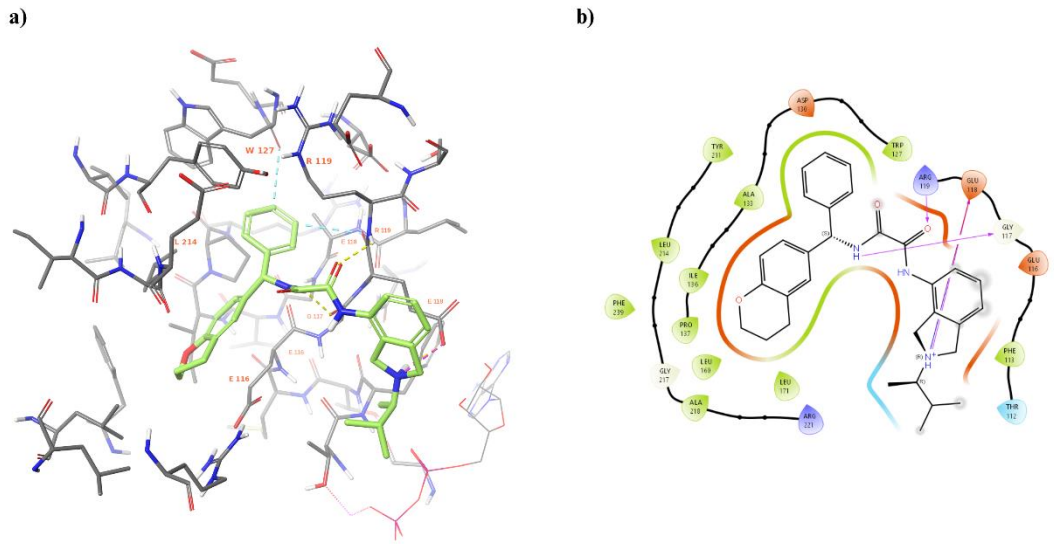

Figure 4.24: Binding pose representations of ZINC410422515.

A 100 ns MD simulation was employed to test the dynamic viability of this pose. The ligand RMSD plot (Figure 4.25a) shows that the compound remained stably bound within the active site throughout the simulation, with RMSD values generally fluctuating between 2.0 and 3.0 Å after an initial equilibration, confirming that the compound did not dissociate and maintained a consistent binding orientation. Interestingly, the interaction analysis over the trajectory (Figure 4.25b) revealed a dynamic adaptation from the initial pose to a more focused and persistent set of interactions. The primary anchor points shift to a consistent hydrogen bond with the backbone of Gly117 (62% occupancy) and a stable water-bridged hydrogen bond with Glu116 (49% occupancy). The initial dense interactions with Glu118 and Arg119 evolved into a stabilizing π-cation interaction with Arg119 (38% occupancy) and an additional water-bridged contact with Glu215 (31% occupancy). This transition from a broad initial contact map to a refined, dynamically stable network realistically represents ligand binding and provides strong evidence for durable interactions. This unique combination of a rigid hydrophobic body and a dynamically adapted network of polar interactions resulted in a highly favorable calculated MM/GBSA binding free energy of -73.79 ± 4.78 kcal/mol (Figure 4.25c). This energy is not only significantly better than that of the reference inhibitor K858 (–53.66 kcal/mol) but also validates the potential of targeting the L5 pocket with such structurally complex and rigid amine-containing scaffolds, which are capable of forming stable, solvent-mediated interaction networks.

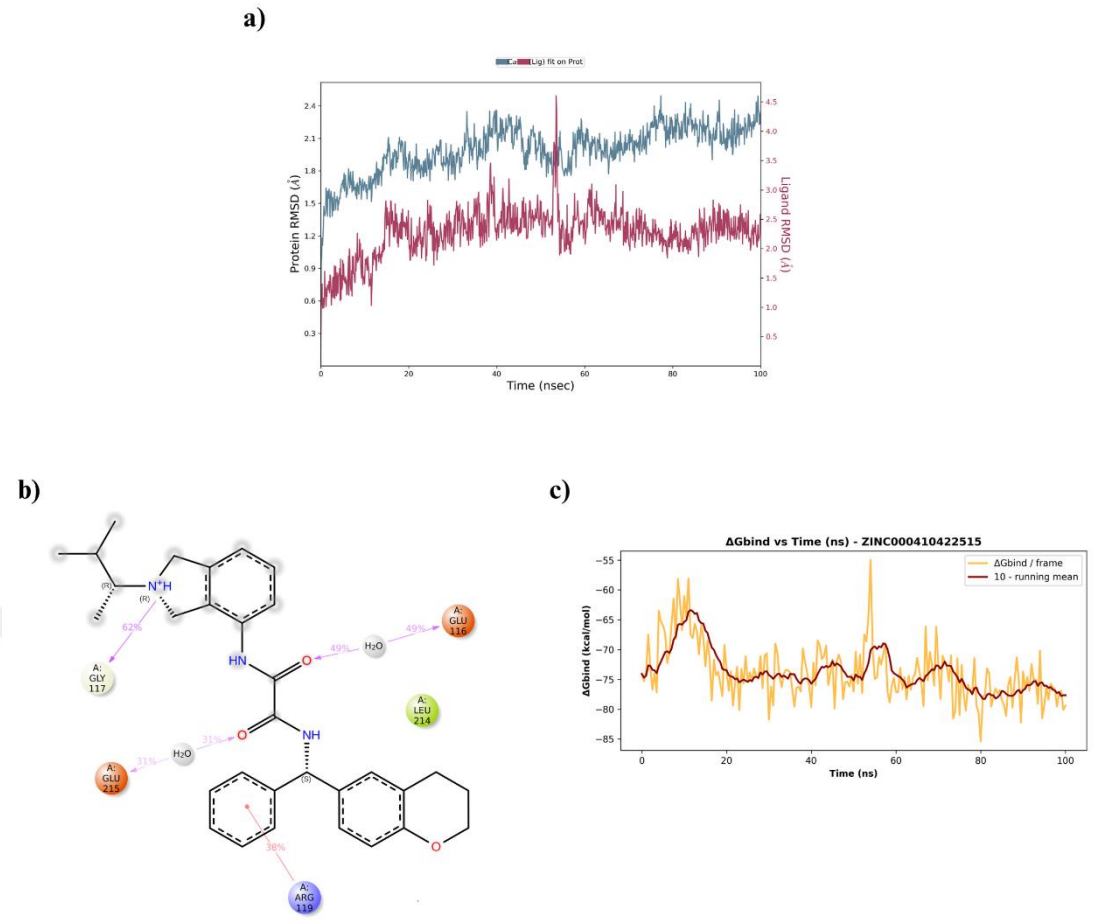

Figure 4.25: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC410422515.

Analysis of ZINC28804061 (Pyrazole-indole derivative)

This candidate, featuring a pyrazole-indole core linked via an amide moiety, represents another novel and distinct scaffold identified in the DeepDocking campaign. Its structural characteristics suggest a unique way to satisfy the pharmacophoric requirements of the L5 pocket. A comprehensive analysis is presented in Figure 4.26.

The initial binding pose (Figure 4.26a and b) revealed the capacity of the compound to form an extensive network of interactions. The pyrazole group is positioned to form hydrogen bonds with Glu116 and Gly117, whereas the indole and phenyl rings engage in a rich network of hydrophobic and aromatic contacts, including π–π stacking with Tyr211, an aromatic hydrogen bond with Trp127, and a hydrogen bond with Leu214. This dense web of initial interactions strongly suggests a high potential for binding.

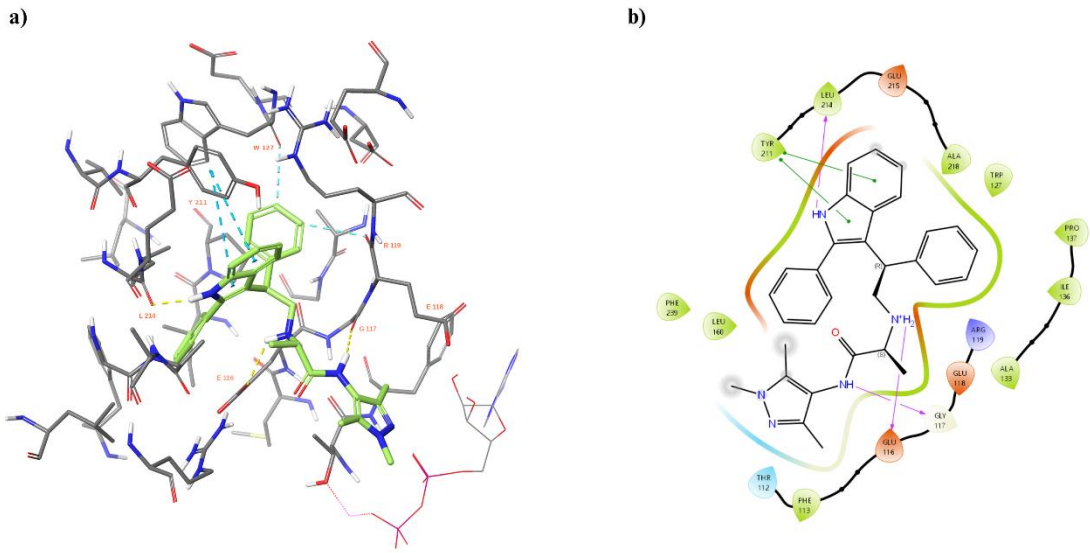

Figure 4.26: Binding pose representations of ZINC28804061.

The 100 ns MD simulation confirmed that the compound maintained a stable binding mode. The ligand RMSD plot (Figure 4.27a) shows that although the compound exhibits some flexibility, it remains well-seated within the binding pocket without any signs of dissociation, with RMSD values stabilizing after an initial equilibration period.

The dynamic interaction analysis over the trajectory (Figure 4.27b) is particularly insightful, revealing a highly stable and persistent network of contacts anchoring the compound. Two extremely high-occupancy hydrogen bonds were formed: one between the ligand's amine group and the side chain of Glu116 (99% occupancy) and another between the amide oxygen and the side chain of Arg221 (90% occupancy). These two interactions form a powerful electrostatic "clamp" that locks the compound into place. This is further reinforced by another strong and persistent hydrogen bond with the backbone of Leu214 (99% occupancy), which provides a robust hydrophobic anchor.

This robust network of dynamically stable interactions resulted in a highly favorable calculated MM/GBSA binding free energy of -71.65 ± 4.89 kcal/mol (Figure 4.27c). This energy is significantly better than that of the reference inhibitor K858 (–53.66 kcal/mol), further validating the potential of this unique pyrazole-indole scaffold as a high-affinity Eg5 inhibitor.

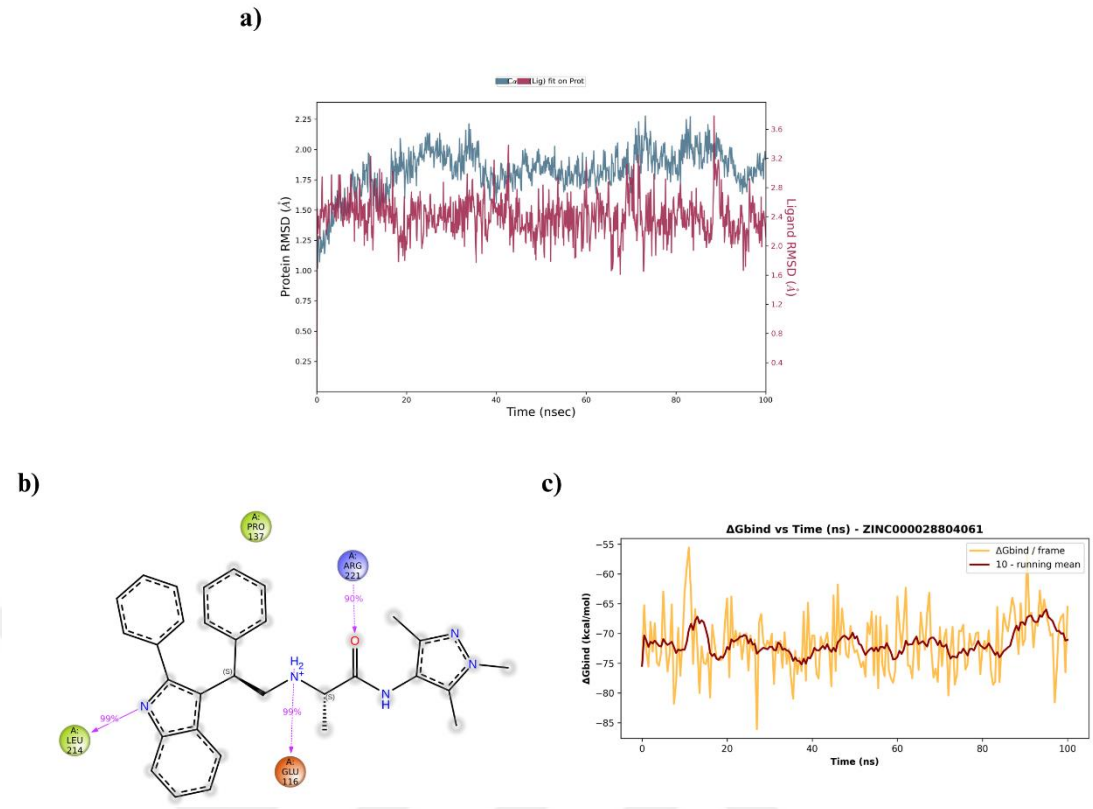

Figure 4.27: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC28804061.

Analysis of ZINC484655776 (Indole-piperidine derivative)

This compound represents another highly novel chemotype characterized by a central indole core linked to a large, conformationally restricted piperidine-containing ring system. This unique architecture provides a rigid scaffold that pre-organizes key pharmacophoric elements for interaction with the L5 pocket. A detailed analysis is presented in Figure 4.28.

The initial binding pose (Figure 4.28a, b) revealed a rich network of initial contacts anchoring the compound. These include direct hydrogen bonds to the side chain of Glu116 and the backbone of Gly117, an aromatic hydrogen bond with Trp127, and another with Glu118, respectively. The indole ring itself forms a further key hydrogen bond with the backbone of Leu214 while also engaging in hydrophobic contacts. This dense web of interactions rationalizes the high initial docking score of the compound and suggests a strong potential for a stable complex.

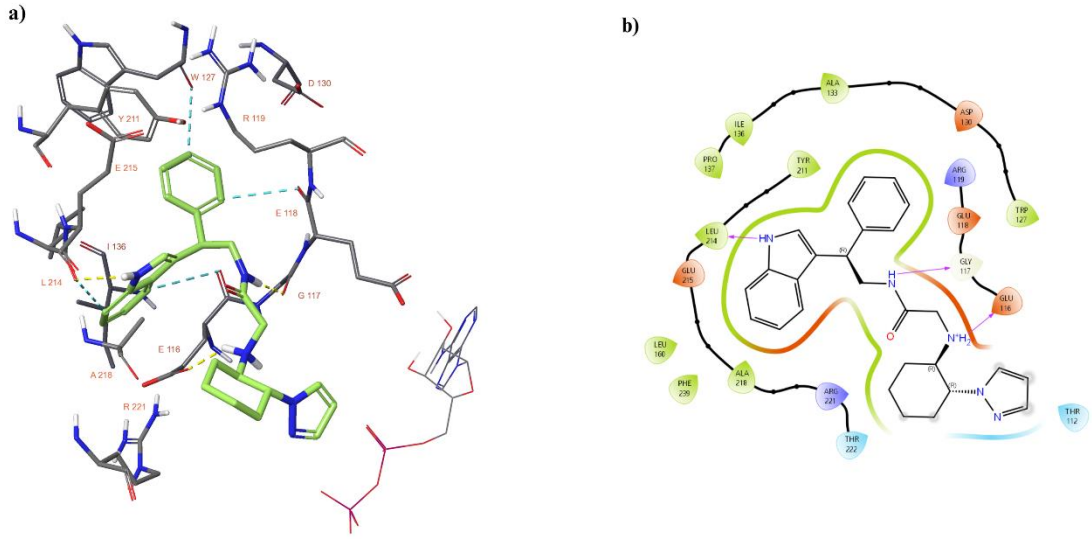

Figure 4.28: Binding pose representations of ZINC484655776.

The 100 ns MD simulation provided strong evidence of the exceptional dynamic stability of this complex. As shown in the ligand RMSD plot (Figure 4.29a), after a brief initial equilibration, the compound achieved a very stable pose, with RMSD values remaining consistently low (generally below 2.5 Å) for the duration of the trajectory. This indicates a tight and well-defined binding mode with minimal conformational fluctuations.

Dynamic interaction analysis (Figure 4.29b) revealed the specific high-occupancy contacts responsible for this stability. Binding is dominated by a powerful and persistent network of hydrogen bonds. An extremely durable hydrogen bond is maintained between the amide hydrogen of the ligand and the backbone of Gly117 (98% occupancy). This is complemented by a second, very strong hydrogen bond between the protonated amine and the side chain of Glu116 (82% occupancy). This primary dual-anchor system is further stabilized by a persistent hydrophobic contact between the indole ring and Leu214 (97% occupancy), effectively "clamping" the compound into the active pocket. In addition to these dominant interactions, the stability of the compound is reinforced by a network of moderately persistent water-bridged hydrogen bonds, most notably with Arg221 (46% occupancy) and Gly217 (33% occupancy), which help to correctly orient the bulky ring system and satisfy additional polar contacts.

The energetic consequence of this durably anchored and well-fitted pose is a calculated MM/GBSA binding free energy of -69.26 ± 4.41 kcal/mol (Figure 4.29c). This value

is substantially more favorable than that of the reference inhibitor K858 and is competitive with the other top-tier candidates from this screening track. The discussion for ZINC484655776 revolves around the concept of "scaffold rigidity" and "pre-organization." The fused-ring system enforces a specific 3D conformation, which minimizes the entropic penalty upon binding and ensures that the key interacting groups (indole, amide, and basic amine) are precisely positioned for optimal contact. This pre-organization is a classic medicinal chemistry strategy for enhancing the affinity. The discovery of a novel, rigid scaffold that achieves a stable, multipoint binding mode and favorable energy profile is a significant outcome of the DeepDocking campaign, presenting a distinct and highly promising avenue for lead optimization.

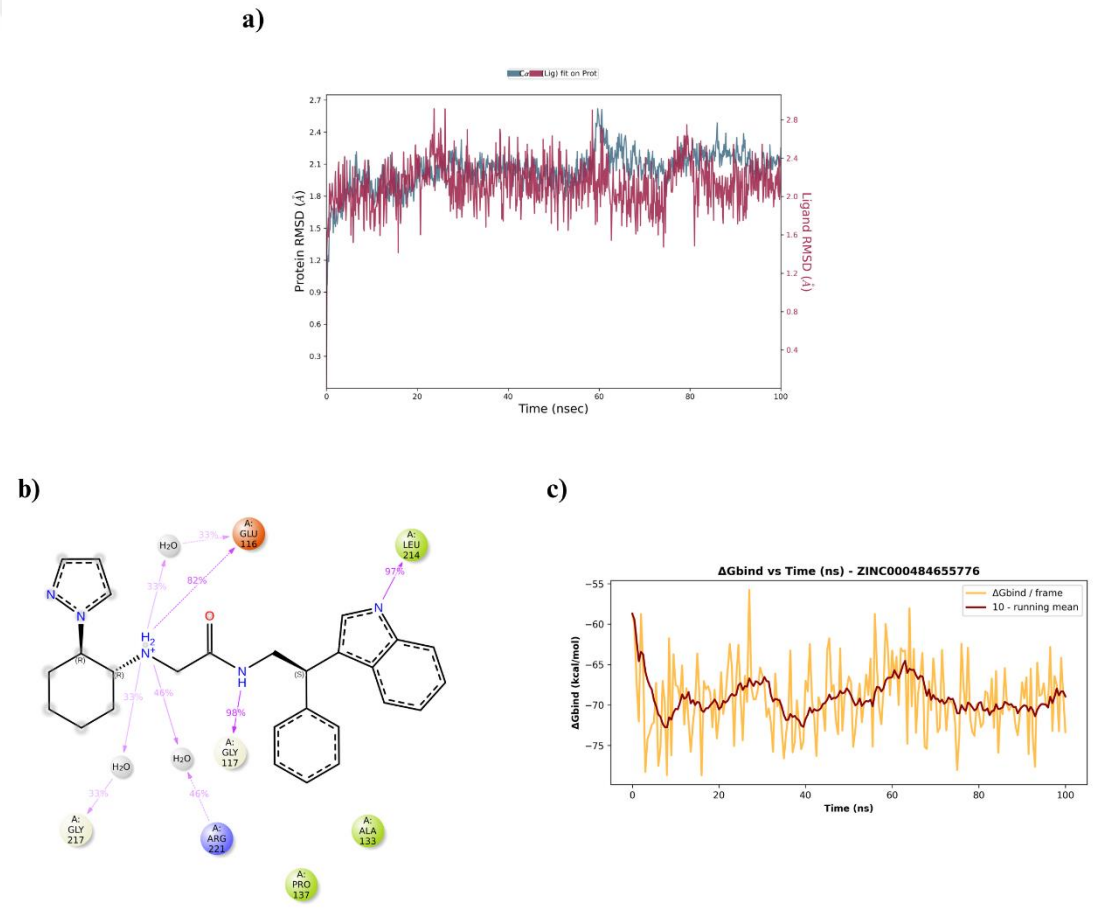

Figure 4.29: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC484655776.

Analysis of ZINC1159968549 (Biphenyl-azole derivative)

This candidate introduces a biphenyl-azole scaffold, a classic motif in medicinal chemistry, which is a novel entry for the Eg5 L5 pocket in the context of our screen.

The structure consists of two phenyl rings linked by a single bond, providing rotational freedom, and connected to a nitrogen-containing azole ring that serves as a polar anchor. This combination of a flexible hydrophobic body and polar head is a well-established strategy for achieving high affinity. A detailed analysis is presented in Figure 4.30.

The initial binding pose (Figure 4.30a and b) revealed a complex network of interactions. The compound is anchored by a direct hydrogen bond to Glu118 and another to Arg119. The aromatic systems engage in multiple favorable contacts, including aromatic hydrogen bonds with Trp127, Leu214, and Arg119, which effectively position the biphenyl moiety within the hydrophobic channels.

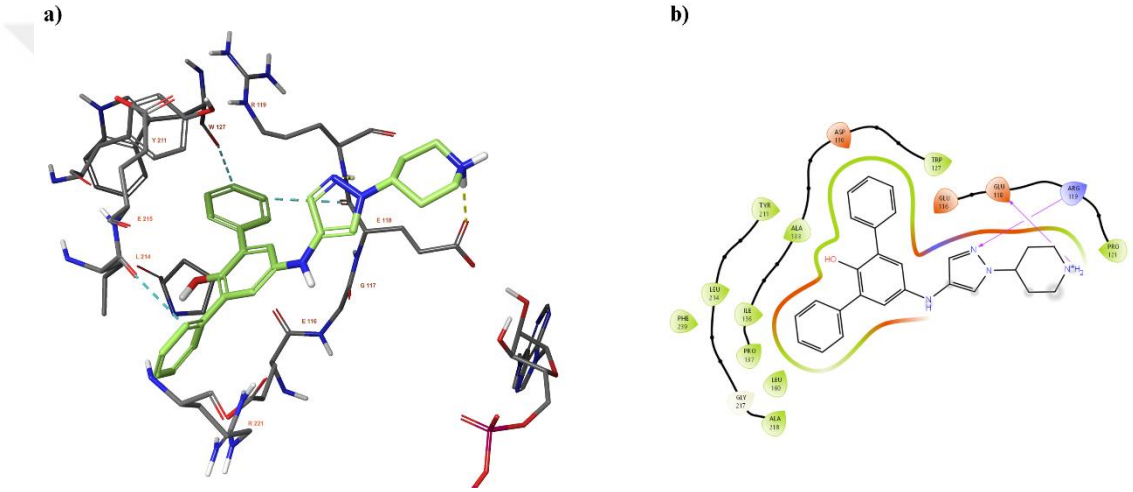

Figure 4.30: Binding pose representations of ZINC1159968549.

The 100 ns MD simulation demonstrated that the compound achieved a remarkably stable binding mode. The ligand RMSD plot (Figure 4.31a) is exceptionally flat, with fluctuations well below 0.5 Å for most of the trajectory. This indicates a highly rigid and well-defined binding pose with almost no conformational freedom, contrary to what might be expected from a flexible biphenyl linker. This suggests that the compound adopts a single, highly favorable conformation upon binding.

The interaction timeline (Figure 4.31b) confirms that this exceptional stability is maintained by a strong and persistent anchoring interaction. A direct hydrogen bond between the protonated piperidine and the side chain of Glu118 was maintained with an exceptionally high occupancy of 98%. This is supplemented by a direct, moderately persistent hydrogen bond with Arg119 (45% occupancy). The polar network is further stabilized by water-bridged hydrogen bonds with Glu116 (54%) and Gly117 (34%),

which collectively tether the polar head of the compound, allowing the biphenyl tail to settle into a stable, optimal position.

The overall energetic profile reflected this successful binding strategy. The calculated MM/GBSA binding free energy was -68.87 ± 4.58 kcal/mol (Figure 4.31c), which was highly competitive and significantly better than that of the reference inhibitor. The discussion for this candidate is particularly interesting because the MD results challenge the initial assumption of flexibility. Instead of dynamic fitting, the biphenyl-azole scaffold adopts a single, "locked" conformation that satisfies all key interactions simultaneously. This "conformationally-selected" binding is highly advantageous because it minimizes entropic penalties and leads to a very stable complex. The success of this scaffold demonstrates that the L5 pocket can accommodate and lock moieties with rotatable bonds, provided they are flanked by a sufficiently strong polar anchor. This presents another promising design philosophy for medicinal chemistry exploration.

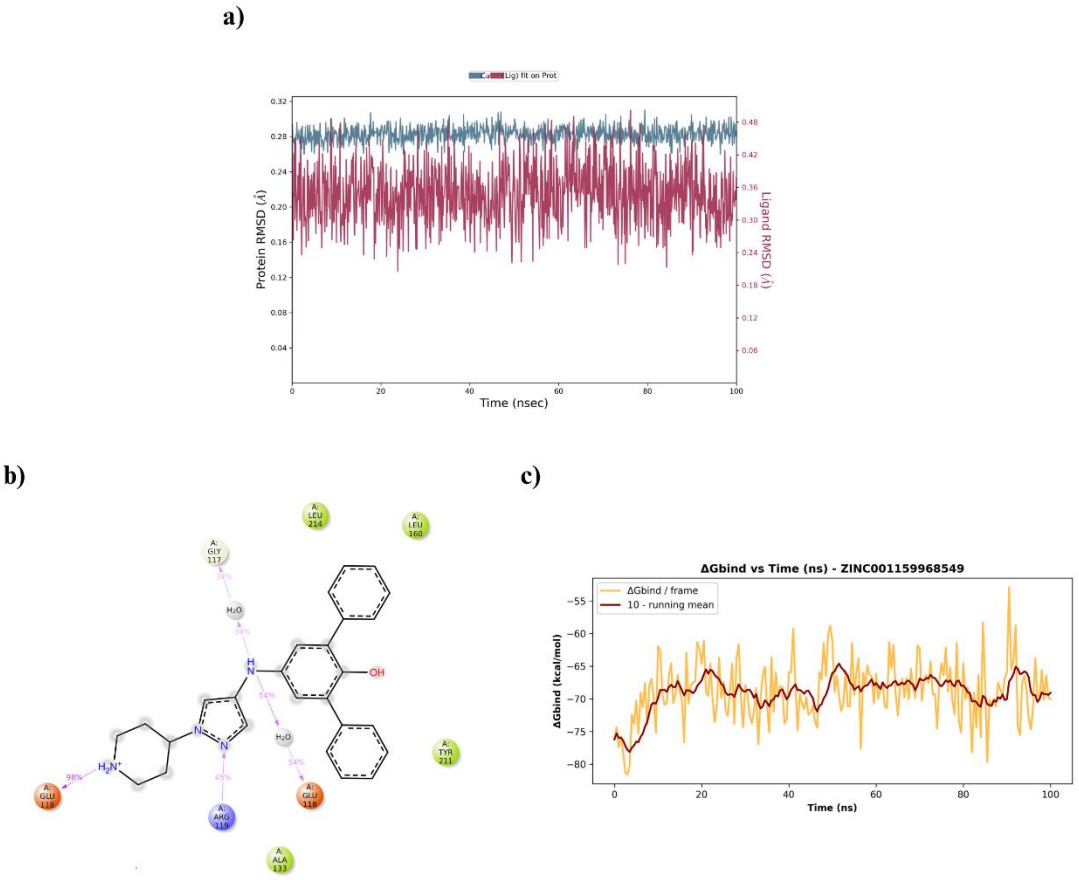

Figure 4.31: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC1159968549.

Analysis of ZINC1817238515 (Diaryl-heterocycle piperidine derivative)

The final candidate for detailed analysis, ZINC1817238515, features a complex diaryl heterocyclic core appended with a flexible piperidine-containing chain. This structure combines a large, relatively rigid aromatic system designed to occupy the core hydrophobic regions of the pocket with a flexible polar tail capable of seeking key electrostatic interactions. A comprehensive analysis of this compound is presented in Figure 4.32.

The initial binding pose (Figure 4.32a, b) places the diaryl-heterocycle deep within the L5 pocket, where it establishes an extensive network of favorable contacts. These include direct hydrogen bonds to Glu118 and Arg119 and aromatic hydrogen bonds to Trp127 and Arg119. The large aromatic core also establishes a π-π stacking interaction with Tyr211, effectively occupying the hydrophobic space.

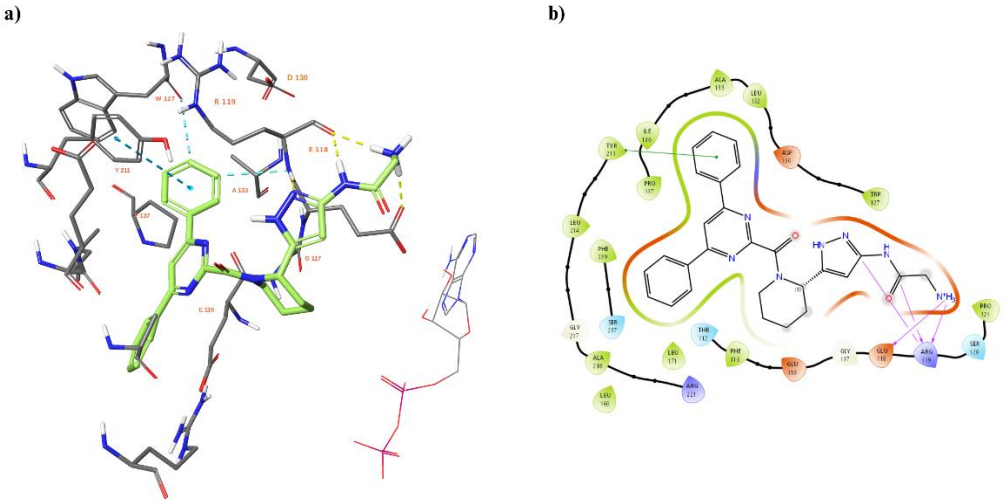

Figure 4.32: Binding pose representations of ZINC1817238515.

The 100 ns MD simulation reveals a stable binding mode characterized by a "tethered flexibility." The ligand RMSD plot (Figure 4.33a) was particularly informative. It shows an initial period of rearrangement (up to ~15 ns), where the ligand RMSD reaches ~2.5 Å, after which it settles into a much more stable state, fluctuating consistently between 1.0 Å and 2.0 Å for the remainder of the simulation. This behavior suggests that the ligand finds a more stable and energetically favorable pose upon initial binding than that predicted by static docking.

The dynamic interaction analysis (Figure 4.33b) provides a clear picture of the binding mechanism. The complex is anchored by a highly persistent salt bridge and hydrogen

bond between the flexible tail and the side chain of Arg119 (40% direct H-bond plus significant water-bridged interactions). This interaction with Arg119 is the primary tether. This is complemented by a consistent water-bridged hydrogen bond with Glu118 (33%) and another water-bridged contact with Arg221, which collectively stabilize the polar head of the compound. The diaryl core maintained extensive and stable hydrophobic contacts with the nonpolar residues lining the pocket, including Pro137 and Ile136, throughout the simulation.

The calculated MM/GBSA binding free energy for this complex was -67.60 ± 5.78 kcal/mol (Figure 4.33c). This highly favorable energy, which is competitive with the other top hits and superior to that of the reference inhibitor, validates the binding hypothesis. The discussion for this candidate centers on the "anchor and explore" binding strategy. The compound leverages a potent polar anchor, primarily Arg119, to secure itself to the protein. This allows its large hydrophobic component to establish favorable contacts within a broader pocket. This strategy is distinct from the rigid, pre-organized binders or multi-point H-bond networks observed in other candidates. This represents a viable alternative path to high-affinity inhibition and offers unique opportunities for optimization, such as modifying the flexible linker to alter the pharmacokinetic properties without disturbing the core binding interactions. The success of this scaffold further enriches the chemical diversity of our validated hits and underscores the power of our screening workflow in identifying inhibitors with varied and sophisticated binding mechanisms.

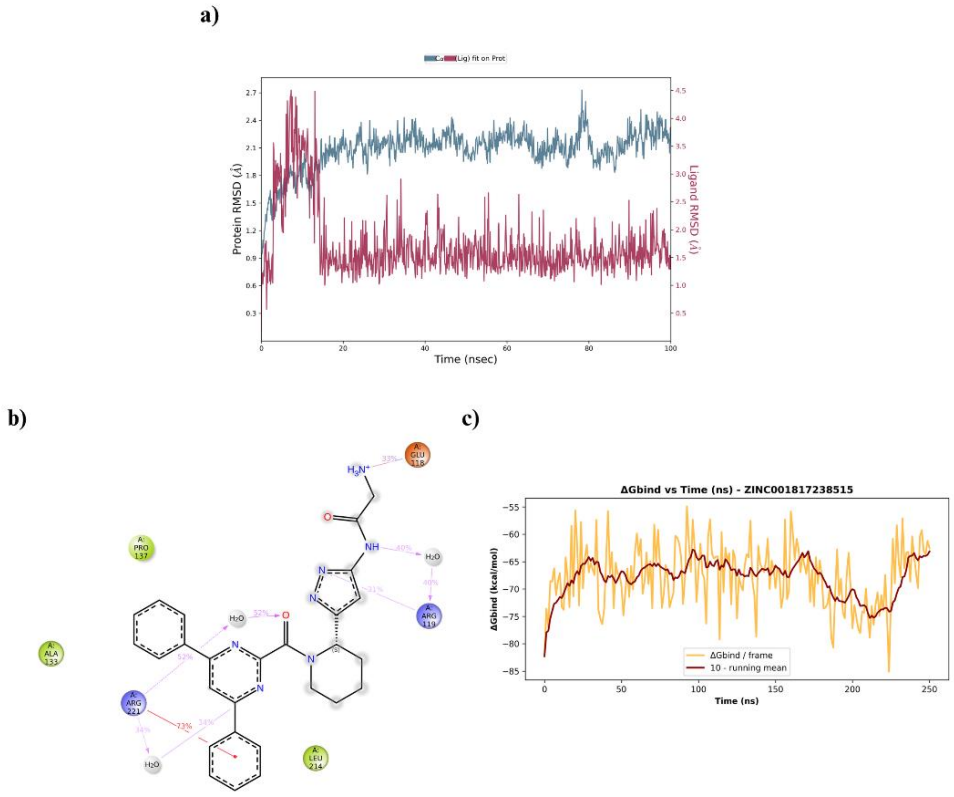

Figure 4.33: Molecular dynamics simulations and MM/GBSSA binding free energy analysis of Eg5 - ZINC1817238515.

4.2 In Vitro biological evaluation of lead candidates

The ultimate validation of any in silico drug discovery campaign lies in experimental testing. In this study, the four most promising candidates identified from the more conservative Track 1 screening pipeline (ZINC102735853, ZINC299785885, ZINC1857672049, and ZINC243914928) were advanced for in vitro evaluation of their Eg5 ATPase-inhibitory activity. The results of these assays not only provide a crucial reality check for our computational predictions but also establish a vital benchmark against which the extraordinary potential of the Track 2 candidates can be contextualized in the future.

4.2.1 Kinesin ATPase activity evaluation and IC_{50} estimation

To quantify the inhibitory capacity of the lead compounds from Track 1, their effect on the microtubule-activated ATPase activity of the Eg5 motor domain was assessed. The reference inhibitor, K858, was used to validate the assay setup. As shown in Figure 4.23, K858 induced a clear dose-dependent reduction in Eg5 ATPase activity,

yielding an IC_{50} value of 3.78 µM. This result is in good agreement with the literature value (1.3 µM), confirming the reliability and accuracy of our experimental setup.

Among the four tested candidates, ZINC243914928 was the most effective inhibitor. A clear dose-dependent reduction in phosphate release was observed, with a calculated IC_{50} value of 28.97 µM (Figure 4.24). This confirms that the compound is an active Eg5 inhibitor, with the strongest effect observed at higher concentrations (50 and 100 µM).

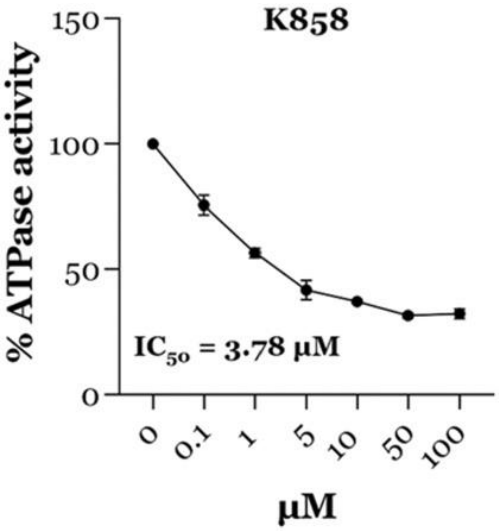

Figure 4.34: Dose-response curve for the inhibition of Eg5 ATPase activity by the reference compound K858.

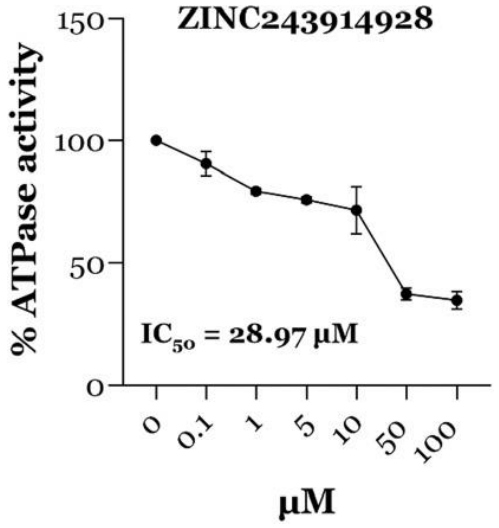

Figure 4.35: Dose-response curve for the inhibition of Eg5 ATPase activity by the lead candidate ZINC243914928.

The other three compounds, ZINC1857672049, ZINC299785885, and ZINC102735853, exhibited more variable and complex responses to Eg5 inhibition

(Figure 4.36). ZINC1857672049 displayed a dose-dependent effect from 0 to 5 µM; however, from 10 to 100 µM, a defined response was absent, with ATPase activity recorded at 74% at the highest concentration tested. ZINC299785885 slightly inhibited Eg5 at a very low concentration (0.1 µM, ~70% activity), after which the inhibitory capacity plateaued, reaching 80% activity at 100 µM concentration. Similarly, ZINC102735853 reduced Eg5 activity by approximately 40% at 1 µM, but showed limited further inhibition, with 77% ATPase activity remaining at the maximum concentration. These complex dose-response profiles may indicate several factors, including lower binding affinity, solubility issues at higher concentrations, or alternative binding kinetics not fully captured by this endpoint assay.

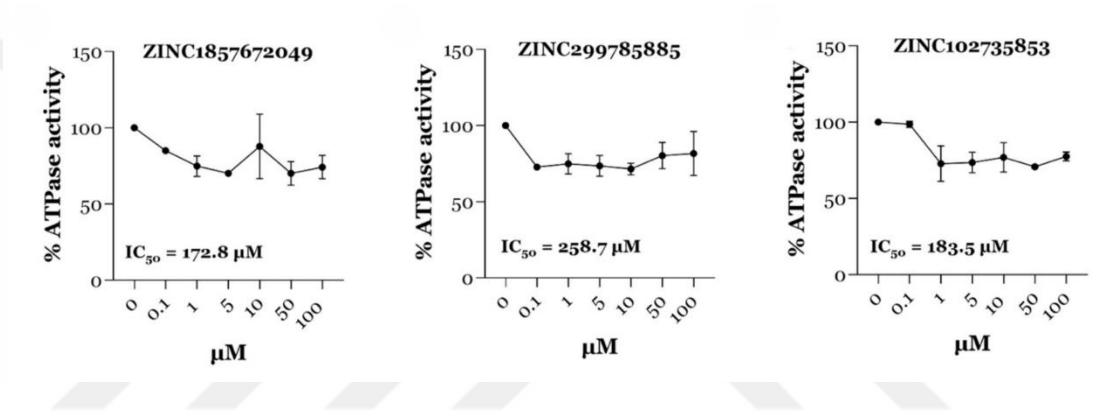

Figure 4.36: Dose-response curves for the inhibition of Eg5 ATPase activity by the candidates; ZINC1857672049, ZINC299785885, ZINC102735853.

4.2.2 Synthesis of computational and experimental findings

The in vitro results from Track 1 provide a powerful lens through which to interpret the entirety of the computational campaign. The experimental validation of four novel hits confirmed that the screening and filtering workflow was effective in identifying biologically active compounds. However, a deeper comparative analysis of the hits from both tracks reveals a clear and compelling hierarchy of potential, underscoring the strategic power of our dual-track approach. It is critical to note that the subsequent energetic comparison relies on the mean binding free energies (ΔGbind) calculated from all molecular dynamics trajectories, a metric far more robust than the initial static MM/GBSA rescoring of single binding poses.

Track 1 candidates, while successful, represent a proficient and solid incremental advancement. Their trajectory-averaged binding free energies (e.g., ΔGbind = -49.34 kcal/mol for ZINC243914928) were comparable to or slightly better than that of the

reference inhibitor K858 (-53.66 kcal/mol), and their *in vitro* activity was confirmed in the micromolar range. Their MD simulations demonstrated their competence, confirming stable binding within the L5 pocket, although often anchored by a few primary interactions. This track successfully fulfilled its goal of identifying novel active scaffolds through a validated, structure-guided approach and establishing a reliable baseline for comparing computational and experimental data.

In contrast, candidates from Track 2 represent a potential paradigm shift in potency and design. This superiority is not merely captured by the dramatically more favorable binding energies but is also evident in the quality and robustness of the binding modes observed during MD simulations.

Energetic Dominance: The top Track 2 candidates, with mean trajectory-averaged ΔGbind values ranging from -69 to -79 kcal/mol, occupy a completely different energetic landscape compared to the reference inhibitor. Although a direct correlation is not always linear, this substantial energetic gap of approximately 20-25 kcal/mol suggests the potential for a significant improvement in intrinsic binding affinity and potency. However, it is crucial to interpret these findings with scientific caution. These values represent computational predictions that provide a strong rationale for prioritizing these compounds; however, they must be empirically validated through experimental assays. Nevertheless, such a profound energetic advantage, coupled with the observed stability of their binding modes, strongly indicates that the Track 2 candidates are exceptionally promising for further development.

S**uperior Dynamic Stability and Interaction Networks:** Beyond the energy values, the MD simulations of Track 2 hits revealed sophisticated and durable binding mechanisms. Unlike the Track 1 hits, which were often anchored by a few primary contacts, candidates such as ZINC28804061 and ZINC000484655776 exhibited interaction redundancy, ” forming multiple persistent hydrogen bonds and extensive hydrophobic contacts simultaneously (e.g., with Glu116, Gly117, Arg221, and Leu214). This creates an exceptionally stable "clamp" on the L5 pocket, which is less likely to be disrupted by thermal fluctuations. Others, such as the rigid scaffold of ZINC000484655776, demonstrated the principle of "pre-organization," locking into the pocket with a minimal conformational penalty. Even flexible compounds, such as ZINC001159968549, were tethered by an extremely high-occupancy (98%) interaction, demonstrating a stable "locked" conformation. This represents a

significant improvement in the binding quality and predicted durability over the profiles of the Track 1 hits.

Greater Chemical Novelty: The AI-driven exploration of the billion-compound library in Track 2 was profoundly successful in its primary mission of discovering entirely new chemical matter. The identified scaffolds (indole-ureas, polycyclic amines, and novel biphenyl-azoles) are structurally distinct from the Track 1 hits and the well-trodden chemical classes of previously reported Eg5 inhibitors. This novelty is invaluable, as it offers the potential to circumvent class-specific liabilities, such as P-gp efflux or off-target effects, and may provide new avenues to overcome known resistance to mutations.

In summary, the *in vitro* validation of Track 1 hits serves as a crucial calibration point for our computational models. This proves that our methodology works. This, in turn, provides high confidence in the computational predictions for Track 2. The synthesis of all our findings leads to an inescapable conclusion: the candidates from Track 2 are not just marginally better; they are superior across every critical metric, namely, predicted binding affinity, dynamic stability, interaction complexity, and chemical novelty. They represent a true breakthrough in this research and hold the most realistic promise for developing a next-generation Eg5 inhibitor that could finally succeed where others have fallen short.

5. CONCLUSIONS AND RECOMMENDATIONS

This dissertation embarked on an ambitious dual-track computational quest to identify novel, potent, and drug-like allosteric inhibitors of human kinesin Eg5, a well-validated but clinically challenging anticancer target. This study was predicated on the hypothesis that the limitations of previous inhibitors stemmed not from the target itself but from a lack of chemical diversity in the compounds tested. By leveraging a synergistic combination of classical structure-based design and state-of-the-art AI-driven screening, this study successfully addressed this challenge and achieved its primary objectives.

Conservative Track 1 screening, which applied a hierarchical filtering cascade to the 11-million-compound ZINC Anodyne library, successfully identified four novel active chemotypes. The experimental validation of these hits, particularly ZINC243914928 (IC_{50} = 28.97 μM), provided crucial proof-of-concept for the entire computational workflow and established an essential internal benchmark for correlating computational energetics with real-world biological activity.

The more exploratory Track 2 campaign, which deployed a DeepDocking active learning strategy on the 1-billion-compound ZINC-ML library, represents the most significant contribution of this study. This approach demonstrated the remarkable capability of AI in navigating ultra-large chemical spaces, leading to the discovery of at least four new, highly diverse chemical classes of potential inhibitors. The computational profiles of these candidates, such as indole-urea ZINC580729109 and polycyclic amine ZINC410422515, are extraordinary. The calculated binding free energies (ΔGbind ≈ -70 to -80 kcal/mol) were vastly superior to those of both the reference inhibitor K858 and the hits from Track 1. Based on the correlation

established in Track 1, this strongly predicts that these novel scaffolds possess the potential for nanomolar or even picomolar potency, representing a substantial improvement over previously reported inhibitors.

In conclusion, this thesis makes the following key contributions.

Validation of a dual-track CADD workflow: This demonstrated that combining a conservative, validated pipeline with a high-risk, AI-driven exploration is a powerful and effective strategy for modern drug discovery.

Discovery of novel active scaffolds: Multiple chemical series with confirmed (Track 1) or exceptionally high predicted (Track 2) activity against Eg5 have been identified in this study. Crucially, these novel scaffolds demonstrate favorable drug-like properties and low similarity to known chemotypes, addressing the key challenges of resistance and patentability.

Establishment of a strong in silico-in vitro correlation: This provides a clear, data-driven rationale for prioritizing computationally superior Track 2 candidates for future development.

The journey from a computational hypothesis to a promising lead compound is long and challenging. However, the novel high-potential compounds identified in this study, which were discovered through the intelligent exploration of vast chemical possibilities, represent a promising new dawn in the quest for a clinically successful Eg5 inhibitor. They offer the potential to overcome the hurdles of potency and resistance that have hindered their predecessors and to finally translate the therapeutic promise of targeting this critical mitotic motor into a tangible benefit for cancer patients.

Based on the compelling findings of this dissertation, the following strategic steps are recommended to advance these novel candidates toward therapeutic development.

Synthesize and test top Track 2 candidate: The most critical next step is the synthesis and in vitro biological evaluation of the top six candidates identified from the DeepDocking track (including indole-urea ZINC580729109, polycyclic amine ZINC410422515, triphenylmethane analogs ZINC1189728758 and ZINC1189725451, pyrazole-indole ZINC28804061, and indole-piperidine ZINC000484655776). Performing Eg5 ATPase inhibition assays is essential to

confirm the high predicted potency and validate the success of the active learning campaign.

Evaluate cellular activity: For potent compounds identified in vitro, subsequent studies in relevant cancer cell lines (e.g., breast, colon, and pancreatic cancer lines with known high Eg5 expression) are crucial. These experiments should aim to:

- Confirm the on-target mechanism by observing the characteristic monopolar spindle phenotype using immunofluorescence microscopy.
- Determine cellular potency (e.g., GI_{50} values) in proliferation/viability assays.
- Assess the induction of apoptosis following mitotic arrest.

Lead optimization: For the most promising validated lead scaffold(s), a focused medicinal chemistry program should be initiated. The detailed interaction data from the MD simulations provide a clear roadmap for rational, structure-guided optimization to improve potency and selectivity while maintaining the favorable drug-like properties identified in the initial profiling.

Pharmacokinetic and resistance profiling:h Preliminary in silico and subsequent in vitro ADMET profiling (e.g., metabolic stability in microsomes and Caco-2 permeability) should be conducted for the lead series. Given their confirmed structural novelty, testing the most potent compounds against cell lines engineered to express known Eg5 resistance mutations (e.g., D130V) is a high priority, as it would provide early insights into their potential to overcome clinical resistance.

Exploration of advanced modalities: The novel high-affinity scaffolds discovered in this study could serve as ideal "warheads" for advanced therapeutic modalities. Their favorable physicochemical profiles and high predicted potency make them excellent candidates for incorporation into Antibody-Drug Conjugates (ADCs) or Proteolysis-Targeting Chimeras (PROTACs) to further enhance tumor selectivity and overcome systemic toxicity, directly addressing the key challenges that have historically plagued Eg5 inhibitor development.

REFERENCES

[1] **Sung H., Ferlay J., Siegel R. L., Laversanne M., Soerjomataram I., Jemal A., et al.** (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249

[2] **Jordan M. A. and Wilson L.** (2004). Microtubules as a target for anticancer drugs. *Nature Reviews Cancer*, 4(4), 253–265

[3] **Dumontet C. and Jordan M. A.** (2010). Microtubule-binding agents: a dynamic field of cancer therapeutics. *Nature Reviews Drug Discovery*, 9(10), 790–803

[4] **Mayer T. U., Kapoor T. M., Haggarty S. J., King R. W., Schreiber S. L. and Mitchison T. J.** (1999). Small Molecule Inhibitor of Mitotic Spindle Bipolarity Identified in a Phenotype-Based Screen. *Science*, 286(5441), 971–974

[5] **Kapoor T. M., Mayer T. U., Coughlin M. L. and Mitchison T. J.** (2000). Probing Spindle Assembly Mechanisms with Monastrol, a Small Molecule Inhibitor of the Mitotic Kinesin, Eg5. *The Journal of Cell Biology*, 150(5), 975–988

[6] **Zou L., Su L., Qi R., Bao F., Fang X., Wang L., et al.** (2017). Abnormal topological organization in white matter structural networks in survivors of acute lymphoblastic leukaemia with chemotherapy treatment. *Oncotarget*, 8(36), 60568–60575

[7] **Kaan H. Y. K., Ulaganathan V., Hackney D. D. and Kozielski F.** (2010). An allosteric transition trapped in an intermediate state of a new kinesin–inhibitor complex. *Biochemical Journal*, 425(1), 55–61

[8] **Purcell J. W., Davis J., Reddy M., Martin S., Samayoa K., Vo H., et al.** (2010). Activity of the Kinesin Spindle Protein Inhibitor Ispinesib (SB-715992) in Models of Breast Cancer. *Clinical Cancer Research*, 16(2), 566–576

[9] **Raza A. and Galili N.** (2012). The genetic basis of phenotypic heterogeneity in myelodysplastic syndromes. *Nature Reviews Cancer*, 12(12), 849–859

[10] **Gao C., Lowndes N. F. and Eriksson L. A.** (2017). Analysis of Biphenyl-Type Inhibitors Targeting the Eg5 α4/α6 Allosteric Pocket. *ACS omega*, 2(5), 1836–1849

[11] **Irwin J. J., Tang K. G., Young J., Dandarchuluun C., Wong B. R., Khurelbaatar M., et al.** (2020). ZINC20—A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *Journal of Chemical Information and Modeling*, 60(12), 6065–6073

[12] **Gentile F., Yaacoub J. C., Gleave J., Fernandez M., Ton A.-T., Ban F., et al.** (2022). Artificial intelligence–enabled virtual screening of ultra-large chemical libraries with deep docking. *Nature Protocols*, 17(3), 672–697

[13] **Irwin J. J.** *ZINC20-ML (Machine-Learning-ready subset, 2D/SMILES).*

[14] **Nirmaladevi R.** (2020). Epigenetic alterations in cancer. *Frontiers in Bioscience*, 25(6), 1058–1109

[15] **Bray F., Laversanne M., Sung H., Ferlay J., Siegel R. L., Soerjomataram I., et al.** (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 74(3), 229–263

[16] **Tran K. B., Lang J. J., Compton K., Xu R., Acheson A. R., Henrikson H. J., et al.** (2022). The global burden of cancer attributable to risk factors, 2010–19: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 400(10352), 563–591

[17] **Bizuayehu H. M., Ahmed K. Y., Kibret G. D., Dadi A. F., Belachew S. A., Bagade T., et al.** (2024). Global Disparities of Cancer and Its Projected Burden in 2050. *JAMA Network Open*, 7(11), e2443198

[18] **Sanchez-Vega F., Mina M., Armenia J., Chatila W. K., Luna A., La K. C., et al.** (2018). Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell*, 173(2), 321-337.e10

[19] **Bunning A. R. and Gupta Jr. M. L.** (2023). The importance of microtubule-dependent tension in accurate chromosome segregation. *Frontiers in Cell and Developmental Biology*, 11, 1096333

[20] **Matković J., Ghosh S., Ćosić M., Eibes S., Barišić M., Pavin N., et al.** (2022). Kinetochore- and chromosome-driven transition of microtubules into bundles promotes spindle assembly. *Nature Communications*, 13(1), 7307

[21] *Asexual Reproduction Fragmentation Diagram Illustration Stok Vektör (Telifsiz) 2143491021*.

[22] **Prosser S. L. and Pelletier L.** (2017). Mitotic spindle assembly in animal cells: a fine balancing act. *Nature Reviews Molecular Cell Biology*, 18(3), 187–201

[23] **Dominguez-Brauer C., Thu K. L., Mason J. M., Blaser H., Bray M. R. and Mak T. W.** (2015). Targeting Mitosis in Cancer: Emerging Strategies. *Molecular Cell*, 60(4), 524–536

[24] *Spindle apparatus Centriole Cell division Chromosome, spindle transparent background PNG clipart | HiClipart*.

[25] **Decelle J., Veronesi G., Gallet B., Stryhanyuk H., Benettoni P., Schmidt M., et al.** (2020). Subcellular Chemical Imaging: New Avenues in Cell Biology. *Trends in Cell Biology*, 30(3), 173–188

[26] **Weaver B. A.** (2014). How Taxol/paclitaxel kills cancer cells. *Molecular Biology of the Cell*, 25(18), 2677–2681

[27] **Kavallaris M.** (2010). Microtubules and resistance to tubulin-binding agents. *Nature Reviews Cancer*, 10(3), 194–204

[28] **Robey R. W., Pluchino K. M., Hall M. D., Fojo A. T., Bates S. E. and Gottesman M. M.** (2018). Revisiting the role of ABC transporters in multidrug-resistant cancer. *Nature Reviews Cancer*, 18(7), 452–464

[29] **Bar-Klein G., Klee R., Brandt C., Bankstahl M., Bascuñana P., Töllner K., et al.** (2016). Isoflurane prevents acquired epilepsy in rat models of temporal lobe epilepsy. *Annals of Neurology*, 80(6), 896–908

[30] **Jaffrès P.-A., Gajate C., Bouchet A. M., Couthon-Gourvès H., Chantôme A., Potier-Cartereau M., et al.** (2016). Alkyl ether lipids, ion channels and lipid raft reorganization in cancer therapy. *Pharmacology & Therapeutics*, 165, 114–131

[31] **Liu Y., Han J., Li Z., Jiang Z., Luo L., Zhang Y., et al.** (2022). Heterologous Expression of Jatropha curcas Fatty Acyl-ACP Thioesterase A (JcFATA) and B (JcFATB) Affects Fatty Acid Accumulation and Promotes Plant Growth and

Development in Arabidopsis. *International Journal of Molecular Sciences*, 23(8), 4209

[32] **Tyrtyshnaia A., Konovalova S., Bondar A., Ermolenko E., Sultanov R. and Manzhulo I.** (2021). Anti-Inflammatory Activity of N-Docosahexaenoylethanolamine and N-Eicosapentaenoylethanolamine in a Mouse Model of Lipopolysaccharide-Induced Neuroinflammation. *International Journal of Molecular Sciences*, 22(19), 10728

[33] **Shahin R. and Aljamal S.** (2022). Kinesin Spindle Protein Inhibitors in Cancer: From High Throughput Screening to Novel Therapeutic Strategies. *Future Science OA*, 8(3), FSO778

[34] **Kelly K. R., Nawrocki S. T., Espitia C. M., Zhang M., Yang J. J., Padmanabhan S., et al.** (2012). Targeting Aurora A kinase activity with the investigational agent alisertib increases the efficacy of cytarabine through a FOXO-dependent mechanism. *International Journal of Cancer*, 131(11), 2693–2703

[35] **Gjertsen B. T. and Schöffski P.** (2015). Discovery and development of the Polo-like kinase inhibitor volasertib in cancer therapy. *Leukemia*, 29(1), 11–19

[36] **Payton M., Belmontes B., Hanestad K., Moriguchi J., Chen K., McCarter J. D., et al.** (2024). Small-molecule inhibition of kinesin KIF18A reveals a mitotic vulnerability enriched in chromosomally unstable cancers. *Nature Cancer*, 5(1), 66–84

[37] **Klimaszewska-Wiśniewska A., Neska-Długosz I., Buchholz K., Durślewicz J., Grzanka D., Kasperska A., et al.** (2021). Prognostic Significance of KIF11 and KIF14 Expression in Pancreatic Adenocarcinoma. *Cancers*, 13(12), 3017

[38] **Wojcik E. J., Buckley R. S., Richard J., Liu L., Huckaba T. M. and Kim S.** (2013). Kinesin-5: Cross-bridging mechanism to targeted clinical therapy. *Gene*, 531(2), 133–149

[39] **Vale R. D. and Milligan R. A.** (2000). The Way Things Move: Looking Under the Hood of Molecular Motor Proteins. *Science*, 288(5463), 88–95

[40] **Miki H., Okada Y. and Hirokawa N.** (2005). Analysis of the kinesin superfamily: insights into structure and function. *Trends in Cell Biology*, 15(9), 467–476

[41] **Hirokawa N., Noda Y., Tanaka Y. and Niwa S.** (2009). Kinesin superfamily motor proteins and intracellular transport. *Nature Reviews Molecular Cell Biology*, 10(10), 682–696

[42] **Rice S., Lin A. W., Safer D., Hart C. L., Naber N., Carragher B. O., et al.** (1999). A structural change in the kinesin motor protein that drives motility. *Nature*, 402(6763), 778–784

[43] **Hirokawa N. and Noda Y.** (2008). Intracellular Transport and Kinesin Superfamily Proteins, KIFs: Structure, Function, and Dynamics. *Physiological Reviews*, 88(3), 1089–1118

[44] **Lawrence C. J., Dawe R. K., Christie K. R., Cleveland D. W., Dawson S. C., Endow S. A., et al.** (2004). A standardized kinesin nomenclature. *The Journal of Cell Biology*, 167(1), 19–22

[45] **Brenner D., Yilmaz R., Müller K., Grehl T., Petri S., Meyer T., et al.** (2018). Hot-spot KIF5A mutations cause familial ALS. *Brain*, 141(3), 688–697

[46] **Reid E., Kloos M., Ashley-Koch A., Hughes L., Bevan S., Svenson I. K., et al.** (2002). A Kinesin Heavy Chain (KIF5A) Mutation in Hereditary Spastic

Paraplegia (SPG10). *The American Journal of Human Genetics*, 71(5), 1189–1194

[47] Nair A., Greeny A., Rajendran R., Abdelgawad M. A., Ghoneim M. M., Raghavan R. P., et al. (2023). KIF1A-Associated Neurological Disorder: An Overview of a Rare Mutational Disease. *Pharmaceuticals*, 16(2), 147

[48] Willemsen M. H., Ba W., Wissink-Lindhout W. M., De Brouwer A. P. M., Haas S. A., Bienek M., et al. (2014). Involvement of the kinesin family members *KIF4A* and *KIF5C* in intellectual disability and synaptic function. *Journal of Medical Genetics*, 51(7), 487–494

[49] Ricci A., Carradori S., Cataldi A. and Zara S. (2024). Eg5 and Diseases: From the Well-Known Role in Cancer to the Less-Known Activity in Noncancerous Pathological Conditions. *Biochemistry Research International*, 2024, 3649912

[50] McNeely K. C., Cupp T. D., Little J. N., Janisch K. M., Shrestha A. and Dwyer N. D. (2017). Mutation of Kinesin-6 Kif20b causes defects in cortical neuron polarization and morphogenesis. *Neural Development*, 12(1), 5

[51] Mirzaa G. M., Vitre B., Carpenter G., Abramowicz I., Gleeson J. G., Paciorkowski A. R., et al. (2014). Mutations in CENPE define a novel kinetochore-centromeric mechanism for microcephalic primordial dwarfism. *Human Genetics*, 133(8), 1023–1039

[52] Payton M., Belmontes B., Hanestad K., Moriguchi J., Chen K., McCarter J. D., et al. (2023). Small-molecule inhibition of kinesin KIF18A reveals a mitotic vulnerability enriched in chromosomally unstable cancers. *Nature Cancer*, 5(1), 66–84

[53] Meng Z., Meng Q., Gao T., Zhou H., Xue J., Li H., et al. (2023). Identification of bi-allelic KIF9 loss-of-function variants contributing to asthenospermia and male infertility in two Chinese families. *Frontiers in Endocrinology*, 13, 1091107

[54] Min B.-J., Kim N., Chung T., Kim O.-H., Nishimura G., Chung C. Y., et al. (2011). Whole-Exome Sequencing Identifies Mutations of KIF22 in Spondyloepimetaphyseal Dysplasia with Joint Laxity, Leptodactylic Type. *The American Journal of Human Genetics*, 89(6), 760–766

[55] Kalantari S. and Filges I. (2020). 'Kinesinopathies': emerging role of the kinesin family member genes in birth defects. *Journal of Medical Genetics*, 57(12), 797–807

[56] Mi J., Ma S., Chen W., Kang M., Xu M., Liu C., et al. (2022). Integrative Pan-Cancer Analysis of KIF15 Reveals Its Diagnosis and Prognosis Value in Nasopharyngeal Carcinoma. *Frontiers in Oncology*, 12, 772816

[57] Poirier K., Lebrun N., Broix L., Tian G., Saillour Y., Boscheron C., et al. (2013). Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly. *Nature Genetics*, 45(6), 639–647

[58] Zhi E., Li P., Chen H., Xu P., Zhu X., Zhu Z., et al. (2016). Decreased Expression of KIFC1 in Human Testes with Globozoospermic Defects. *Genes*, 7(10), 75

[59] Friel C. T. and Welburn J. P. (2018). Parts list for a microtubule depolymerising kinesin. *Biochemical Society Transactions*, 46(6), 1665–1672

[60] Verhey K. J., Dishinger J. and Kee H. L. (2011). Kinesin motors and primary cilia. *Biochemical Society Transactions*, 39(5), 1120–1125

[61] Wandke C., Barisic M., Sigl R., Rauch V., Wolf F., Amaro A. C., et al. (2012). Human chromokinesins promote chromosome congression and spindle microtubule dynamics during mitosis. *Journal of Cell Biology*, 198(5), 847–863

[62] Nislow C., Lombillo V. A., Kuriyama R. and Mclntosh J. R. (1992). A plus-end-directed motor enzyme that moves antiparallel microtubules in vitro localizes to the interzone of mitotic spindles. *Nature*, 359(6395), 543–547

[63] Weaver B. A. A., Bonday Z. Q., Putkey F. R., Kops G. J. P. L., Silk A. D. and Cleveland D. W. (2003). Centromere-associated protein-E is essential for the mammalian mitotic checkpoint to prevent aneuploidy due to single chromosome loss. *The Journal of Cell Biology*, 162(4), 551–563

[64] Brown J. M., Green J., Das Neves R. P., Wallace H. A. C., Smith A. J. H., Hughes J., et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *The Journal of Cell Biology*, 182(6), 1083–1097

[65] Mountain V., Simerly C., Howard L., Ando A., Schatten G. and Compton D. A. (1999). The Kinesin-Related Protein, Hset, Opposes the Activity of Eg5 and Cross-Links Microtubules in the Mammalian Mitotic Spindle. *The Journal of Cell Biology*, 147(2), 351–366

[66] Myers S. M. and Collins I. (2016). Recent Findings and Future Directions for Interpolar Mitotic Kinesin Inhibitors in Cancer Therapy. *Future Medicinal Chemistry*, 8(4), 463–489

[67] Kapitein L. C., Peterman E. J. G., Kwok B. H., Kim J. H., Kapoor T. M. and Schmidt C. F. (2005). The bipolar mitotic kinesin Eg5 moves on both microtubules that it crosslinks. *Nature*, 435(7038), 114–118

[68] Raaijmakers J. A., Van Heesbeen R. G. H. P., Meaders J. L., Geers E. F., Fernandez-Garcia B., Medema R. H., et al. (2012). Nuclear envelope-associated dynein drives prophase centrosome separation and enables Eg5-independent bipolar spindle formation: NE-dynein promotes prophase centrosome separation. *The EMBO Journal*, 31(21), 4179–4190

[69] Sawin K. E. and Mitchison T. J. (1995). Mutations in the kinesin-like protein Eg5 disrupting localization to the mitotic spindle. *Proceedings of the National Academy of Sciences*, 92(10), 4289–4293

[70] Krzysiak T. C., Wendt T., Sproul L. R., Tittmann P., Gross H., Gilbert S. P., et al. (2006). A structural model for monastrol inhibition of dimeric kinesin Eg5. *The EMBO Journal*, 25(10), 2263–2273

[71] Liu M., Ran J. and Zhou J. (2018). Non-canonical functions of the mitotic kinesin Eg5. *Thoracic Cancer*, 9(8), 904–910

[72] Miyamoto D. T., Perlman Z. E., Burbank K. S., Groen A. C. and Mitchison T. J. (2004). The kinesin Eg5 drives poleward microtubule flux in *Xenopus laevis* egg extract spindles. *The Journal of Cell Biology*, 167(5), 813–818

[73] Cahu J., Olichon A., Hentrich C., Schek H., Drinjakovic J., Zhang C., et al. (2008). Phosphorylation by Cdk1 Increases the Binding of Eg5 to Microtubules In Vitro and in Xenopus Egg Extract Spindles. *PLoS ONE*, 3(12), e3936

[74] Enos A. P. and Morris N. R. (1990). Mutation of a gene that encodes a kinesin-like protein blocks nuclear division in A. nidulans. *Cell*, 60(6), 1019–1027

[75] **Heck M. M., Pereira A., Pesavento P., Yannoni Y., Spradling A. C. and Goldstein L. S.** (1993). The kinesin-like protein KLP61F is essential for mitosis in Drosophila. *The Journal of cell biology*, 123(3), 665–679

[76] **Tsai T. Y.-C., Theriot J. A. and Ferrell J. E.** (2014). Changes in Oscillatory Dynamics in the Cell Cycle of Early Xenopus laevis Embryos. *PLoS Biology*, 12(2), e1001788

[77] **Gergely Z. R., Jones M. H., Zhou B., Cash C., McIntosh J. R. and Betterton M. D.** (2023). Distinct regions of the kinesin-5 C-terminal tail are essential for mitotic spindle midzone localization and sliding force. *Proceedings of the National Academy of Sciences*, 120(39), e2306480120

[78] **Yukawa M., Yamada Y. and Toda T.** (2019). Suppressor Analysis Uncovers That MAPs and Microtubule Dynamics Balance with the Cut7/Kinesin-5 Motor for Mitotic Spindle Assembly in *Schizosaccharomyces pombe*. *G3 Genes|Genomes|Genetics*, 9(1), 269–280

[79] **Paillard L.** (1998). EDEN and EDEN-BP, a cis element and an associated factor that mediate sequence-specific mRNA deadenylation in Xenopus embryos. *The EMBO Journal*, 17(1), 278–287

[80] **Blangy A., Lane H. A., d'Hérin P., Harper M., Kress M. and Niggt E. A.** (1995). Phosphorylation by p34cdc2 regulates spindle association of human Eg5, a kinesin-related motor essential for bipolar spindle formation in vivo. *Cell*, 83(7), 1159–1169

[81] **Hagan I. and Yanagida M.** (1992). Kinesin-related cut 7 protein associates with mitotic and meiotic spindles in fission yeast. *Nature*, 356(6364), 74–76

[82] **Yu W.-X., Li Y.-K., Xu M.-F., Xu C.-J., Chen J., Wei Y.-L., et al.** (2022). Kinesin-5 Eg5 is essential for spindle assembly, chromosome stability and organogenesis in development. *Cell Death Discovery*, 8(1), 490

[83] **Kashlna A. S., Baskin R. J., Cole D. G., Wedaman K. P., Saxton W. M. and Scholey J. M.** (1996). A bipolar kinesin. *Nature*, 379(6562), 270–272

[84] **Jin Q., Huang F., Wang X., Zhu H., Xian Y., Li J., et al.** (2017). High Eg5 expression predicts poor prognosis in breast cancer. *Oncotarget*, 8(37), 62208–62216

[85] **UniProt Consortium.** *Kinesin-like protein KIF11 – Homo sapiens (Human)*.

[86] **Waitzman J. S. and Rice S. E.** (2014). Mechanism and regulation of kinesin-5, an essential motor for the mitotic spindle. *Biology of the Cell*, 106(1), 1–12

[87] **Scholey J. E., Nithianantham S., Scholey J. M. and Al-Bassam J.** (2014). Structural basis for the assembly of the mitotic motor Kinesin-5 into bipolar tetramers. *eLife*, 3, e02217

[88] **Turner J., Anderson R., Guo J., Beraud C., Fletterick R. and Sakowicz R.** (2001). Crystal Structure of the Mitotic Spindle Kinesin Eg5 Reveals a Novel Conformation of the Neck-linker. *Journal of Biological Chemistry*, 276(27), 25496–25502

[89] **Jon Kull F., Sablin E. P., Lau R., Fletterick R. J. and Vale R. D.** (1996). Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature*, 380(6574), 550–555

[90] **Waitzman J. S. and Rice S. E.** (2014). Mechanism and regulation of kinesin-5, an essential motor for the mitotic spindle. *Biology of the Cell*, 106(1), 1–12

[91] **Sindelar C. V. and Downing K. H.** (2007). The beginning of kinesin's force-generating cycle visualized at 9-Å resolution. *The Journal of Cell Biology*, 177(3), 377–385

[92] **Gilbert S. P. and Sardar H. S.** (2012). *4.17 Kinesin Structure and Biochemistry*. Elsevier

[93] **Parke C. L., Wojcik E. J., Kim S. and Worthylake D. K.** (2010). ATP Hydrolysis in Eg5 Kinesin Involves a Catalytic Two-water Mechanism. *Journal of Biological Chemistry*, 285(8), 5859–5867

[94] **Uchimura S., Oguchi Y., Hachikubo Y., Ishiwata S. and Muto E.** (2010). Key residues on microtubule responsible for activation of kinesin ATPase. *The EMBO Journal*, 29(7), 1167–1175

[95] **Sablin E.** (2001). Nucleotide switches in molecular motors: structural analysis of kinesins and myosins. *Current Opinion in Structural Biology*, 11(6), 716–724

[96] **Parke C. L., Wojcik E. J., Kim S. and Worthylake D. K.** (2010). ATP Hydrolysis in Eg5 Kinesin Involves a Catalytic Two-water Mechanism. *Journal of Biological Chemistry*, 285(8), 5859–5867

[97] **Behnke-Parks W. M., Vendome J., Honig B., Maliga Z., Moores C. and Rosenfeld S. S.** (2011). Loop L5 Acts as a Conformational Latch in the Mitotic Kinesin Eg5. *Journal of Biological Chemistry*, 286(7), 5242–5253

[98] **Peña A., Sweeney A., Cook A. D., Locke J., Topf M. and Moores C. A.** (2020). Structure of Microtubule-Trapped Human Kinesin-5 and Its Mechanism of Inhibition Revealed Using Cryoelectron Microscopy. *Structure (London, England: 1993)*, 28(4), 450-457.e5

[99] **Jana S. C., Marteil G. and Bettencourt-Dias M.** (2014). Mapping molecules to structure: unveiling secrets of centriole and cilia assembly with near-atomic resolution. *Current Opinion in Cell Biology*, 26, 96–106

[100] **Krzysiak T. C., Grabe M. and Gilbert S. P.** (2008). Getting in Sync with Dimeric Eg5. *Journal of Biological Chemistry*, 283(4), 2078–2087

[101] **Shang Z., Zhou K., Xu C., Csencsits R., Cochran J. C. and Sindelar C. V.** (2014). High-resolution structures of kinesin on microtubules provide a basis for nucleotide-gated force-generation. *eLife*, 3, e04686

[102] **Cao L., Cantos-Fernandes S. and Gigant B.** (2017). The structural switch of nucleotide-free kinesin. *Scientific Reports*, 7(1), 42558

[103] **Chen G.-Y., Kang Y. J., Gayek A. S., Youyen W., Tüzel E., Ohi R., et al.** (2017). Eg5 Inhibitors Have Contrasting Effects on Microtubule Stability and Metaphase Spindle Integrity. *ACS Chemical Biology*, 12(4), 1038–1046

[104] **Olmsted Z. T., Colliver A. G., Riehlman T. D. and Paluh J. L.** (2014). Kinesin-14 and kinesin-5 antagonistically regulate microtubule nucleation by γ-TuRC in yeast and human cells. *Nature Communications*, 5(1), 5339

[105] **Jana B. and Onuchic J. N.** (2016). Strain Mediated Adaptation Is Key for Myosin Mechanochemistry: Discovering General Rules for Motor Activity. *PLOS Computational Biology*, 12(8), e1005035

[106] **Larson A. G., Naber N., Cooke R., Pate E. and Rice S. E.** (2010). The Conserved L5 Loop Establishes the Pre-Powerstroke Conformation of the Kinesin-5 Motor, Eg5. *Biophysical Journal*, 98(11), 2619–2627

[107] **Muretta J. M., Jun Y., Gross S. P., Major J., Thomas D. D. and Rosenfeld S. S.** (2015). The structural kinetics of switch-1 and the neck linker explain the functions of kinesin-1 and Eg5. *Proceedings of the National Academy of Sciences*, 112(48)

[108] **Zheng W.** (2012). Coarse-grained modeling of the structural states and transition underlying the powerstroke of dynein motor domain. *The Journal of Chemical Physics*, 136(15), 155103

[109] **Valentine M. T. and Block S. M.** (2009). Force and Premature Binding of ADP Can Regulate the Processivity of Individual Eg5 Dimers. *Biophysical Journal*, 97(6), 1671–1677

[110] **Llinas P., Isabet T., Song L., Ropars V., Zong B., Benisty H., et al.** (2015). How Actin Initiates the Motor Activity of Myosin. *Developmental Cell*, 33(4), 401–412

[111] **Ulaganathan V., Talapatra S. K., Rath O., Pannifer A., Hackney D. D. and Kozielski F.** (2013). Structural Insights into a Unique Inhibitor Binding Pocket in Kinesin Spindle Protein. *Journal of the American Chemical Society*, 135(6), 2263–2272

[112] **Ogunwa T. H., Laudadio E., Galeazzi R. and Miyanishi T.** (2019). Insights into the Molecular Mechanisms of Eg5 Inhibition by (+)-Morelloflavone. *Pharmaceuticals*, 12(2), 58

[113] **Belsham H. R. and Friel C. T.** (2019). Identification of key residues that regulate the interaction of kinesins with microtubule ends. *Cytoskeleton*, 76(7–8), 440–446

[114] **Blasius T. L., Yue Y. and Verhey K.** (2019). *Microtubule binding of the kinesin-4 KIF7 and its regulation by autoinhibition*. Cell Biology

[115] **Barlan K., Lu W. and Gelfand V. I.** (2013). The Microtubule-Binding Protein Ensconsin Is an Essential Cofactor of Kinesin-1. *Current Biology*, 23(4), 317–322

[116] **Indorato R.-L., DeBonis S., Garcia-Saez I. and Skoufias D. A.** (2022). Drug resistance dependent on allostery: A P-loop rigor Eg5 mutant exhibits resistance to allosteric inhibition by STLC. *Frontiers in Oncology*, 12, 965455

[117] **Maliga Z. and Mitchison T. J.** (2006). Small-molecule and mutational analysis of allosteric Eg5 inhibition by monastrol. *BMC Chemical Biology*, 6(1), 2

[118] **Garcia-Saez I., DeBonis S., Lopez R., Trucco F., Rousseau B., Thuéry P., et al.** (2007). Structure of Human Eg5 in Complex with a New Monastrol-based Inhibitor Bound in the R Configuration. *Journal of Biological Chemistry*, 282(13), 9740–9747

[119] **Talapatra S. K., Schüttelkopf A. W. and Kozielski F.** (2012). The structure of the ternary Eg5–ADP–ispinesib complex. *Acta Crystallographica Section D Biological Crystallography*, 68(10), 1311–1319

[120] **Yokoyama H., Sawada J., Sato K., Ogo N., Kamei N., Ishikawa Y., et al.** (2018). Structural and Thermodynamic Basis of the Enhanced Interaction between Kinesin Spindle Protein Eg5 and STLC-type Inhibitors. *ACS Omega*, 3(9), 12284–12294

[121] **Indorato R.-L., Talapatra S. K., Lin F., Haider S., Mackay S. P., Kozielski F., et al.** (2019). Is the Fate of Clinical Candidate Arry-520 Already Sealed? Predicting Resistance in Eg5–Inhibitor Complexes. *Molecular Cancer Therapeutics*, 18(12), 2394–2406

[122] **Yokoyama H., Sawada J., Katoh S., Matsuno K., Ogo N., Ishikawa Y., et al.** (2015). Structural Basis of New Allosteric Inhibition in Kinesin Spindle Protein Eg5. *ACS Chemical Biology*, 10(4), 1128–1136

[123] **Talapatra S. K., Tham C. L., Guglielmi P., Cirilli R., Chandrasekaran B., Karpoormath R., et al.** (2018). Crystal structure of the Eg5 - K858 complex and implications for structure-based design of thiadiazole-containing inhibitors. *European Journal of Medicinal Chemistry*, 156, 641–651

[124] **Yan Y., Sardana V., Xu B., Homnick C., Halczenko W., Buser C. A., et al.** (2004). Inhibition of a Mitotic Motor Protein: Where, How, and Conformational Consequences. *Journal of Molecular Biology*, 335(2), 547–554

[125] **Cox C. D., Torrent M., Breslin M. J., Mariano B. J., Whitman D. B., Coleman P. J., et al.** (2006). Kinesin spindle protein (KSP) inhibitors. Part 4: Structure-based design of 5-alkylamino-3,5-diaryl-4,5-dihydropyrazoles as potent, water-soluble inhibitors of the mitotic kinesin KSP. *Bioorganic & Medicinal Chemistry Letters*, 16(12), 3175–3179

[126] **Fraley M. E., Garbaccio R. M., Arrington K. L., Hoffman W. F., Tasber E. S., Coleman P. J., et al.** (2006). Kinesin spindle protein (KSP) inhibitors. Part 2: The design, synthesis, and characterization of 2,4-diaryl-2,5-dihydropyrrole inhibitors of the mitotic kinesin KSP. *Bioorganic & Medicinal Chemistry Letters*, 16(7), 1775–1779

[127] **Kim K. S., Lu S., Cornelius L. A., Lombardo L. J., Borzilleri R. M., Schroeder G. M., et al.** (2006). Synthesis and SAR of pyrrolotriazine-4-one based Eg5 inhibitors. *Bioorganic & Medicinal Chemistry Letters*, 16(15), 3937–3942

[128] **Atherton J., Farabella I., Yu I.-M., Rosenfeld S. S., Houdusse A., Topf M., et al.** (2014). Conserved mechanisms of microtubule-stimulated ADP release, ATP binding, and force generation in transport kinesins. *eLife*, 3, e03680

[129] **Liu M., Li D., Sun L., Chen J., Sun X., Zhang L., et al.** (2014). Modulation of Eg5 activity contributes to mitotic spindle checkpoint activation and Tat-mediated apoptosis in CD4 -positive T-lymphocytes. *The Journal of Pathology*, 233(2), 138–147

[130] **Hackney D. D. and McGoff M. S.** (2016). Nucleotide-free kinesin motor domains reversibly convert to an inactive conformation with characteristics of a molten globule. *Archives of Biochemistry and Biophysics*, 608, 42–51

[131] **Zhang W.** (2011). Exploring the Intermediate States of ADP−ATP Exchange: A Simulation Study on Eg5. *The Journal of Physical Chemistry B*, 115(5), 784–795

[132] **Garcia-Saez I. and Skoufias D. A.** (2021). Eg5 targeting agents: From new anti-mitotic based inhibitor discovery to cancer therapy and resistance. *Biochemical Pharmacology*, 184, 114364

[133] **Cross R. A.** (2015). *Mechanisms of Action of Eg5 Inhibitors*. Springer Netherlands, Dordrecht

[134] **Ogunwa T. H., Taii K., Sadakane K., Kawata Y., Maruta S. and Miyanishi T.** (2019). Morelloflavone as a novel inhibitor of mitotic kinesin Eg5. *The Journal of Biochemistry*, 166(2), 129–137

[135] **Jia N., Zhang B., Huo Z., Qin J., Ji Q. and Geng Y.** (2024). Binding patterns of inhibitors to different pockets of kinesin Eg5. *Archives of Biochemistry and Biophysics*, 756, 109998

[136] **Tcherniuk S., Van Lis R., Kozielski F. and Skoufias D. A.** (2010). Mutations in the human kinesin Eg5 that confer resistance to monastrol and S-trityl-l-cysteine in tumor derived cell lines. *Biochemical Pharmacology*, 79(6), 864–872

[137] **Luo L., Parrish C. A., Nevins N., McNulty D. E., Chaudhari A. M., Carson J. D., et al.** (2007). ATP-competitive inhibitors of the mitotic kinesin KSP that function via an allosteric mechanism. *Nature Chemical Biology*, 3(11), 722–726

[138] Wacker S. A., Houghtaling B. R., Elemento O. and Kapoor T. M. (2012). Using transcriptome sequencing to identify mechanisms of drug action and resistance. *Nature Chemical Biology*, 8(3), 235–237

[139] Sturgill E. G., Norris S. R., Guo Y. and Ohi R. (2016). Kinesin-5 inhibitor resistance is driven by kinesin-12. *Journal of Cell Biology*, 213(2), 213–227

[140] Tarui Y., Chinen T., Nagumo Y., Motoyama T., Hayashi T., Hirota H., et al. (2014). Terpendole E and its Derivative Inhibit STLC- and GSK-1-Resistant Eg5. *ChemBioChem*, 15(7), 934–938

[141] Zhang W., Zhai L., Lu W., Boohaker R. J., Padmalayam I. and Li Y. (2016). Discovery of Novel Allosteric Eg5 Inhibitors Through Structure-Based Virtual Screening. *Chemical Biology & Drug Design*, 88(2), 178–187

[142] Fakhri N. and Schmidt C. F. (2014). A surprising twist. *eLife*, 3, e02715

[143] Hildebrandt E. R., Gheber L., Kingsbury T. and Hoyt M. A. (2006). Homotetrameric Form of Cin8p, a Saccharomyces cerevisiae Kinesin-5 Motor, Is Essential for Its in Vivo Function. *Journal of Biological Chemistry*, 281(36), 26004–26013

[144] Pandey H., Popov M., Goldstein-Levitin A. and Gheber L. (2021). Mechanisms by Which Kinesin-5 Motors Perform Their Multiple Intracellular Functions. *International Journal of Molecular Sciences*, 22(12), 6420

[145] Bodrug T., Wilson-Kubalek E. M., Nithianantham S., Thompson A. F., Alfieri A., Gaska I., et al. (2020). The kinesin-5 tail domain directly modulates the mechanochemical cycle of the motor domain for anti-parallel microtubule sliding. *eLife*, 9, e51131

[146] Castillo A. and Justice M. J. (2007). The Kinesin Related Motor Protein, Eg5, is Essential for Maintenance of Pre-implantation Embryogenesis. *Biochemical and biophysical research communications*, 357(3), 694–699

[147] Valentine M. T. and Gilbert S. P. (2007). To step or not to step? How biochemistry and mechanics influence processivity in Kinesin and Eg5. *Current Opinion in Cell Biology*, 19(1), 75–81

[148] Chen G.-Y., Mickolajczyk K. J. and Hancock W. O. (2016). The Kinesin-5 Chemomechanical Cycle Is Dominated by a Two-heads-bound State♦. *The Journal of Biological Chemistry*, 291(39), 20283–20294

[149] Goulet A., Behnke-Parks W. M., Sindelar C. V., Major J., Rosenfeld S. S. and Moores C. A. (2012). The Structural Basis of Force Generation by the Mitotic Motor Kinesin-5. *The Journal of Biological Chemistry*, 287(53), 44654–44666

[150] Lanzetta P. A., Alvarez L. J., Reinach P. S. and Candia O. A. (1979). An improved assay for nanomole amounts of inorganic phosphate. *Analytical Biochemistry*, 100(1), 95–97

[151] Gicking A. M., Qiu W. and Hancock W. O. (2018). Mitotic kinesins in action: diffusive searching, directional switching, and ensemble coordination. *Molecular Biology of the Cell*, 29(10), 1153–1156

[152] Gennerich A. and Vale R. D. (2009). Walking the walk: how kinesin and dynein coordinate their steps. *Current Opinion in Cell Biology*, 21(1), 59–67

[153] Cochran J. C. and Gilbert S. P. (2005). ATPase Mechanism of Eg5 in the Absence of Microtubules: Insight into Microtubule Activation and Allosteric Inhibition by Monastrol. *Biochemistry*, 44(50), 16633–16648

[154] Harrington T., Naber N., Larson A., Cooke R., Rice S. and Pate E. (2011). Analysis of the interaction of the Eg5 Loop5 with the nucleotide site. *Journal of theoretical biology*, 289, 107–15

[155] **Hariharan V. and Hancock W. O.** (2009). Insights into the Mechanical Properties of the Kinesin Neck Linker Domain from Sequence Analysis and Molecular Dynamics Simulations. *Cellular and Molecular Bioengineering*, 2(2), 177–189

[156] **Kaan H. Y. K., Major J., Tkocz K., Kozielski F. and Rosenfeld S. S.** (2013). "Snapshots" of Ispinesib-induced Conformational Changes in the Mitotic Kinesin Eg5. *The Journal of Biological Chemistry*, 288(25), 18588–18598

[157] **Tomishige M.** (2008). Activation of mitotic kinesin by microtubule bundling. *The Journal of Cell Biology*, 182(3), 417–419

[158] **Valentine M. T., Fordyce P. M. and Block S. M.** (2006). Eg5 steps it up! *Cell Division*, 1, 31

[159] **Valentine M. T., Fordyce P. M., Krzysiak T. C., Gilbert S. P. and Block S. M.** (2006). Individual dimers of the mitotic kinesin motor Eg5 step processively and support substantial loads in vitro. *Nature Cell Biology*, 8(5), 470–476

[160] **Shimamoto Y., Forth S. and Kapoor T. M.** (2015). Measuring Pushing and Braking Forces Generated by Ensembles of Kinesin-5 Crosslinking Two Microtubules. *Developmental Cell*, 34(6), 669–681

[161] **Rapley J., Nicolàs M., Groen A., Regué L., Bertran M. T., Caelles C., et al.** (2008). The NIMA-family kinase Nek6 phosphorylates the kinesin Eg5 at a novel site necessary for mitotic spindle formation. *Journal of Cell Science*, 121(23), 3912–3921

[162] **Muretta J. M., Reddy B. J. N., Scarabelli G., Thompson A. F., Jariwala S., Major J., et al.** (2018). A posttranslational modification of the mitotic kinesin Eg5 that enhances its mechanochemical coupling and alters its mitotic function. *Proceedings of the National Academy of Sciences*, 115(8), E1779–E1788

[163] **Smith E., Hégarat N., Vesely C., Roseboom I., Larch C., Streicher H., et al.** (2011). Differential control of Eg5-dependent centrosome separation by Plk1 and Cdk1: Plk1- and Cdk1-triggered centrosome separation. *The EMBO Journal*, 30(11), 2233–2245

[164] **Dumontet C. and Sikic B. I.** (1999). Mechanisms of Action of and Resistance to Antitubulin Agents: Microtubule Dynamics, Drug Transport, and Cell Death. *Journal of Clinical Oncology*, 17(3), 1061–1061

[165] **Mukhtar E., Adhami V. M. and Mukhtar H.** (2014). Targeting Microtubules by Natural Agents for Cancer Therapy. *Molecular Cancer Therapeutics*, 13(2), 275–284

[166] **Jackson J. R., Patrick D. R., Dar M. M. and Huang P. S.** (2007). Targeted anti-mitotic therapies: can we improve on tubulin agents? *Nature Reviews Cancer*, 7(2), 107–117

[167] **Rath O. and Kozielski F.** (2012). Kinesins and cancer. *Nature Reviews Cancer*, 12(8), 527–539

[168] **Maliga Z., Kapoor T. M. and Mitchison T. J.** (2002). Evidence that Monastrol Is an Allosteric Inhibitor of the Mitotic Kinesin Eg5. *Chemistry & Biology*, 9(9), 989–996

[169] **DeBonis S., Skoufias D. A., Lebeau L., Lopez R., Robin G., Margolis R. L., et al.** (2004). *In vitro* screening for inhibitors of the human mitotic kinesin Eg5 with antimitotic and antitumor activities. *Molecular Cancer Therapeutics*, 3(9), 1079–1090

[170] Skoufias D. A., DeBonis S., Saoudi Y., Lebeau L., Crevel I., Cross R., et al. (2006). S-Trityl-L-cysteine Is a Reversible, Tight Binding Inhibitor of the Human Kinesin Eg5 That Specifically Blocks Mitotic Progression. *Journal of Biological Chemistry*, 281(26), 17559–17569

[171] Burris H. A., Jones S. F., Williams D. D., Kathman S. J., Hodge J. P., Pandite L., et al. (2011). A phase I study of ispinesib, a kinesin spindle protein inhibitor, administered weekly for three consecutive weeks of a 28-day cycle in patients with solid tumors. *Investigational New Drugs*, 29(3), 467–472

[172] Shah J. J., Kaufman J. L., Zonder J. A., Cohen A. D., Bensinger W. I., Hilder B. W., et al. (2017). A Phase 1 and 2 study of Filanesib alone and in combination with low-dose dexamethasone in relapsed/refractory multiple myeloma. *Cancer*, 123(23), 4617–4630

[173] Blagden S. P., Molife L. R., Seebaran A., Payne M., Reid A. H. M., Protheroe A. S., et al. (2008). A phase I trial of ispinesib, a kinesin spindle protein inhibitor, with docetaxel in patients with advanced solid tumours. *British Journal of Cancer*, 98(5), 894–899

[174] Parrish C. A., Adams N. D., Auger K. R., Burgess J. L., Carson J. D., Chaudhari A. M., et al. (2007). Novel ATP-Competitive Kinesin Spindle Protein Inhibitors. *Journal of Medicinal Chemistry*, 50(20), 4939–4952

[175] Huszar D., Theoclitou M.-E., Skolnik J. and Herbst R. (2009). Kinesin motor proteins as targets for cancer therapy. *Cancer and Metastasis Reviews*, 28(1–2), 197–208

[176] Cochran J. C., Gatial J. E., Kapoor T. M. and Gilbert S. P. (2005). Monastrol inhibition of the mitotic kinesin Eg5. *The Journal of Biological Chemistry*, 280(13), 12658–12667

[177] Nikam D. and Jain A. (2023). Advances in the discovery of DHPMs as Eg5 inhibitors for the management of breast cancer and glioblastoma: A review. *Results in Chemistry*, 5, 100718

[178] Paull K. D., Lin C. M., Malspeis L. and Hamel E. (1992). Identification of Novel Antimitotic Agents Acting at the Tubulin Level by Computer-assisted Evaluation of Differential Cytotoxicity Data. *Cancer Research*, 52(14), 3892–3900

[179] Brier S., Lemaire D., DeBonis S., Forest E. and Kozielski F. (2004). Identification of the Protein Binding Region of *S* -Trityl- L -cysteine, a New Potent Inhibitor of the Mitotic Kinesin Eg5. *Biochemistry*, 43(41), 13072–13082

[180] El-Nassan H. B. (2013). Advances in the discovery of kinesin spindle protein (Eg5) inhibitors as antitumor agents. *European Journal of Medicinal Chemistry*, 62, 614–631

[181] Fukai R., Ogo N., Ichida T., Yamane M., Sawada J., Miyoshi N., et al. (2021). Design, synthesis, and evaluation of a novel prodrug, a S-trityl--cysteine derivative targeting kinesin spindle protein. *European Journal of Medicinal Chemistry*, 215, 113288

[182] Nakai R., Iida S., Takahashi T., Tsujita T., Okamoto S., Takada C., et al. (2009). K858, a Novel Inhibitor of Mitotic Kinesin Eg5 and Antitumor Agent, Induces Cell Death in Cancer Cells. *Cancer Research*, 69(9), 3901–3909

[183] Taglieri L., Rubinacci G., Giuffrida A., Carradori S. and Scarpa S. (2018). The kinesin Eg5 inhibitor K858 induces apoptosis and reverses the malignant invasive phenotype in human glioblastoma cells. *Investigational New Drugs*, 36(1), 28–35

[184] **Ricci A., Cataldi A., Carradori S. and Zara S.** (2022). Kinesin Eg5 Selective Inhibition by Newly Synthesized Molecules as an Alternative Approach to Counteract Breast Cancer Progression: An In Vitro Study. *Biology*, 11(10), 1450

[185] **Nicolai A., Taurone S., Carradori S., Artico M., Greco A., Costi R., et al.** (2022). The kinesin Eg5 inhibitor K858 exerts antiproliferative and proapoptotic effects and attenuates the invasive potential of head and neck squamous carcinoma cells. *Investigational New Drugs*, 40(3), 556–564

[186] **Marconi G. D., Carradori S., Ricci A., Guglielmi P., Cataldi A. and Zara S.** (2019). Kinesin Eg5 Targeting Inhibitors as a New Strategy for Gastric Adenocarcinoma Treatment. *Molecules*, 24(21), 3948

[187] **Novais P., Silva P. M. A., Amorim I. and Bousbaa H.** (2021). Second-Generation Antimitotics in Cancer Clinical Trials. *Pharmaceutics*, 13(7), 1011

[188] **Infante J. R., Patnaik A., Verschraegen C. F., Olszanski A. J., Shaheen M., Burris H. A., et al.** (2017). Two Phase 1 dose-escalation studies exploring multiple regimens of litronesib (LY2523355), an Eg5 inhibitor, in patients with advanced cancer. *Cancer Chemotherapy and Pharmacology*, 79(2), 315–326

[189] **Wakui H., Yamamoto N., Kitazono S., Mizugaki H., Nakamichi S., Fujiwara Y., et al.** (2014). A phase 1 and dose-finding study of LY2523355 (litronesib), an Eg5 inhibitor, in Japanese patients with advanced solid tumors. *Cancer Chemotherapy and Pharmacology*, 74(1), 15–23

[190] **Szeliga M.** (2020). Thiadiazole derivatives as anticancer agents. *Pharmacological Reports*, 72(5), 1079–1100

[191] **Elseginy S. A.** (2024). Identifying and characterising promising small molecule inhibitors of kinesin spindle protein using ligand-based virtual screening, molecular docking, molecular dynamics and MM-GBSA calculations. *Journal of Computer-Aided Molecular Design*, 38(1), 16

[192] **Chen L. C., Rosen L. S., Iyengar T., Goldman J. W., Savage R., Kazakin J., et al.** (2011). First-in-human study with ARQ 621, a novel inhibitor of Eg5: Final results from the solid tumors cohort. *Journal of Clinical Oncology*, 29(15_suppl), 3076–3076

[193] **Schiemann K., Finsinger D., Zenke F., Amendt C., Knöchel T., Bruge D., et al.** (2010). The discovery and optimization of hexahydro-2H-pyrano[3,2-c]quinolines (HHPQs) as potent and selective inhibitors of the mitotic kinesin-5. *Bioorganic & Medicinal Chemistry Letters*, 20(5), 1491–1495

[194] **Konjikusic M. J., Gray R. S. and Wallingford J. B.** (2021). The developmental biology of kinesins. *Developmental Biology*, 469, 26–36

[195] **Alrazi I. M. D., Ogunwa T. H., Kolawole A. O., Elekofehinti O. O., Omotuyi O. I., Miyanishi T., et al.** (2021). Kolaflavanone, a biflavonoid derived from medicinal plant *Garcinia* , is an inhibitor of mitotic kinesin Eg5. *The Journal of Biochemistry*, 170(5), 611–622

[196] **Nakazawa J., Yajima J., Usui T., Ueki M., Takatsuki A., Imoto M., et al.** (2003). A Novel Action of Terpendole E on the Motor Activity of Mitotic Kinesin Eg5. *Chemistry & Biology*, 10(2), 131–137

[197] **Sheff J. G., Farshidfar F., Bathe O. F., Kopciuk K., Gentile F., Tuszynski J., et al.** (2017). Novel Allosteric Pathway of Eg5 Regulation Identified through Multivariate Statistical Analysis of Hydrogen-Exchange Mass Spectrometry (HX-MS) Ligand Screening Data. *Molecular & Cellular Proteomics*, 16(3), 428–437

[198] **Ma Y., Chen S.-S., Feng Y.-Y. and Wang H.-L.** (2020). Identification of novel biomarkers involved in pulmonary arterial hypertension based on multiple-microarray analysis. *Bioscience Reports*, 40(9), BSR20202346

[199] **Nyamaa B., Kim H. K., Jeong Y. J., Song I.-S. and Han J.** (2014). Kinesin Spindle Protein Inhibition in Translational Research. *Journal of Lipid and Atherosclerosis*, 3(2), 63

[200] **Chourasiya R. K., Hourani W., Deb P. K. and Venugopala K. N.** (2024). *Mitotic kinesin spindle protein (KSP/Eg5 ATPase) inhibitors*. Elsevier

[201] **Woessner R., Tunquist B., Lemieux C., Chlipala E., Jackinsky S., Dewolf W., et al.** (2009). ARRY-520, a novel KSP inhibitor with potent activity in hematological and taxane-resistant tumor models. *Anticancer Research*, 29(11), 4373–4380

[202] **Ocio E. M., Motlló C., Rodríguez-Otero P., Martínez-López J., Cejalvo M. J., Martín-Sánchez J., et al.** (2021). Filanesib in combination with pomalidomide and dexamethasone in refractory MM patients: safety and efficacy, and association with alpha 1-acid glycoprotein (AAG) levels. Phase Ib/II Pomdefil clinical trial conducted by the Spanish MM group. *British Journal of Haematology*, 192(3), 522–530

[203] **Gerecitano J. F., Stephenson J. J., Lewis N. L., Osmukhina A., Li J., Wu K., et al.** (2013). A Phase I trial of the kinesin spindle protein (Eg5) inhibitor AZD4877 in patients with solid and lymphoid malignancies. *Investigational New Drugs*, 31(2), 355–362

[204] **Jones R., Vuky J., Elliott T., Mead G., Arranz J. A., Chester J., et al.** (2013). Phase II study to assess the efficacy, safety and tolerability of the mitotic spindle kinesin inhibitor AZD4877 in patients with recurrent advanced urothelial cancer. *Investigational New Drugs*, 31(4), 1001–1007

[205] **Wang Y., Wu X., Du M., Chen X., Ning X., Chen H., et al.** (2017). Eg5 inhibitor YL001 induces mitotic arrest and inhibits tumor proliferation. *Oncotarget*, 8(26), 42510–42524

[206] **Mross K. B., Scharr D., Richly H., Frost A., Bauer S., Krauss B., et al.** (2014). First-in-human study of 4SC-205 (AEGIS), a novel oral inhibitor of Eg5 kinesin spindle protein. *Journal of Clinical Oncology*, 32(15_suppl), 2564–2564

[207] **Ye X. S., Fan L., Van Horn R. D., Nakai R., Ohta Y., Akinaga S., et al.** (2015). A Novel Eg5 Inhibitor (LY2523355) Causes Mitotic Arrest and Apoptosis in Cancer Cells and Shows Potent Antitumor Activity in Xenograft Tumor Models. *Molecular Cancer Therapeutics*, 14(11), 2463–2472

[208] **Toudjarska I., Judge A., Buck T., McClintock K., De Jong S., Ambegia E., et al.** (2009). Abstract B204: Development of ALN-VSP: An RNAi therapeutic for liver malignancies. *Molecular Cancer Therapeutics*, 8(12_Supplement), B204–B204

[209] **Tabernero J., Shapiro G. I., LoRusso P. M., Cervantes A., Schwartz G. K., Weiss G. J., et al.** (2013). First-in-Humans Trial of an RNA Interference Therapeutic Targeting VEGF and KSP in Cancer Patients with Liver Involvement. *Cancer Discovery*, 3(4), 406–417

[210] **Kasap C., Elemento O. and Kapoor T. M.** (2014). DrugTargetSeqR: a genomics- and CRISPR-Cas9–based method to analyze drug targets. *Nature Chemical Biology*, 10(8), 626–628

[211] Wacker S. A., Houghtaling B. R., Elemento O. and Kapoor T. M. (2012). Using transcriptome sequencing to identify mechanisms of drug action and resistance. *Nature Chemical Biology*, 8(3), 235–237

[212] Indorato R.-L., DeBonis S., Kozielski F., Garcia-Saez I. and Skoufias D. A. (2013). STLC-resistant cell lines as tools to classify chemically divergent Eg5 targeting agents according to their mode of action and target specificity. *Biochemical Pharmacology*, 86(10), 1441–1451

[213] Dumas M. E., Chen G.-Y., Kendrick N. D., Xu G., Larsen S. D., Jana S., et al. (2019). Dual inhibition of Kif15 by oxindole and quinazolinedione chemical probes. *Bioorganic & Medicinal Chemistry Letters*, 29(2), 148–154

[214] Kaan H. Y. K., Weiss J., Menger D., Ulaganathan V., Tkocz K., Laggner C., et al. (2011). Structure−Activity Relationship and Multidrug Resistance Study of New *S* -trityl- L -Cysteine Derivatives As Inhibitors of Eg5. *Journal of Medicinal Chemistry*, 54(6), 1576–1586

[215] Szakács G., Paterson J. K., Ludwig J. A., Booth-Genthe C. and Gottesman M. M. (2006). Targeting multidrug resistance in cancer. *Nature Reviews Drug Discovery*, 5(3), 219–234

[216] Tanenbaum M. E., Macůrek L., Janssen A., Geers E. F., Alvarez-Fernández M. and Medema R. H. (2009). Kif15 Cooperates with Eg5 to Promote Bipolar Spindle Assembly. *Current Biology*, 19(20), 1703–1711

[217] Milic B., Chakraborty A., Han K., Bassik M. C. and Block S. M. (2018). KIF15 nanomechanics and kinesin inhibitors, with implications for cancer chemotherapeutics. *Proceedings of the National Academy of Sciences*, 115(20)

[218] Mardin B. R., Isokane M., Cosenza M. R., Krämer A., Ellenberg J., Fry A. M., et al. (2013). EGF-Induced Centrosome Separation Promotes Mitotic Progression and Cell Survival. *Developmental Cell*, 25(3), 229–240

[219] Mouchlis V. D., Afantitis A., Serra A., Fratello M., Papadiamantis A. G., Aidinis V., et al. (2021). Advances in De Novo Drug Design: From Conventional to Machine Learning Methods. *International Journal of Molecular Sciences*, 22(4), 1676

[220] Orth J. D., Loewer A., Lahav G. and Mitchison T. J. (2012). Prolonged mitotic arrest triggers partial activation of apoptosis, resulting in DNA damage and p53 induction. *Molecular Biology of the Cell*, 23(4), 567–576

[221] Mousavi S. F. and Fatemi M. H. (2019). A combination of molecular docking, receptor-guided QSAR, and molecular dynamics simulation studies of S-trityl-l-cysteine analogues as kinesin Eg5 inhibitors. *Structural Chemistry*, 30(1), 115–126

[222] Jiang C., Chen Y., Wang X. and You Q. (2007). Docking studies on kinesin spindle protein inhibitors: an important cooperative ‘minor binding pocket’ which increases the binding affinity significantly. *Journal of Molecular Modeling*, 13(9), 987–992

[223] Nagarajan S., Skoufias D. A., Kozielski F. and Pae A. N. (2012). Receptor–Ligand Interaction-Based Virtual Screening for Novel Eg5/Kinesin Spindle Protein Inhibitors. *Journal of Medicinal Chemistry*, 55(6), 2561–2573

[224] Bodun D. S., Omoboyowa D. A., Omotuyi O. I., Olugbogi E. A., Balogun T. A., Ezeh C. J., et al. (2023). QSAR-based virtual screening of traditional Chinese medicine for the identification of mitotic kinesin Eg5 inhibitors. *Computational Biology and Chemistry*, 104, 107865

[225] **Sebastian J., Raghav D. and Rathinasamy K.** (2023). MD simulation-based screening approach identified tolvaptan as a potential inhibitor of Eg5. *Molecular Diversity*, 27(3), 1203–1221

[226] **Maiti P., Sharma P., Nand M., Bhatt I. D., Ramakrishnan M. A., Mathpal S., et al.** (2022). Integrated Machine Learning and Chemoinformatics-Based Screening of Mycotic Compounds against Kinesin Spindle ProteinEg5 for Lung Cancer Therapy. *Molecules*, 27(5), 1639

[227] **Bodun D., Ibrahim I., Bashiru M., Abolade R. O., Ndarawit W. K., Owolade A. J.-J., et al.** (2025). *De Novo Design and Bioactivity Prediction of Mitotic Kinesin Eg5 Inhibitors Using MPNN and LSTM-Based Transfer Learning*. Chemistry

[228] **Oguievetskaia K., Martin-Chanas L., Vorotyntsev A., Doppelt-Azeroual O., Brotel X., Adcock S. A., et al.** (2010). Computational fragment-based drug design to explore the hydrophobic subpocket of the mitotic kinesin Eg5 allosteric binding site. *Journal of Cheminformatics*, 2(S1), P29, 1758-2946-2-S1-P29

[229] **Naseer Q. A., Xuexian C., Yimai D., Khan M. A. and Chen S.** (2025). Fragment-Based Drug Discovery of KIF11 Inhibitors for Glioblastoma Treatment: Molecular Insights and Therapeutic Potential. *Drug Research*, 75(03/04), 114–124

[230] **Karpov A. S., Nieto-Oberhuber C. M., Abrams T., Beng-Louka E., Blanco E., Chamoin S., et al.** (2019). Discovery of Potent and Selective Antibody–Drug Conjugates with Eg5 Inhibitors through Linker and Payload Optimization. *ACS Medicinal Chemistry Letters*, 10(12), 1674–1679

[231] **Zhao D.-G., Liu J., Su Z., Zou W., Zhou Q., Yin T., et al.** (2025). Discovery of novel KSP-targeting PROTACs with potent antitumor effects in vitro and in vivo. *European Journal of Medicinal Chemistry*, 282, 117052

[232] **Békés M., Langley D. R. and Crews C. M.** (2022). PROTAC targeted protein degraders: the past is prologue. *Nature Reviews Drug Discovery*, 21(3), 181–200

[233] **Liu X., Gong H. and Huang K.** (2013). Oncogenic role of kinesin proteins and targeting kinesin therapy. *Cancer Science*, 104(6), 651–656

[234] **Pérez-Melero C.** (2014). KSP inhibitors as antimitotic agents. *Current Topics in Medicinal Chemistry*, 14(20), 2286–2311

[235] **Chen Y., Chow J. P. H. and Poon R. Y. C.** (2012). Inhibition of Eg5 acts synergistically with checkpoint abrogation in promoting mitotic catastrophe. *Molecular cancer research: MCR*, 10(5), 626–635

[236] **Zheng J., Tan Y., Liu X., Zhang C., Su K., Jiang Y., et al.** (2022). NAT10 regulates mitotic cell fate by acetylating Eg5 to control bipolar spindle assembly and chromosome segregation. *Cell Death & Differentiation*, 29(4), 846–860

[237] **Castillo A., Morse H. C., Godfrey V. L., Naeem R. and Justice M. J.** (2007). Overexpression of *Eg5* Causes Genomic Instability and Tumor Formation in Mice. *Cancer Research*, 67(21), 10138–10147

[238] **Balasundaram A. and C G. P. D.** (2023). In silico analysis revealed the potential circRNA-miRNA-mRNA regulative network of non-small cell lung cancer (NSCLC). *Computers in Biology and Medicine*, 152, 106315

[239] **Liu J., Feng Y., Zeng X., He M., Gong Y. and Liu Y.** (2021). LncRNA VPS9D1-AS1 Promotes Malignant Progression of Lung Adenocarcinoma by Targeting miRNA-30a-5p/KIF11 Axis. *Frontiers in Genetics*, 12, 807628

[240] **Piñero F., Dirchwolf M. and Pessôa M. G.** (2020). Biomarkers in Hepatocellular Carcinoma: Diagnosis, Prognosis and Treatment Response Assessment. *Cells*, 9(6), 1370

[241] **Neska-Długosz I., Buchholz K., Durślewicz J., Gagat M., Grzanka D., Tojek K., et al.** (2021). Prognostic Impact and Functional Annotations of KIF11 and KIF14 Expression in Patients with Colorectal Cancer. *International Journal of Molecular Sciences*, 22(18), 9732

[242] **Kim M., Jeong H. J., Ju H., Song J., Jang S. J. and Choi J.** (2023). Overexpression of the NEK9–EG5 axis is a novel metastatic marker in pathologic stage T3 colon cancer. *Scientific Reports*, 13(1), 342

[243] **Zhou Y., Yang L., Xiong L., Wang K., Hou X., Li Q., et al.** (2021). KIF11 is upregulated in colorectal cancer and silencing of it impairs tumor growth and sensitizes colorectal cancer cells to oxaliplatin via p53/GSK3β signaling. *Journal of Cancer*, 12(12), 3741–3753

[244] **Gu X., Zhu Q., Tian G., Song W., Wang T., Wang A., et al.** (2022). KIF11 manipulates SREBP2-dependent mevalonate cross talk to promote tumor progression in pancreatic ductal adenocarcinoma. *Cancer Medicine*, 11(17), 3282–3295

[245] **Shi H., Xu H., Chai C., Qin Z. and Zhou W.** (2022). Integrated bioinformatics analysis of potential biomarkers for pancreatic cancer. *Journal of Clinical Laboratory Analysis*, 36(5), e24381

[246] **Ricci A., Gallorini M., Del Bufalo D., Cataldi A., D'Agostino I., Carradori S., et al.** (2022). Negative Modulation of the Angiogenic Cascade Induced by Allosteric Kinesin Eg5 Inhibitors in a Gastric Adenocarcinoma In Vitro Model. *Molecules (Basel, Switzerland)*, 27(3), 957

[247] **Zeng X., Shi G., He Q. and Zhu P.** (2021). Screening and predicted value of potential biomarkers for breast cancer using bioinformatics analysis. *Scientific Reports*, 11(1), 20799

[248] **Kaur B., Mukhlis Y., Natesh J., Penta D. and Musthapa Meeran S.** (2022). Identification of hub genes associated with EMT-induced chemoresistance in breast cancer using integrated bioinformatics analysis. *Gene*, 809, 146016

[249] **Dong C., Tian X., He F., Zhang J., Cui X., He Q., et al.** (2021). Integrative analysis of key candidate genes and signaling pathways in ovarian cancer by bioinformatics. *Journal of Ovarian Research*, 14(1), 92

[250] **Zhao Y., Pi J., Liu L., Yan W., Ma S. and Hong L.** (2021). Identification of the Hub Genes Associated with the Prognosis of Ovarian Cancer Patients via Integrated Bioinformatics Analysis and Experimental Validation. *Cancer Management and Research*, 13, 707–721

[251] **Hao M., Zhang J., Sun M., Diao K., Wang J., Li S., et al.** (2022). TRAF4 Inhibits the Apoptosis and Promotes the Proliferation of Breast Cancer Cells by Inhibiting the Ubiquitination of Spindle Assembly-Associated Protein Eg5. *Frontiers in Oncology*, 12, 855139

[252] **Guido B. C., Brandão D. C., Barbosa A. L. A., Vianna M. J. X., Faro L., Ramos L. M., et al.** (2021). Exploratory comparisons between different anti-mitotics in clinically-used drug combination in triple negative breast cancer. *Oncotarget*, 12(19), 1920–1936

[253] **Pei Y., Li G., Ran J., Wan X., Wei F. and Wang L.** (2019). Kinesin Family Member 11 Enhances the Self-Renewal Ability of Breast Cancer Cells by Participating in the Wnt/β-Catenin Pathway. *Journal of Breast Cancer*, 22(4), 522

[254] **Pudova E. A., Krasnov G. S., Kobelyatskaya A. A., Savvateeva M. V., Fedorova M. S., Pavlov V. S., et al.** (2020). Gene Expression Changes and Associated Pathways Involved in the Progression of Prostate Cancer Advanced Stages. *Frontiers in Genetics*, 11, 613162

[255] **Wang H., Li S., Liu B., Wei S., Wang T., Li T., et al.** (2022). KIF11: A potential prognostic biomarker for predicting bone metastasis-free survival of prostate cancer. *Oncology Letters*, 24(3), 312

[256] **Kowalewski A., Jaworski D., Antosik P., Smolińska M., Ligmanowska J., Grzanka D., et al.** (2022). Overexpression of kif11 is a poor prognostic factor in clear cell renal cell carcinoma. *Polish Journal of Pathology*, 73(2), 82–87

[257] **Kenchappa R. S., Dovas A., Argenziano M. G., Meyer C. T., Stopfer L. E., Banu M. A., et al.** (2022). Activation of STAT3 through combined SRC and EGFR signaling drives resistance to a mitotic kinesin inhibitor in glioblastoma. *Cell Reports*, 39(12), 110991

[258] **Venere M., Horbinski C., Crish J. F., Jin X., Vasanji A., Major J., et al.** (2015). The mitotic kinesin KIF11 is a driver of invasion, proliferation, and self-renewal in glioblastoma. *Science Translational Medicine*, 7(304)

[259] **Algarín E. M., Hernández-García S., Garayoa M. and Ocio E. M.** (2020). Filanesib for the treatment of multiple myeloma. *Expert Opinion on Investigational Drugs*, 29(1), 5–14

[260] **Ostergaard P., Simpson M. A., Mendola A., Vasudevan P., Connell F. C., van Impel A., et al.** (2012). Mutations in KIF11 Cause Autosomal-Dominant Microcephaly Variably Associated with Congenital Lymphedema and Chorioretinopathy. *The American Journal of Human Genetics*, 90(2), 356–362

[261] **Johnson K., Moriarty C., Tania N., Ortman A., DiPietrantonio K., Edens B., et al.** (2014). Kif11 dependent cell cycle progression in radial glial cells is required for proper neurogenesis in the zebrafish neural tube. *Developmental Biology*, 387(1), 73–92

[262] **Jones G. E., Ostergaard P., Moore A. T., Connell F. C., Williams D., Quarrell O., et al.** (2014). Microcephaly with or without chorioretinopathy, lymphoedema, or mental retardation (MCLMR): review of phenotype associated with KIF11 mutations. *European journal of human genetics: EJHG*, 22(7), 881–887

[263] **Robitaille J. M., Gillett R. M., LeBlanc M. A., Gaston D., Nightingale M., Mackley M. P., et al.** (2014). Phenotypic overlap between familial exudative vitreoretinopathy and microcephaly, lymphedema, and chorioretinal dysplasia caused by KIF11 mutations. *JAMA ophthalmology*, 132(12), 1393–1399

[264] **Freund R. K., Gibson E. S., Potter H. and Dell'Acqua M. L.** (2016). Inhibition of the Motor Protein Eg5/Kinesin-5 in Amyloid β-Mediated Impairment of Hippocampal Long-Term Potentiation and Dendritic Spine Loss. *Molecular Pharmacology*, 89(5), 552–559

[265] **Borysov S. I., Granic A., Padmanabhan J., Walczak C. E. and Potter H.** (2011). Alzheimer Aβ disrupts the mitotic spindle and directly inhibits mitotic microtubule motors. *Cell Cycle (Georgetown, Tex.)*, 10(9), 1397–1410

[266] **Bougé A.-L. and Parmentier M.-L.** (2016). Tau excess impairs mitosis and kinesin-5 function, leading to aneuploidy and cell death. *Disease Models & Mechanisms*, 9(3), 307–319

[267] **Lucero E. M., Freund R. K., Smith A., Johnson N. R., Dooling B., Sullivan E., et al.** (2022). Increased KIF11/kinesin-5 expression offsets Alzheimer Aβ-mediated toxicity and cognitive dysfunction. *iScience*, 25(11), 105288

[268] **Ogbomo H., Timm-McCann M., Barnes T., Xiang R. F., Jamil K., Ganguly A., et al.** (2018). Granule-Dependent NK Cell Killing of Cryptococcus Requires Kinesin to Reposition the Cytolytic Machinery for Directed Cytotoxicity. *Cell Reports*, 24(11), 3017–3032

[269] **Auwul M. R., Rahman M. R., Gov E., Shahjaman M. and Moni M. A.** (2021). Bioinformatics and machine learning approach identifies potential drug targets and pathways in COVID-19. *Briefings in Bioinformatics*, 22(5), bbab120

[270] **Chen L., K. Morrow J., T. Tran H., S. Phatak S., Du-Cuny L. and Zhang S.** (2012). From Laptop to Benchtop to Bedside: Structure-based Drug Design on Protein Targets. *Current Pharmaceutical Design*, 18(9), 1217–1239

[271] **Martin R. L., Heifetz A., Bodkin M. J. and Townsend-Nicholson A.** (2024). *High-Throughput Structure-Based Drug Design (HT-SBDD) Using Drug Docking, Fragment Molecular Orbital Calculations, and Molecular Dynamic Techniques*. Springer US, New York, NY

[272] **Xiang M., Cao Y., Fan W., Chen L. and Mo Y.** (2012). Computer-Aided Drug Design: Lead Discovery and Optimization. *Combinatorial Chemistry & High Throughput Screening*, 15(4), 328–337

[273] **Yu W. and MacKerell A. D.** (2017). *Computer-Aided Drug Design Methods*. Springer New York, New York, NY

[274] **Leach A. R. and Gillet V. J.** (2007). *An Introduction To Chemoinformatics*. Springer Netherlands, Dordrecht

[275] **Sun H., Tawa G. and Wallqvist A.** (2012). Classification of scaffold-hopping approaches. *Drug Discovery Today*, 17(7–8), 310–324

[276] **Rush T. S., Grant J. A., Mosyak L. and Nicholls A.** (2005). A Shape-Based 3-D Scaffold Hopping Method and Its Application to a Bacterial Protein−Protein Interaction. *Journal of Medicinal Chemistry*, 48(5), 1489–1495

[277] **Grant J. A., Gallardo M. A. and Pickup B. T.** (1996). A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape. *Journal of Computational Chemistry*, 17(14), 1653–1666

[278] *FastROCS Shape based virtual screening software*.

[279] **Hawkins P. C. D., Skillman A. G. and Nicholls A.** (2007). Comparison of Shape-Matching and Docking as Virtual Screening Tools. *Journal of Medicinal Chemistry*, 50(1), 74–82

[280] **Giordano D., Biancaniello C., Argenio M. A. and Facchiano A.** (2022). Drug Design by Pharmacophore and Virtual Screening Approach. *Pharmaceuticals*, 15(5), 646

[281] **Chen Z., Li H., Zhang Q., Bao X., Yu K., Luo X., et al.** (2009). Pharmacophore-based virtual screening versus docking-based virtual screening: a benchmark comparison against eight targets. *Acta Pharmacologica Sinica*, 30(12), 1694–1708

[282] **Chiang Y.-K., Kuo C.-C., Wu Y.-S., Chen C.-T., Coumar M. S., Wu J.-S., et al.** (2009). Generation of Ligand-Based Pharmacophore Model and Virtual Screening for Identification of Novel Tubulin Inhibitors with Potent Anticancer Activity. *Journal of Medicinal Chemistry*, 52(14), 4221–4233

[283] **Friesner R. A., Murphy R. B., Repasky M. P., Frye L. L., Greenwood J. R., Halgren T. A., et al.** (2006). Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *Journal of Medicinal Chemistry*, 49(21), 6177–6196

[284] **Hecker E. A., Duraiswami C., Andrea T. A. and Diller D. J.** (2002). Use of Catalyst Pharmacophore Models for Screening of Large Combinatorial Libraries. *Journal of Chemical Information and Computer Sciences*, 42(5), 1204–1211

[285] **Bauer M. R., Ibrahim T. M., Vogel S. M. and Boeckler F. M.** (2013). Evaluation and Optimization of Virtual Screening Workflows with DEKOIS 2.0 – A Public Library of Challenging Docking Benchmark Sets. *Journal of Chemical Information and Modeling*, 53(6), 1447–1462

[286] **Vogel S. M., Bauer M. R. and Boeckler F. M.** (2011). DEKOIS: Demanding Evaluation Kits for Objective *in Silico* Screening — A Versatile Tool for Benchmarking Docking Programs and Scoring Functions. *Journal of Chemical Information and Modeling*, 51(10), 2650–2665

[287] **Boeckler F. M., Bauer M. R., Ibrahim T. M. and Vogel S. M.** (2014). Use of DEKOIS 2.0 to gain insights for virtual screening. *Journal of Cheminformatics*, 6(Suppl 1), O24

[288] **Janssens A. C. J. W. and Martens F. K.** (2020). Reflection on modern methods: Revisiting the area under the ROC Curve. *International Journal of Epidemiology*, 49(4), 1397–1403

[289] **Cebi E., Lee J., Subramani V. K., Bak N., Oh C. and Kim K. K.** (2024). Cryo-electron microscopy-based drug design. *Frontiers in Molecular Biosciences*, 11, 1342179

[290] **Van Montfort R. L. M. and Workman P.** (2017). Structure-based drug design: aiming for a perfect fit. *Essays in Biochemistry*, 61(5), 431–437

[291] **Pinzi L. and Rastelli G.** (2019). Molecular Docking: Shifting Paradigms in Drug Discovery. *International Journal of Molecular Sciences*, 20(18), 4331

[292] **Guedes I. A., De Magalhães C. S. and Dardenne L. E.** (2014). Receptor–ligand molecular docking. *Biophysical Reviews*, 6(1), 75–87

[293] **Anighoro A. and Bajorath J.** (2016). Three-Dimensional Similarity in Molecular Docking: Prioritizing Ligand Poses on the Basis of Experimental Binding Modes. *Journal of Chemical Information and Modeling*, 56(3), 580–587

[294] **Eberhardt J., Santos-Martins D., Tillack A. F. and 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

[295] **Trott O. and 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

[296] **Halgren T. A., Murphy R. B., Friesner R. A., Beard H. S., Frye L. L., Pollard W. T., et al.** (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*, 47(7), 1750–1759

[297] **Verdonk M. L., Cole J. C., Hartshorn M. J., Murray C. W. and Taylor R. D.** (2003). Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics*, 52(4), 609–623

[298] **McGann M.** (2011). FRED Pose Prediction and Virtual Screening Accuracy. *Journal of Chemical Information and Modeling*, 51(3), 578–596

[299] **Meli R., Morris G. M. and Biggin P. C.** (2022). Scoring Functions for Protein-Ligand Binding Affinity Prediction Using Structure-based Deep Learning: A Review. *Frontiers in Bioinformatics*, 2, 885983

[300] **Wang J.-C. and Lin J.-H.** (2013). Scoring Functions for Prediction of Protein-Ligand Interactions. *Current Pharmaceutical Design*, 19(12), 2174–2182

[301] **Alhossary A., Handoko S. D., Mu Y. and Kwoh C.-K.** (2015). Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics*, 31(13), 2214–2216

[302] **Friesner R. A., Banks J. L., Murphy R. B., Halgren T. A., Klicic J. J., Mainz D. T., et al.** (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, 47(7), 1739–1749

[303] **Meng X.-Y., Zhang H.-X., Mezei M. and Cui M.** (2011). Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer Aided-Drug Design*, 7(2), 146–157

[304] **Truchon J.-F. and Bayly C. I.** (2007). Evaluating Virtual Screening Methods: Good and Bad Metrics for the "Early Recognition" Problem. *Journal of Chemical Information and Modeling*, 47(2), 488–508

[305] **Ibrahim T. M., Bauer M. R., Dörr A., Veyisoglu E. and Boeckler F. M.** (2015). pROC-Chemotype Plots Enhance the Interpretability of Benchmarking Results in Structure-Based Virtual Screening. *Journal of Chemical Information and Modeling*, 55(11), 2297–2307

[306] **Alonso H., Bliznyuk A. A. and Gready J. E.** (2006). Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*, 26(5), 531–568

[307] **Abraham M. J., Murtola T., Schulz R., Páll S., Smith J. C., Hess B., et al.** (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1–2, 19–25

[308] **Case D. A., Aktulga H. M., Belfon K., Cerutti D. S., Cisneros G. A., Cruzeiro V. W. D., et al.** (2023). AmberTools. *Journal of Chemical Information and Modeling*, 63(20), 6183–6191

[309] **Phillips J. C., Hardy D. J., Maia J. D. C., Stone J. E., Ribeiro J. V., Bernardi R. C., et al.** (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. *The Journal of Chemical Physics*, 153(4), 044130

[310] **Brooks B. R., Brooks C. L., Mackerell A. D., Nilsson L., Petrella R. J., Roux B., et al.** (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, 30(10), 1545–1614

[311] **Bowers K. J., Chow D. E., Xu H., Dror R. O., Eastwood M. P., Gregersen B. A., et al.** (2006). *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters*. IEEE, Tampa, FL

[312] **De Vivo M., Masetti M., Bottegoni G. and Cavalli A.** (2016). Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of Medicinal Chemistry*, 59(9), 4035–4061

[313] **Hollingsworth S. A. and Dror R. O.** (2018). Molecular Dynamics Simulation for All. *Neuron*, 99(6), 1129–1143

[314] **Durrant J. D. and McCammon J. A.** (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, 9(1), 71

[315] **Genheden S. and Ryde U.** (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, 10(5), 449–461

[316] **Hou T., Wang J., Li Y. and Wang W.** (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free

energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, 51(1), 69–82

[317] Reichel A. and Lienau P. (2015). *Pharmacokinetics in Drug Discovery: An Exposure-Centred Approach to Optimising and Predicting Drug Efficacy and Safety*. Springer International Publishing, Cham

[318] Lin J., Sahakian D., De Morais S., Xu J., Polzer R. and Winter S. (2003). The Role of Absorption, Distribution, Metabolism, Excretion and Toxicity in Drug Discovery. *Current Topics in Medicinal Chemistry*, 3(10), 1125–1154

[319] Young R. J. (2023). Today's drug discovery and the shadow of the rule of 5. *Expert Opinion on Drug Discovery*, 18(9), 965–972

[320] Vamathevan J., Clark D., Czodrowski P., Dunham I., Ferran E., Lee G., et al. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, 18(6), 463–477

[321] Kim J., Park S., Min D. and Kim W. (2021). Comprehensive Survey of Recent Drug Discovery Using Deep Learning. *International Journal of Molecular Sciences*, 22(18), 9983

[322] Zhang Y., Ye T., Xi H., Juhas M. and Li J. (2021). Deep Learning Driven Drug Discovery: Tackling Severe Acute Respiratory Syndrome Coronavirus 2. *Frontiers in Microbiology*, 12, 739684

[323] Geoffrey A S B., Madaj R., Sanker A., Valluri P. P., Gracia J. and Singh H. (2020). *Target2drug : A Novel Programmatic Workflow to Automate in Silico Drug Discovery*. Chemistry

[324] Amara K., Rodriguez-Perez R. and Jiménez Luna J. (2023). *Explaining compound activity predictions with a substructure-aware loss for graph neural networks*. Chemistry

[325] Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589

[326] Wu L., Gray M., Dang O., Xu J., Fang H. and Tong W. (2023). RxBERT: Enhancing drug labeling text mining and analysis with AI language modeling. *Experimental Biology and Medicine*, 248(21), 1937–1943

[327] Futia G. and Vetrò A. (2020). On the Integration of Knowledge Graphs into Deep Learning Models for a More Comprehensible AI—Three Challenges for Future Research. *Information*, 11(2), 122

[328] Nussinov R., Zhang M., Liu Y. and Jang H. (2022). AlphaFold, Artificial Intelligence (AI), and Allostery. *The Journal of Physical Chemistry B*, 126(34), 6372–6383

[329] Zhang Y., Wang Y., Zhou W., Fan Y., Zhao J., Zhu L., et al. (2019). A combined drug discovery strategy based on machine learning and molecular docking. *Chemical Biology & Drug Design*, 93(5), 685–699

[330] Zhou G., Rusnac D.-V., Park H., Canzani D., Nguyen H. M., Stewart L., et al. (2024). An artificial intelligence accelerated virtual screening platform for drug discovery. *Nature Communications*, 15(1), 7761

[331] Sousa T., Correia J., Pereira V. and Rocha M. (2021). Generative Deep Learning for Targeted Compound Design. *Journal of Chemical Information and Modeling*, 61(11), 5343–5361

[332] Yi J., Shi S., Fu L., Yang Z., Nie P., Lu A., et al. (2024). OptADMET: a web-based tool for substructure modifications to improve ADMET properties of lead compounds. *Nature Protocols*, 19(4), 1105–1121

[333] **Radchenko E. V., Rulev Yu. A., Safanyaev A. Ya., Palyulin V. A. and Zefirov N. S.** (2017). Computer-aided estimation of the hERG-mediated cardiotoxicity risk of potential drug components. *Doklady Biochemistry and Biophysics*, 473(1), 128–131

[334] **Cai H., Zhang H., Zhao D., Wu J. and Wang L.** (2022). FP-GNN: a versatile deep learning architecture for enhanced molecular property prediction. *Briefings in Bioinformatics*, 23(6), bbac408

[335] **Grygorenko O. O., Radchenko D. S., Dziuba I., Chuprina A., Gubina K. E. and Moroz Y. S.** (2020). Generating Multibillion Chemical Space of Readily Accessible Screening Compounds. *iScience*, 23(11), 101681

[336] **Murugan N. A., Podobas A., Gadioli D., Vitali E., Palermo G. and Markidis S.** (2022). A Review on Parallel Virtual Screening Softwares for High-Performance Computers. *Pharmaceuticals*, 15(1), 63

[337] **Ebalunode J. O., Ouyang Z., Liang J. and Zheng W.** (2008). Novel Approach to Structure-Based Pharmacophore Search Using Computational Geometry and Shape Matching Techniques. *Journal of Chemical Information and Modeling*, 48(4), 889–901

[338] **Peach M. L. and Nicklaus M. C.** (2009). Combining docking with pharmacophore filtering for improved virtual screening. *Journal of Cheminformatics*, 1(1), 6

[339] **Kalliokoski T.** (2021). Machine Learning Boosted Docking (HASTEN): An Open-source Tool To Accelerate Structure-based Virtual Screening Campaigns. *Molecular Informatics*, 40(9), 2100089

[340] **Gentile F., Agrawal V., Hsing M., Ton A.-T., Ban F., Norinder U., et al.** (2020). Deep Docking: A Deep Learning Platform for Augmentation of Structure Based Drug Discovery. *ACS Central Science*, 6(6), 939–949

[341] **Gleave J.** (2025). *jamesgleave/Deep-Docking-NonAutomated*.

[342] **Ton A., Gentile F., Hsing M., Ban F. and Cherkasov A.** (2020). https://files.docking.org/zinc20-ML/. *Molecular Informatics*, 39(8), 2000028

[343] **Sastry G. M., Dixon S. L. and Sherman W.** (2011). Rapid Shape-Based Ligand Alignment and Virtual Screening Method Based on Atom/Feature-Pair Similarities and Volume Overlap Scoring. *Journal of Chemical Information and Modeling*, 51(10), 2455–2466

[344] **Dixon S. L., Smondyrev A. M., Knoll E. H., Rao S. N., Shaw D. E. and Friesner R. A.** (2006). PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *Journal of Computer-Aided Molecular Design*, 20(10–11), 647–671

[345] **Jacobson M. P., Pincus D. L., Rapp C. S., Day T. J. F., Honig B., Shaw D. E., et al.** (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function, and Bioinformatics*, 55(2), 351–367

[346] **Johnston R. C., Yao K., Kaplan Z., Chelliah M., Leswing K., Seekins S., et al.** (2023). Epik: p K_a and Protonation State Prediction through Machine Learning. *Journal of Chemical Theory and Computation*, 19(8), 2380–2388

[347] **Yoo A. B., Jette M. A. and Grondona M.** (2003). *SLURM: Simple Linux Utility for Resource Management*. Springer Berlin Heidelberg, Berlin, Heidelberg

[348] *Schrödinger - Physics-based Software Platform for Molecular Discovery & Design*.

[349] **Duan J., Dixon S. L., Lowrie J. F. and Sherman W.** (2010). Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods. *Journal of Molecular Graphics and Modelling*, 29(2), 157–170

[350] *Python Release Python 3.12.3*.

[351] **Harris C. R., Millman K. J., Van Der Walt S. J., Gommers R., Virtanen P., Cournapeau D., et al.** (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362

[352] **Virtanen P., Gommers R., Oliphant T. E., Haberland M., Reddy T., Cournapeau D., et al.** (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272

[353] **Hunter J. D.** (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90–95

[354] *Bash - GNU Project - Free Software Foundation*.

[355] **Ravindranath P. A., Forli S., Goodsell D. S., Olson A. J. and Sanner M. F.** (2015). AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. *PLOS Computational Biology*, 11(12), e1004586

[356] **Schrödinger, LLC.** (2015). *The PyMOL Molecular Graphics System, Version 1.8*.

[357] **Holcomb M., Santos-Martins D., Tillack A. F. and Forli S.** (2022). Performance evaluation of flexible macrocycle docking in AutoDock. *QRB Discovery*, 3, e18

[358] (2025). *forlilab/Meeko*. Forli Lab

[359] **Halgren T. A.** (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5–6), 490–519

[360] *RDKit*.

[361] **Sud M.** (2016). MayaChemTools: An Open Source Package for Computational Drug Discovery. *Journal of Chemical Information and Modeling*, 56(12), 2292–2297

[362] *Docker Engine | Docker Docs*.

[363] *Home - GraphPad*.

[364] *RCSB PDB: Homepage*.

[365] **Madhavi Sastry G., Adzhigirey M., Day T., Annabhimoju R. and Sherman W.** (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, 27(3), 221–234

[366] **Lu C., Wu C., Ghoreishi D., Chen W., Wang L., Damm W., et al.** (2021). OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *Journal of Chemical Theory and Computation*, 17(7), 4291–4300

[367] **Zdrazil B., Felix E., Hunter F., Manners E. J., Blackshaw J., Corbett S., et al.** (2024). The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Research*, 52(D1), D1180–D1192

[368] **Wang S., Witek J., Landrum G. A. and Riniker S.** (2020). Improving Conformer Generation for Small Rings and Macrocycles Based on Distance Geometry and Experimental Torsional-Angle Preferences. *Journal of Chemical Information and Modeling*, 60(4), 2044–2058

[369] **Rogers D. and Hahn M.** (2010). Extended-Connectivity Fingerprints. *Journal of Chemical Information and Modeling*, 50(5), 742–754
[370] **PubChem.** *PubChem*.
[371] *ATPase/GTPase ELIPA™ (Enzyme Linked Inorganic Phosphate Assay)*.

CURRICULUM VITAE

Name SURNAME : Ahmet Fatih Sahin

Place and Date of Birth :

E-mail :

EDUCATION:

- **B.Sc. :** 2020, Bezmialem Vakif University, Pharmacy

PROFESSIONAL EXPERIENCE AND REWARDS:

PUBLICATIONS, PRESENTATIONS AND PATENTS PRODUCED FROM THE THESIS:

- **Sahin, A. F.** (2024). The Quest for Ligands Against Kinesin Motor Proteins Using CADD. Poster presented at *Drug Discovery Chemistry Europe,* Cambridge Healthtech Institute, 3–5 December 2024, Barcelona, Spain.

- **Sahin, A. F., Carradori, S., Ricci, A., Rashad, M., & Akdemir, A.** (2025). The quest for ligands against kinesin motor protein Eg5. *Archiv der Pharmazie*, 358(5), e70010.

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- **Carradori, S., Ammazzalorso, A., De Filippis, B., Şahin, A. F., Akdemir, A., Orekhova, A., Bonincontro, G., & Simonetti, G.** (2022). Azole-Based Compounds That Are Active against Candida Biofilm: In Vitro, In Vivo and In Silico Studies. *Antibiotics*, 11(10), 1375.

- **Sahin, A. F., Küçükgüzel, Ş. G., & Akdemir, A.** (2021). Molecular modelling studies to suggest novel scaffolds against SARS-CoV-2 target enzymes. *Journal of Research in Pharmacy,* 25(6), 1010–1017