Computational Modeling of Kinases in the DLK-JNK Signaling Pathway: From Conformational Dynamics to Inhibition Mechanism

Ph.D. Thesis

By

SUMAN KOIRALA

MEHTA FAMILY SCHOOL OF BIOSCIENCES AND BIOMEDICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

AUGUST 2025

Computational Modeling of Kinases in the DLK-JNK Signaling Pathway: From Conformational Dynamics to Inhibition Mechanism

A THESIS

Submitted in partial fulfillment of the

requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

by

SUMAN KOIRALA

MEHTA FAMILY SCHOOL OF BIOSCIENCES AND BIOMEDICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

AUGUST 2025

INDIAN INSTITUTE OF TECHNOLOGY INDORE

I hereby certify that the work which is being presented in the thesis entitled **Computational Modeling of Kinases in the DLK-JNK Signaling Pathway: From Conformational Dynamics to Inhibition Mechanism** in the partial fulfillment of the requirements for the award of the degree of **DOCTOR OF PHILOSOPHY** and submitted in the **MEHTA FAMILY SCHOOL OF BIOSCIENCES AND BIOMEDICAL ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from **August 2021** to **August 2025** under the supervision of **Dr. Parimal Kar**, Associate Professor, Indian Institute of Technology Indore.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Suman Koirala
08/05/2026

Signature of the student with date

(Suman Koirala)

This is to certify that the above statement made by the candidate is correct to the best of my/our knowledge.

ParimalKar

Signature of Thesis supervisor

(Dr. Parimal Kar)

Suman Koirala has successfully given his Ph.D. Oral Examination held on **08-05-2026**.

ParimalKar

Signature of Thesis Supervisor

(Dr. Parimal Kar)

ACKNOWLEDGEMENTS

I am sincerely grateful to everyone who has supported me, both directly and indirectly, throughout my academic journey. This thesis would not have been possible without their encouragement and assistance. Over the past four years, many individuals have played important roles in my progress, but none have been more influential than my thesis supervisor ***Dr. Parimal Kar.*** *His exceptional expertise, encouraging mentorship, and insightful contributions have guided my research and motivated me to pursue the highest quality in my work. With his support, I have grown in a motivating and friendly atmosphere that inspired me to pursue innovative ideas and develop on multiple levels. Our discussions, whether academic or personal, have broadened my perspective and enriched my experience. His constant support has played a crucial role in directing my academic path and has strengthened my confidence to carry out independent research.*

I am truly thankful to my PSPC members, ***Dr. Abhijit Joshi*** *and* ***Prof. Biswarup Pathak****, whose guidance and valuable advice have been crucial throughout my research journey. I express my sincere gratitude to the Director of the Indian Institute of Technology Indore for giving me the opportunity to become a member of this prestigious institution. I would also like to extend my sincere gratitude to the Head of the Department of Biosciences and Biomedical Engineering (BSBE), the Convener of the Discipline Post-Graduate Committee (DPGC), and all the faculty members of BSBE for their constant support during my research endeavors. I would like to express my gratitude to the University Grants Commission (UGC) for the doctoral fellowship during my Ph.D. I would also like to thank BSBE office staff,* ***Mr. Gaurav Kumar****,* ***Mr. Amit Mishra****,* ***Mr. Arif Patel****, and* ***Mr. Nilesh Omprakash*** *for providing all facilities in the department. I also acknowledge the sincere efforts of all the academic staff, library staff, and other technical and non-technical staff of IIT Indore for their constant support and help. I also wish to extend my gratitude to my collaborators:* ***Dr. Hem Chandra Jha*** *(IIT Indore),* ***Prof. Kiran Bala*** *(IIT Indore), and* ***Dr. Chandan Tamuly*** *(CSIR-NEIST). Furthermore, I extend my thanks to* ***Mr. Yogendra Singh, Mr.***

***Prahalad Singh Panwar, Mr. Subha Jana**, and the entire team at the computer and information technology center at IIT Indore for their invaluable technical assistance.*

*I would like to express my gratitude to all my present lab members: **Mr. Kapil Ursal, Ms. Subhasmita Mahapatra, Mr. Sunanda Samanta, Mr. Kesarpu Uday Kumar, Ms. Anamika Shukla, Mr. Pradeep Kumar, Ms. Ahana Chakraborty and Ms. Dibyanka Dalai.** I would like to extend my special thanks to **Mr. Sunanda Samanta** for being actively involved in every project we undertake. I wish to thank all my past lab members, particularly **Dr. Nisha Amarnath Jonniya, Dr. Md Fulbabu Sk, Dr. Rajarshi Roy, Dr. Sayan Poddar,** and **Mr. Naveen Kumar** for sharing their expertise in lab work and other valuable experiences. Their guidance has equipped me with the essential skills to carry out my research work. I would also like to thank my former M.Sc. lab members, **Ms. Trupti U Rathod** and **Ms. Mohini**, for sharing valuable time and memories with us. I will always be grateful to each of you for your support and friendship. The memories we have created together will stay with me forever. Working with such dedicated and talented people has truly been a privilege. You all have made my journey enjoyable and meaningful. These moments spent in the lab are unforgettable and will always hold a special place in my heart. I am lucky to have been part of such an amazing group. Thank you for being so supportive and inspiring every day. You have made even the toughest days easier to face. I will always cherish the bonds we built and the successes we celebrated together, and I hope you all continue to find happiness and success in everything you do, because you truly are an incredible group of people.*

*I want to thank my friends here at IIT Indore who have made my life much easier and more enjoyable. The arrival at IIT Indore in October 2021 remains unforgettable. Due to the post COVID-19 protocols, we had to quarantine for 15 days. I still remember spending those days with **Mr. Kapil Ursal, Mr. Lekhnath Sharma**, and **Mr. Pramod Patidar**, as our rooms were next to each other, we stayed connected and created some beautiful memories. Time passed so quickly that the quarantine hardly felt like a wait. I would like to extend my heartfelt gratitude to **Mr. Khandu Wadhonkar**, **Mr. Kapil Ursal, Mr. Sunanda Samanta, Ms. Meenakshi Kandpal**, and **Ms. Surbhi Jaiswal.***

They are more than just colleagues or friends, they are the ones who made my journey at IIT Indore truly memorable. I am deeply thankful to each of them for being such an integral part of my life. Whenever I think of IIT Indore, their names are the first that come to mind, and they will always hold a special place in my heart. I would also like to extend my gratitude to my dear friends and former seniors at IIT Indore, who truly enriched my time there. A special mention goes to ***Ms. Anshul, Ms. Kritika Malik, Mr. Sagnik Mitra, Mr. Aritra Chakraborty, Mr. Pranit H Bagde, Mr. Tileshwar Sahare, Mr. Rahul Chaudhari, Mr. Krishna Singh, Dr. Satyam Singh, Mr. Sibi Karthik****. Their cheerful conversations and valuable discussions offered both relaxation from research and a broader perspective to my journey.*

I would like to sincerely thank my parents, ***Mr. Parsu Ram Koirala*** *(Father) and* ***Mrs. Devi Maya Koirala*** *(Mother), whose constant love and support have been with me from the very first day of my journey. Their sacrifices and affection have given me the strength to pursue my dreams, and everything I am today is because of them. I would also like to express my special thanks to my wife,* ***Mrs. Priyanka Devi****; she has supported me in every situation with great care, patience, and understanding. I also want to express my heartfelt gratitude to my beloved grandmother (late* ***Mrs. Subhadra Koirala****), my sisters (****Mrs. Sunita Devi*** *and* ***Mrs. Sangita Koirala****), my cousins (****Mr. Kishore Koirala*** *and* ***Mr. Deepankar Sharma****), for their love and support throughout my life in general.*

Finally, I want to thank God for everything and express my gratitude to all my well-wishers who have supported me, directly or indirectly, throughout this journey. Your blessings mean a lot to me and I hope to always receive them.

….. Suman Koirala

Dedicated with deepest love and gratitude to my dear parents

SYNOPSIS

Introduction

Neurological diseases can be defined as a diverse group of disorders that affect the central and peripheral nervous systems, including conditions such as epilepsy, multiple sclerosis, stroke, neuropathies, and neurodegenerative diseases. These conditions account for a considerable proportion of global mortality, impacting millions of individuals worldwide. Neurological diseases are becoming more common due to an aging population, environmental exposures, and lifestyle modifications, typically resulting in chronic disability, cognitive decline, and reduced life quality[1]. Treatments available today for a variety of neurological disorders predominantly provide symptomatic relief, with limited capacity to influence disease progression. Furthermore, the presence of drug-related side effects and therapeutic resistance adds complexity to prolonged treatment management. Hence, the development of advanced and focused therapies is crucial for effectively addressing neurological diseases and improving clinical outcomes.

Protein kinases are enzymes that regulate other proteins by transferring phosphate group to specific amino acid residues, through a process called phosphorylation. This type of post-translational modification is crucial for regulating signal transduction pathways involved in cell growth, specialization, survival, and programmed cell death. Kinase dysregulation contributes to the progression of various diseases, such as cancer, inflammation, neurodegenerative disorders along with various other diseases, making kinases important targets for drug discovery[2].

The Dual Leucine Zipper Kinase (DLK) to c-Jun N-terminal Kinase (JNK) (DLK-JNK) signalling pathway serves as a central regulator of neuronal development, axon regeneration, and stress-induced apoptosis. DLK (MAP3K12) is positioned at the initial stage of this signalling pathway and is primarily found in neurons, where it functions as a master switch which initiates series of signals that regulate the cell's response to stress, injury, or growth[3]. When neurons encounter stress conditions like axonal damage, oxidative stress, or deprivation of growth-promoting factors, DLK is

activated and subsequently initiates the phosphorylation of downstream Mitogen-Activated Protein Kinase Kinases (MAP2Ks), with Mitogen-Activated Protein Kinase Kinase-7 (MKK7) being the main target and Mitogen-Activated Protein Kinase Kinase-4 (MKK4) activated to a lesser degree[4]. MKK7, characterized as a dual-specificity kinase, acts as a central and highly selective activator of JNKs at the upstream level. It phosphorylates JNK at the threonine and tyrosine residues located within the conserved Thr-Pro-Tyr motif in the activation loop, leading to its activation. In contrast to MKK4, which also targets p38 and MAPK, MKK7 exhibits a high degree of specificity for JNK signalling, highlighting its importance in DLK-JNK signalling pathway. **Figure 1** shows the sequential activation of kinases in the DLK–JNK signalling pathway. The JNK isoforms JNK1, JNK2, and JNK3 respond to stress by activating and translocating into the nucleus, where they regulate transcription factors such as c-Jun, ATF2, and Elk-1. By phosphorylating c-Jun, JNK drives the transcription of pro-apoptotic genes and controls processes like neuronal plasticity, regeneration, and apoptosis[5]. The dysregulation of this pathway has been linked to multiple neuropathological disorders such as Alzheimer's disease, Parkinson's disease, and traumatic brain injury, highlighting it as a promising target for therapeutic intervention.

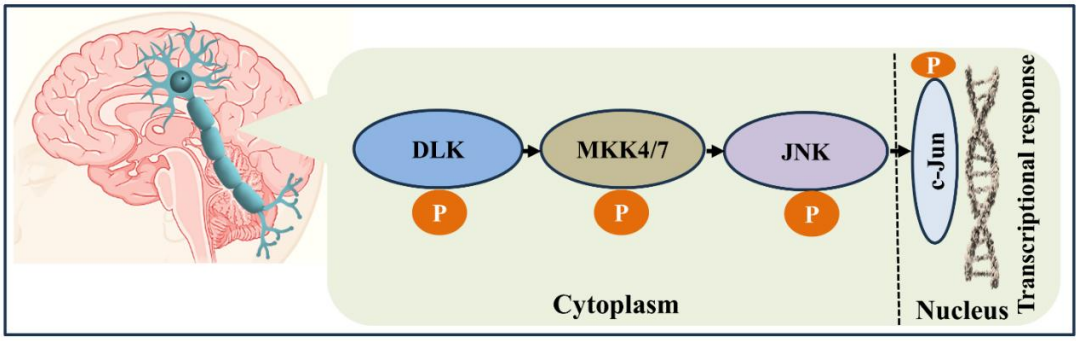


Figure 1: The DLK-JNK pathway showing sequential phosphorylation events from DLK to MKK4/7 to JNK, resulting in the activation of the transcription factor c-Jun and nuclear gene expression.

Objective of the work

DLK-JNK signaling plays a crucial role in neuronal stress responses, yet the detailed molecular processes governing kinase regulation, including activation and structural transitions, are still not fully understood. Although these kinases are known to undergo structural remodeling upon phosphorylation and interaction with regulatory factors, the precise mechanisms behind these changes are still not well known. How does phosphorylation, as a post-translational modification, affect the structural conformation and dynamics of proteins involved in the DLK-JNK signalling pathway? Furthermore, the molecular mechanisms of inhibition remain poorly understood because of the lack of crystal structures of kinases bound to small molecules. There is limited research focused on identifying potential inhibitors targeting DLK. Additionally, the selectivity profiles of allosteric inhibitors, as well as type I and type II inhibitors against MKK7, have not been comprehensively examined. Furthermore, the phosphorylation-induced conformational dynamics of JNK isoforms remain unexplored. Moreover, the interaction between MKK7 and JNK, along with its dynamic behaviour, are also yet to be investigated.

Therefore, this thesis broadly focuses on the following objectives:

1. **Computer-aided drug discovery for neurodegenerative diseases targeting DLK.**
    - To identify Saraswatharishta (SWRT)-derived phytochemicals that inhibit DLK, thereby exploring their potential role in neuroprotection and providing a molecular rationale for the observed neurodegenerative benefits of SWRT.
    - Identifying potential lead compounds against DLK by targeting the ATP-binding site.
2. **Theoretical investigation of MKK7 conformational transitions between active and inactive states through enhanced sampling techniques.**
    - Investigating and characterizing the conformational behaviour of phosphorylated MKK7 bound to type I, type II, and allosteric inhibitors.

3. **Atomistic insights into how post-translational modifications influence the structure and dynamics of JNK isoforms.**
    - Computational exploration of phosphorylation-induced conformational dynamics in JNK isoforms to understand isoform-specific regulation**.**
    - Studying the dynamic interaction patterns of inhibitor SP600125 across different isoforms.
4. **To gain atomistic-level understanding of JNK1 structural transition upon ATP and MKK7 (protein-protein) interactions.**
    - Identifying structural adaptations in JNK1 upon ATP and Mg^{2+} ion interaction at the active site.
    - Characterizing the structural basis of JNK1-MKK7 interaction and its influence on JNK1 dynamics.
    - Characterizing the influence of cooperative ATP and MKK7 binding on the global and local motions of JNK1.

Chapter 1: Overview and literature review of the DLK-JNK signalling pathway involved in neurological disorders

This chapter presents a comprehensive overview of the DLK-JNK signalling pathway and its involvement in various neurological disorders. It outlines the fundamental roles of DLK as an upstream regulator and its downstream activation of JNKs through MKK7/MKK4 in stress signalling, neuronal apoptosis, and neuroinflammation. The isoform-specific functions of JNK kinases and their implications in neuronal degeneration are discussed in detail. Furthermore, the chapter highlights key regulatory mechanisms, including phosphorylation events and scaffold proteins that influence pathway specificity. In addition, we examine the current challenges associated with targeting DLK-JNK signalling therapeutically, particularly with regard to isoform selectivity and context-dependent activity.

Chapter 2: Description of the methodology used for the studies

This chapter deals with the methodology used in the present studies. Here, the molecular dynamics (MD) simulations have been explained in detail. This chapter explains how MD simulations contribute to our understanding of the structural features, dynamic behaviour, and thermodynamics properties of biological macromolecules of interest at the atomistic level. Along with that, we have described the components of the force field, various water models, simulation packages, the basic workflow of MD, and binding free energy calculations using the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA)[6] method. Additionally, we have discussed enhanced simulation techniques, such as accelerated molecular dynamics (aMD) and Gaussian accelerated molecular dynamics (GaMD)[7], along with their advantages in conformational sampling with reduced computational expense. Further, we have provided insights into the different post-simulation analysis techniques of the MD trajectories to explain the molecular mechanism and binding free energy of small molecules.

Chapter 3: Ayurvedic phytochemicals from neurological formulation Saraswatharishta (SWRT) as a Potential Inhibitors of dual leucine zipper kinase

Globally, neurological disorders represent a major risk factor and rank as the second leading cause of death. Over the past three decades, there has been a significant global rise in deaths and disability caused by neurological disorders, with further increases expected due to aging populations and demographic growth. Ayurveda, a traditional medication system in India, has been utilized for centuries to treat various diseases, including neurological disorders. Saraswatharishta, an Ayurvedic formulation, has shown therapeutic potential in the treatment and management of various neurological disorders. The neuroprotective impact of SWRT in neurological disorders underscores the importance of examining its action on key molecular targets, especially DLK. In this work, we focus on elucidating the detailed molecular mechanisms by which SWRT phytochemicals interact with DLK. Initially, we downloaded the phytochemicals from the IMPPAT (Indian medicinal plants phytochemistry and therapeutics) database and curated a library for virtual screening against DLK. From the docking results, the top

ten compounds were shortlisted, of which four leads were selected based on both high docking scores and molecular weights under 500 Da and subjected them to MD simulation. To evaluate the stability and binding efficiency of the lead compounds with DLK, we analyzed ligand-protein distance, hydrogen bond formation, binding site RMSD, and calculated binding free energy using the MM-PBSA approach. The four identified lead compounds are sourced from different plants, specifically *Anethum sowa*, *Acorus calamus*, *Zingiber officinale*, and *Syzygium aromaticum*. By narrowing down the essential phytoconstituents of this ancient formulation, this study may *contribute to* the design of potent inhibitors against DLK. This study identified key phytochemicals of SWRT formulations and characterized their binding interactions with DLK. To the best of our knowledge, this is one of the first studies to provide molecular insight into the potential of Ayurvedic compounds for DLK-targeted neurodegenerative therapy.

Chapter 4: Identifying novel inhibitors and exploring how inhibitor binding influences conformational dynamics in dual leucine zipper kinase

The dual leucine zipper kinase (DLK) pathway plays a crucial role in regulating neuronal responses to stress and injury, and its dysregulation has been linked to the development and progression of various neurodegenerative diseases. Belonging to the MLK subfamily, DLK is a serine/threonine kinase involved in multiple cellular functions, including apoptosis, stress responses, and cell differentiation. The lack of FDA-approved drugs targeting DLK highlights the necessity for the development of potent inhibitors to address neurological pathologies. Therefore, the design of targeted therapies that are both potent and selective, while avoiding cytotoxic effects, has attracted considerable attention. This chapter focuses on the conventional systematic approach of high-throughput virtual screening of natural product libraries (NPAtlas) and FDA-approved compounds (MedChemExpress), followed by molecular dynamics simulations integrated with free energy calculations. We carried out virtual screening of natural products and FDA-approved drugs in combination with pharmacokinetic and pharmacological profiling, followed by unbiased molecular dynamics simulations to evaluate the stability of the complexes. We have also evaluated the ADMET properties

of the selected compounds. We have also explored the conformational dynamics of the kinase upon inhibitor binding. Finally, we identified two compounds from the NPAtlas database exhibiting favorable ADMET profiles, along with two FDA-approved drugs danthron and dithranol showing blood-brain barrier (BBB) permeability based on virtual screening and subjected to MD simulation. Among the natural compounds, CID156581477 originates from the fungus *Pestalotiopsis vaccinii*, while CID139591660 is derived from the marine fungus *Zopfiella marina*. Binding free energy estimates using MM-PBSA revealed higher affinity for CID139591660, dithranol, and danthron compared to control compound sunitinib, whereas CID156581477 exhibited reduced affinity. Furthermore, employing the deep learning-based DeLA Drug server, we identified an analogue of CID156581477 (CID156581477-ANLG) with higher affinity compared to both the parent compound and the control. Based on MM-PBSA and hydrogen bonding results, the inhibitor affinities followed the order: CID139591660 > dithranol > CID156581477-ANLG > danthron > sunitinib > CID156581477. Thus, four compounds were found to exhibit greater affinity than the control compound, sunitinib. However, further molecular and pharmacokinetics are required to test the selectivity and affinity *in vitro* and *in vivo*.

Chapter 5: Theoretical investigation of MKK7 structural transitions between active and inactive states upon binding to distinct inhibitors using enhanced sampling techniques

Mitogen-activated protein kinase kinase 7 (MKK7) is a dual-specificity kinase that plays a critical role in the MAPK pathway by acting as an upstream activator of c-Jun N-terminal kinase (JNK), promoting its phosphorylation and activation. JNK activation is triggered by stress stimuli, including inflammatory cytokines, UV radiation, and environmental stress, leading to cellular responses like apoptosis and inflammation. Upon activation, JNK moves into the nucleus and phosphorylates key transcription factors, such as c-Jun, regulating the expression of genes involved in apoptosis, inflammation, and the cellular stress response. MKK7, in conjunction with MKK4, enhances JNK activation, though it is more specific in JNK phosphorylation under stress conditions.

This chapter mainly comprises the conformational dynamics induced by type I and type II and allosteric inhibitors and their free energy calculations. Although assessing the influence of structural rearrangements on kinase inhibitor binding is challenging due to the high conformational flexibility of their catalytic kinase domain (KD). Four major flexible structural motifs constitute the dynamic kinase domain (KD): the P-loop and αC-helix from the N-terminal lobe, and the DFG motif and activation loop (A-loop) from the C-terminal lobe. These specific structural states reflect a complicated transduction pathway that determines the A-loop's dynamic behaviour and conformational selection necessary for MKK7 function. However, the specific dynamic behaviours of MKK7 in response to different inhibitor-bound states are still not well understood. We performed microsecond-long Gaussian accelerated MD (GaMD) simulations to study four phosphorylated forms of MKK7: apo KD, and KD in complex with Type I, Type II, and allosteric inhibitors. Our conformational analyses along with Markov State Model (MSM) indicates that binding of inhibitors results in conformational changes in mobile regions like the A-loop, αC-helix, and hydrophobic spines. We observed that the inactive state is characterized by a closed A-loop, outward displacement of the αC-helix, disruption of the hallmark salt bridge, and misalignment of the hydrophobic spines. Finally, we computed the binding free energy and compared the binding configurations using the MM/PBSA approach. According to the binding free energy calculations, type II inhibitors show the highest affinity, followed by N-terminal allosteric inhibitors, and then type I inhibitors. Overall, these structural and energetics analyses could serve as foundation for the future development of more effective type I/II and N-terminal allosteric inhibitors, with significant implications for treating MKK7-related diseases.

Chapter 6: Atomic-level investigation of post-translational modification (phosphorylation) effects on the structural dynamics of JNK isoforms

The c-Jun N-terminal kinase (JNK) signalling pathway is an essential part of the mitogen-activated protein kinase (MAPK) signalling pathway, functioning as an important link between external signals and cellular responses like extracellular stress, gene expression, and cell growth. JNK activation is mediated by two MAPK kinases

MKK4 and MKK7, which selectively phosphorylate tyrosine and threonine residues within the activation loop. The three isoforms of the JNK family, namely JNK1, JNK2, and JNK3, are encoded by different genes, and they vary in their expression patterns, substrate specificity, and individual functions within the cell. JNK1 and JNK2 are widely expressed in various tissues, contributing to metabolism, cell cycle control, and immune responses. In contrast, JNK3 is primarily expressed in the brain, where it regulates neuronal functions and stress responses, along with low expression in the heart and testis.

This chapter mainly comprises to explore the comparative structural dynamics and energetic characteristics of the three JNK isoforms JNK1, JNK2, and JNK3 bound with the inhibitor SP600125 in both phosphorylated and unphosphorylated forms. A detailed analysis of the structural dynamics of all three JNK isoforms including the effects of phosphorylation and inhibitor selectivity, has not yet been previously reported in the literature. Investigating the structural behaviour of these isoforms in phosphorylation states is crucial for understanding their functional variety and regulatory mechanism. We observe that the structural integrity of isoforms remains largely conserved, but phosphorylation leads to conformational changes in the activation loop and some other key regions, affecting accessibility, flexibility, and compactness. The phosphorylation affects the β3-Lys and αC-helix Glu salt bridge in JNK1 and JNK3, while JNK2 remains unchanged, and no isoform exhibits a stable β3-Lys and DFG-Asp salt bridge interaction. Finally, we computed the binding free energy and compared the binding configurations using the MM-PBSA approach. The results suggest that SP600125 shows significant affinity for all JNK isoforms. Overall, our findings offer valuable insights into the structural and conformational dynamics of JNK isoforms in complex with a competitive inhibitor, revealing fundamental challenges in isoform specific drug discovery.

Chapter 7: Investigating the Role of ATP Binding and MKK7 Interaction in Modulating the Structural and Dynamic Landscape of JNK1

Within the MAPK signalling pathway, JNK1 serves as a crucial mediator of cell response to stress, inflammation, and apoptotic stimuli. Its activation and regulation

are influenced by ATP binding and interaction with the upstream kinase MKK7, both of which can induce distinct structural and dynamic changes that shape JNK1's functional state. MKK7 activates the JNK pathway and is uniquely characterized by having three motifs in its regulatory domain. This chapter primarily focuses on the conformational dynamics of JNK1 in its apo form, ATP-bound state, MKK7-bound complex, and the dual ATP-MKK7 bound state, along with their associated binding free energy calculations.

The GaMD studies revealed conformational variability among the different states. Binding of ATP resulted in a restricted mobility of the P-loop, leading to a rigid free energy landscape in terms of its up and down motion towards the conserved HRD-motif. The activation segment adopted a range of conformations, including states where the loop shifted toward the αC-helix. Throughout all systems, the alignment of the hydrophobic spine architecture remained consistently intact. The salt bridge between β3-Lys55 and Asp169 of the DFG-motif alternated between formed and broken states. In contrast, the key interaction between β3-Lys55 and αC-helix Glu73 remained mostly intact across systems, except in the ATP-MKK7 bound complexes, where an additional prominent broken conformation was also observed. Furthermore, the binding mechanism, residue-specific energy contributions, and interaction profiles were thoroughly analysed. Overall, the study provided critical mechanistic insights into JNK1 in complex with MKK7 and ATP-bound states, emphasizing how binding influences their dynamic behaviour and interaction patterns.

Chapter 8: Conclusions and future scope

The advancement of computational techniques and algorithms has made it possible to study protein dynamics and structures at an atomic scale. While experimental 3D structures give precise structural information, they capture only a static image of molecular processes that are actually time-dependent and dynamic. In this regard, molecular dynamics and various simulation strategies are key to understanding proteins' spatial and temporal behaviour within timescales. In this thesis, a detailed exploration is provided, starting with DLK inhibitor development, followed by the structural dynamics of MKK7 in its various states, the effect of phosphorylation on

JNK isoforms, and the behavior of JNK1 when bound to ATP and MKK7. Firstly, our primary investigation centered on understanding SWRT's mechanism of action via the DLK pathway by screening its key phytochemical components for their effectiveness in treating neurological diseases and their potential as drug candidates. Following that, we screened two different databases to search for DLK inhibitors and successfully identified four promising compounds with high binding affinity. Additionally, we investigated the essential characteristics of type I, type II, and allosteric inhibitors bound to MKK7 in its active and inactive states using the recently developed GaMD enhanced sampling method. Furthermore, we explored the effects of phosphorylation on different JNK isoforms and examined how it influences their distinct dynamic behaviours, along with investigating the role of ATP binding and MKK7 interaction in modulating the structural and dynamic landscape of JNK1.

Since there are currently no FDA-approved drugs targeting any of the kinases in this pathway, there exists a significant opportunity to discover potent inhibitors for each kinase involved. Furthermore, identifying allosteric pockets within these kinases could aid in the development of novel inhibitors. In addition, designing inhibitors with high selectivity among JNK isoforms offers a promising approach to overcome issues related to isoform-specific selectivity. There is currently limited understanding of DLK kinase dynamics, particularly the distortion of the αC-helix. With no known phosphorylating residues identified in human DLK, this presents a promising area for detailed structural and functional investigation. This thesis opens up a new opportunity in the structure-based drug development field, where drug affinity and selectivity toward a given target enzyme can be improved through experimental structural biology and medicinal chemistry.

References

(1) Chin, J. H.; Vora, N. The Global Burden of Neurologic Diseases. *Neurology* **2014**, *83* (4), 349–351. https://doi.org/10.1212/WNL.0000000000000610.

(2) Schwartz, P. A.; Murray, B. W. Protein Kinase Biochemistry and Drug Discovery. *Bioorganic Chemistry* **2011**, *39* (5–6), 192–210. https://doi.org/10.1016/j.bioorg.2011.07.004.

(3) Tedeschi, A.; Bradke, F. The DLK Signalling Pathway—a Double-edged Sword in Neural Development and Regeneration. *EMBO Reports* **2013**, *14* (7), 605–614. https://doi.org/10.1038/embor.2013.64.

(4) Yamasaki, T.; Kawasaki, H.; Nishina, H. Diverse Roles of JNK and MKK Pathways in the Brain. *Journal of Signal Transduction* **2012**, *2012*, 1–9. https://doi.org/10.1155/2012/459265.

(5) Bogoyevitch, M. A.; Kobe, B. Uses for JNK: The Many and Varied Substrates of the c-Jun N-Terminal Kinases. *Microbiol Mol Biol Rev* **2006**, *70* (4), 1061–1095. https://doi.org/10.1128/MMBR.00025-06.

(6) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33* (12), 889–897. https://doi.org/10.1021/ar000033j.

(7) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11* (8), 3584–3595. https://doi.org/10.1021/acs.jctc.5b00436.

LIST OF PUBLICATIONS

Publication from thesis

1. **Koirala, S.**; Roy, R.; Samanta, S.; Mahapatra, S.; Kar, P. Plant Derived Active Compounds of Ayurvedic Neurological Formulation, Saraswatharishta as a Potential Dual Leucine Zipper Kinase Inhibitor: An in-Silico Study. *Journal of Biomolecular Structure and Dynamics* 2024, 42 (20), 11201–11214. https://doi.org/10.1080/07391102.2023.2260892.
2. **Koirala, S.**; Samanta, S.; Kar, P. Identification of Inhibitors for Neurodegenerative Diseases Targeting Dual Leucine Zipper Kinase through Virtual Screening and Molecular Dynamics Simulations. *SAR and QSAR in Environmental Research* 2024, 35 (6), 457–482. https://doi.org/10.1080/1062936X.2024.2363195.
3. **Koirala, S.**; Samanta, S.; Mahapatra, S.; Ursal, K. D.; Poddar, S.; Kar, P. Molecular Level Investigation for Identifying Potential Inhibitors against Thymidylate Kinase of Monkeypox through *in Silico* Approaches. *Journal of Biomolecular Structure and Dynamics* **2024**, *42* (23), 13247–13260. https://doi.org/10.1080/07391102.2023.2274982.
4. **Koirala, S.**; Samanta, S.; Kar, P. Characterizing the Conformational Dynamics and Inhibition Mechanism of MKK7 by Different Inhibitor Types via Gaussian Accelerated Molecular Dynamics Simulations. International Journal of Biological Macromolecules (**Submitted**)
5. **Koirala, S.**; Samanta, S.; Ursal, KD.; Kar, P. Unravelling the Molecular Choreography of JNK Isoforms to Understand the Structural Dynamics and Phosphorylation Driven Conformational Transitions. Archives of Biochemistry and Biophysics (**Submitted**)
6. **Koirala, S.**; Samanta, S.; Kar, P. Dynamic Variability in JNK1: The Role of ATP and MKK7-Mediated Interactions in Structural Regulation and Energy Landscape. **(Manuscript under preparation)**

Publication apart from the thesis

1. Samanta, S.; Sk, M. F.; **Koirala, S.**; Kar, P. Exploring Molecular Interactions of Potential Inhibitors against the Spleen Tyrosine Kinase Implicated in Autoimmune Disorders via Virtual Screening and Molecular Dynamics Simulations. *SAR and QSAR in Environmental Research* **2023**, *34* (11), 869–897. https://doi.org/10.1080/1062936X.2023.2266364.

2. Kashyap, D.; **Koirala, S.**; Roy, R.; Saini, V.; Varshney, N.; Bagde, P. H.; Samanta, S.; Kar, P.; Jha, H. C. Computational Insights into VacA Toxin Inhibition: Harnessing FDA-Approved Antibiotics against *Helicobacter Pylori*. *Journal of Biomolecular Structure and Dynamics* **2024**, *42* (24), 13725–13737. https://doi.org/10.1080/07391102.2023.2278080.

3. Kashyap, D.; **Koirala, S.**; Saini, V.; Bagde, P. H.; Samanta, S.; Kar, P.; Jha, H. C. Prediction of Rab5B Inhibitors through Integrative in Silico Techniques. *Mol Divers* **2024**, *28* (4), 2547–2562. https://doi.org/10.1007/s11030-023-10693-9.

4. Mahapatra, S.; Jonniya, N. A.; **Koirala, S.**; Ursal, K. D.; Kar, P. The FGF/FGFR Signalling Mediated Anti-Cancer Drug Resistance and Therapeutic Intervention. *Journal of Biomolecular Structure and Dynamics* **2023**, *41* (22), 13509–13533. https://doi.org/10.1080/07391102.2023.2191721.

5. Phukan, A.; Baruah, D.; **Koirala, S.**; Kar, P.; Boni, T.; Tamuly, C. 'Bhim Kol (*Musa Balbisiana*)' Wine: Chemical Profiling and Antidiabetic Properties with MD Simulation Insights. *Chemistry & Biodiversity* **2025**, *22* (1), e202401855. https://doi.org/10.1002/cbdv.202401855.

6. Mahapatra, S.; Jonniya, N. A.; **Koirala, S.**; Kar, P. Molecular Dynamics Simulations Reveal Phosphorylation-Induced Conformational Dynamics of the Fibroblast Growth Factor Receptor 1 Kinase. *Journal of Biomolecular Structure and Dynamics* **2024**, *42* (6), 2929–2941. https://doi.org/10.1080/07391102.2023.2209189.

7. Samanta, S.; Sk, M. F.; **Koirala, S.**; Kar, P. Dynamic Interplay of Loop Motions Governs the Molecular Level Regulatory Dynamics in Spleen Tyrosine Kinase: Insights from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2024**, *128* (43), 10565–10580. https://doi.org/10.1021/acs.jpcb.4c03217.

8. Sonia, H.; Boni, T.; Chelleng, N.; Afzal, N. U.; Manna, P.; **Koirala, S.**; Kar, P.; Tamuly, C. Unraveling the Bioactive Profile of Finger Millet Wine: In-Vitro Hypoglycemic, Anti-Urease Activities, and In-Silico Molecular Docking Approach. *South African Journal of Botany* **2025**, *180*, 820–836. https://doi.org/10.1016/j.sajb.2025.03.058.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Appendix A

Appendix B

Appendix C

Appendix D

Appendix E

LIST OF TABLES

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Appendix A

Appendix B

Appendix C

Appendix D

ACRONYMS

EPKs	Eukaryotic kinases
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
TPKs	Tyrosine protein kinases
STPKs	Serine/Threonine protein kinases
MAK	Mitogen-activated protein kinases
DLK	Dual Leucine Zipper Kinase
MKK4	Mitogen-activated protein kinase kinase 4
MKK7	Mitogen-activated protein kinase kinase 7
JNK	c-Jun N-terminal Kinase
PKA	Protein kinase A
P-loop	Phosphate binding loop
G-loop	Glycine-rich loop
A-loop	Activation loop
C-Spine	Catalytic spine
R-Spine	Regulatory spine
FDA	Food and Drug Administration
FF	Forcefield
DFT	Density Functional Theory
PME	Particle mesh Ewald
AMBER	Assisted Model Building with Energy Refinement

CHARMM	Chemistry at HARward Macromolecular Mechanics
OPLS	Optimized Potentials for Liquid Simulations
GROMOS	GROningen MOlecular Simulations
GAFF	Generalized Amber Force Field
SANDER	Simulated Annealing with NMR-Derived Energy Restraints
MD	Molecular Dynamics Simulations
PBC	Periodic Boundary Condition
TIP3P	Transferable Intermolecular Potential with three-Point charges
SASA	Solvent-Accessible-Surface-Area
PB	Poisson-Boltzmann
GB	Generalized Born
aMD	Accelerated molecular dynamic simulations
CV	Collective Variables
RMSD	Root-mean-square-deviations
RMSF	Root-mean-square-fluctuations
DCCM	Dynamic Correlation matrix
PCA	Principal component analysis
dPCA	Dihedral Principal Component Analysis
FES	Free energy surface
FEP	Free Energy Perturbation
TI	Thermodynamic Integration
MM-PBSA	Molecular Mechanics Poisson-Boltzmann/Generalized-Born Surface Area

PDB	Protein Data Bank
PMF	Potential of Mean Force
PSN	Protein Structure Network

NOMENCLATURE

ns	nanosecond
μs	microsecond
ps	picosecond
NVT	number, volume, temperature
NPT	number, pressure, temperature
Cα	carbon alpha atom
Cβ	carbon beta atom
PC	principal components

Chapter 1

Chapter 1

Introduction and Literature Review

1.1. Overview of Protein Kinases

Eukaryotic protein kinases (EPK) were discovered in 1943, and their crucial role in phosphorylation was identified in 1956 [1,2]. Understanding protein kinases has provided deeper insights into diverse cellular processes such as signaling cascades [3,4], transcription of DNA [5], DNA replication [6], intracellular organelle transport [7,8], smooth muscle contraction [9,10], cellular differentiation [11], different neurological regulations and transmission [12,13], etc. Human protein kinases play a crucial role, comprising approximately 2% to 3% of all genes in the eukaryotic genome. The human genome includes 518 kinases [14] and 106 pseudogenes, collectively making up about 2% of all human genes. Because of overexpression and genetic alterations such as mutations and translocations, dysfunctional protein kinase activity contributes to the development of autoimmune, inflammatory, cardiovascular, neurological diseases, and various malignancies [15,16]. Consequently, targeting this enzyme family has become a major focus in the development of modern therapeutics in the 21st century [17,18]. One-fourth of global drug discovery programs focus on protein kinases. Therefore, detailed insights into their mechanisms and structures are crucial for designing more effective and selective therapeutics [14]. Atomic-level structural data are informative for understanding biomolecular mechanisms [19]. Yet, because atoms are always in motion, molecular function and interactions are also governed by these dynamic properties. Rather than relying only on static structures, one aims to visualize biomolecules dynamically, interact with them at the atomic level, and observe their subsequent behaviour. Unfortunately, visualizing and manipulating individual atoms in this manner is extremely difficult. As an alternative, atomic-level simulations provide a powerful tool to study the dynamics of biomolecules [19]. Thus, this thesis primarily employs computational strategies, including conventional molecular dynamics (MD) simulations [20] and enhanced sampling methods, to

analyse conformational variability and inhibitory mechanisms of various kinases of DLK-JNK signalling pathways.

1.2. Kinase Function

Protein kinases have the ability to alter protein function through a wide range of mechanisms. Post-translational modifications [21] like phosphorylation [22], glycosylation, methylation, ubiquitination, acylation, nitrosylation, lipidation, and proteolysis modulate enzyme activity and regulate biological events such as gene expression, which are crucial to both normal physiology and disease mechanisms [23,24]. Kinases catalyse phosphorylation by transferring the γ-phosphate group from ATP to hydroxyl groups on substrates such as sugars, lipids, and amino acids, with ADP being released as a byproduct (**Figure 1.1**). Phosphorylation is a reversible process regulated by phosphatase enzymes, which remove phosphate groups from the substrate.

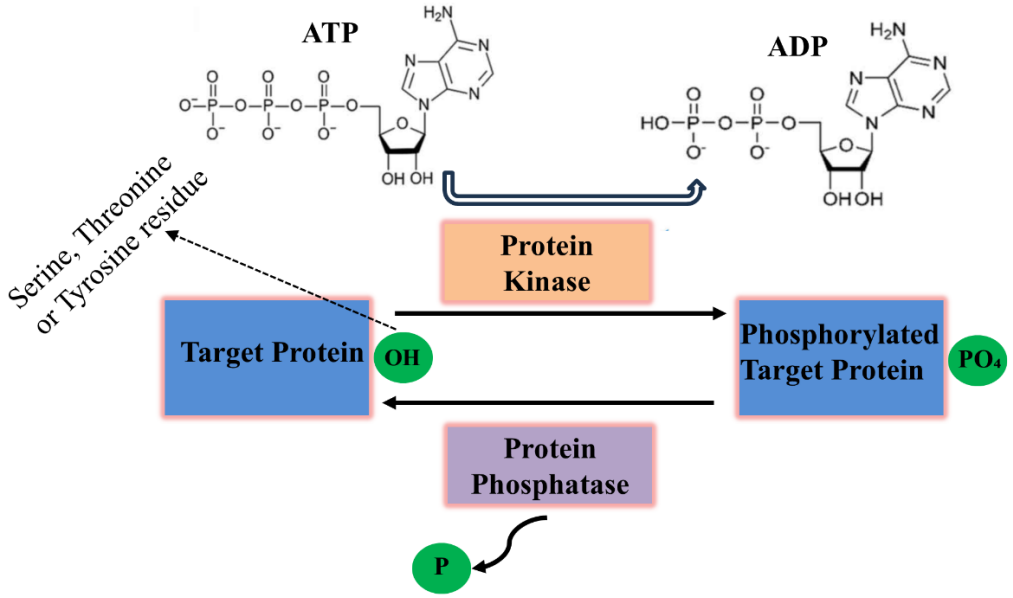


Figure 1.1: The schematic representation of the function of a protein kinase.

Phosphorylation plays a significant role in many cellular processes [21]. Eukaryotic protein kinases (ePKs) are classified according to the amino acid residue they phosphorylate, such as tyrosine (TPKs; tyrosine-protein kinases), serine/threonine

(STPKs; serine/threonine protein kinases), or both (dual-specificity kinases) [23,25]. A total of 385 serine/threonine kinases, 90 tyrosine kinases, and 43 tyrosine-like kinases have been identified. Of the protein-tyrosine kinases, 58 are classified as transmembrane receptor kinases, while 32 function as intracellular non-receptor kinases. Some kinases, like MEK1 and MEK2, structurally resemble serine/threonine kinases but possess the ability to phosphorylate both threonine and tyrosine residues, classifying them as dual-specificity kinases [12]. The activity of protein kinases is often modulated by their own phosphorylation. Dysregulated phosphorylation of cellular proteins has been linked to numerous diseases, including cancer, inflammation-related conditions, cardiovascular disorders, diabetes, hypertension, and abnormalities in cell growth and survival [21,22,26].

1.3. Architecture of the Kinase Domain

Protein kinase A (PKA), a cyclic-AMP dependent kinase, was the first protein kinase whose three-dimensional structure was resolved using crystallography [27]. Being a well-characterized serine/threonine kinase, it serves as a model for understanding the broader kinase family [28]. The catalytic domain is composed of ~350 residues and is organized into a bilobular structure [29,30]. The conserved catalytic core of a kinase domain is generally composed of two lobes: an N-terminal lobe and a C-terminal lobe. The kinase domain and its critical regions are highlighted in **Figure 1.2**. The small N-lobe (residues 40-120) consists of an antiparallel β sheet (β1-β5) and a prominent α-helix called the αC-helix. The large C-lobe (residues 128-300) is mainly composed of α-helices (αD-αI). The C-lobe contains short conserved β-strands (β6-β9) containing catalytic residues. Followed by the αE-helix is the β6-strand, the catalytic loop, the β7 and β8 strands, and the activation segment with the β9-strand. The activation segment in the majority of protein kinases begins with the DFG (Asp-Phe-Gly) motif and terminates at the APE (Ala-Pro-Glu) sequence. However, in EGFR, the activation segment starts with the DFG motif but terminates with ALE (Ala-Leu-Glu) instead of the typical APE. Within the kinase structure, ATP is located in the cleft between the N- and C-lobes, also interacting with the loop linking the β1 and β2 strands. The connecting loop, which is also called the P-loop, contains a highly conserved glycine-

rich motif designated as GXGXφG, where the φ residue may be either tyrosine or phenylalanine, and is alternatively known as the glycine-rich loop or G-loop. The presence of glycine residues in this loop enables it to interact and coordinate with the phosphate groups of ATP through backbone-mediated interactions.

Following the G-loop, the β2 strand includes a conserved valine residue that establishes a hydrophobic interaction with the adenine ring of ATP. In the absence of ATP, the glycine within the P-loop enhances its conformational flexibility, which supports small-molecule inhibitor binding. Two hydrophobic structural motifs spanning the N- and C-lobes are the Catalytic spine (C-spine) and the Regulatory spine (R-spine). The Regulatory spine (R-spine) is composed of four key residues: two from the C-lobe, Y164 from the Y/HRD motif in the catalytic loop and F185 from the DFG motif within the activation loop (A-loop), and two from the N-lobe, L95 from the αC-helix and L106 from the β4-strand. The R-spine is stabilized further by hydrophobic amino acids such as V104 and the gatekeeper residue M120 from the αC–β4 loop region, along with M118 located in the β5-strand [31,32]. The functional regulation of EPKs involves transitions between inactive and active conformations, marked by disassembly and assembly of the R-spine [33]. During the catalytic cycle, the functional conformation of EPKs shifts from an open to a closed state to enable proper enzymatic activity. The binding of ATP induces a conformational shift in the kinase core from an open state to an intermediate one, mainly due to contacts between the adenine moiety of ATP and the hinge region and C-spine [34,35]. As a result, the fusion of the N-terminal and C-terminal lobes enables the kinase to move from an intermediate conformation to a fully closed state during substrate binding [36].

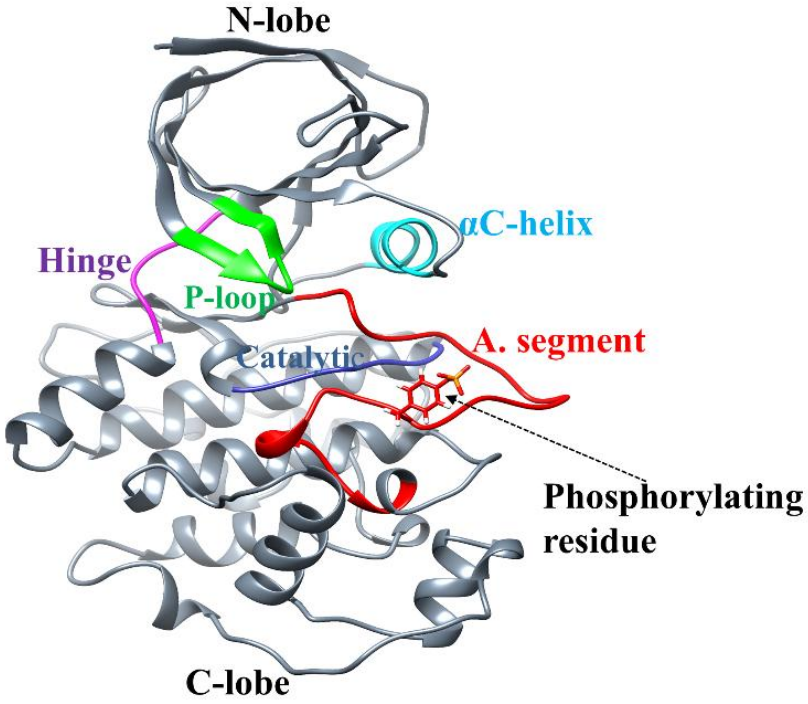


Figure 1.2: The structural features of a protein kinase, depicting different important regions like α-helix, activation segment, catalytic loop, hinge, P-loop, and phosphorylating residue.

The negative charge in the active site cleft of the C-lobe in EPKs is countered by Mg^{2+}. These ions bind the DFG-motif (D184 in PKA) of the activation loop, and an asparagine residue (N171) from the catalytic loop contributes to aligning the γ-phosphate of ATP for phosphorylation. The α and β phosphates of ATP form interactions with the highly conserved catalytic lysine (K72) located in the β3-strand, and this lysine also forms a hallmark salt-bridge with a conserved glutamate (E91) in the αC-helix [37,38]. Although both serine/threonine and tyrosine kinases catalyze phosphorylation, the catalytic loop structure differs between the two classes. Primarily, the protein kinase AGC family contains the YRDLKPEN motif, while the HRDLKPQN sequence is typical of serine/threonine kinases, and HRDLAARN and HRDLRAAN are found in receptor and non-receptor tyrosine kinases, respectively. Earlier studies indicated that K72 located in the β3-strand is essential for ATP binding, while E91 in the αC-helix plays a stabilizing role for K72. According to mutational

and biochemical investigations, a K72 mutation leads to a complete loss of catalytic function without altering ATP binding, indicating that the specific role of K72 is not yet clearly defined [39]. Furthermore, the closed conformation of the active site is attained through interactions between the G-loop and ATP, the substrates, and the C-lobe. The transfer of the phosphate group happens when the substrate (either serine, threonine, or tyrosine) is properly positioned through the action of D166, the aspartate residue within the HRD-motif. Finally, the phosphorylated substrates are released from the active site along with ADP as a byproduct. Again, EPKs return to their open conformation, allowing the binding of a new ATP molecule, and the phosphorylation cycle continues [40].

1.4. The Functional States of Kinase: Active and Inactive Forms

Eukaryotic kinases are flexible and structurally dynamic proteins, playing essential roles in regulating biological signaling pathways [33]. Protein kinases function as molecular switches that alternate between active "on" states and inactive "off" states [41,42]. The kinase is highly active in the 'on' conformation, but shows minimal activity in its 'off' state. All protein kinases catalyse substrate phosphorylation by transferring the γ-phosphate group from ATP, resulting in a structurally conserved active conformation across almost all kinases [43,44]. In contrast, the inactive or 'off' state of kinases is not structurally conserved across all types. As a result, both different kinase classes and even the same kinase can exhibit varied inactive conformations. Active kinases share a conserved tertiary structure, whereas Huse and Kuriyan reported that inactive kinases display structurally diverse conformations [33]. The conformational dynamics of regions like the activation loop, G-loop, and αC-helix play a key role in defining kinase activity states. The protein undergoes major structural rearrangements when it becomes phosphorylated, binds ATP, or interacts with inhibitors. Hence, a detailed comprehension of such conformational rearrangements could facilitate the rational design of novel therapeutic strategies.

The activation segment plays a central role in regulating catalytic activity and substrate binding [32]. Nolan et al. reported that autophosphorylation of residues within the activation segment is critical for the activation of various kinases [45]. However, some

kinases, such as EGFR, do not require such phosphorylation for their activation [46]. A loop within the activation segment plays a central role in controlling conformational transitions in numerous protein kinases. Tournier et al. reported that phosphorylation at Thr221 alone, despite the presence of a dual site (Thr221 and Tyr223), is sufficient to activate JNK [47]. JNK is expressed in neuronal cells and is a well-studied therapeutic target for small-molecule inhibitors in treating neurological disorders. pThr221 interacts with basic A-loop residues, Arg227 and Arg188, stabilizing the protein in its active, open conformation. According to Xie et al. [48] and Mishra et al. [49], basic residues assist in mediating the conformational rearrangements of the A-loop in JNK3. The conformational state of phosphorylated, unphosphorylated, or ATP-bound kinases significantly influences their structural and functional behaviour. It regulates the different open/closed states of a kinase. For example, substantial structural shifts occur during the open-to-closed transition in unphosphorylated JNK3 (uJNK3) and kinases such as CDK5 [50]. A similar mechanism has been reported in Abl kinase 1 and Src kinase, where the αC-helix, DFG motif, and A-loop coordinate to enable open-to-closed transitions [51].

Furthermore, Shaw and colleagues investigated the DFG-in to DFG-out transition in Abl kinase using long-timescale MD simulations (2.2 μs) [52]. Moreover, Meng and Roux found that phosphorylation at Tyr416 within the activation loop of c-Src kinase enhances the stability of multiple structural components, thereby locking the enzyme in a catalytically active conformation [53]. According to Musavizadeh et al. [54], phosphorylation at a single site (Thr288) within Aurora-A kinase promotes stabilization of the A-loop in an open conformation. In contrast, a second phosphorylation at Thr287 reduces this stability and enhances loop flexibility. These results indicate that the conformational behaviour of Aurora-A changes notably between its singly and doubly phosphorylated forms and is distinct from the unphosphorylated kinase. Similarly, studies have shown that the αC-helix can undergo structural shifts between αC-in and αC-out conformations, which is a key regulator in controlling the transition between active and inactive forms of various kinases [55]. Therefore, the fundamental difference between the active and inactive state of the

kinase is defined by the motion between the lobes, positioning of the αC-helix, and the configuration of the activation loop.

1.5. Kinase Inhibitors

In recent decades, studies focusing on kinase-directed drug development have expanded significantly due to their therapeutic potential in treating a wide range of diseases [56,57]. By 2024, the United States Food and Drug Administration (FDA) had approved 85 compounds targeting kinases. Almost all kinase-targeting drugs approved to date are ATP-competitive inhibitors that target the ATP pocket of the kinase. The high similarity of ATP-binding sites across the kinome poses a significant challenge in achieving selective inhibition of the desired kinase [58,59]. However, the FDA-approved drug Imatinib, against the protein kinase for treating chronic myelogenous leukemia (CML), has eliminated this notion. Later structural studies examined distinct areas of the ATP-binding site and its adjacent regions to achieve specificity in kinase inhibitor development. The list of FDA-approved protein kinase inhibitors specific to their respective targets is shown in **Table 1.1**.

Table 1.1. Selected FDA-approved protein kinase inhibitors against various kinases. Adapted from www.brimr.org/PKI/PKIs.htm.

Drug Target	Protein substrate	Receptor	Drug
ALK	Tyrosine	Yes	Crizotinib, ceritinib, Alectinib, Brigatinib, Entrectinib, Lorlatinib
BCR-Abl	Tyrosine	No	Bosutinib, dasatinib, imatinib, nilotinib, ponatinib, Asciminib,
EGFR family	Tyrosine	Yes	Gefitinib, erlotinib, lapatinib,vandetanib, afatinib, Brigatinib, Lazertinib, Dacomitinib

PDGFRα/β	Tyrosine	Yes	Axitinib, gefitinib, imatinib, Lenvatinib, nintedanib, pazopanib, regorafenib, sorafenib, sunitinib, Avapritinib, Ripretinib
VEGFR family	Tyrosine	Yes	Axitinib, lenvatinib, nintedanib, pazopanib, regorafenib, sorafenib, sunitinib, Fruquintinib, Tivozanib, Midostaurin
c-Met	Tyrosine	Yes	Crizotinib, cabozantinib, Tepotinib, Capmatinib
RET	Tyrosine	Yes	Vandetanib, Alectinib, Pralsetinib, Selpercatinib
BTK	Tyrosine	No	Ruxolitinib, tofacitinib, Acalabrutinib, Pirtobrutinib, Zanubrutinib
JAK family	Tyrosine	No	Ruxolitinib, tofacitinib, Abrocitinib, Baricitinib, Momelotinib, Ritlecitinib, Pacritinib, Fedratinib, Upadacitinib
Src family	Tyrosine	No	Bosutinib, dasatinib, ponatinib, vandetanib,
CDK family	Serine/threonine	No	Palbociclib, sorafenib, Abemaciclib, Trilaciclib, Ribociclib
B-Raf	Serine/threonine	No	Vemurafenib, dabrafenib, Tovorafenib, Encorafenib
MEK1/2	Dual specificity	No	Trametinib, Avutometinib, Binimetinib, Selumetinib,

ROCK family	Serine/threonine	No	Belumosudil, Netarsudil
ROS1	Tyrosine	Yes	Brigatinib, Repotrectinib, Entrectinib,
FLT3	Tyrosine	Yes	Quizatinib, Gilteritinib, Midostaurin
AKT family	Serine/threonine	No	Capivasertib
FGFR family	Tyrosine	Yes	Pemigatinib, Erdafitinib

Based on the binding mechanisms, kinase inhibitors are categorized into four types: Type I, Type II, Type III, and Type IV [60–62].

1.5.1 Type I Inhibitors

Type I kinase inhibitors bind within the conserved ATP-binding pocket and are classified as ATP-competitive inhibitors. These inhibitors bind within the ATP-binding site and further extend into adjacent regions, including the front pocket, hydrophobic pocket, DFG motif, and P-loop. In the active kinase state, Type I inhibitors bind the "DFG-in" conformation, where the aspartic acid in the DFG motif aligns toward the hinge and participates in hydrogen bonding. Most kinase inhibitors approved by the FDA belong to this particular category. Type I kinase inhibitors include baricitinib, EGFR-targeting drugs like gefitinib, afatinib, erlotinib, and lapatinib, as well as crizotinib and ceritinib (ALK), and others such as ruxolitinib, palbociclib, and dasatinib.

1.5.2 Type II Inhibitors

The structural plasticity of kinases has been elucidated through analysis of their interactions with inhibitors of distinct binding modes. In 2001, the discovery of Imatinib confirmed the structural binding mode attributed to Type II kinase inhibitors. Structural analysis of the Abl-imatinib complex revealed that the inhibitor occupies the ATP-binding site and extends into the adjacent allosteric pocket. Additionally, Type II inhibitors bind to the "DFG-out" conformation, where the DFG motif's phenylalanine flips toward the hinge, representing an inactive state of the kinase [62]. Type II

inhibitors exhibit enhanced selectivity by targeting the distinctive features of inactive kinase states, unlike inhibitors that bind the standard active conformation. Examples of Type II inhibitors include imatinib (Abl/Kit), ponatinib (Kit/B-Raf), cabozantinib, nilotinib (Abl), regorafenib, sorafenib (CDK8), and axitinib (VEGFR).

1.5.3 Type III Inhibitors

Type III kinase inhibitors bind at an allosteric pocket near the ATP-binding pocket, where ATP continues to occupy its ATP-binding pocket. They are considered ATP non-competitive inhibitors since they do not prevent ATP from binding to the kinase. Cobimetinib, which targets MEK, is a well-known example of a Type III inhibitor.

1.5.4 Type IV Inhibitors

These Type IV inhibitors target allosteric sites that are structurally distant from the ATP-binding pocket, often located on the C-lobe, N-lobe, or surface of the kinase domain. The inhibitor maintains and stabilizes the inactive state of the enzyme. Examples include the Abl-GNF complex, where the allosteric pocket is located on the C-lobe, the AKT-MK-2206 complex at the kinase-PH domain interface, and the CDK2-2AN complex on the N-lobe.

In general, the majority of kinase-focused drug discovery programs have focused on developing inhibitors that specifically target the ATP-binding pocket. However, the high degree of conservation in the ATP-binding site, particularly among kinase isoforms and the existence of flexible segments such as the nucleotide-binding loop, activation loop, and catalytic loop, which often undergo ligand-induced conformational shifts, make it challenging to selectively target the ATP-binding site of kinase isoforms. Moreover, due to their stringent active function, the conformation of isoforms' active states is more or less similar. In contrast, their inactive conformations exhibit significant differences. Thus, binding to a particular conformational state of kinases can help clarify ligand-induced rearrangements in conformationally adaptable regions. This strategy could enable the rational development of kinase inhibitors with improved affinity and selectivity by targeting novel non-ATP competitive binding sites in isoforms. For example, Shukla et al. [63] identified intermediate states along the Src

kinase activation pathway and anticipated potential sites for novel allosteric inhibitor binding. Furthermore, Knight et al. [64] proposed that targeting allosteric sites could facilitate selective kinase inhibition and enhance understanding of their roles within dynamic cellular signaling pathways. In addition, Yan et al. [65] suggested that the inhibitory mechanism of allosteric inhibitors may act via either conformational selection or induced-fit binding models. They observed a DFG-flip transition upon tivantinib binding. The mechanistic detail offered by molecular dynamics simulations can support experimental studies in elucidating the structure and function of complex biomolecules. Thus, it can offer valuable insights for the design of novel allosteric inhibitors. Altogether, the detailed atomistic insights provided by MD simulations can suggest essential pharmacophoric features for designing bioactive allosteric kinase inhibitors.

1.6. The DLK-JNK Signalling Pathway

The dual leucine zipper kinase (DLK) and c-Jun N-terminal kinase (JNK) (DLK-JNK) signalling pathway is a crucial mitogen-activated protein kinase (MAPK) cascade that governs neuronal responses to diverse cellular stresses [66]. As an evolutionarily conserved mechanism, this pathway integrates external and internal stress signals to modulate transcriptional programs that influence cell survival and apoptosis. Unlike classical MAPK pathways that primarily govern mitogenic signaling, the DLK-JNK pathway is uniquely specialized for regulating axonal regeneration, apoptosis, and neural development, highlighting its essential role in the nervous system. This pathway is structured around a conserved three-tiered MAPK signalling cascade. DLK, acting as the initiating MAP3K under stress conditions, phosphorylates MKK4 and MKK7, the MAP2Ks responsible for activating JNK via dual phosphorylation on conserved residues [67]. Activated JNK phosphorylates downstream targets, especially c-Jun, which governs the transcription of genes involved in neuronal fate decisions ranging from survival to apoptosis. The DLK-JNK pathway involved in triggering neuronal cell death is visualized in **Figure 1.3**. A complex regulatory network governs the DLK-JNK pathway, integrating post-translational modifications, scaffold proteins, subcellular signalling, and feedback control [68]. JNK-interacting proteins (JIPs),

functioning as scaffold molecules, play an indispensable role in DLK-JNK signalling, forming tightly regulated complexes that promote efficient and specific transmission.

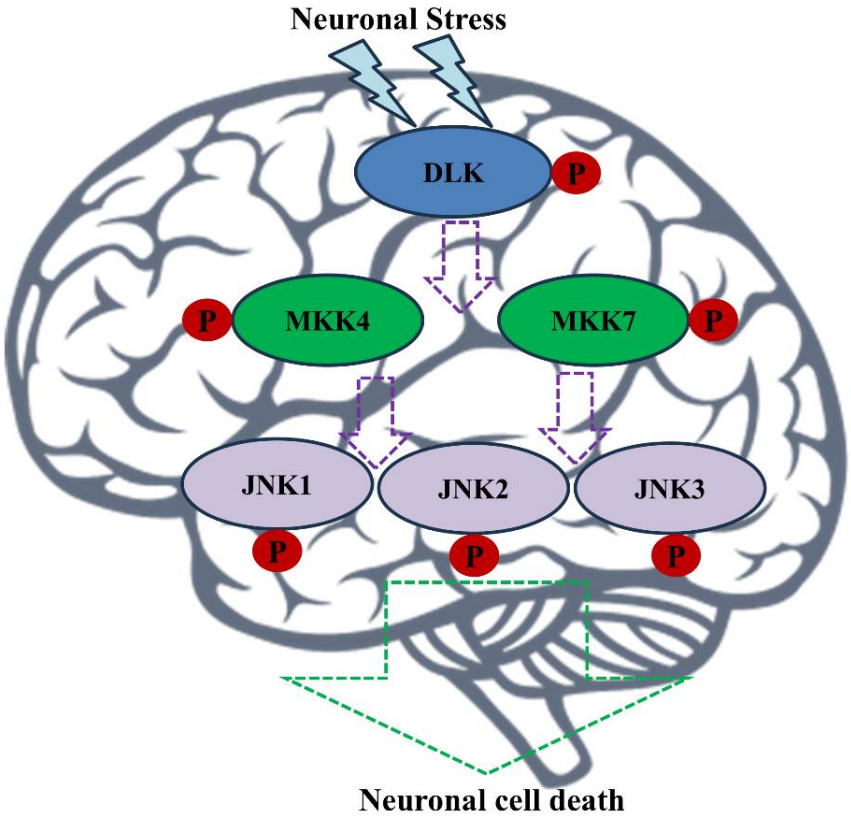


Figure 1.3: A schematic representation of the stress-induced signaling cascade is shown, where neuronal stress initiates the activation of DLK. Once activated, DLK phosphorylates downstream kinases MKK4 and MKK7, which then activate JNKs through phosphorylation. Persistent JNK activation leads to neuronal cell death.

The DLK-JNK pathway serves functionally distinct and sometimes opposite roles within physiological systems. During neural development, the DLK-JNK pathway plays a vital role in regulating axon guidance, dendritic branching, and developmental apoptosis that shapes the architecture of the nervous system [69]. In mature neurons, pathway activation following axonal injury can promote regenerative responses in peripheral neurons or trigger apoptosis in central neurons, depending on the molecular milieu and cellular context. This contrasting behavior reflects the remarkable flexibility of the pathway and its crucial function in regulating neuronal homeostasis

during dynamic cellular states [70]. Given its highly conserved and crucial function, dysregulation of the DLK-JNK pathway is involved in the onset and progression of multiple neurological conditions. Due to its involvement in neuronal development, injury response, and disease progression, the DLK-JNK pathway has emerged as a major focus for therapeutic intervention [71].

1.6.1. The Dual Leucine Zipper Kinase (DLK)

DLK, also designated as MAP3K12, is a serine/threonine kinase with evolutionarily conserved function, playing important roles in neural development, stress adaptation, and neuropathological processes. It was first identified by Holzman et al. in 1994 as a member of the mixed-lineage kinase (MLK) subfamily [72]. Since its discovery, DLK has emerged as a central regulator of axonal injury responses, coordinating the balance between axon regeneration and degeneration, and is considered a promising therapeutic target in various neurological and metabolic disorders [73]. Functionally, DLK becomes activated in response to cellular stress and transmits signals via the downstream kinases MKK4/MKK7, leading to JNK pathway activation and the regulation of diverse transcriptional programs [74].

Structurally, DLK is composed of 859 amino acids in humans and 888 in mice, and features four primary functional domains, an N-terminal glycine-proline-proline-rich (GPP) region, a catalytic serine/threonine kinase domain (residues 127-375 in humans), a central dual leucine zipper motif, and a C-terminal glycine-serine-proline-rich (GSP) region [66,75,76]. The two leucine zipper domains, connected by a 25-residue linker, play a key role in promoting DLK homodimerization, which is necessary for its kinase activity and signaling through stress-activated protein kinases (SAPKs) [75]. Within the catalytic domain, several highly conserved motifs are critical for ATP coordination and catalysis, including Lys185 (Lys152 in humans), which plays a central role in ATP stabilization at the active site. DLK undergoes various post-translational modifications that regulate its enzymatic activity and subcellular localization. Palmitoylation at Cys127 is essential for targeting DLK to membrane vesicles and enabling interaction with scaffold proteins such as JIPs [74]. Additionally, phosphorylation at key residues, including Ser272 and Ser302 within the kinase

domain and Ser584 and Thr659 in the C-terminal region, modulates stability and kinase activity. Other regulatory phosphorylation sites, such as Thr43 and Ser533, further contribute to the dynamic control of DLK signaling [77,78].

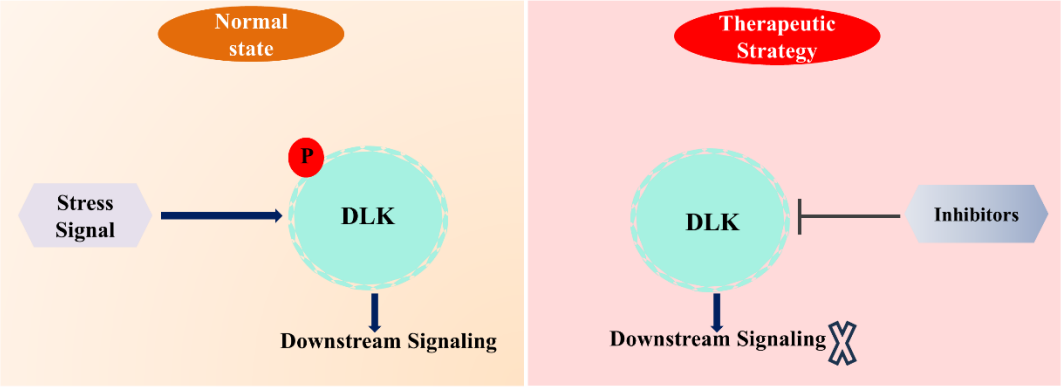


Figure 1.4: The schematic outlines of DLK signaling pathway and highlights a strategy to therapeutically inhibit DLK, thereby preventing activation of downstream effectors.

Dysregulation of DLK has been implicated in numerous neurodegenerative disorders, including Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis, where its upregulation drives persistent JNK signaling, resulting in synaptic degradation and neuronal cell death [79]. DLK knockout or its pharmacological inhibition reduces neurodegenerative effects and enhances neuronal survival, supporting its pathogenic role in disease progression. Although DLK kinase is emerging as a key therapeutic target, the lack of FDA-approved drugs and limited data present a major research gap. This provides an exciting opportunity to characterize the kinase and identify new inhibitors. **Figure 1.4** presents a therapeutic approach focused on targeting DLK to interrupt downstream signaling cascades.

1.6.2. The Mitogen-Activated Protein Kinase Kinase 4 (MKK4)

Mitogen-activated protein kinase kinase 4 (MKK4) was initially discovered via a PCR-based technique in a cDNA library developed from *Xenopus laevis* embryonic tissue and given the name XMEK2 [80]. Located on chromosome 17 in humans (and on chromosome 11 in mice, 10 in rats, and 3R in *Drosophila*), the **MKK4** gene encodes a protein composed of 399 amino acids, sharing 94% sequence identity with the

homologous proteins in both mouse and rat. MKK4, a serine/threonine-specific kinase, possesses two essential phosphorylation sites within its kinase domain, specifically situated in the S-I-A-K-T sequence motif [81]. A broad range of MKKKs, including ASK, MEKK, MLK, DLK, and TAK, are known to activate MKK4 [82]. These activations are facilitated by specific interactions that occur through the domain for versatile docking (DVD) situated at the C-terminal region of the MKK4 protein. Binding to upstream kinases through the domain for versatile docking (DVD) is essential for the phosphorylation of MKKs and plays a critical role in determining the specificity of MKKK-MKK interactions [83]. MKK4 further binds with scaffold proteins through the DVD domain. Under stressful cellular environments, it forms complexes with scaffolds like JNK-interacting protein 3 (JIP-3) or plenty of SH3 (POSH). The D-domain-type docking site located in the N-terminal region of MKK4 is responsible for mediating interactions with its downstream substrates, JNKs and p38 [84]. In vitro experiments demonstrate that MKK4 preferentially targets the tyrosine residue of JNKs for phosphorylation [85], which is critical for full JNK activation. A limited number of studies have reported that MKK4 alone is capable of inducing JNK activation [86]. However, alternative studies revealed that similar stimuli predominantly activate MKK7, with several reports suggesting that JNK phosphorylation relies on MKK7 function [47], supporting that MKK7 plays a central role in JNK activation.

MKK4 contributes significantly to both normal biological functions and disease progression. During early embryonic development in mice, mkk4 mRNA is primarily detected in the central nervous system, but by embryonic (day 12 mice), its expression becomes widespread, with higher levels in the embryonic liver [87]. Within the nervous system, loss of MKK4 function results in cerebellar malformations, impaired cortical neuron migration, neurological dysfunctions, and early death following birth [88]. Targeted deletion of MKK4 in cardiac tissue inhibits the development of pathological hypertrophy, consistent with prior studies showing that a kinase-inactive form of MKK4 suppresses JNK pathway activation [89]. Given MKK4's crucial role in embryonic development, neural function, cardiac remodelling, and immune

regulation, and the absence of any FDA-approved drugs targeting it, MKK4 represents a promising yet underexplored candidate for therapeutic studies.

1.6.3. The Mitogen-Activated Protein Kinase Kinase 7 (MKK7)

The mitogen-activated protein kinase kinase 7 (MAP2K7) plays an integral role in the JNK signalling pathway, which predominantly governs stress-induced and inflammatory responses [90]. Due to its involvement in multiple pathological conditions, MKK7 is considered a promising target for developing novel drugs to treat diseases like arthritis, liver carcinoma, neurological disease, and cardiac hypertrophy [91,92]. In MKK7, the serine and threonine residues within the S-K-A-K-T motif of the kinase domain are phosphorylated by multiple MAP3Ks, which bind to MKK7 through its C-terminal DVD domain. Similar to MKK4, the DVD domain in MKK7 plays a role in promoting its phosphorylation and the specificity of interactions between MKKKs and MKK7 [83]. The three isoforms of MKK7 generated by alternative splicing vary in their capacity to activate JNK, where MKK7α exhibits significantly lower activity than MKK7γ and MKK7β [93]. The difference in activity results from the N-terminal extension found in MKK7α and MKK7β, which contain three distinct JNK docking sites that facilitate stronger binding. The relatively low binding affinity of individual D domains, along with evidence from substitution mutations, suggests that the three D sites in MKK7α and MKK7β act cooperatively to achieve full binding affinity [94].

MKK7 functions as a central signalling hub within stress-activated pathways, coordinating cellular responses that underlie both physiological regulation and the development of pathological conditions [95]. By activating JNK, it plays a key role in controlling apoptosis, especially when cells are exposed to DNA damage, oxidative stress, or cytokines [96]. For example, when MKK7 activates JNK, it leads to the phosphorylation of pro-apoptotic proteins such as Bim and Bax, which helps open the mitochondrial outer membrane and triggers caspase activation [97]. Under certain circumstances, MKK7-JNK signaling favours cell survival through activation of transcription factors such as c-Jun and ATF2, which stimulate anti-apoptotic gene expression. This dual nature of MKK7 activity highlights its dual role in cellular

context, signalling duration, and interactions with other pathways like NF-κB and ERK [47]. In neural tissues, MKK7 supports processes like neuronal development and synaptic remodeling, but its overactivation contributes to diseases like Alzheimer's and Parkinson's by promoting JNK-driven tau phosphorylation and neuroinflammation [98]. MKK7 can be a promising target, especially in diseases like cancer, neurological diseases and inflammation. Till now, only a limited number of inhibitors have been reported that directly target MKK7, and in most cases, MKK7 is inhibited unintentionally while targeting other kinases. In a recent study, Shraga et al. used docking methods to identify compounds that bind covalently to MKK7 and reduce JNK activity in cells, with micromolar range EC50 values [99]. Therefore, more research is needed to develop highly potent and selective inhibitors to better understand MKK7 biology.

1.6.4. The c-Jun N-terminal Kinase (JNK)

The c-Jun N-terminal kinase (JNK) is a member of the mitogen-activated protein kinase (MAPK) family that regulates key cellular processes such as apoptosis, inflammation, and stress responses. However, dysregulation of MAPK signaling has been linked to the progression of several diseases, including metabolic syndromes, neurodegenerative conditions, and several types of cancer [100]. JNK, also referred to as stress-activated protein kinase (SAPK), is named after its ability to phosphorylate serine residues at the N-terminal region of the c-Jun transcription factor, leading to its activation [101,102]. The JNK family consists of three isoforms: JNK1, JNK2, and JNK3, which are encoded by the *jnk1*, *jnk2*, and *jnk3* genes, respectively [103]. The three JNK isoforms exist in two molecular weight forms, 46 kDa and 55 kDa. Structural determination of the JNK isoforms showed that they are highly conserved, with more than 90% sequence homology[48,104].

Structurally, the JNK protein is composed of three main regions: an N-terminal lobe, a C-terminal lobe, and a flexible linker that connects them. The N-lobe primarily consists of β-strands, while the C-lobe is predominantly made up of α-helices [105]. In addition, the former region is rich in glutamate-aspartate (ED) sequences, while the latter is characterized by the presence of a common docking (CD) domain [106]. The

flexible peptide region contains the catalytic center of JNK, which includes a Thr-Pro-Tyr motif that is phosphorylated by MAP2Ks. Additionally, JNK possesses a kinase activation site and a conserved ATP-binding pocket, contributing to poor selectivity for small-molecule inhibitors [105]. The crystal structures of several JNK complexes were recently resolved, including JNK1 with the MKK7 docking motif (D2 peptide) [107], JNK1-ATF2 complex [108], AMP-bound JNK2 [109], and JNK3 bound to thiophene-pyrazolourea inhibitors [110].

JNK is a serine/threonine kinase that can be activated by a variety of external stimuli, including cytokines, growth factors, reactive oxygen species (ROS), heat shock, shear stress, pathogens, and pharmacological agents [111]. Similar to other MAPK cascades, the JNK signaling pathway primarily consists of three categories: MAP3Ks (such as ASK1), MAP2Ks (including MKK4 and MKK7), and MAPKs (JNK1, JNK2, and JNK3). Upon stimulation by extracellular stress, the JNK pathway is initiated with the activation of a MAP3K, such as MEKK1. This activated MAP3K then phosphorylates the MAP2Ks, MKK4 and MKK7, which subsequently phosphorylate and activate the MAPKs JNK1, JNK2, and JNK3 [112]. Similar to other members of the MAPK family, JNK becomes activated when the Thr-Pro-Tyr tripeptide motif (T-X-Y motif) within its activation loop is phosphorylated at both threonine and tyrosine residues by MAP2Ks. For instance, JNK1 is activated upon dual phosphorylation at Thr-183 and Tyr-185 [113].

JNK signalling contributes to disease through distinct isoform-dependent mechanisms. In metabolic syndromes such as obesity and type 2 diabetes, free fatty acids and cytokines like TNF-α activate JNK1, which impairs insulin signalling by phosphorylating IRS1/2 [114,115]. JNK1 knockout mice demonstrate greater insulin sensitivity and show resistance to obesity under high-fat dietary conditions [116]. In neurodegenerative disorders, hyperactivation of JNK3 facilitates neuronal apoptosis in Alzheimer's disease by promoting Tau phosphorylation and enhancing amyloid-β toxicity [117], whereas in Parkinson's disease, it contributes to the degeneration of dopaminergic neurons [118]. In cancer, the functions of JNK isoforms are context-dependent; JNK1/2 facilitate tumor growth in glioblastoma and lung cancer [119], but

trigger apoptosis in prostate cancer [103]. In chronic inflammatory disorders such as rheumatoid arthritis, JNK signaling contributes to tissue destruction by upregulating matrix metalloproteinases [120].

Due to its critical involvement in the pathogenesis of various diseases, JNK has emerged as a highly attractive target for therapeutic intervention. Inhibitors like SP600125, CEP-1347, and AS601245 act by occupying the ATP-binding site of JNK, functioning as competitive inhibitors and reducing kinase activity [121–123]. Natural compounds like the curcumin derivative C66 have shown the ability to inhibit JNK activity, resulting in attenuation of inflammation and metabolic disruption [124]. Although the therapeutic potential of targeting JNK is well-recognized, developing inhibitors that are both isoform-specific and have minimal off-target consequences remains a significant challenge. Hence, future investigations should prioritize the design of inhibitors for individual isoforms and clarify the distinct biological functions of JNK1, JNK2, and JNK3 in the pathogenesis of various diseases. As no FDA-approved drugs have yet been developed against JNK or its individual isoforms, conducting detailed studies on isoform-specific dynamics and performing inhibitor screening may open new avenues for targeted therapeutic discovery.

1.7. Inhibitors Targeting the DLK-JNK Signaling Pathway

As a fundamental regulator of stress responses in neuronal and various other cell types, the DLK-JNK signalling pathway offers strong promise as a druggable target for several major pathological conditions including neurodegeneration, inflammation, and cancers. Starting with DLK, this MAP3K activates downstream MAP2Ks such as MKK4 and MKK7, which in turn activate JNK1, JNK2, and JNK3 through phosphorylation. Earlier studies have shown multiple compounds as potential DLK inhibitors. Genentech has introduced GNE-3511 and GNE-8505 as highly potent DLK inhibitors with desirable central nervous system (CNS) penetration and optimal drug metabolism and pharmacokinetics (DMPK) features [125]. In addition, GDC-0134 entered clinical trials for ALS treatment as DLK inhibitor but was terminated due to safety issues [126].

In recent developments, Denali Therapeutics reported DN-1289 [127], and Kang Le et al. reported IACS-52825 and IACS-8287 [128] as strong DLK inhibitors. In the years 2003 and 2004, Bayer described a compound bearing a 9-H-pyrimido[4,5-b]ind-6-ol scaffold with dual MKK4 and MKK7 inhibitory activity [129]. Compounds such as daidzein and 7,3′,4′-trihydroxyisoflavone have been suggested to inhibit MKK4 and may be effective against skin cancer [130]. Further studies identified genistein, Dehydroglyasperin C, HWY336, 3-Arylindazoles, BSJ-04-122, and PQA-11 as potent MKK4 inhibitors [131]. Research to find selective MKK7 inhibitors started when 5Z7O was shown to bind uniquely to the ATP binding site of MKK7 [132]. Through covalent docking and subsequent hit refinement and validation, Shraga et al. identified several MKK7 covalent inhibitors targeting a non-conserved cysteine residue [99]. ATP-competitive inhibitors like SP600125, CEP-1347, and AS601245 inhibit JNK and have been primarily investigated for therapeutic intervention for various diseases [115]. The ATP-non-competitive JNK inhibitors BI-78D3, BI87G3, and compound 9 show potent therapeutic results. BI-78D3, first developed by Pellecchia and coworkers, blocks JNK1 activity by interfering with its binding to a JIP-derived peptide [133].

Although multiple inhibitors have been developed against kinases within the DLK-JNK signalling pathway, including DLK, MKK4, MKK7, and JNK, the lack of FDA approval highlights a major gap from preclinical insights to clinical application. Although GNE-3511, GDC-0134, 5Z7O, and SP600125 have exhibited promising activity in vitro and in animal models, their clinical advancement has been restricted by limited selectivity, poor pharmacokinetics, off-target toxicity, and poor understanding of isoform-specific roles. This highlights the urgent need for comprehensive structural and conformational studies to better understand the dynamic behaviour of these kinases, which frequently adopt various activation states and engage in complex protein-protein interactions. Techniques such as docking, molecular dynamics simulations, and structural bioinformatics offer valuable tools for developing selective and potent inhibitors to modulate kinase activity. Therefore, it is essential that current efforts prioritize both the discovery of new inhibitors and the comprehensive understanding of activation mechanisms and regulatory dynamics of each kinase in this signaling pathway.

1.8. Scope and Aim of the Thesis

This thesis aims to characterize the conformational dynamics of proteins in the DLK-JNK signalling pathway and to investigate the molecular basis of their inhibition using advanced computational methods and algorithms. The DLK-JNK pathway is a primary signalling cascade involved in cellular stress responses, particularly in the nervous system. The DLK kinase along with MKK4, MKK7 and JNK, acts as the master regulator of this pathway. Previous experimental studies have demonstrated that dysfunction of this pathway plays an important role in the pathogenesis of various neurodegenerative diseases, inflammatory conditions, and certain cancers. Kinases involved in this pathway play a vital role and are considered attractive drug targets for modulating the dysregulated DLK-JNK pathway. During the past decade, efforts have led to the development and clinical evaluation of inhibitors targeting different molecules within this pathway for diverse therapeutic applications.

Although studies have established the involvement of the DLK-JNK signalling in various cellular pathways, its underlying mechanisms remain poorly understood. Furthermore, the limited availability of crystal structures in diverse conformations of poorly studied protein targets has limited the progress in identifying specific small-molecule inhibitors. The main challenges in structure-based drug discovery (SBDD) include the high structural conservation of kinase domains and the resulting lack of target specificity. Therefore, identifying unique conformations is essential in SBDD to improve both the affinity and specificity of inhibitors for their target protein. Therefore, a deep understanding of conformational plasticity is crucial for identifying novel structural states of kinases. However, investigating these conformational dynamics at atomic resolution through experimental techniques remains challenging.

In continuation of the above challenges in kinase drug development, we have successfully addressed them by applying advanced computational approaches and GPU-enhanced MD simulations that allow for long-timescale dynamics and break energy barriers through enhanced sampling methods using Gaussian accelerated molecular dynamics (GaMD). The dynamic and adaptable structure of protein kinases allows them to undergo major structural transformations. Despite its importance,

detailed experimental and theoretical investigations into the conformational plasticity of DLK-JNK pathway kinases are lacking, largely because of the influence of phosphorylation, interacting proteins, and ligands binding to both ATP-binding sites and allosteric sites. Broadly, the following aims have been covered in this thesis:

Aim 1: In this aim, we have focused on investigating the plant-derived active compounds from the Ayurvedic formulation Saraswatharishta (SWRT) as potential inhibitors of DLK. Ayurveda, one of the oldest traditional medicinal systems, offers a rich source of bioactive compounds, and scientific validation of its formulations is crucial for integrating traditional knowledge with modern drug discovery. Over 500 phytochemicals from the constituent plants of SWRT were screened using a systematic *in-silico* approach involving molecular docking and molecular dynamics simulations to identify promising DLK-targeting compounds. These promising compounds may serve as potential DLK inhibitors, offering a molecular basis for the neuroprotective effects of SWRT in neurological disorders.

Aim 2: In this aim, we have explored identifying novel inhibitors targeting DLK, a key mediator in neurodegenerative disease pathways. We employed a systematic virtual screening strategy using natural product (NPAtlas) and FDA-approved (MedChemExpress) libraries. This was followed by pharmacokinetic profiling, molecular dynamics simulations, and deep learning-based analogue prediction to identify promising ATP-competitive DLK inhibitors for further investigation.

Aim 3: In this aim, we explored the conformational behavior of MKK7 kinase under the influence of diverse inhibitor types using enhanced sampling methods (GaMD). Our focus was on capturing the dynamic transitions in mobile structural regions and comparing the binding modes and affinities across the inhibitor classes. This objective is crucial for understanding the mechanistic basis of MKK7 inhibition and offers a foundation for the development of more selective therapeutic agents.

Aim 4: In this aim, we focused on deciphering the phosphorylation-induced conformational dynamics of JNK isoforms, which are involved in stress response, apoptosis, and neuroinflammation. Despite their high sequence similarity, the isoforms exhibit distinct tissue distributions and functional roles, highlighting a significant

challenge for isoform-selective drug development. To address this, we employed an integrated in-silico approach combining molecular modeling and long-timescale molecular dynamics simulations to investigate the structural organization and dynamic behavior of each isoform in both phosphorylated and unphosphorylated states when bound to the ATP-competitive inhibitor SP600125. By comparing global and local conformational shifts across the three isoforms, we sought to identify the molecular determinants that influence ligand accommodation and conformational plasticity. This objective further aims to uncover the structural basis of isoform-specific regulation to guide the rational design of selective JNK inhibitors for neurological and inflammatory disorders.

Aim 5: In this aim, we focus on elucidating the conformational dynamics of JNK1, a crucial mediator in the MAPK signaling pathway that governs cellular responses to stress, inflammation, and apoptosis. The functional state of JNK1 is intricately regulated by its binding to ATP and the upstream kinase MKK7, yet the precise structural and dynamic changes induced by these interactions are not fully understood. To address this, we employed Gaussian accelerated molecular dynamics (GaMD) simulations coupled with binding free energy calculations.

The study was designed to systematically investigate and compare the conformational landscapes of JNK1 in four key states: its apo form, the ATP-bound state, the MKK7-bound complex, and the dual ATP-MKK7 bound state. By analysing global and local conformational shifts, residue-specific energy contributions, and key interaction profiles, this work seeks to uncover the molecular determinants of JNK1's dynamic behaviour. The overall objective is to provide critical mechanistic insights into how ATP and MKK7 binding modulate the interaction patterns and conformational plasticity of JNK1, thereby advancing our understanding of its regulatory mechanisms within cellular signaling.

1.9. Organization of The Thesis

Based on the objectives stated above, this thesis has been organized into eight chapters. The first chapter provides a comprehensive introduction and a review of the relevant literature, establishing the context and significance of the DLK-JNK signaling pathway

and its key kinases (DLK, MKK4, MKK7, and JNK) in neurodegeneration, stress, and inflammation. **Chapter 2** then details the theoretical foundations and protocols of the computational methodologies used throughout the investigation, including molecular docking, high-throughput virtual screening, pharmacokinetic profiling, and both conventional and Gaussian accelerated molecular dynamics (GaMD) simulations. The research work begins in **Chapter 3**, which focuses on investigating active compounds from the traditional Ayurvedic formulation Saraswatharishta (SWRT) as potential DLK inhibitors, aiming to provide a molecular rationale for its neuroprotective effects. **Chapter 4** follows with the discovery of novel inhibitors for DLK through a systematic virtual screening of natural product and FDA-approved libraries. In **Chapter 5**, we explore the intricate conformational dynamics of MKK7 kinase as influenced by various types of inhibitors. This investigation, carried out using enhanced sampling simulations, provides crucial insights into the molecular mechanisms that govern selective inhibition. Following this, **Chapter 6** discusses the phosphorylation-induced conformational dynamics of JNK isoforms. It compares the structural behavior of the isoforms in their phosphorylated and unphosphorylated states when bound to the inhibitor SP600125, aiming to reveal the basis for their unique regulation and to guide the design of selective inhibitors. **Chapter 7** provides a mechanistic investigation into JNK1's conformational dynamics across its apo, ATP-bound, MKK7-bound, and dual-bound states. Using GaMD simulations and free energy calculations, the study analyzes key conformational shifts and interactions to clarify how binding by ATP and MKK7 modulates the kinase's structural plasticity. The last chapter (**Chapter 8**) summarises the entire study and explores future research possibilities.

Chapter 2

Chapter 2

Theoretical Background and Methodology

2.1 Computational Modelling of Protein Molecules

A well-established term in science, molecular modeling refers to using computational techniques to model functions and behaviors of molecules. In most biological contexts, proteins function through conformational flexibility, resulting from their dynamic polymeric nature. In the field of structural biology, protein 3D structures are typically resolved using X-ray crystallography [134], nuclear magnetic resonance (NMR) spectroscopy [135] and cryo-electron microscopy [136]. These experimental methods mostly give static snapshots, but to truly understand how biological processes work, we need detailed information over time and space, which remains challenging to obtain through these methods [19]. In recent decades, molecular modeling has emerged as a vital technique for studying biomolecular dynamics with atomic-level accuracy across time and space. As stated by *Klaus Schulten* in his 2015 national lecture, these modeling approaches function like "computational microscopy," generating data that experimental techniques could not previously access. Key molecular modelling techniques include QM approaches (such as ab initio molecular orbital or DFT calculations) [137], classical molecular dynamics (MD) simulations, and coarse-grained MD [138]. The correlation between different temporal biological events and the appropriate molecular modeling techniques is illustrated in **Figure 2.1**.

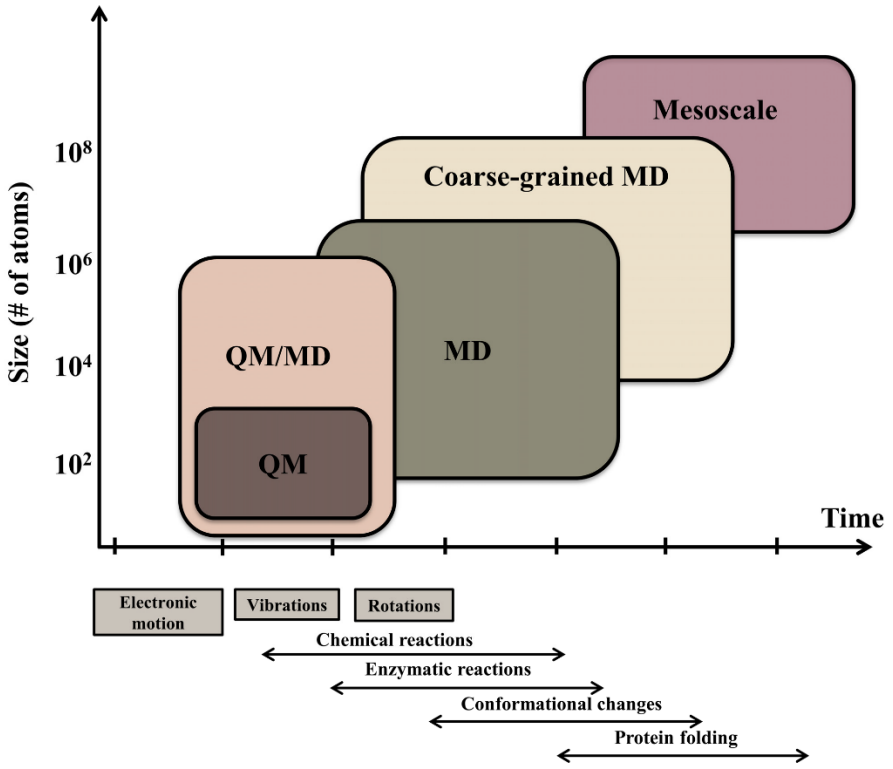


Figure 2.1: This schematic representation of the spatio-temporal resolutions of different molecular modeling techniques. Image sourced from Aminpour et al. [139]

2.1.1. Quantum Mechanical Methods

During the early 20th century, physicists developed quantum mechanics (QM), a set of rules that explained the motion of tiny particles like electrons and atomic nuclei. In chemistry, QM techniques allow detailed determination of thermodynamic quantities (heat capacity) and structural properties, including molecular geometry, dipole moment, internal energy barriers, and conformational stability. While QM delivers an exceptionally accurate depiction of atomic-level systems, extending its use to larger, more complex systems is still a significant challenge.

The time-dependent **Schrödinger** equation governs QM computations for complete systems. For a 1D system with a single particle, the future state vector is derived from its current state, as expressed by

$$-\frac{\hbar\, d\psi(x,t)}{i\, dt} = \frac{\hbar^2 d^2\psi(x,t)}{2m\, d^2x} + V(x,t)\psi(x,t) \qquad [2.1]$$

Here, ψ signifies the wave function, $\hbar$ is the Planck's constant, m is the particle mass, t is time, x denotes the particle's position, and V represents the total potential energy. In atoms with larger atomic numbers, the Hartree-Fock method [140] is often used initially to generate an approximate wave function, which is then refined using electron correlation techniques such as Møller-Plesset (MP) [141] perturbation theory or coupled-cluster theory [142]. For polyatomic molecules like biomolecules, methods such as density functional and semi-empirical calculations are often used. However, these approaches are computationally expensive and inadequate for long-timescale simulations.

2.1.2. Molecular Mechanics Method

The QM method is highly accurate in analysing electronic structures but is unsuitable for very large atomic systems, as noted in Section 2.1.1. Molecular mechanics (MM) is different from quantum mechanics (QM) because it does not use wave functions or electron density in its calculations. MM simplifies molecular modeling by excluding electron-related dynamics and focuses on predicting molecular functions through classical mechanics applied to atomic coordinates. MM balances the need for accuracy with the demands of computing power, allowing it to efficiently and reliably model large systems such as enzymes. Since MM relies on parameters suited to ground-state configurations, it lacks the ability to simulate structural changes that occur during bond-making and bond-breaking events [143]. By applying the Born-Oppenheimer approximation [144], the calculations estimate molecular energy using only the positions of the nuclei, resulting in a classical potential energy surface. MM is widely used in receptor-ligand docking approaches, particularly because it captures the vital role of conformational flexibility in complex formation.

Furthermore, the hybrid QM/MM method integrates the high accuracy of ab initio quantum mechanical calculations [145] with the computational efficiency of molecular mechanics. Thus, these methods were largely developed to study biochemical reactions and to characterize transition states, with QM/MM approaches mostly applied to reaction mechanisms rather than docking studies. Warshel and Levitt were the first to introduce the QM/MM approach [146]. The high accuracy of QM is applied to the

reactive core of the enzyme, including the catalytic site and interacting residues. In contrast, MM handles other bulk regions of the protein, which have minimal involvement in the reaction mechanism.

2.2. Force Fields

Proteins are biologically important macromolecules with diverse structures and functions. For a protein to function properly within a biological system, it must fold into a precise three-dimensional structure known as its native conformation. The native structure represents the most stable conformation of a protein, corresponding to the lowest free energy state. Protein folding occurs over time scales from milliseconds (*ms*) to microseconds *(*μs). A widely accepted computational method for studying this process is the funnel theory, which suggests that the protein's free-energy landscape is rugged, with several local minima and a global minimum representing the folded structure. The folding process of proteins can be studied through high-resolution experimental techniques like NMR spectroscopy or MD simulations.

MM is often called the force field (FF) method, as it is fundamentally based on classical mechanics using force fields. A force field is used to determine the potential energy of a molecular system by evaluating the spatial arrangement of atoms. When empirical models use well-optimized parameters, they can closely replicate the accuracy of quantum mechanical methods. FF accurately estimates a system's free energy and enables a quick and efficient exploration of protein conformations.

In molecular mechanics, total potential energy (V_{total}) is expressed as a combination of bonded interactions among covalently connected atoms and non-bonded interactions, which include long-range electrostatics and short-range van der Waals forces, as shown below:

$$V_{total} = V_{bonds} + V_{angles} + V_{dihedrals} + V_{improper} + V_{vdW} + V_{elec} \qquad [2.2]$$

$$V_{bonds} = \sum_{bonds} K_r (r - r_{eq})^2 \qquad [2.3]$$

$$V_{angles} = \sum_{angles} K_\theta (\theta - \theta_{eq})^2 \qquad [2.4]$$

$$V_{improper} = \sum_{improper} \frac{1}{2} K_{\psi}(\psi - \psi_0)^2 \quad [2.5]$$

$$V_{dihedrals} = \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\chi - \gamma)] \quad [2.6]$$

$$V_{vdW} = 4\epsilon_{ij} \sum_{LJ,i<j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) \quad [2.7]$$

$$V_{elec} = \sum_{Coul.,i<j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad [2.8]$$

Here, r denotes bond length, θ is the bond angle involving three atoms, and χ represents the dihedral angle among four atoms. Parameters like r_{eq}, K_r, θ_{eq}, K_θ, K_ψ, V_n, γ, and n are constants for bonded interactions. The terms ϵ_{ij}, A_{ij}, and B_{ij} are Lennard-Jones parameters, r_{ij} is the interatomic distance, and ϵ_0 is the permittivity of free space.

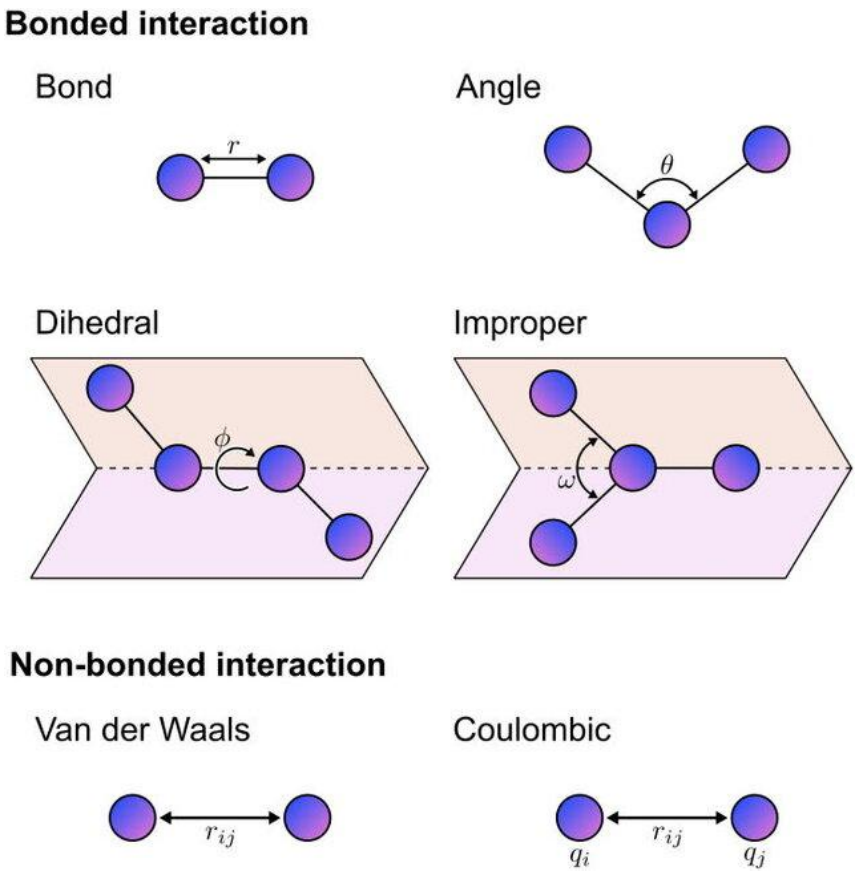


Figure 2.2: Schematic representation of typical force field components: bonded interactions (bond, angle, dihedral, and improper torsion) and non-bonded interactions (van der Waals and Coulombic). The figure was inspired by Yu et. al. [147].

The internal molecular contributions to total energy are mathematically represented in equations 2.2 to 2.6. This includes contributions from bonded terms like bond stretching, angle bending, and torsions. Meanwhile, equations 2.7 and 2.8 deal with long-range interactions, namely, van der Waals and electrostatic interactions (**Figure 2.2**). Additional terms are introduced in the force field to improve correlation with experimental results and treatment with hydrogen bond modeling.

2.2.1 Renowned Molecular Force Field

Several widely adopted force fields are used in biomolecular simulations, including AMBER [148] (Assisted Model Building with Energy Refinement), CHARMM [149] (Chemistry at HARvard Macromolecular Mechanics), OPLS [150] (Optimized Potentials for Liquid Simulations), and GROMOS [151] (GROningen MOlecular Simulations), Each provides specific parameter sets optimized for various types of biomolecular systems. Since a universal force field (FF) cannot accurately model all molecular types, we opted for system-specific FFs to enhance precision in biomolecular simulations. The most used FFs for various biomolecular systems are listed below.

• Proteins [ff19SB [152], **ff14SB** [153], ff99SB [154], **OPLS3** [155], CHARMM36m [156]]

• Phosphorylated Protein [phosaa19SB, phosaa14SB, **phosaa10**]

• Nucleic Acids [OL15 [157], OL21 [158], CHARMM36 DNA [159], CHARMM36 RNA [160]]

• Carbohydrates [GLYCAM_06j [161]]

• Lipids [lipids21, CHARMM36 lipids [162]]

• Ligands [**GAFF2** [163], OPLS-AA [164], CGenFF [165]]

In my thesis, I have used the **ff14SB** and **phosaa10** force fields for proteins and **GAFF2** for ligands, all available within the AMBER package. Therefore, a detailed discussion of AMBER protein force fields is presented in the following section.

2.2.2. AMBER Protein Force Field

AMBER, CHARMM, and OPLS represent the most commonly used families of protein force fields. Although their functional forms are largely similar, they differ primarily in specific parameters. For my study, I used the AMBER18 molecular dynamics suite and ff14SB for protein and the updated generalized Amber force field (GAFF2) for small molecules or inhibitors. The AMBER force field includes five primary energy components: bond stretching energy, angle bending energy, torsional energy for bond rotation, van der Waals interactions, and electrostatic energy. Hence, the total potential energy of the system follows:

$$E_{AMBER} = E_{bond} + E_{angle} + E_{torsional} + E_{vdW} + E_{electrostatic}$$

$$= \sum_{bond} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angle} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{torsion} \frac{1}{2} V_n [1 + \cos(n\omega - \gamma)] +$$

$$\sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left(4\varepsilon_{i,j} \left[\left(\frac{\delta_{i,j}}{r_{i,j}} \right)^{12} - \left(\frac{\delta_{i,j}}{r_{i,j}} \right)^{6} \right] \right) + \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left(\frac{q_i q_j}{4\pi\varepsilon_0 r_{i,j}} \right) \quad [2.9]$$

AMBER force fields use fixed atomic charges, which can reduce their accuracy compared to more advanced polarizable models like AMOEBA [166] and DRUDE [167]. As shown in equation 2.9, the vibrational mode of a bond is modeled through bond stretching, with the underlying assumption that the bond stays unbroken throughout the simulation. Therefore, this model is unsuitable for studying chemical reactions, as it cannot account for the breaking and formation of covalent bonds.

2.3. Molecular Dynamics (MD) Simulations

Experimental methods like X-ray crystallography and NMR spectroscopy offer static snapshots of proteins, yet the true understanding of biological function requires insights into molecular dynamics. To gain insights into the dynamic aspects of biomolecules, simulation techniques are required to extend static structural views into time dependent representations. Multiscale modelling techniques, including both all-

atom and coarse-grained molecular dynamics, are crucial for studying complex biological phenomena. MD simulation is a computational technique in biophysics that offers detailed insights into the structure, function, dynamics, mechanisms, and thermodynamics of biomolecules at the atomic level. MD serves as a bridge between theoretical models and experimental observations, connecting microscopic behaviour with macroscopic properties. It enables the observation of time-dependent evolution of the system, offering valuable insights into conformational fluctuations and changes that are crucial for understanding binding and folding mechanisms.

MD approaches are broadly divided into classical and quantum mechanical methods, with classical MD often modelled as a "ball and spring" model. In the classical MD, atoms are represented as balls, and strings correspond to bonds. The motion and behaviour of the system are described by classical mechanics. However, quantum MD captures the dynamics of systems by accounting the electronic structure underlying chemical bonds. Quantum mechanical equations compute the electron density of valence electrons to characterize chemical bonding. This approach provides a more refined and accurate description than classical MD but is more computationally demanding. Hence, classical MD is considered more suitable for studying large biological systems containing thousands of atoms.

2.3.1 Theoretical Foundations of MD Simulations

In this section, I will briefly discuss the fundamentals of MD simulations, starting from Newton's equations of motion and how these equations are involved in the molecular modeling of biological macromolecules.

For a given particle i in a system of N particles:

$$\boldsymbol{F}_i = m_i \boldsymbol{a}_i \qquad [2.10]$$

where $\boldsymbol{a_i}$ and m_i are the acceleration and mass of the i^{th} particle, respectively and $\boldsymbol{F}_i$ is a force working upon i^{th} particle due to the interaction with the rest of the particles present in the system. Here, bold letters correspond to a vector quantity and otherwise to a scalar. We know that force can also be estimated from the gradient of the potential, as shown below:

$$\boldsymbol{F}_i = -\frac{\delta V(\boldsymbol{r}^N)}{\delta \boldsymbol{r}_i} \quad [2.11]$$

In general, the **Hamiltonian** $\mathcal{H}$ relates the position and momenta of the particles to the system's total energy. By which the instantaneous positions ($\boldsymbol{r}_i$) and momenta ($\boldsymbol{p}_i$) of the i^{th} particle can be described as follows:

$$\boldsymbol{p}_i = -\frac{\delta \mathcal{H}}{\delta \boldsymbol{r}_i} \quad [2.12]$$

$$\boldsymbol{r}_i = -\frac{\delta \mathcal{H}}{\delta \boldsymbol{p}_i} \quad [2.13]$$

$$\mathcal{H}(\boldsymbol{p}_i, \boldsymbol{r}_i) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m_i} + U(\boldsymbol{r}_i) \quad [2.14]$$

Also, for a closed system, the **Hamiltonian** can be written as a sum of **kinetic energy** (E_K) and **potential energy** (V) as a function of coordinates and momenta.

$$\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) \quad [2.15]$$

$$\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_N) \quad [2.16]$$

$$\mathcal{H}(\boldsymbol{r}, \boldsymbol{p}) = E_K(\boldsymbol{p}) + V(\boldsymbol{r}) \quad [2.17]$$

The kinetic energy can be expressed as:

$$E_K = \sum_{i=1}^{N} \frac{1}{2m_i} (p_{ix}^2 + p_{iy}^2 + p_{iz}^2) \quad [2.18]$$

where m_i is the mass of the i^{th} particle and p_{ix}, p_{iy}, and p_{iz} are the momenta of the same particle along x, y, and z directions. So, now it is possible to integrate the entire trajectory over time as we have the potential function. However, most cases cannot be solved for an N-body problem. Instead, an approximation of the solution is obtained using several finite-difference integration methods. All are based on the Taylor series expansion, as shown below:

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \Delta t \boldsymbol{v}(t) = \frac{1}{2}\Delta t^2 \boldsymbol{a}(t) + \frac{1}{6}\Delta t^3 \boldsymbol{b}(t) + \cdots \quad [2.19]$$

$$\boldsymbol{v}(t + \Delta t) = \boldsymbol{v}(t) + \Delta t \boldsymbol{a}(t) + \frac{1}{2}\Delta t^2 \boldsymbol{b}(t) \quad [2.20]$$

$$\boldsymbol{a}(t + \Delta t) = \boldsymbol{a}(t) + \Delta t \boldsymbol{b}(t) \quad [2.21]$$

Here, $\boldsymbol{r}$, $\mathbf{v}$, $\boldsymbol{a}$, and $\boldsymbol{b}$ are the position, velocity, acceleration, and third derivative of position, respectively. Several algorithms exist to perform these calculations, which need to be computationally efficient and conserve energy and momentum. Among these, the Verlet algorithm is one of the well-established and reliable algorithms for these integrations. Here, the Taylor series expansions for the positions at time t and $t - \Delta t$ are combined.

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \Delta t\boldsymbol{v}(t) + \frac{1}{2}\Delta t^2\boldsymbol{a}(t) + \cdots \qquad [2.22]$$

$$\boldsymbol{r}(t - \Delta t) = \boldsymbol{r}(t) - \Delta t\boldsymbol{v}(t) + \frac{1}{2}\Delta t^2\boldsymbol{a}(t) + \cdots \qquad [2.23]$$

Adding equations 2.22 and 2.23 yields:

$$\boldsymbol{r}(t + \Delta t) = 2\boldsymbol{r}(t) - \boldsymbol{r}(t - \Delta t) + \Delta t^2\boldsymbol{a}(t) \qquad [2.24]$$

So, by estimating the positions at the current and previous timestep and accelerations at the current timestep, the Verlet algorithm can predict the new positions of the system. The velocity of the same system can be computed using the following formula.

$$\boldsymbol{v}(t) = \frac{\boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t-\Delta t)}{2\Delta t} \qquad [2.25]$$

The Verlet algorithm is a two-step method as it calculates the position at two different times. Additional computer memory is required for running the same as storing positions from three consecutive timesteps. So, there is an improvement of the same algorithm, termed as the velocity Verlet algorithm, where the velocity is calculated as step $\left(n + \frac{1}{2}\right)$, and then the coordinates at step $n + 1$. So, in this algorithm, positions and velocities are calculated using the formula stated below.

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \Delta t\boldsymbol{v}(t) + \frac{1}{2}\Delta t^2\boldsymbol{a}(t) \qquad [2.26]$$

$$\boldsymbol{v}(t + \Delta t) = \boldsymbol{v}(t) + \frac{1}{2}\Delta t[\boldsymbol{a}(t) + \boldsymbol{a}(t + \Delta t)] \qquad [2.27]$$

Here, the position of the next time step is estimated using the equation, which is further used to calculate the acceleration at the new time. Also, the velocity is obtained from the positions and acceleration using equation 2.27.

Another modification of the Verlet algorithm is the leapfrog algorithm. Here, the velocities $\boldsymbol{v}$ are calculated at time $(t + \frac{1}{2}\Delta t)$, which are used to calculate positions at a time $(t + \Delta t)$. So, velocities leap over the positions, and the positions leap over the velocities, giving the advantage of explicitly calculating the velocity, not in sync with the positions.

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t + \tfrac{1}{2}\Delta t)\Delta t \qquad [2.28]$$

$$\boldsymbol{v}\left(t + \tfrac{1}{2}\Delta t\right) = \boldsymbol{v}\left(t - \tfrac{1}{2}\Delta t\right) + \boldsymbol{a}(t)\Delta t \qquad [2.29]$$

2.3.2. Simulation Time-Step

One of the key components found and discussed in algorithms is the timestep, Δt. The accuracy of energy conservation and the numerical stability during the integration of motion equations directly depend on how one selects the timestep. A timestep smaller than the shortest dynamic timescale, typically bonds vibrations at ($\sim 10^{-13}$ sec), is generally applied in MD simulations to maintain integration stability. When a too large timestep is applied, energy in the system tends to increase sharply, resulting in the instability of simulations. All bonds involving hydrogen atoms are typically constrained using the LINCS [168] or SHAKE [169] algorithm. This helps us to use a larger time step of 2 fs. A timestep of 1-2 fs is generally used in biomolecular simulations when bonds involving hydrogen atoms are restrained.

2.3.3. The Solvent Model

The solvent model influences the stability, dynamic flexibility, and intermolecular interactions of both chemical and biomolecular systems. Two distinct solvent models are commonly used in MD simulations of biomolecules: the implicit solvent model, also called the continuum approach, and the explicit solvent model. The explicit solvent model provides a more detailed and accurate representation of a biomolecular system. Therefore, the following section will provide a detailed discussion of both the implicit and explicit solvent models.

2.3.3.1. Implicit Solvent Model

The implicit solvent model treats the surrounding solvent as a continuous medium rather than representing individual solvent molecules, using a dielectric continuum that reflects the average bulk properties of the solvent. In this model, the solvation free energy (ΔG_{solv}), representing the solvent energy contribution, is composed of three components: the cavity formation energy (ΔG_{cav}), the electrostatic interaction between solute and solvent (ΔG_{elec}), and the van der Waals interaction energy (ΔG_{vdW}).

$$\Delta G_{solv} = \Delta G_{cav} + \Delta G_{elec} + \Delta G_{vdW} \quad [2.30]$$

Several continuum models have been used in calculations, including Solvent-Accessible-Surface-Area (SASA), Poisson-Boltzmann Equation (PB/PBE), Generalized Born (GB), and Generalized Born Surface Area (GBSA). The PBE provides a high level of accuracy but is computationally demanding. Therefore, the approximation methods like GB model [170,171] are generally used. This model introduces an electrostatic interaction term between each solute and solvent atom, where the interatomic distance is replaced with a parameter called the Born radius (f_{GB}). It determines whether each solute atom resides in a buried region or is exposed to solvent molecules. The electrostatic interaction term is mathematically defined as follows.

$$E_{GB} = -\frac{1}{8\pi\epsilon_0}\left(1-\frac{1}{\epsilon}\right)\sum_{i,j}^{N}\frac{q_i q_j}{f_{GB}} \quad [2.31]$$

where ϵ_0 is the permittivity of free space, and ϵ is the permittivity of the solvent, q_i and q_j are the charges on atoms i and j, respectively. This method exhibits high computational efficiency. The model is fast, but it does not capture all the important solute-solvent interactions. Therefore, its accuracy is somewhat reduced. This study used a combination of PB(GB) and SASA to calculate solvation energies, both polar and non-polar for predicting the binding free energy.

2.3.3.2. Explicit Solvent Model

For most biomolecular simulation studies, we generally treated water molecules as the solvent that surrounds the biomolecules. The explicit solvent model includes the effects of solvent viscosity and the formation of hydrogen bond bridges. Therefore, the explicit

model provides greater accuracy when compared to the implicit solvent model. In this study, we selected water as the explicit solvent to accurately capture solute-solvent interactions. There are several water models commonly used in biomolecular simulations. These could be 3-point, 4-point, 5-point, and polarizable models, as presented in **Figure 2.3**. The commonly used TIP3P [172], TIPS [173], SPC/E [174], and SPC [175] models follow the 3-site framework, with their primary distinction being the HOH bond angle. In the case of SPC, the HOH angle is set to 109° to reflect a tetrahedral arrangement, while TIP3P preserves the actual water angle of 105°. Among all 3-point water models, TIP3P has been widely reported to reproduce several key thermodynamic features.

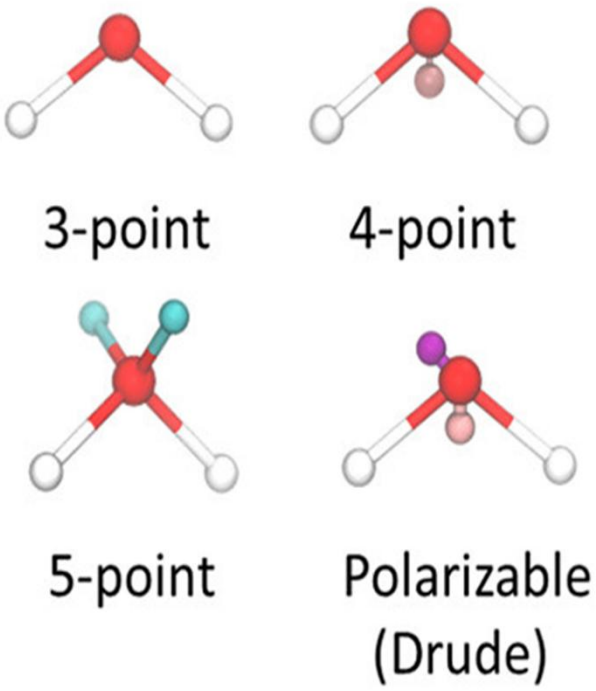


Figure 2.3: The TIP3P model uses three points to define water with parameters including bond length and HOH angle. The figure is inspired by Ref. [176].

The limited representation of viscosity in the 3-point model results in unrealistically rapid diffusion. TIP4P [177] addresses this by positioning a dummy site behind the oxygen atom in its 4-point water model. The added atom lacks mass and holds an

electrostatic charge that is typically attributed to the oxygen atom. This adjustment enhances the water molecule's polarity and leads to more accurate diffusion properties. Beyond TIP3P and TIP4P, models like TIP5P [178] and TIP6P [179] have been introduced. These achieve higher accuracy at a greater computational expense.

2.3.4. Periodic Boundary Conditions

Periodic boundary conditions help in more accurately estimating bulk system properties. In realistic physiological environments, the system contains not only solutes but also numerous solvent molecules. Because of computational limitations and the exclusion of surface effects of solvent molecules, we must restrict the number of atoms in simulations. Therefore, periodic boundary conditions (PBC) are applied to address these challenges by mimicking a bulk environment. Under PBC, the system is enclosed in a simulation box containing solute and solvent molecules, which is then replicated in all directions of the Cartesian coordinate system to mimic an space. The use of PBC helps to eliminate surface effects with a limited number of atoms in the system. During simulation, only one unit cell is modelled directly, while surrounding effects are generated using image particles. The main principle is that as an atom shifts inside the primary simulation box, its corresponding images follow the same motion in all periodic replicas. The benefit is that we need to track only the original image as it represents all other images computationally. As the simulation progresses, atoms may move across the boundary of the simulation box. When a particle exits the box, an identical image particle enters from the opposite side to maintain continuity. It preserves the system's dynamical balance and keeps the particle count consistent throughout the simulation. Among all the identical periodic boxes, only the atomic coordinates within the central box are saved (**Figure 2.4**).

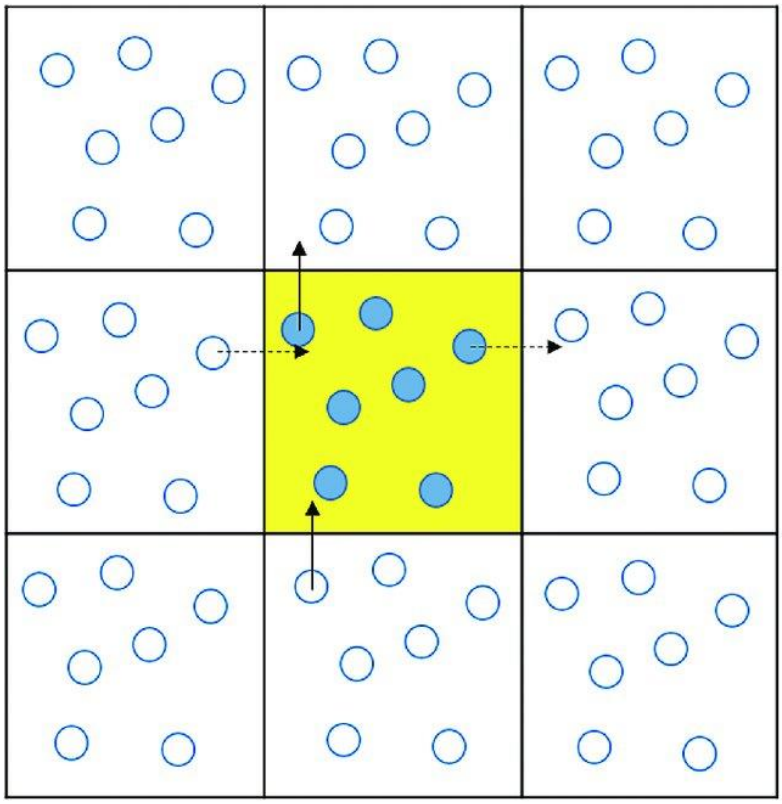


Figure 2.4: Two-dimensional representation of periodic boundary condition. The central cell (filled with yellow) represents the simulation box. Filled circles represent particles in the simulation box and open circles represent their periodic image in other cells. Bold and dashed lines show movement of two particles near the boundary; as a particle leaves the simulation box, its image enters the box from the opposite end.

The simulation box must be large enough to allow efficient calculation of long-range interactions. When periodic boundary conditions are applied, the number of interacting atom pairs increases significantly. In an N-particle system, pairwise interaction calculations require time proportional to $\frac{1}{2}N(N-1)$ ½ N (N–1). This is due to the fact that particles interact with both their neighbors in the simulation box and their image counterparts. This challenge is tackled by using a predefined interaction distance, considering only immediate neighbors of particle images. Interactions at longer ranges are effectively neglected. The maximum interaction distance must not exceed half the length of the simulation box (L). A cutoff radius (RC), which is smaller than or equal to L/2, defines the point beyond which two particles no longer interact. It will help to

avoid expensive force calculations. This truncation technique is applied only to short-distance interactions such as the Lennard-Jones potential. Electrostatic interactions, which are long-ranged, are usually treated using the PME method, a version of the Ewald sum methods.

2.3.5. Statistical Ensembles

MD simulations are designed to replicate and model the thermodynamic behaviour of actual systems studied experimentally. This usually refers to systems on a macroscopic scale, containing approximately 10^{23} particles or greater. Hence, ensembles represent a set of independent systems that are macroscopically identical. Macroscopically identical refers to the fact that each system forming the ensemble is defined by the same macroscopic variables, including volume (V), energy (E), pressure (P), temperature (T), and the total number of particles (N). Statistical mechanics serves as a bridge connecting the microscopic behavior of particles to the system's macroscopic properties.

Running MD simulations with a macroscopic number of particles is not computationally feasible. Therefore, applying statistical mechanical methods to extract system-level information becomes essential. A macroscopic observable arises from the contribution of numerous distinct microscopic states. An ensemble represents the collection of all possible microstates under a fixed set of parameters. For instance, isolated systems fall under the microcanonical ensemble (NVE), where the number of particles (N), volume (V), and total energy (E) are constant. In this ensemble, each microstate must maintain a constant energy. As a result, the system's ability to explore the potential energy surface is limited due to free energy barriers that exceed the system's total energy. An alternative to the microcanonical ensemble includes the canonical (NVT) and isobaric-isothermal (NPT) ensembles. In the NVT ensemble, volume and temperature remain fixed, while in the NPT ensemble, pressure and temperature are constant. The production run in MD simulations typically uses the NPT ensemble as it reflects realistic thermodynamic environments. In these ensembles, energy is not fixed, enabling the system to spontaneously cross large potential energy barriers.

2.3.6. Temperature and Pressure Regulation

MD simulations help in visualizing molecular motion and deriving key characteristics from the studied systems. Typically, thermodynamic properties of interest are measured in laboratory settings under ambient conditions. To replicate such conditions in simulations, it is ideal to sample from the canonical (constant temperature) or isobaric-isothermal (constant pressure and temperature) ensembles. Hence, temperature or pressure control in MD simulations is achieved by applying appropriate thermostat or barostat algorithms. The time average of kinetic energy serves as the basis for measuring temperature in MD simulations, as described by the equipartition theorem.

$$< E_k >= \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 = \frac{3}{2} N k_B T \quad [2.32]$$

Therefore, we define the time-averaged temperature as the instantaneous temperature. Under the NVT ensemble, this temperature should exhibit fluctuations around the preset target temperature, T.

2.3.6.1. Popular Thermostats

Multiple thermostats have been developed to regulate temperature during MD simulations, including the Berendsen [180], Andersen [181], Langevin [182], and Nose-Hoover thermostats [183]. For example, Langevin and Andersen thermostats regulate a system variable by enforcing a predefined distribution function. The Andersen thermostat, an early method in MD studies, assigns a random particle a velocity from the Boltzmann distribution at the chosen temperature through simulated collisions. I used a Langevin thermostat in my thesis that operates based on a general equation of the form-

$$F = F_{interaction} + F_{friction} + F_{random} \quad [2.33]$$

Where $F_{interaction}$ is the standard interactions calculated during the simulation, $F_{friction}$ is the damping used to tune the "viscosity" of the implicit bath, and F_{random} effectively gives random collisions with the solvent molecules. A user-defined friction damping parameter is used to couple the frictional and random forces during simulation.

2.3.6.2. Popular Barostat

Just as thermostats regulate temperature, there are multiple barostats which help maintain constant pressure (P) during simulations in the NPT ensemble. These include Berendsen [180], Andersen [181], Parrinello-Rahman [184], and Martyna-Tuckerman-Tobias-Klein [185] (MTTK) barostats. Barostat algorithms regulate the system's pressure but do not affect its temperature. Thus, when simulating under the NPT ensemble, then a barostat in combination with a thermostat must be used. In this thesis, we selected the Berendsen barostat and the Langevin thermostat to regulate the pressure and temperature, respectively.

The widely used Berendsen barostat couples the system pressure with a pressure bath, as shown below:

$$\lambda = 1 - \kappa \frac{\delta t}{\tau_P}(P - P_{bath}) \quad [2.34]$$

$$\boldsymbol{r}_i = \lambda^{\frac{1}{3}} \boldsymbol{r}_i \quad [2.35]$$

Here, the rescaled coordinates are represented by $\boldsymbol{r}_i$, the time constant by τP, and the isothermal compressibility (κ) controls the strength of the pressure bath coupling. The isothermal compressibility is also related to the volume fluctuations.

$$\kappa = \frac{1}{V}\left(\frac{\delta V}{\delta P}\right) T = \frac{1}{k_B T} \frac{<V^2> - <V>^2}{<V^2>} \quad [2.36]$$

We can scale the pressure using several options, like isotropic pressure scaling (which simultaneously scales all dimensions) and semi-isotropic pressure scaling (which scales selected dimensions).

2.3.7. Long-Range Interactions

The Coulombic interactions among charged particles in molecular systems fall off proportionally to 1/r, with *r* representing the distance between the particles. In these systems, atom-atom forces are categorized as long-range because other interactions decay much more quickly, usually as $1/r^3$ or beyond. This implies that even when atoms or molecules are far apart, they may still exhibit significant electrostatic interactions.

However, this effect is influenced by the shielding properties of the surrounding medium, specifically its relative permittivity or dielectric constant.

This indicates that electrostatic interactions are 80 times weaker in water compared to in a vacuum. Bonded interactions stay constant throughout the MD simulation and are relatively simple to evaluate. In contrast, non-bonded interactions are typically the most computationally demanding part of molecular mechanics, with their strength and relevance fluctuating over time. The stability of protein tertiary and quaternary structures largely depends on non-bonded interactions. They are the key forces responsible for preserving the structural integrity of proteins. Though long-range interactions might be significant, determining the distances between all atom pairs becomes increasingly expensive as number of atoms increase. The main idea here is to neglect certain interaction pairs to increase the simulation efficiency and maintain the authenticity of the produced results. The impact of Lennard-Jones and electrostatic interactions becomes insignificant at greater distances and can be ignored compared to the simulation error. To address this, a cut-off scheme is implemented to exclude all long-range interactions beyond a predefined distance. Due to their longer range compared to Lennard-Jones interactions, electrostatic forces pose a greater challenge for simplification in simulations. To address this, the Particle Mesh Ewald (PME) [186] method is commonly employed to compute long-range electrostatic interactions in periodic systems during MD simulations.

To understand the PME technique, it is useful to examine the relationship between charge distribution and Coulombic potential as described by the Poisson equation in differential form:

$$\nabla^2 \varphi(\boldsymbol{x}) = -(\tfrac{1}{\varepsilon})\rho(\boldsymbol{x}) \qquad [2.37]$$

Here, $\boldsymbol{\varphi(x)}$ represents the potential at point $\boldsymbol{x}$, $\boldsymbol{\rho(x)}$ denotes the charge density at the same location, and ε is the medium's permittivity. To solve this equation, we assume that ρ and φ are smooth functions; however, for a point charge, ρ becomes a delta function. Additionally, electrostatic charges are mapped onto a grid to compute the

corresponding potentials, and the force on each particle is then determined based on its position within this grid.

2.4. Basic Workflows of MD

MD simulations are widely used in biology to investigate the structural and dynamic behavior of biomolecular systems at the atomic level. The basic workflow involves several key steps, including system preparation, energy minimization, equilibration, production simulation, and trajectory analysis. These steps are discussed in detail in the subsequent sections.

2.4.1. System Topology Preparation

System preparation is perhaps the most crucial phase of a simulation, yet it often gets the least focus. Mistakes made here can critically compromise the entire simulation. To prepare an MD simulation, one must construct the initial structure (using X-ray, NMR, Cryo-EM, or homology modeling), solvate if required, apply a current force field, determine protonation states for ionizable residues, and ensure system neutrality and proper salt concentration. To mimic biological conditions, systems are solvated by introducing water molecules. The computational cost and the calculation of pairwise long-range interactions are both directly affected by the number of solvent molecules included. Hence, the size of the solvent box is critical for simulations, and it is generally recommended that its length be at least twice the long-range interaction cut-off.

2.4.2 Minimization

Energy minimization serves as a fundamental step in molecular modelling and is essential for exploring conformational space. Energy minimization is commonly used as a preparatory step in MD simulations to optimize the initial structure by removing steric strain [187]. Minimization is used to optimize the molecular structure by adjusting atomic positions to achieve the lowest possible potential energy. Various energy minimization algorithms exist, among which the steepest descent, conjugate gradient, and Newton-Raphson are widely used. Such methods are designed to locate

nearby local minima in the potential energy landscape. They play a critical role in eliminating structural irregularities like steric clashes or stereochemical errors.

2.4.3. Heating

At the beginning, the system is at zero temperature, which suggests there is no thermal energy present. We slowly increase the kinetic motion of atoms to bring the system to the desired temperature. To prevent instability in the simulation, the temperature must be raised gradually over an appropriate duration or in discrete steps to avoid any abrupt spike in kinetic energy. The NVE ensemble does not permit the addition of energy, and in the NPT ensemble, a rise in kinetic energy causes the system to expand to uphold constant pressure. Hence, NVT is used for the heating process. The system temperature increases as the particle velocities are reassigned from a fresh Maxwell-Boltzmann distribution at every integration step. The heating process helps prevent the system from crashing by allowing it to equilibrate gradually at each temperature level. Alternatively, applying weight restraints can regulate the heating procedure.

2.4.4. System Equilibration

As the production simulation is carried out using the NPT ensemble, a brief equilibration period is essential after heating to allow proper ensemble switching. The purpose of the equilibration phase is to transition the system from its starting conformation to a thermodynamically stable equilibrium state. In this phase, key thermodynamic variables like energy, temperature, and pressure are closely observed. Hence, equilibration is carried out under the NVT ensemble to balance kinetic and potential energies. That is, the kinetic energy added during heating should be properly shared among all the system's degrees of freedom to ensure uniform motion. This often indicates that the potential energy responds more slowly and needs to be equilibrated to match the kinetic energy of the system. The system is considered equilibrated when the potential energy becomes constant and exhibits minimal variation over time. Following that, the initially sharp drop in potential energy stabilizes near a constant value, while other essential properties remain unchanged over time. This pattern indicates that the system has equilibrated within the defined time frame.

2.4.5. Production Run

Once equilibration has been successfully achieved, we proceed to collect simulation data. This part of the simulation is typically called the production phase. In both equilibration and production, the simulation conditions are identical, but the significant difference is that data from the production run is saved and analyzed. Before starting a production simulation, appropriate equilibration consistent with the chosen ensemble must be conducted. Production data must not be recorded immediately after significant system changes, such as sudden temperature adjustments, box scaling, or minimization, unless these transitions are intentionally part of the study. In certain cases, the initial segment of production data is treated as additional equilibration and excluded from analysis. This approach is typically used after changing the simulation setup, such as altering the ensemble from NVT to NPT or vice versa. Production runs can range from a few nanoseconds to several microseconds, and with the enhanced computational power of modern Graphics Processing Units (GPUs), microsecond-scale simulations are now routinely conducted [19].

2.5. Enhanced Sampling Method

Atomistic MD simulations are among the most popular techniques, offering a balance of simplicity and precision in exploring biomolecular structures and their conformational dynamics [20]. Because of the presence of high free energy barriers, many biologically significant processes such as protein-ligand binding and conformational transitions occur over timescales that exceed the capabilities of conventional MD simulations. These processes are referred to as rare events [188,189]. A variety of methods exist to increase the probability of crossing energy barriers, thereby enabling broader conformational sampling within the same simulation time. These are collectively known as “enhanced sampling” techniques. A number of widely used enhanced sampling techniques include umbrella sampling [190,191], metadynamics [192], adaptive biasing force (ABF) [193], steered molecular dynamics (sMD) [194], replica exchange MD (REMD) [195], conformational flooding [196], parallel tempering [197], accelerated MD (aMD) [198], Gaussian accelerated MD (GaMD) [199], and Ligand Gaussian accelerated MD (LiGaMD) [200], etc. In this

study, I used GaMD simulations to investigate the conformational dynamics of kinases in DLK-JNK signaling pathway, as well as their respective protein-ligand complexes.

2.5.1. Gaussian Accelerated Molecular Dynamic Simulations (GaMD)

Gaussian accelerated molecular dynamics (GaMD) is a powerful computational technique that enables both unconstrained enhanced sampling and accurate free energy estimation of biomolecular systems. It works with the same theoretical background as aMD. GaMD significantly speeds up biomolecular simulations by several orders of magnitude. Similar to accelerated MD (aMD), it does not require the use of predefined collective variables (CVs). Additionally, the harmonic boost potential applied in GaMD follows a Gaussian distribution, allowing accurate reconstruction of the original biomolecular free energy surface using a second order cumulant expansion based on Gaussian approximation [201]. This effective approach significantly lowers statistical noise, thus addressing the limitations encountered in the aMD method.

If the system's potential energy (V) is lower than a set threshold energy (E), then a harmonic boost potential (ΔV) is applied, and it is defined by the following expression:

$$\Delta V = \frac{1}{2}k(E - V)^2, if\ V < E \quad [2.38]$$

where k is the harmonic force constant. Then the modified potential (V^*) is given by:

$$V^* = V + \frac{1}{2}k(E - V)^2, if\ V < E \quad [2.39]$$

When the system potential energy (V) exceeds the threshold energy (E), the boost potential (ΔV) is not applied and is therefore set to zero, i.e., $\Delta V = 0$.

$$V^* = V, if\ V \geq E \quad [2.40]$$

Two conditions must be fulfilled by the boost potential to ensure a smooth potential energy surface. The first states that if V_1 and V_2 are any two values on the energy surface where $V_1 < V_2$, then their boosted equivalents must satisfy $V_1^* < V_2^*$. Consequently, Equation (2.40) can be reformulated as:

$$E < \frac{1}{2}(V_1 + V_2) + \frac{1}{k} \quad [2.41]$$

Secondly, for any two potential energies where $V_1 < V_2$, the difference between their boosted potentials should be smaller than that on the original surface, i.e., $V_2^* - V_1^* < V_2 - V_1$. By substituting V^* into Equation (2.40), we obtain:

$$E > \frac{1}{2}(V_1 + V_2) \quad [2.42]$$

By combining Equation (2.40) and Equation (2.41), and applying the condition $V_{min} \leq V_1 < V_2 \leq V_{max}$, the threshold energy ($E$) is determined to lie within the following range:

$$V_{max} \leq E \leq V_{min} + \frac{1}{k} \quad [2.43]$$

where $Vmin$ and $Vmax$ are the minimum and maximum potential energy of the system, respectively. Further to satisfy the Eq (2.43) is valid, $V_{max} \leq V_{min} + \frac{1}{k}$, and k have to satisfy this:

$$\frac{1}{k} \leq \frac{1}{V_{max} - V_{min}} \quad [2.44]$$

where k is defined as

$$k \equiv k_0 \left(\frac{1}{V_{max} - V_{min}}\right), 0 < k_0 \leq 1 \quad [2.45]$$

where k_0 is the magnitude of the applied boost potential. Higher k_0, reflecting that for decreasing the energy barrier, a large boost potential is added.

To ensure accurate reweighting of potential surface, the standard deviation of the $\Delta V(\sigma_{\Delta V})$ must be significantly small (i.e., narrow distribution)

$$\sigma_{\Delta V} = k(E - V_{avg})\sigma_V \leq \sigma_0 \quad [2.46]$$

where V_{avg} and σV are the average and standard deviation of the system potential energy, respectively. $\sigma_{\Delta V}$ is the standard deviation of the ΔV with σ_0 is the user-defined upper limit (e.g., 10 k_BT) for accurate reweighting.

If a threshold (E) is set to lower bound V_{max} i.e., $E = V_{max}$, then Eq (2.44) gives the following relation:

$$k_0 \leq \frac{\sigma_0}{\sigma_V} . \frac{V_{max}-V_{min}}{V_{max}-V_{avg}} \quad [2.47]$$

If $E=V_{max}+ (1/k)$, then we can obtain the following relation:

$$k_0 \geq (1 - \frac{\sigma_0}{\sigma_V}) \frac{V_{max}-V_{min}}{V_{max}-V_{avg}} \quad [2.48]$$

The boost potential can be calculated for a given E and k_0 using the following equation:

$$\Delta V = \frac{1}{2} k_0 \frac{1}{V_{max}-V_{min}} (E - V)^2, if\ V < E \quad [2.49]$$

The GaMD also provide options to apply total potential boost ΔV_P, the dihedral potential boost ΔV_D, or the dual potential boost (ΔV_P and ΔV_D). Mainly, the dual-boost simulations provide a higher acceleration than others. The values of V_{max}, V_{min}, V_{avg} and σ_V are taken from the initial equilbrated simulations. Here, we applied the dual potential boost for GaMD simulations.

2.6. Molecular Docking: Principles and Algorithms

In structural molecular biology and drug design, molecular docking is a critical *in silico* technique. It is used to determine the probable binding orientations of a ligand or another protein with a protein whose three-dimensional structure is already known. Molecular docking serves as a powerful tool for virtual screening of extensive ligand libraries, ranking potential inhibitors, and proposing how they may bind to and inhibit the target protein, aiding lead optimization. Docking protocols consist of two main components: searching algorithms and scoring functions. The search algorithm first explores various ligand orientations and positions within the binding site, and then the scoring function evaluates and ranks these conformations based on their predicted binding affinities [202]. Primarily, these methods account for the relative flexibility of both the receptor and the ligand, as binding often triggers conformational adjustments consistent with the “induced-fit” mechanism.

2.6.1. Search Algorithm

Ligand conformational variability can be analyzed through root-mean-square deviations (RMSD) calculations, with smaller RMSD values indicating closer resemblance to the native or reference binding pose. The accuracy of any searching algorithm largely depends on the number of conformational degrees of freedom explored during the search process. It can be reduced by:

a) Incorporating implicit solvent effects into the scoring function.

b) limiting flexibility to the ligand only while treating the protein as rigid, thereby enabling extensive sampling of ligand conformational space.

The algorithms used for ligand searching can be classified into **systematic**, **stochastic**, and **simulation** based strategies.

Systematic Method: It explores all degrees of freedom using exhaustive, fragmentation, and ensemble approaches. The exhaustive method systematically evaluates every possible orientation of the ligand's rotatable bonds. With more rotatable bonds, the number of conformations to be sampled also increases. To enhance computational efficiency, geometric constraints are often applied during docking. Further conformational states of the ligand are processed through refinement. This technique is employed in docking programs like Glide [203] and FRED [204], whereas DOCK [205], LUDI [206], and FlexX [207] utilize the fragmentation approach. Ligand flexibility is modeled using the "place and join" strategy, where multiple ligand fragments are docked into the binding site and subsequently connected through covalent bonds. For instance, the ensemble method, such as that used in FLOG, utilizes libraries of 3D coordinates to identify the ligand that best fits the receptor.

Stochastic Method: This approach modifies ligand conformations and employs a probability-based criterion to accept or reject them. It falls under two major types: Monte Carlo (MC) and Genetic Algorithm (GA). MC uses the Boltzmann probability criterion to accept or reject conformations, whereas GA employs evolutionary concepts, as seen in docking programs such as AutoDock and GOLD [208]. Each initial ligand pose is associated with a specific phenotype and genotype. The genotype

describes the ligand's translational and orientational parameters relative to the target protein, while the phenotype corresponds to its atomic coordinates.

2.6.2. Scoring Functions

Accurate docking studies depend on a strong scoring function that assigns numerical values to each protein-ligand complex generated during the sampling process. Scoring functions enable the ranking of ligands by estimating their relative binding affinities to the target protein. In molecular docking, scoring functions are broadly categorized into force field, empirical, and knowledge-based methods.

Force field (FF) based scoring functions calculate the total internal energy of the ligand as well as the interaction energy between the ligand and the protein. The disadvantages of FF scoring include limited ability to capture long-range interactions and entropy, along with significant computational expense. However, more recent scoring functions like AutoDock and GoldScore have been developed to account for such factors. Additionally, molecular mechanics force fields such as AMBER and OPLS are also applicable for scoring purposes. The empirical-based scoring approach relies on fitting parameters to experimental data in order to predict protein-ligand binding affinity, as seen in tools like ChemScore [209], Score [210], and X-score [211]. These scoring functions use statistical data derived from known structures to assess interatomic interactions, assigning the highest scores to poses that closely match experimentally observed geometries. A major limitation of this method is its dependence on the type of data available from experimental techniques like X-ray crystallography or NMR, which often lack critical details such as protonation states. DrugScore [212] and SMoG [213] are two examples of knowledge-based scoring functions. The three scoring approaches discussed above provide a foundational framework for developing improved strategies aimed at more accurate scoring and ranking of docked ligand poses.

2.7. Techniques for Analysing Simulation Trajectories

Molecular dynamics simulations generate vast and intricate datasets. The Cartesian coordinates of every atom often numbering in the thousands or millions are saved at

each time step over trajectories that can span thousands to millions of steps. Consequently, post-simulation processing is essential to derive meaningful insights from this raw data. In this section, we present analytical tools designed to examine conformational changes in biomolecules from both short and long timescale simulations.

2.7.1. Stability and Flexibility Analyses

The root-mean-square deviation (RMSD) provides a statistical measure for assessing structural similarities, making it essential for monitoring biomolecular stability during simulations. RMSD is used to assess the overall deviation of a molecular conformation from its reference by comparing atomic coordinates. It measures the mean distance between the atoms of the reference structure and the selected atoms in the target conformation. RMSD is defined as:

$$RMSD = \sqrt{\frac{\sum_{i=0}^{N}[m_i.(X_i - Y_i)^2]}{M}} \quad [2.50]$$

Here, N denotes the total number of atoms, m_i represents the mass of the i, X_i refers to the coordinates of the atom in the target structure, and Y_i indicates the coordinates of the same atom in the reference structure. In the case of non-mass-weighted RMSD, each m_i is set to 1, making the total mass M equal to N.

Another additional valuable quantity, root-mean-square fluctuation (RMSF), is commonly used to evaluate residue-level flexibility. Similar to RMSD, it quantifies the deviation of particle i from its reference position over time.

$$RMSF_i = \sqrt{\frac{\sum_{t_{j=1}}^{T}[x_{i(t_j)} - < x_i >]^2}{T}} \quad [2.51]$$

Here, T represents the total number of time steps in the simulation, $x_{i(t_j)}$ denotes the coordinate of atom x_i at time point t_j, and $< x_i >$ is the average position of that atom over the entire simulation time. RMSF provides insight into the time-dependent positional variation of atoms across the complete structure. RMSF values are often compared with B-factors to assess the reliability of the simulation, as B-factors indicate

atomic thermal motion observed during crystallographic analysis and are included in PDB files along with the atomic coordinates. Atoms exhibiting high B-factor values typically correspond to the most flexible regions of a protein, such as loop segments, and should show good agreement with the regions of high RMSF in simulations. The B-factor is defined as:

$$B = \frac{8}{3N}\pi^2 RMSF^2 \quad [2.52]$$

We additionally analyzed the compactness of the system by determining the radius of gyration (Rg), calculated according to the following expression:

$$R_g = \left(\frac{\sum_i |r_i|^2 m_i}{\sum_i m_i}\right)^2 \quad [2.53]$$

The term m_i refers to the mass of atom i, and r_i describes the coordinate of atom i relative to the center of mass of the entire system.

2.7.2. Correlation Matrix Analysis of Molecular Dynamics Trajectories

The degree of correlation within a system can be assessed by calculating the cross-correlation coefficients between atomic pairs, which are graphically represented in matrix form as a dynamical cross-correlation matrix (DCCM). It has been extensively utilized to investigate and quantify correlations in atomic motions across biomolecular systems. The dynamical cross-correlation (DCC) between the i^{th} and j^{th} atoms is mathematically defined as follows:

$$DCC(i,j) = \frac{<\Delta r_i(t).\Delta r_j(t)>_t}{(\sqrt{<|\Delta r_i(t)|^2>_t})\sqrt{<|\Delta r_j(t)|^2>_t}} \quad [2.54]$$

here $r_i(t)$ denotes the atom's coordinates as a function of time t, t denotes the time ensemble average and $\Delta r_i(t) = r_i(t) - < r_i(t) >_t$. DCC has been employed to investigate the dynamic cross-correlations that occur within a molecular system. An $N \times N$ heatmap is generated by the DCC, where N is the number of atoms (most often alpha carbons), and each entry indicates the cross-correlation in motion between atoms i and j. Correlation values in the DCC matrix range from −1 to +1, where +1 indicates perfect correlation, −1 denotes perfect anti-correlation, and 0 signifies no correlation

at all. Completely correlated motions have identical timing and phase, while anti-correlated motions occur over the same time interval but in opposite phases. Strong correlations are typically observed along the matrix diagonal. The maximum diagonal value is observed when the indices are equal ($i = j$), which results in $DCC(i, j)$ being 1.00. Positive correlation along the diagonal is a result of the coordinated movement of adjacent residues within defined secondary structures. Off-diagonal regions with either positive or negative values signify correlated motions between non-contiguous residues.

2.7.3 Essential Dynamics via Principal Component Analysis (PCA)

Principal component analysis PCA [214,215] reduces the complexity of a dataset by transforming it into a lower-dimensional form composed of principal components, which are eigenvectors associated with the highest eigenvalues and explain most of the variance. By applying PCA to molecular dynamics trajectories, one can identify a simplified configurational landscape governed by a small number of dominant motions. Cartesian principal component analysis removes overall rotation and translation by aligning the structures using a least-squares fit approach. A covariance matrix of size $3N \times 3N$ is constructed using the Cartesian coordinates. Diagonalizing the covariance matrix yields a set of eigenvectors and corresponding eigenvalue. The eigenvectors indicate the motion directions, while the eigenvalues represent the magnitude of those motions.

Let C be the covariance matrix, where each element C_{ij} is defined as follows:

$$C_{ij} =< (x_i - < x_i >)(x_j - < x_j >) > \qquad [2.55]$$

Here, x_i and x_j represent the coordinates of the i^{th} and j^{th} atoms, while $\langle x_i \rangle$ and $\langle x_j \rangle$ are their respective mean coordinates averaged over the ensemble. In three-dimensional space, the covariance matrix is constructed for the (x, y), (x, z), and (y, z) coordinate pairs, resulting in a $3N \times 3N$ matrix C, where N represents the total number of atoms.

The covariance matrix is subsequently diagonalized to extract its eigenvalues:

$$A^T C\ A = \lambda \qquad [2.56]$$

Here, A denotes the matrix of eigenvectors, and λ represents the corresponding eigenvalues. The eigenvectors are arranged in descending order based on their corresponding eigenvalues, with the eigenvector having the largest eigenvalue representing the first principal component (PC), followed sequentially by others. Principal components define a new coordinate system; thus, the simulation trajectories are projected onto the eigenvectors to obtain the PCs:

$$q(t) = A^T M^{\frac{1}{2}}(x(t) - < x >) \qquad [2.57]$$

Trajectory data can be filtered along specific principal components (PCs) to separate particular motions from those associated with other components. Typically, only the top few PCs, out of the entire set of eigenvectors, capture the most meaningful motions and are used for detailed analysis.

2.7.4. Dihedral Principal Component Analysis (dPCA)

PCA is commonly employed to simplify high-dimensional systems by reducing them to a smaller number of informative variables. However, Mu et al. [216] demonstrated that internal coordinates, such as backbone dihedral angles, can be used in place of Cartesian coordinates for PCA in MD simulations, particularly when dealing with complex biomolecules with many degrees of freedom. This may lead to avoiding mixing overall motions with the internal motions. The primary advantage of using backbone dihedral angles is that they exhibit greater variability compared to other internal coordinates, such as bond lengths and bond angles, in a molecule. As a result, a variant of PCA called dihedral angle principal component analysis (dPCA) [216], which relies on internal coordinates like the backbone dihedral angles (ϕ_n, ψ_n), can help in separating internal conformational changes from global motions.

However, angular data always poses a PCA problem or analyses other statistical processes due to their circular nature. As an example, the average of 10° and 350° using standard arithmetic gives 180°, even though the proper circular mean is actually 0°. To circumvent such difficulties due to the circularity of variables, the space of dihedral angle (φn) is transformed to the linear metric coordinate space using the trigonometric functions sin φn and $cos\varphi n$.

2.7.5. Potential of Mean Force and Free Energy Landscapes

Potential of mean force PMF [217] represents the free energy profile as a function of a selected coordinate, which may be geometrical, like the distance between two atoms, or energetic, such as a molecular torsion angle. This energy landscape represents the mean force resulting from the sum of forces across all configurations of the molecular system. To express the potential of mean force in its simplest form, one may use a reaction coordinate. The PMF can be expressed in terms of the radial distribution function via the Helmholtz free energy equation.

$$A(r) = -k_B T \, ln[g(r)] + const. \qquad [2.58]$$

The constant is selected in such a way that the most probable configuration aligns with zero free energy. In this context, the radial distribution function indicates only a slight variation in free energy.

Free energy landscape (FEL) [218,219] theory serves as a conceptual basis for analyzing both the thermodynamics and kinetics of intricate systems, and also for predicting the probability of individual events. For complex systems such as proteins, the potential energy function spans multiple dimensions and often presents a highly convoluted surface. When the system is at a nonzero temperature, entropic effects become significant, and thus the free energy landscape controls the system's thermodynamics and kinetics. A widely used method to study and represent the free energy changes during protein folding and biomolecular isomerization is by plotting it against one or more order parameters, meaning macroscopic indicators that effectively differentiate conformational states of the systems. Each state is defined by a local minimum in the free energy of the reduced-dimensional landscape. Minima depth in the free-energy surface correlates with the stability of associated states, whereas the energy barriers between them represent the activation energies required for state transitions [220]. If the energy barriers separating states are small, transitions occur more frequently. Nevertheless, due to the intricate nature of the system and its many degrees of freedom, order parameters are often selected arbitrarily.

The free energy landscape can be drawn based on the following equation:

$$G_i \approx -k_B T \, ln(\frac{N_i}{N_m}) \quad [2.59]$$

Here, k_B denotes the Boltzmann constant, T represents the absolute temperature, N_i is the number of configurations in the i^{th} bin, and N_m refers to the population of the most occupied bin. In this analysis, a two-dimensional free energy landscape (FEL) was generated using reaction coordinates like PC1 vs. PC2, dPC1 vs. dPC2, RMSD vs. distance, and others.

2.7.6. Protein Structure Network (PSN) Analysis

An alternative approach to visualizing protein structures apart from examining secondary structures and folding patterns is through mapping residue interactions as networks, which helps reveal critical information about their structure-function relationship. Here, we constructed protein networks by applying two web-based tools: WebPSN [221] and the Network Analysis of Protein Structures (NAPS) portal [222]. Through the webserver, users can interactively visualize contacts between residues in predicted protein structures and MD trajectory data. Networks can be constructed using criteria such as Cα, Cβ, atoms, atom pairs, centroids, or interaction strength. Here, nodes in the network correspond to the Cα atoms of residues, and an edge is drawn between two nodes if the inter-residue Cα–Cα distance is within a cutoff radius of approximately 7 Å.

NAPS allows users to select nodes based on centrality scores, physicochemical attributes, or residue clusters. It also evaluates node and edge relevance using several network properties, including centrality, shortest path length, betweenness, hub classification, link strength, and community structure. Overall network visualization supports multiple analyses, including identifying functionally important residues, predicting coevolving sites, and investigating protein-protein or domain-domain interaction mechanisms involved in intra- or inter-molecular communication.

2.8. Calculation of Binding Free Energy

In computational biology, free energy calculations play an important role in various applications, such as in rational drug design and structural modelling of proteins. Free energies can be expressed in two forms: the Helmholtz (F) and Gibbs free energy (G).

$$F = U - TS \quad [2.60]$$

$$G = H - TS \quad [2.61]$$

Here, U denotes the internal energy, T represents the absolute temperature, H refers to the system's enthalpy, and S indicates its entropy. Where H is the enthalpy of the system, as given by $\boldsymbol{H} = \boldsymbol{U} + \boldsymbol{PV}$, where P is the pressure and V is the volume of the system. This indicates that free energy is determined by both the system's internal energy and its entropy. Thus, both the strength of atomic interactions and entropic shifts influence system stability and binding affinity, and solvent entropy becomes particularly important during the formation of hydrophobic protein interiors or membrane cores.

In molecular systems, Gibbs free energy determines the binding affinity for protein-ligand as well as protein-protein interactions. According to statistical mechanics, the probability of a specific state x correlates with its free energy:

$$p(x) \propto e^{-G(x)/k_B T} \quad [2.62]$$

Where kB is the Boltzmann constant and $e^{-G(x)/k_B T}$ is called the Boltzmann factor. Hence, the free energy difference for two states A and B can be estimated by:

$$\frac{p(A)}{p(B)} = e^{-(G(A)-G(B))/k_B T} = e^{-\Delta G/k_B T} \quad [2.63]$$

In this way, free energy estimation becomes possible from well-equilibrated simulations, and ΔG can be directly derived from the ratio of populations in states A and B.

The use of free energy calculations has become fundamental in computational biochemistry, particularly for predicting binding affinities of small candidate drugs and evaluating their structural stability with biomolecular systems. Free energy prediction methods can be broadly distinguished as either rigorous or approximate, depending on

their level of accuracy and computational demand. Among the rigorous approaches are free energy perturbation (FEP) and thermodynamic integration (TI). Among available approximate methods, the Molecular Mechanics-Generalized Born/Poisson-Boltzmann Surface Area (MM-GB/PBSA) method is most widely used.

2.8.1 Free Energy Perturbation (FEP)

Free energy perturbation (FEP) [223,224] is a statistical mechanics approach used to calculate free energy differences by introducing intermediate states between the initial reference and the final perturbed configuration [225,226]. The mathematical formulation of FEP is stated below:

$$\Delta F = F_1 - F_0 = -kT\, ln < \exp[\frac{-<E_1 - E_0>}{kT}] > \quad [2.64]$$

In this method, the free energy difference between the initial (reference) and final (target) states is determined by averaging a function of their energy difference, with sampling performed from the initial state's ensemble. The equation is further expressed using a power series expansion to obtain the free energy change. Here, E_1 denotes the energy of the target state, E_0 represents the energy of the reference state, and V corresponds to a small perturbation. The resulting expressions are derived through a second-order expansion:

$$E_1 = E_0 + V \quad [2.65]$$

$$\Delta F = F_1 - F_0 = < V > - \frac{1}{2kT}(< V^2 > - < V >^2) \quad [2.66]$$

Free energy methods like this are most appropriate when system fluctuations are primarily described by a Gaussian distribution.

2.8.2 Thermodynamics Integration (TI)

Another important and extensively used method for free energy estimation is thermodynamic integration (TI) [227,228]. The comparison of binding affinities between ligands L1 and L2 involves an alchemical transformation, where a set of intermediate states is generated using a coupling parameter λ. The coupling parameter is defined such that when λ equals 0, the system represents ligand L1 (initial

configuration), and when λ reaches 1, it corresponds into ligand L2 (final configuration). The total energy (V) of the system can be defined as:

$$V(\lambda, x) = (1 - \lambda)V_1(\lambda, x) + \lambda V_2(\lambda, x) \quad [2.67]$$

Here, V_1 and V_2 denote the potential energies of ligands L1 and L2, respectively. The derivative of the potential energy with respect to λ is then utilized to calculate the free energy difference as follows:

$$\Delta G_{alch} = \int_0^1 < \frac{\partial V(\lambda,x)}{\partial \lambda} >_\lambda \, d\lambda \quad [2.68]$$

Here, $\langle \cdots \rangle \lambda$ represents the ensemble average at a given λ state. Typically, each ensemble corresponding to a λ window is derived from a single molecular dynamics (MD) simulation.

To estimate the relative binding affinities (ΔΔG) of two ligands bound to a protein, the thermodynamic cycle method is applied using the equation provided below:

$$\Delta\Delta G = \Delta G_1 - \Delta G_2 = \Delta G_{alch}^{aq} - \Delta G_{alch}^{bound} \quad [2.69]$$

Here, ΔG_1 and ΔG_2 represent the binding free energies corresponding to ligands L1 and L2, respectively.

2.8.3 Molecular Mechanics (Generalized Born) Poisson Boltzmann Surface Area (MM/(GB)PBSA)

The binding free energy of protein-ligand systems is frequently estimated using the molecular mechanics Poisson-Boltzmann (generalized Born) surface area MM/PB(GB)SA method, which integrates molecular mechanics and either Poisson-Boltzmann or generalized Born models with surface area calculations [229,230]. This method utilizes classical force fields along with continuum solvation models to compute the binding free energies of complex molecular systems. Alchemical techniques like exponential averaging and thermodynamic integration (TI) often yield higher accuracy, while MM/PB(GB)SA offers a favourable balance between computational efficiency and accuracy [231,232].

The following is a concise explanation of the MMPB(GB)SA method as applied to the protein (*P*) and ligand (*L*) complex (*PL*). The overall binding free energy is shown below:

$$\Delta G_{bind} = G(PL) - G(P) - G(L) \quad [2.70]$$

The free energy values corresponding to the protein, ligand, and their complex (denoted collectively as *X*) are expressed as follows:

$$G(X) = E_{MM}(X) + G_{solv}(X) - TS(X) \quad [2.71]$$

$$E_{MM}(X) = E_{bonded}(X) + E_{nonbonded}(X) \quad [2.72]$$

In this expression, E_{MM} represents the total molecular mechanics energy, G_{solv} is the solvation free energy, T stands for absolute temperature, and S denotes the conformational entropy. E_{MM} includes both bonded (E_{bonded}(X)) and nonbonded ($E_{nonbonded}$(X)) contributions:

$$E_{bonded}(X) = E_{bond}(X) + E_{angle}(X) + E_{dihedral}(X) \quad [2.73]$$

$$E_{nonbonded}(X) = E_{electrostatics}(X) + E_{vdW}(X) \quad [2.74]$$

Furthermore, the bonded energy term further incorporates contributions from bond lengths (E_{bond}), bond angles (E_{angle}), and dihedral angles ($E_{dihedral}$), whereas the nonbonded term consists of electrostatic forces ($E_{electrostatics}$) and van der Waals interactions (E_{vdW}). Combing all the equations it will gives:

$$\Delta G_{bind} = \Delta G_{MM} - \Delta G_{solv} - T\Delta S \quad [2.75]$$

Where;

$$\Delta E_{MM} = \Delta E_{bonded} + \Delta E_{electrostatic} + \Delta E_{vdW} \quad [2.76]$$

$$\Delta G_{solv} = \Delta G_{cav} + \Delta G_{vdW} + \Delta G_{elec} = \Delta G_{SASA} + \Delta G_{PB/GB} \quad [2.77]$$

$$\Delta G_{SASA} = \gamma . SASA + \beta \quad [2.78]$$

Solvation free energy is divided into G_{cav}, which corresponds to the energy for solvent cavity creation, G_{vdW}, which represent van der Waals interactions, and G_{elec},

representing electrostatic effects. The first two components combined to give G_{SASA}. An estimate of this parameter is obtained by integrating SASA with two empirically determined coefficients (γ and β), which are based on linear regression analyses of the solvation energies of small apolar compounds in water. In this context, γ is the proportionality constant for surface tension, set at 0.00542 $kcal{\cdot}mol^{-1}{\cdot}Å^{-2}$, while β is the empirical offset value, set to 0.92 $kcal{\cdot}mol^{-1}$. SASA values are estimated through the linear combination of pairwise overlap (LCPO) method, utilizing a spherical probe of radius 1.4 Å.

Electrostatic contributions to solvation energy are typically calculated using PB or GB models. Specifically, in the PB model, the electrostatic potential is obtained by solving the Poisson-Boltzmann equation [233]:

$$\nabla\left[\varepsilon(\boldsymbol{r})\nabla_{\varphi}(\boldsymbol{r})\right] = -4\pi\rho(\boldsymbol{r}) - 4\pi\lambda(\boldsymbol{r})\sum_i z_i c_i \exp\left(-\frac{z_i\varphi(\boldsymbol{r})}{k_B T}\right) \quad [2.79]$$

Here, ε(**r**) refers to the dielectric constant, φ(**r**) corresponds to the electrostatic potential, and λ(**r**) serves as the masking function for the Stern layer. Analytical solutions are not feasible for the PB equation, and numerical methods remain computationally expensive.

According to the GB model, the same is given by:

$$\Delta G_{GB} = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i,j}\frac{q_i q_j}{\sqrt{r_{ij}^2 + R_i R_j \exp\left(-\gamma\frac{r_{ij}^2}{R_i R_j}\right)}} \quad [2.80]$$

Here, ε is the dielectric property of the bulk solvent, q_i denotes the charge carried by atom i, and r_{ij} represents its effective Born radius. Multiple GB model formulations exist, with their primary differences arising from how the Born radii are calculated [234,235].

2.8.4. Conformational Entropy Calculation

The estimation of configurational entropy was performed by applying the NMA (normal mode analysis) method [236,237]. The following equation was used to calculate the vibrational entropy (S_{vib}) through the NMA technique.

$$S_{vib} = R\left[\frac{x}{e^{x}-1} - \ln(1-e^{-x})\right]; x = \frac{h\vartheta}{kT} \quad [2.81]$$

Here, R is the universal gas constant, k is the Boltzmann constant, and T is the absolute temperature. The symbols h and ϑ are used to denote the Planck constant and vibrational frequency, respectively.

Chapter 3

Chapter 3

Plant Derived Active Compounds of Ayurvedic Neurological Formulation, Saraswatarishta as a Potential Dual Leucine Zipper Kinase Inhibitor: An *in-silico* Study

AIM: To validate the therapeutic relevance of SWRT molecules in inhibiting DLK-mediated neurodegenerative signaling.

This chapter is reproduced from our original article:

Koirala, S.; Roy, R.; Samanta, S.; Mahapatra, S.; Kar, P. Plant Derived Active Compounds of Ayurvedic Neurological Formulation, Saraswatharishta as a Potential Dual Leucine Zipper Kinase Inhibitor: An in-Silico Study. *Journal of Biomolecular Structure and Dynamics* 2024, 42 (20), 11201–11214. https://doi.org/10.1080/07391102.2023.2260892.

3.1. Introduction

Throughout the world, neurological illnesses are the leading risk factor and the second most significant cause of mortality [238]. The absolute number of deaths and individuals disabled by neurological disorders has increased dramatically over the last 30 years, and future rises are projected internationally due to population expansion and aging. According to the World Health Organization (WHO), more than 6 million people die every year due to stroke, and more than 50 million people suffer from epilepsy worldwide [239]. Efficient therapeutics are critically needed, but only if the origin and causes of each disease are well known. Approximately 85% of all medication approaches fail in clinical testing, and only 25 to 30 novel biochemical entities are licensed in the United States each year on average [240].

Protein kinases generally control every constituent of cell life, and mutations or changes in their activity cause various types of disorders. After the G-protein-coupled receptor (GPCR), protein kinase is the second most important class for pharmaceutical targets and has become the primary major drug target of the 21st century [17]. Neurological diseases such as Parkinson's and Alzheimer's need emergency medication, but no specific therapies exist that slow down the loss of neurons. Consequently, interest exists in finding novel molecules that specifically target the molecular pathway which plays a role in neuronal degeneration [241]. Several studies show the role of Dual leucine zipper kinase (DLK), and its downstream kinases in neurological disorders and how targeting this kinase can play a role in preventing neuronal diseases [66]. DLK is a serine/threonine protein kinase, which is a member of mixed lineage kinase [242]. Mixed lineage kinases (MLKs), members of the MAP3K family, are part of a phosphorylation cycle in which phosphorylate MKKs or MAP2Ks, followed by JNKs, causing further phosphorylation of the transcription factors such as c-jun, leading to the degeneration of axons and apoptosis [243]. An earlier study showed that JNK activity is linked to Alzheimer's disease and various other neurological disorders [244]. A team at Genentech conducted several DLK knockdown investigations to support the concept that DLK is involved in the stress-induced degeneration of neurons [126]. DLK knockdown studies suggested that it protects neurons against neurodegeneration [66]. Because of its critical involvement in neuronal degeneration, inhibiting the JNK pathway downstream of DLK has often been a therapeutic target [245]. A massive kinase screening effort led to the discovery of many reported drugs that interact with DLK [246]. However, numerous other protein kinases have a role in central nervous system disorders, such as GSK3, DAPK1, ROCK1, mTOR, KIT, etc. Several potent inhibitors targeting these kinases have already been discovered that are either already FDA-approved or are in preclinical or clinical development [247].

The Indian traditional medicinal system, also known as Ayurveda, is an ancient medical system based on natural products and aims to improve one's mental and physical health. It is one of the oldest medical systems in the world and is India's traditional health care system. According to the WHO, 70-80 percent of the world's

population uses nonconventional medications, primarily from plant preparations, for their treatment. Saraswatharishta (SWRT), one of the several formulations found in the Ayurveda system, has been demonstrated to be effective in treating and controlling neurological diseases and disorders [248]. SWRT is said to be beneficial in the treatment of central nervous system abnormalities. In one investigation, Albino mice were protected against the memory and learning impairment effects of diazepam when they were pre-treated with SWRT for two weeks. As mentioned, the impact of SWRT in neurological diseases shows a neuroprotective effect, making it important to explore its mechanism of action against various targets of neurological disorders, the DLK kinase being a crucial one. Seventeen [249] ingredients compose the SWRT formulation, which is listed in **Table 3.1**.

Table 3.1. List of plants used in making SWRT formulation.

Sl.no	Plant Name (Scientific)	Local name
1	*Bacopa monnieri*	Brahmi
2	*Asparagus racemosus*	Satawar
3	*Pueraria tuberosa*	Kudzu
4	*Terminalia chebula*	Harad
5	*Zinziber officinalis*	Adrack
6	*Anethum sowa*	Indian Dill
7	*Operculina ipomoea*	Nishoth
8	*Piper longum*	Thippali
9	*Syzygium aromaticum*	Laung
10	*Acorus calamus*	Sweet flag
11	*Saussurea lappa*	Costus Root
12	*Withania somnifera*	Ashwagandha
13	*Terminalia belerica*	Haritaki
14	*Tinospora cordifolia*	Guduchi
15	*Elettaria cardamomum*	Elaichi
16	*Embelia ribes*	Vavding
17	*Cinamomum zelonica*	Tejpata

Using modern-day computational facilities, like virtual screening, ADMET analysis, and molecular dynamics simulation studies, will accelerate the drug discovery process and help analyze the molecular interactions between the kinase and the inhibitors. Hence, the investigation of actions of the plant-derived components of SWRT formulation on DLK and its downstream proteins will play a role in its mechanistic understanding and verify its significance in the traditional ayurvedic medication system. In this study, we tried to validate the role of SWRT against neurological diseases and the molecular mechanism of action behind it. We used natural medicinal plant-derived phytochemicals composing the SWRT formulation as inhibitors against DLK. The created database containing 534 compounds (derived from the IMPPAT database and PubChem) was docked against the DLK protein using the Schrodinger suite's virtual screening (VS) workflow. With the aid of molecular dynamics simulations and an estimated binding free energy obtained from the MMPBSA technique, we further explored the structural dynamics and stability of complexes in 500 ns replica runs. Thus, this study marks the first effort to validate the molecular mechanism of action of SWRT-based phytochemicals against DLK. The entire work scheme is shown in **Figure 3.1**.

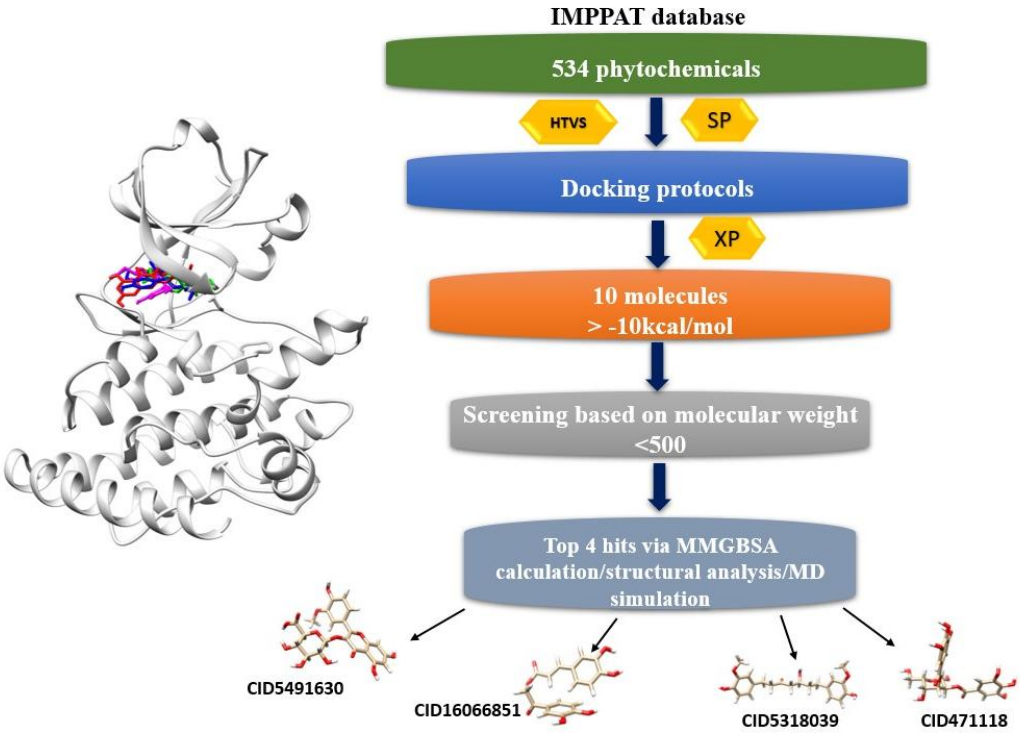


Figure 3.1. Diagrammatic representation of steps involves in detecting important inhibitors against DLK via virtual screening, dynamics study and energy calculation.

3.2. Materials and Methods

3.2.1. Data Collection

The X-ray crystal structure of DLK kinase (5CEN), with a resolution of 1.7 Å [250] was derived from the RCSB protein data bank, which constitutes of the residues 117-397. A visual representation of the kinase domain along with some significant regions have been depicted in **Figure 3.2(A)**. The 3D structure of the phytochemicals of SWRT formulation were downloaded from the IMPPAT (Indian medicinal plants phytochemistry, and therapeutics) and PubChem database and were included to make a database of 534 phytochemicals in SDF format, derived from the fifteen plants. This database was utilized for virtual screening (VS) and identification of potential DLK inhibitors among the 534 phytochemicals.

3.2.2. Ligand Preparation

We downloaded the three-dimensional structures of the phytochemical components from 15 of the 17 plants (available in database and literature), that made up the SWRT formulation. The phytochemicals was downloaded from the IMPPAT database [251] and PubChem using sources from existing literature [252]; [253]; [248]; [254]; [255]. A total of 534 compounds were screened, which were used to create a new database that was used in our subsequent studies. All the inhibitors were downloaded in SDF file format and loaded in Schrödinger suite's LigPrep module [256]. Through the Maestro interface, we established the ligand conformers and their corresponding tautomeric states using a pair of fast rule-based programs: the ionizer and tautomerizer. The ligands were reduced after the hydrogen atoms were added, and the OPLS3 force field was used for optimization. [155].

3.2.3. Receptor Grid Generation and Protein Preparation

The protein preparation wizard module was used for the preparation of receptors using the maestro platform of Schrödinger software. The DLK was first minimized and prepared for docking where OPLS3 force field was used with other default settings [155]. Glide was used to create a docking grid around the binding pocket, conserving the conserved motifs to decrease the search space leaving a 12 Å cubic space surrounding it [257]

3.2.4. Virtual Screening

Virtual screening (VS) screens a large number of ligands in order to find suitable compounds which can fit a pharmacological target. The successive molecular docking from our database was done against the DLK kinase. High throughput virtual screening (HTVS), standard precession (SP), and extra precession (XP) were the three stages of virtual screening that were performed [258]. The workflow's remaining parameters were kept at default values. Based on the extra precession score, we selected the top compounds with the best binding affinity towards DLK for further studies. A semi-flexible docking technique was adopted, where we kept the protein rigid and allowed the ligand to move within the binding pocket of DLK. As a result, a single conformation of the protein was used, and several conformations of each ligand were used for docking.

3.2.5. MD Simulation

The most extensively used technique for investigating protein-ligand binding is molecular dynamics (MD) simulation. The pmemd.cuda module of the AMBER18 was used to run the MD simulations of all systems. To describe the proteins, we used the Amber ff14SB [153] force field. For ligand parameters, the generalized Amber force field (GAFF2) [163] was used, and the Antechamber module [259] of Amber was used to determine the AM1-BCC [260] charge that was assigned to the inhibitors. The TIP3P water model was used to solvate the protein in an octahedron box with size of 10 Å from the protein surface [261]. Next, necessary Na+ and Cl- ions were introduced to neutralize the system, as well as to keep the salt concentration in accordance with the physiological salt concentration. To constrain the hydrogen atoms, the SHAKE algorithm [169] was used. Langevin thermostat was utilized to maintain the system temperature at 300 K [262]. Calculations of the long-range electrostatic interactions were done using the particle-mesh Ewald scheme [263]. A time step of 2 fs was used in the simulations. The first minimization step was conducted in 5000 steps of the steepest descent algorithm followed by 5000 steps of the conjugant gradient algorithm, with a harmonic restraint of 2 $kcal\,mol^{-1}\text{Å}^{-2}$. A subsequent unrestrained minimization was carried out using 100 cycles in the steepest descent algorithm followed by 900 cycles of the conjugant gradient algorithm. Next, systematic heating from 0 K to 300 K was performed in the NVT ensemble. With a 2.0 $kcal\,mol^{-1}\text{Å}^{-2}$ constraint force in the NVT ensemble, we carried out the density equilibration of each complex at 300 K for 50 ps. Following that, a 1.0 ns unrestrained equilibrium was carried out. Finally, at the NPT ensemble, each system was subjected to 2×500 n*s* production simulation run. The *Cpptraj* module was used to do post-MD analyses of the trajectories [264].

3.2.6. Protein Structure Network (PSN)

We also used the webPSN v2.0 [221] webserver, which employs the protein structure network (PSN) [265] and elastic network model-normal mode analysis (ENM-NMA) to compute the intra-protein structural interaction. Numerous characteries of the protein were examined in relation to the network properties, including nodes

representing amino acid residues, node linkers, hubs, communities representing interconnected nodes and various other matrics were studied. The network's hubs and communities were depicted using Visual Molecular Dynamics (VMD) [266]. In our previous studies, we have elaborated on the methodological approach of protein structure network [267].

3.2.7. MM-PBSA Calculation and Protein-Ligand Energy Decomposition Analysis

The MM-PBSA is well established technique was used to determine the binding affinity of the different types of biomolecular complexes. For free energy calculation, the last 300 ns of simulations were taken to calculate binding free energy using the following equations.

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad 1$$

$G\text{-}_{complex}$, $G\text{-}_{receptor}$, and $G\text{-}_{ligand}$ are the average Gibbs free energies of the complex, receptor, and ligand, respectively as shown in eq 1. The Gibbs free energy is defined as the sum of the molecular mechanics free energy (ΔE_{MM}), the solvation free energy (ΔG_{sol}), and the entropy ($T\Delta S$) as shown in equation 2.

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad 2$$

Molecular mechanic's energy (ΔE_{mm}) can further decompose into

$$\Delta E_{MM} = \Delta E_{cov} + \Delta E_{elec} + \Delta E_{vdw} \quad 3$$

Whereas Covalent, electrostatic, and van der Waals interactions are represented by E_{cov}, E_{elec}, and E_{vdw}, respectively. Further, the covalent component includes the bond energy (E_{bond}), the angular vibrational energy (E_{angle}), and the dihedral angle energies ($E_{dihedral}$). The solvation free energy (ΔG_{sol}) comprises polar as well as nonpolar free energy component, which is given by,

$$\Delta G_{solv} = \Delta G_{pol} + \Delta G_{np} \quad 4$$

The polar contributions (ΔG_{pol}) are calculated using the Poisson Boltzmann (PB) equation. The solvent-accessible surface area, or SASA, is used to calculate the

nonpolar solvation free energy (ΔG_{np}) as given in equation 5. Where γ and b are the experimental solvation parameters [268].

$$\Delta G_{np} = \gamma(SASA) + b. \quad (5)$$

Herein, 10000 frames from the final 200 ns of trajectory data were used to estimate the free energy. We skipped the entropy calculations due to the high computational expense. Using the MM-PBSA pairwise decomposition method, we also calculated every residual contribution to the binding free energy [269].

3.3. Results and Discussions

3.3.1. Virtual Screening

The selected 534 compounds from our libraries were docked against the DLK protein using the Schrödinger suite's virtual screening (VS) workflow. A selection of molecules with docking score lower than –10 kcal/mol following the completion of XP docking were selected for further analysis as shown in **Table 3.2**. We selected the top 10 lead compounds based on docking score, which were further screened to the 4 leads having high docking score and molecular weight less than 500 Da. Further these four leads are subjected to toxicological studies using ADMETlab 2.0 [270] webserver and results are shown in **Table A1**. All leads exhibited overall satisfactorily low toxicity profiles. Numerous plants contain these kinds of phytochemicals, that demonstrate antitumor capabilities as reported in literature. Isorhamnetin (lead-1) is our other hit compound among the four studied molecules. Numerous studies on isorhamnetin have reported that it protects the heart, brain, reduces inflammation, oxidative stress, protects organs, and prevents obesity [271]. Jarry et al have reported the effect of petasiphenone (lead2) which prevents the proliferation of different human leukemia cell lines [272]. The medicinal properties of lead-3, our next hit compound, have also been reported in recent literature. Huang et al. has reported the impact of Hexahydrocurcumin (lead-3), which has antitumor, anti-inflammatory, antioxidant, and cardiovascular protective properties [273]. The selected four lead compounds, CID5491630 (lead-1), CID16066851 (lead-2), CID5318039 (lead-3) and CID471118 (lead-4), which were considered for the simulation studies, were derived from four

different ingredients, lead-1 derived from *Anethum Sowa (Indian dill*), lead-2 derived from *Acorus calamus (Sweet flag)*, lead-3 derived from *Zingiber officinale (Adrack),* and lead-4 derived from *Syzygium aromaticum (Laung)*. In Supporting Information **Table A2**, we presented the SMILES of all four of the top leads.

Table 3.2. Top compound after XP-docking with glide score.

Lead molecules with ID	**Molecular weight (Da)**	**G-Score[a]**	**Glide-hbond[c]**	**Glide-evdw[d]**
CID13888122	788.576	-14.66	-6.94	-62.03
CID65238	940.681	-13.84	-7.10	-70.03
CID5491630	**492.389**	**-12.20**	**-3.66**	**-48.23**
CID16066851	**330.29**	**-11.70**	**-4.00**	**-39.46**
CID14032966	610.521	-11.69	-4.91	-44.58
CID442679	938.7	-11.38	-3.36	-39.58
CID102445430	756.534	-10.60	-4.85	-57.13
CID5318039	**374.433**	**-10.48**	**-3.15**	**-39.29**
CID471118	**484.366**	**-10.17**	**-6.01**	**-34.86**
CID85067758	899.10	-10.00	-1.88	-31.16

[a]Glide Score (kcal/mol).

[c]hydrophilic term.

[d]protein-ligand steric contact information.

Furthermore, we conducted a similarity search among these compounds, including FDA-approved drug sunitinib as a reference [274]; [275]. The Tanimoto coefficient gives values ranging from 0 to 1, where values near 1 represent higher similarity between compounds. It is used to calculate atom pair and maximum common substructure (MCS) similarities between compounds in the similarity workbench interface. The similarity scores of the screened compounds have been listed in the supporting information **Table A3**. All compounds show a wide range of dissimilarities among themselves and with sunitinib.

3.3.2. Molecular Dynamic Simulations

After screening the compounds from the database, complex structures of DLK bound to the top four hit molecules were subjected to 2×500 ns of atomistic molecular dynamics simulations. Subsequent analyses for structural stability, flexibility, rigidity, ligand dynamics, interactions profile and binding free energy calculations, were conducted using trajectories obtained after the simulations.

3.3.2.1. Structural Stability and Flexibility Analysis

The best four compounds interacting with the binding pocket of DLK were chosen and 2x500 ns of MD simulations were performed on them to track their dynamic, structural, and energetic properties and the results were compared to the apo. To analyze apo and each complex's dynamic behavior, we measured the root-mean-square deviation (RMSD) of backbone atoms in relation to their starting configurations, as shown in supporting information **Figure A1(A)**. To further investigate the stability of the entire protein backbone and the binding pocket, we used the Kernel Density Estimation (KDE) method to estimate the probability density of RMSD of various important regions of DLK for the entire course of the duplicate simulations as shown in **Figure 3.2 (B)**. The peaks of the plots are representative of the most probable conformation states. RMSD value for apo, DLK/lead1, DLK/lead2, and DLK/lead3 show stability in both runs after ~ 150-200 ns of simulation indicating the convergence of simulation, as shown in **Figure A1(A).** The RMSD average value of the last 300 ns are listed in **Table 3.3**. The average RMSD values lie in the range of 1.71 ± 0.20 Å to 2.49 ± 0.25 Å in case of run1, and for run2 the values lie in 1.51 ± 0.15 Å to 2.41 ± 0.25 Å range. The probability density plot of backbone RMSD as shown **Figure 3.2(B)**, reveals that apo and lead-4 show two peaks indicating two different confirmation states. Lead-2 shows a pronounced single peak and lead-1 gives a single slightly broader peak indicating a single conformation state for each, while lead-3 shows one pronounced peak and one small peak.

By determining the root-mean square-fluctuations (RMSF) of Cα atoms of each amino acids, as shown in **Figure 3.2 (C)**, we then explored the flexibility of DLK mainly focusing on the binding pocket region and compared it with apo. **Figure 3.2(C)** clearly

shows that the residual fluctuation for each complex follows the same pattern. The major residues interacting with the ligand are 130-135, which constitute the P-loop, 190-197, which constitute the hinge region, and several other residues from the pocket. The complex showed less fluctuation than apo in both P-loop and hinge regions. As there are no interacting residues from a-loop, it showed greater fluctuation compared to the interacting regions. The RMSF plot clearly shows that lead-4 complex and apo exhibited a more notable fluctuation in the various regions of DLK than other complexes. In contrast to other inhibitors, lead-2 and lead-3 complexes showed less fluctuation.

The solvent-accessible surface area (SASA) was calculated to determine the accessibility by solvent in the binding cavity of DLK, as depicted in **Figure 3.2(D),** while the time series plot of SASA is shown in **Figure A1(C)**. The surface area of a protein that is available to the solvent is known as the solvent-accessible surface area (SASA). It is one of the key characteristics for determining the intensity of ligand binding in the binding pocket. As shown in plot, the SASA value for complex is slightly lower than apo indicating less area accessible to solvent due solvent displacement by the ligand. In **Table 3.3** we show the average value of SASA and for both runs, which lies in the 451.48 ± 17.92 nm^2 to 460.72 ± 18.18 nm^2 range.

Table 3.3: The average backbone RMSD, binding pocket RMSD, the radius of gyration (Rg), and solvent accessible surface area (SASA) of both production simulation runs. The data are reported as average ± standard error of the mean.

System	Run	RMSD (Å) (backbone)	5Å, RMSD (Å) (Binding-pocket)	Rg (Å)	SASA (nm^2)
Apo	Run1	1.81 ± 0.18	-	19.12 ± 0.11	467.35 ± 18.08
	Run2	2.41 ± 0.25	-	18.89 ± 0.14	443.61 ± 18.82
Lead-1	Run1	1.82 ± 0.15	1.68 ± 0.21	19.23 ± 0.10	432.45 ± 19.95
	Run2	2.29 ± 0.34	1.67 ± 0.27	19.48 ± 0.11	451.48 ± 17.92
Lead-2	Run1	1.71 ± 0.20	1.75 ± 0.24	19.14 ± 0.09	444.69 ± 16.01
	Run2	1.64 ± 0.13	1.66 ± 0.22	19.07 ± 0.10	460.54 ± 17.13
Lead-3	Run1	2.49 ± 0.25	1.26 ± 0.45	19.22 ± 0.15	445.18 ± 17.18

	Run2	2.15 ± 0.15	1.00 ± 0.28	19.09 ± 0.09	449.56 ± 16.69
Lead-4	Run1	2.24 ± 0.23	1.15 ± 0.17	19.36 ± 0.11	461.42 ± 17.68
	Run2	1.51 ± 0.15	0.93 ± 0.20	19.31 ± 0.13	460.72 ± 18.18

We have studied the protein's compactness and dynamic characteristics by examining the radius of gyration (R_g). Numerous prior research have extensively utilized this structural parameter, which provides insights into the protein's overall conformational landscape [276]. By computing the (R_g) from the MD trajectories, as shown in **Figure 3.2(E)** and time series figure shown in **Figure A1(D)**, we determined the compactness of the systems. The R_g values for apo are slightly lower as compared to complexes. The average value of (Rg) is reported in **Table 3.3**, from which it is clearly observed that all the complexes showed similar compactness.

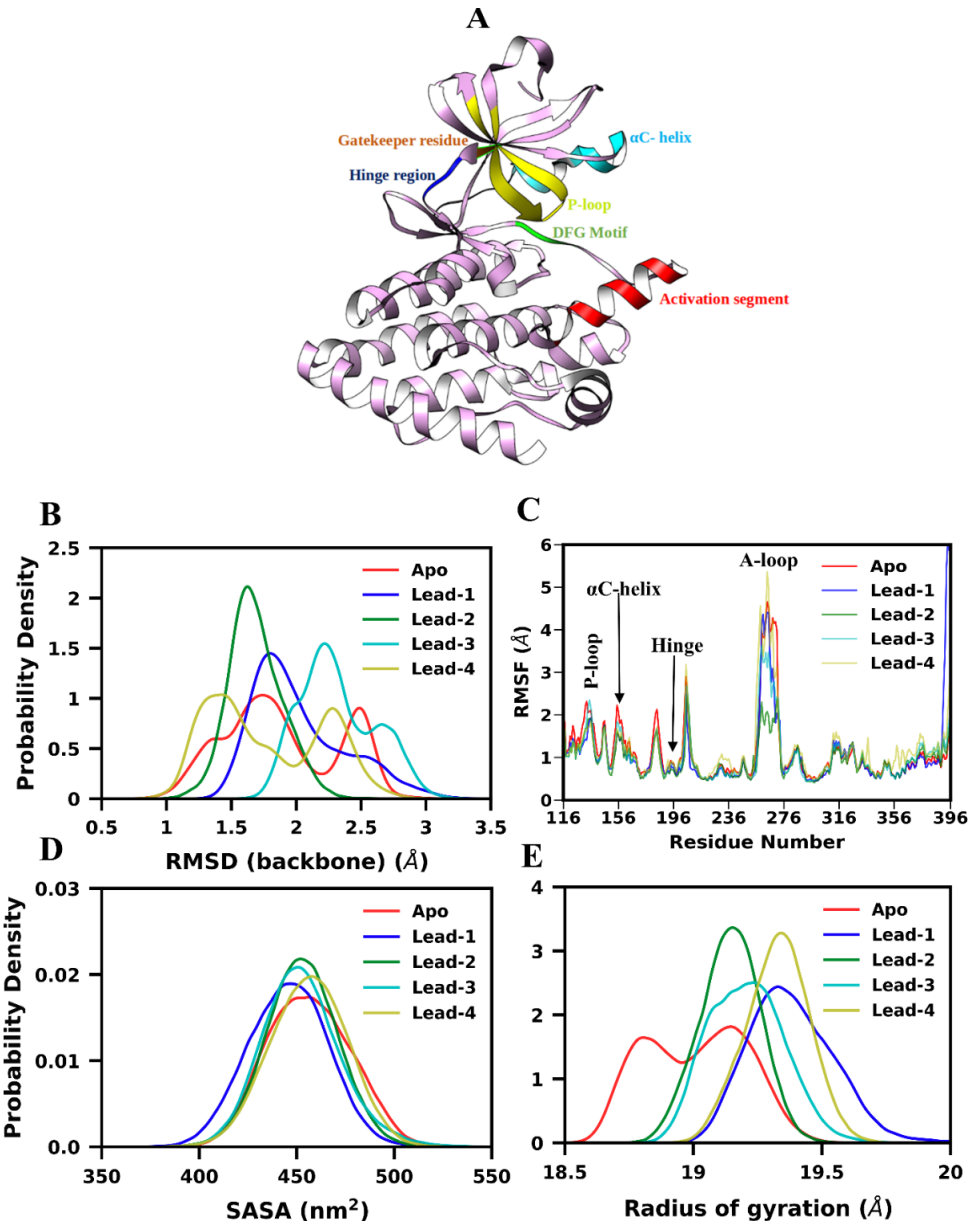


Figure 3.2. (A) Crystal structure of DLK showing several important regions of Kinase domain; **(B)** The probability distribution of root-mean-square deviations of all backbone atoms; **(C)** the root-mean-square fluctuations (RMSF) of DLK for each docked complex; **(D)** probability density of solvent accessible surface area (SASA) of all systems; **(E)** Probability distribution of radius of gyration (Rg) of all systems.

3.3.2.2. Binding Pocket Stability, Ligand Dynamics, and Ligand-Protein Distance Analysis

After examining the overall DLK structure, we investigate the RMSD of backbone atoms residues within a 5 Å distance from the ligands. As the RMSD (pocket) in **Figure 3.3(B)** shows, it fluctuated less in run1 as compared to run2, and the overall system stabilized after ∼200 ns in both runs of simulations. The average RMSD values varied from 0.93 ± 0.20 Å to 1.68 ± 0.21 Å for both MD runs. Overall, the results indicate that the DLK binding pocket in association with inhibitors is relatively rigid and compact with lesser fluctuation, which is advantageous for higher binding.

Similarly, we investigated the ligand dynamics in the bound state with DLK by calculating the RMSD of heavy atoms, as shown in supplementary information **Figure A1(B)**. All the ligands showed stable conformation and did not show much deviation, with the RMSD values staying within 3 Å. Lead-1 shows less deviation than other leads for both runs. The overall RMSD profile of each lead was stable, indicating a good and stable binding with the binding pocket of DLK. Subsequently, we also investigate the potential of mean force (PMF) for all four ligands, as shown in **Figure 3.3(C)**. To determine the low energy states during the simulation the PMF is generated. It is clear from **Figure 3.3(C)** that lead-2 and lead-3 showed a single free energy minimum found in ∼2.7 and ∼1.8, respectively. In cases of lead-1 and lead-4, it shows global minima at ∼0.4 and ∼2.4 and secondary minima at ∼1.6 and ∼1.7. The ligand movement inside the pocket of DLK is dissimilar among each other. In the later section, we also calculated the distance between ligand and protein to study the stability of ligands in the binding pocket of DLK.

To assess the stability of the ligands at the binding site, we calculated the ligand-protein distance from the binding pocket where the binding pocket is comprised of the residues having ≤ 5 Å distance from the ligand as shown in **Figure 3.3(A)**. It is visible from **Figure 3.3(A)** that except lead4 run1 at ∼150-210 ns simulation duration, all other leads showed stable interaction with binding pocket of DLK. Lead-1 and lead-3 had the lowest ligand protein distance indicating stable interaction in both runs. Lead-4 showed stable interaction and lowest ligand protein distance in case of run2, but in case of run1 the distance increased after 150 ns and reached up to 9 Å. after 250 ns the distance decreased again and remained constant around 5.5 Å indicating stable binding

while remaining in association with the binding pocket of DLK for the rest of the simulation. Lead-2 showed stable binding in both runs and ligand-protein distance is around 6 Å which is due to the larger size of it compared to the other ligands. Overall, in the ATP-binding pocket of DLK, all ligands interact closely and bind stably which indicates that all four of these phytochemicals present in SWRT may have a high potency to inhibit DLK, which gives us some mechanistic insight into their therapeutic capability against neurodegeneration.

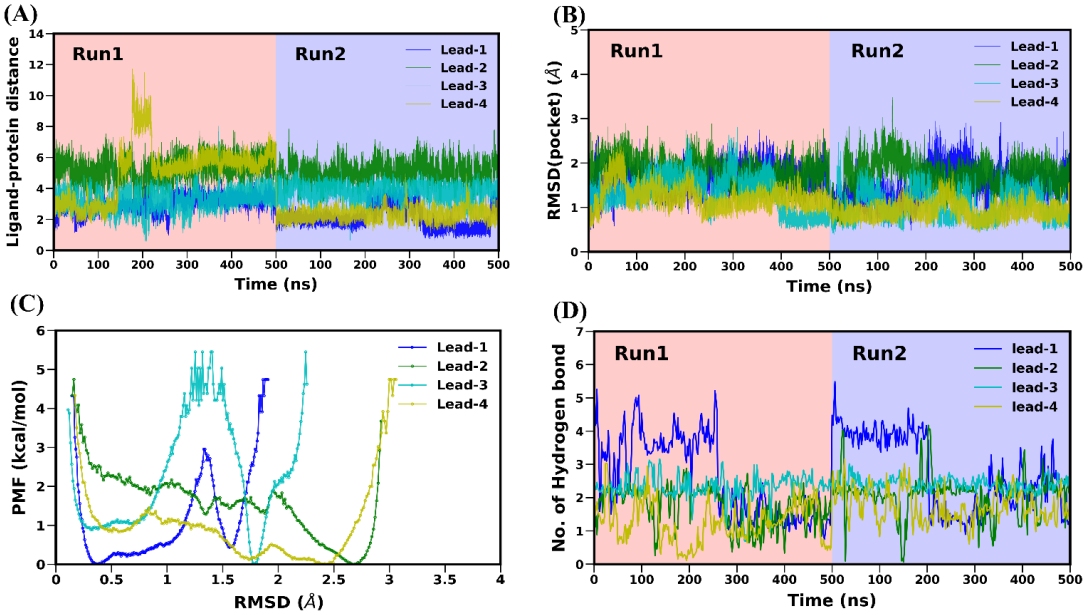


Figure 3.3. (A) Time evolution of distance between the inhibitors and the binding pocket of DLK; **(B)** Time evolution of backbone atoms of the binding pocket of DLK where the residues with < 5 Å distance from the ligands are considered; **(C)** Potential of mean force (PMF) of all the ligands; **(D)** The time evolution of the average hydrogen bonds formed between the inhibitors and DLK for all complexes.

3.3.3. Protein Structure Network Analysis

To understand the changes brought by complexation in the lead complexes, we performed a protein structure network (PSN) study of the apo and lead complexes. PSN gives the network representing protein structure where the single amino acid represents nodes, and these nodes interact among themselves to form links. The number of edges or direct links that a node forms determines its degree and nodes, with

higher degrees referred to as hubs. A community is a set of nodes with more connections among them than with other nodes in the network. Using webPSN v2.0 webserver the network analysis was conducted for both apo and complexed systems. **Table 3.4** represents the network characteristics while **Figure 3.4** displays the communities sorted by rank represented in different colors and the hubs sorted by their strength of interaction.

Table 3.4: Network properties of DLK (apo) and inhibitor bound lead-1, lead-2, lead-3, and lead-4 generated through PSNs.

Parameters	Apo	Lead-1	Lead-2	Lead-3	Lead-4
No. of linked nodes	243	251	246	237	245
No. of links	265	279	266	272	274
No. of hubs	29	30	23	35	29
No. of linked mediated hubs	129	126	114	157	138
No. of communities	4	7	6	6	7
No. of nodes involved in communities	33	42	39	42	35
No. of links involved in communities	51	66	56	61	48

As it is evident from the **Table 3.4** that the complexes form a greater number of links; lead-1 (279), lead2 (266), lead-3 (272), lead-4 (274), as compared to apo (265). Except for lead-2 all complexes show greater number of hubs compared to apo. In case of community, after complexation the community formation increases lead-1 (7), lead-2 (6), lead-3 (6) and lead-4 (7) than in the apo (4) structure. It is also evident from the **Table 3.4** that the number of nodes and hubs involved in forming community is higher in the complexes. The analysis of all these network features indicates the higher stability of protein backbone in case of complexes as compared to the apo structure which is also revealed by our other analyses. After the complexation the network formation increases in the region of binding pocket which is not observed in case of

the apo DLK. In case of all leads the third largest communities were formed near the binding pocket region as indicated by blue color community. These comparative results provide lights to the structural dynamics of DLK that could guide in the DLK specific drug development to treat various neurological diseases.

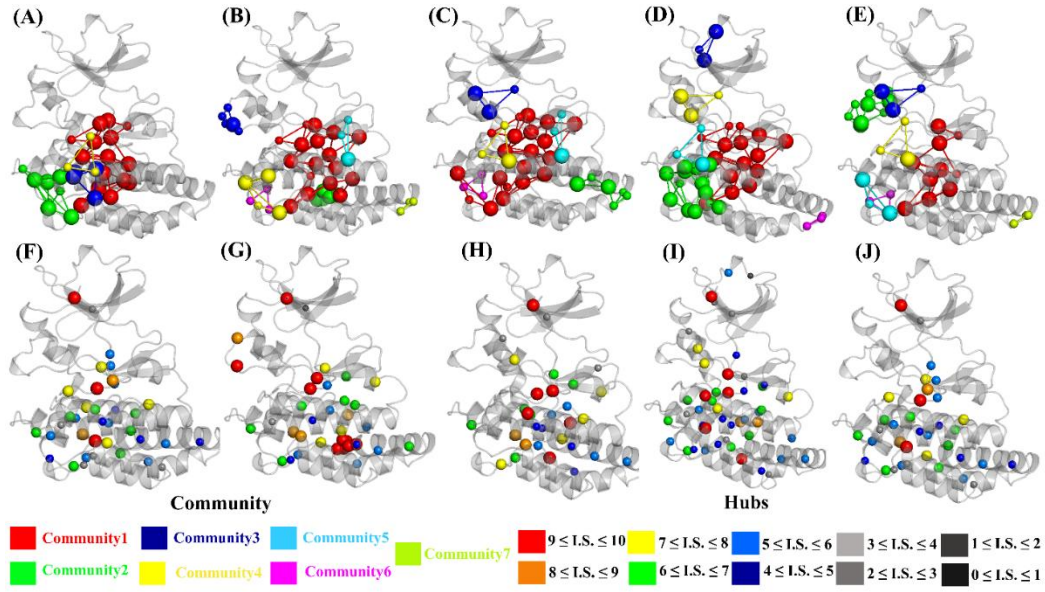


Figure 3.4. The communities and hubs present in DLK, (A, F) Apo, (B, G) lead-1, (C, H) lead-2, (D, I) lead-3 and (E, J) lead-4. The hubs are represented in color codes with their interaction strength and communities are sorted rank wise from highest to lowest (Community1-Community7).

3.3.4. Binding Free Energy Analysis.

The best-screened compounds against the DLK were assessed using the MM-PBSA technique to shed light into their binding affinities. **Table 3.5** shows binding free energy and all its components, and **figure 3.5(A)** represents the same graphically.

Table 3.5. Average Binding free energy and the energetics components calculated from the MM-PBSA scheme in kcal/mol for DLK against all four inhibitors.

System	ΔE_{vdw}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
Lead-1	-41.91	-34.81	58.55	-4.29	-76.74	54.25	-22.48

	(0.043)	(0.105)	(0.090)	(0.002)	(0.105)	(0.090)	(0.048)
Lead-2	-29.28 (0.038)	-38.19 (0.148)	41.20 (0.083)	-3.415 (0.001)	-67.48 (0.132)	37.79 (0.082)	-29.69 (0.064)
Lead-3	-37.74 (0.030)	-25.55 (0.059)	39.89 (0.048)	-4.07 (0.001)	-61.29 (0.065)	35.81 (0.048)	-25.47 (0.041)
Lead-4	-36.26 (0.041)	-63.47 (0.206)	76.34 (0.141)	-4.71 (0.001)	-99.73 (0.191)	71.62 (0.140)	-28.10 (0.077)

Standard errors of the mean (SEM) are provided in parentheses.

$\Delta G_{solv} = \Delta G_{np} + \Delta G_{pol}$

$\Delta G_{bind} = \Delta E_{vdw} + \Delta E_{elec} + \Delta G_{np} + \Delta G_{pol}$

The binding free energy varied from –22.48 kcal/mol to –29.69 kcal/mol. The highest binding free energy against DLK was obtained for lead2 which was –29.69 kcal/mol, followed by lead4 (-28.10 kcal/mol), lead-3 (-25.47 kcal/mol), and lead1 (-22.48 kcal/mol), even though lead-1 demonstrated a higher docking score of –12.20 kcal/mol during virtual screening. It is clear from these results that only docking score alone cannot disclose the protein-ligand interaction strength and their stability. Therefore, to determine the stability of complexes and their binding strengths, molecular dynamic simulations must be used in association with free energy calculations. Overall, from the MM-PBSA results, lead-2 binds most strongly to its target protein DLK. The binding free energy calculation clearly showed that the intermolecular van der Waals (ΔE_{vdw}) and electrostatic interactions (ΔE_{ele}) favored the formation of DLK-ligand complex. Along with these two-component, non-polar solvation energy (ΔG_{np}) also supported the interaction. The complexation was opposed by the polar solvation energy (ΔG_{pol}). **Table 3.5** shows clearly that ΔE_{vdw} varied from –29.28 kcal/mol and –41.91 kcal/mol, while ΔE_{ele} varied from –25.55 kcal/mol and –63.47 kcal/mol for all the complexes. It suggests that for lead-1 and lead-3, van der Waals interaction played a more significant role, while for lead-2 and lead-4 electrostatic interaction contributed to a higher significance in complex formation between DLK and ligand. By using high throughput virtual screening, and MM-PBSA scheme, we presented all four leads as strong phytochemical molecules with good pharmacological characteristics and

potency which are present in SWRT formulation, and act as a potential therapy against DLK based neurodegeneration. Our findings showed that all four phytochemicals and DLK had favorable electrostatic and van der Waals interactions, which may aid the future development of drugs.

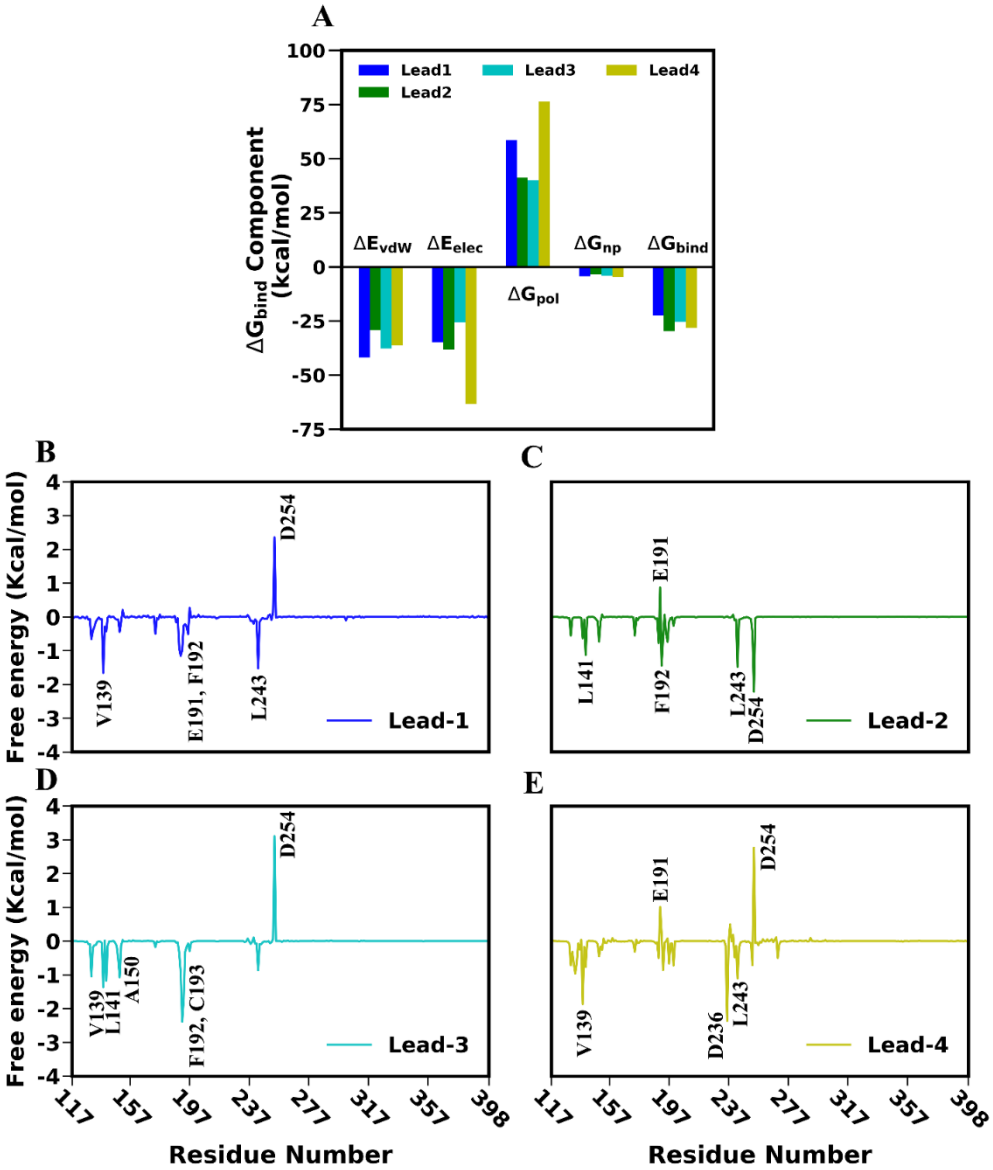


Figure 3.5 (A) Components of binding free energy (kcal/mol) for each protein-inhibitors complex. The per-residue decomposition of the binding free energy is graphically represented for DLK bound to **(B)** lead-1, **(C)** lead-2, **(D)** lead-3, **(E)** lead-4. The residues with energy contributions of < -1 kcal/mol and > 1 kcal/mol are shown in the plot in single letter code.

3.3.5. Per-Residue Energy Contributions to Binding

The contribution of each residue to the total binding free energy was determined using the MM-PBSA technique to understand more about the binding mechanism of the screened inhibitors against DLK. The hotspot residues participating in ligand binding are shown by the per-residue decomposition of free energy. The residue which contributed with energy ≤ -1 kcal/mol are listed in supporting information **Table A4**. In **Figure 3.7**, the interaction spectra of each protein-inhibitor complex are shown, which clearly depicts that out of the four tested compounds three showed the residues Val139, Phe192 and Leu243 as common crucial residues. Besides, Leu141 is also an important residue found in case of lead-2 and lead-3. Interestingly, Asp254 of the DFG motif disfavored the ligand-protein interaction in case of lead-1, lead-3 and lead-4. Along with Asp254, Glu191 also disfavored the interaction in case of lead-2 and lead-4. Overall, the identification of these crucial residues may aid in the search for novel selective ATP-competitive DLK inhibitors for the treatment of neurodegenerative disorders.

3.3.6. Hydrogen Bond Analysis and Protein-Ligand Interactions Profiles

To better understand molecular recognition and molecular interactions, the hydrogen bond profiles between the ligands and DLK were examined. Accordingly, intermolecular hydrogen bond between ligand and DLK was also computed. **Figure 3.3(D)** shows the total number of hydrogen bonds during simulation for all the DLK-ligand complexes. The average number of hydrogen bonds varied from 1 to 4 for all complexes. Lead-1 showed the presence of ~ 4 hydrogen bonds till 250 ns of simulation time, which decreased after 250 ns in case of both the runs. Lead-2 formed ~2 hydrogen bonds with DLK which decreased slightly after 250 ns of simulation. Lead-3 also had ~2 hydrogen bonds which remained stable throughout the simulation in case of both the runs indicating the stable interaction of the complex. In case of lead-4, ~1 hydrogen bond was observed in case of both the runs. All the leads showed more than 1 average hydrogen bond throughout the simulation indicating stable interaction between the ligands and DLK.

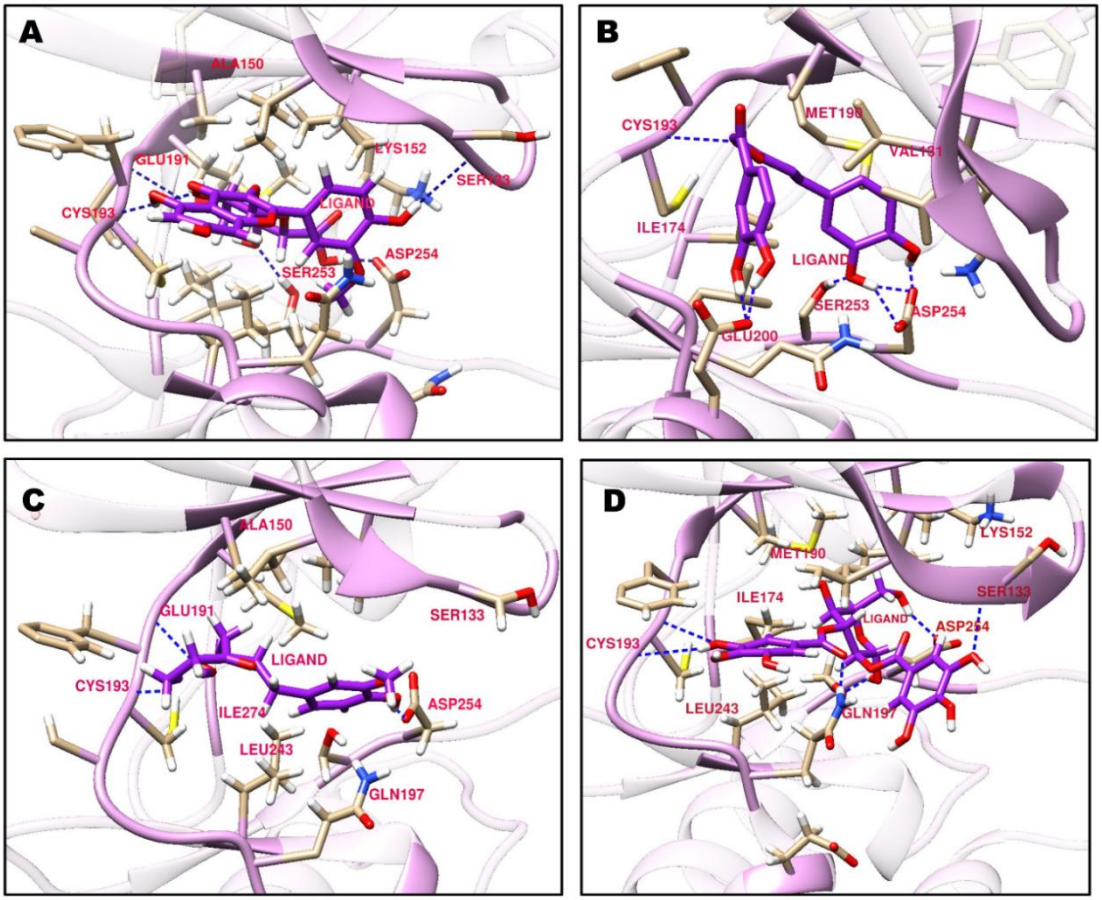


Figure 3.6. 3D interaction profile of DLK-Inhibitor complexes. **A** (lead-1; CID5491630), **B** (lead-2; CID16066851), **C** (lead-3; CID5318039), and **D** (lead-4; CID471118). The key residues are displayed by stick and ball model and the blue dotted lines represent the hydrogen bonds.

Additionally, throughout the simulations the occupancy of hydrogen bond between DLK and the inhibitors was calculated for each complex, as listed in **Table A5** of supporting information. Lead-1 formed a H-bond to DLK mainly with the residue Glu191 (70.39% and 81.68% occupancy in case of run1 and run2), Asp 254 (24.85% and 25.30% occupancy) and Ser 253 (18.32% and 24.26% occupancy); Lead-2 with Cys 193 (59.14% and 52.76%), Asp 254 (39.04% and 47.50%) and Ser 253 (36.33% and 20.68%); Lead-3 with Glu 191 (86.35% and 86.63%) and Cys 193 (51.59% and 57.24%); Lead-4 using Cys 193 (40.41% and 70.04%), Asp 236 (35.60% and 39.71%) and Asp 254 (14.84% and 35.26%). Only the top residues are mentioned here and all

the residue that forms H-bond with the ligands with ≥10% occupancy are listed in **Table A5**. It is evident from the **Table A5** that lead-4 showed the maximum number of residues that interacted with the ligand and is strongly favored by electrostatic interactions (-63.47 kcal/mol) then van der Waals interaction (-36.26 kcal/mol), along with lead-2 as reported in **Table 3.5**. Lead-3 showed minimum number of residues interacting with ligand and is strongly favored by van der Waals interaction (-37.74 kcal/mol) with slightly lesser electrostatic contribution (-25.55 kcal/mol) along with lead-1. Overall, our result suggest that lead-4 and lead-2 form stronger H-bond with DLK than the other two leads, which are mainly favored by van der Waals interaction.

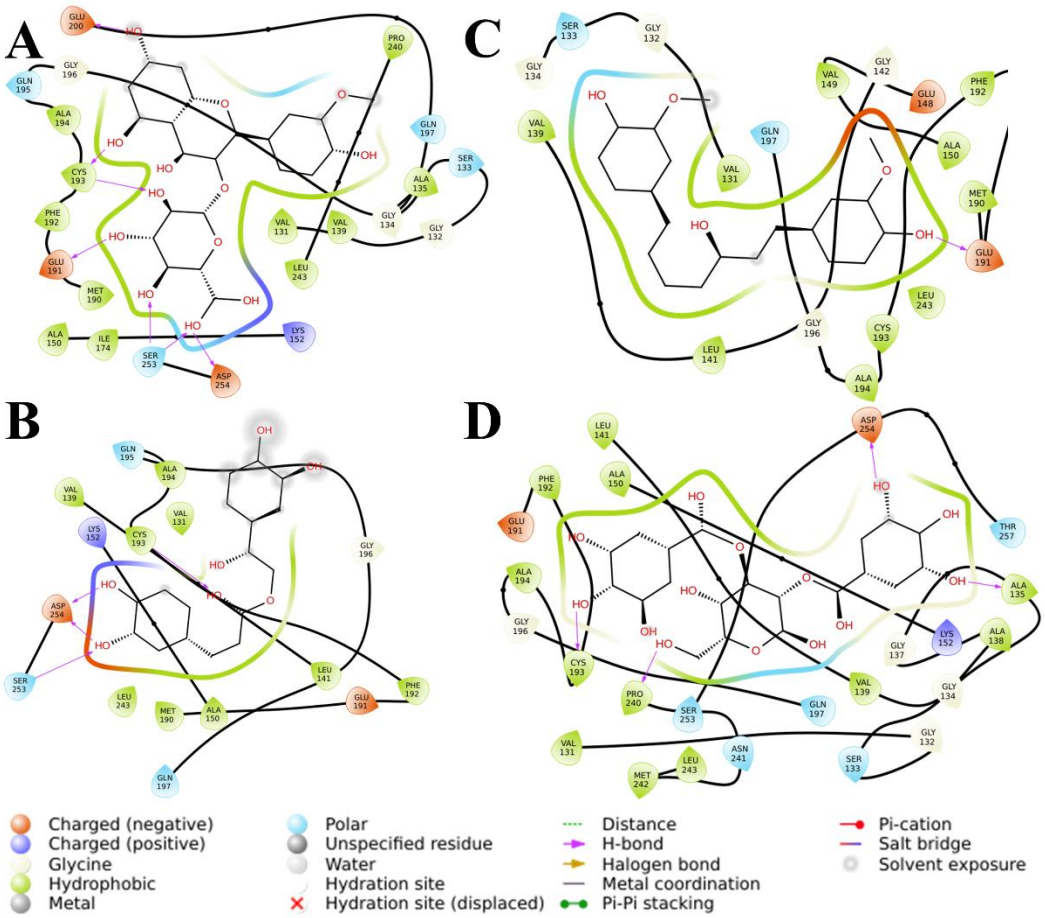


Figure 3.7. 2D interaction profiles of DLK-inhibitor complexes are depicted for **A** (lead-1), **B** (lead-2), **C** (lead-3), and **D** (lead-4). The purple (+ve) and brown (-ve) circles represent amino acids involved in electrostatic interactions. The green circle

indicates amino acids involved in hydrophobic interactions, the blue circle represents polar interactions, and the purple arrow signifies hydrogen bonds.

Additionally, we examined the hydrophobic interactions and hydrogen bond formation in the DLK-inhibitor complexes from structure taken from final simulation trajectory as shown in **Figure 3.6 and 3.7**. The 3D orientation of DLK-inhibitor complexes, which shows the binding residues and hydrogen bonds are shown in **Figure 3.6 (A-D)**. Using Schrodinger maestro, we computed 2D ligand-protein interaction map for all the lead molecules, as depicted in **Figure 3.7 (A-D),** which visualizes the eminent residues taking part in various interactions including hydrophobic interactions, hydrogen bonding, polar contributors among others in the final conformations of the simulations. From **Figure 3.7**, we observed some prominent residues involved in binding with ligands – Phe192, Val139, Leu243, Leu141, and Cys193 that formed the hydrophobic interactions with the ligands. Some common hydrogen bond forming residues on occurred between DLK-inhibitor complexes were as follows: lead-1 (Asp254, Ser253, Glu191, Cys193, Ser133), lead-2 (Glu200, Cys193, Ser253, Asp254), lead-3 (Asp254, Glu191) lead-4 (Gln197, Ser133, Ser253, Asp254). Asp254, Cys193, Glu200, Ser133, Gln197. This analysis also agrees with our more detailed hydrogen bond profile analysis above. Overall, this analysis of residual contribution to binding free energy helps us to find the prominent residues and may aid in finding novel inhibitors against DLK.

3.4. Conclusion

Saraswatarishta is an age-old Ayurvedic formulation primarily known for enhancing cognitive functions. Ayurveda often recommends it for conditions like delayed speech, cognitive deficits, and intellectual fatigue, among others. Our recent study aimed to explore the molecular interactions of phytochemicals in SWRT with the DLK receptor. To determine this, we conducted virtual screening of 534 phytochemicals from SWRT's 15 plant ingredients. This represents one of the first systematic approaches to classifying and identifying the molecular constituents of SWRT. Our workflow identified four lead phytochemicals that also demonstrated stability in interactions during our classical molecular dynamics simulations. We investigated the binding of

these four lead compounds in complex with DLK by assessing various structural parameters, such as ligand-protein distance, H-bond interactions, binding pocket RMSD, and binding free energy using the MMPBSA scheme. It is worth noting that each of the four leads originates from a distinct plant: *Anethum Sowa*, *Acorus calamus*, *Zingiber officinale*, and *Syzygium aromaticum*, respectively. All these molecules have already been identified as having potential in various medicinal conditions, which supports our findings. This research may help us narrow down the key ingredients from this age-old medicine and assist in formulating inhibitors against DLK. Our approach not only pinpointed the key ingredients of SWRT but also characterized the binding mode of DLK. To the best of our knowledge, this is one of the first studies to explore the use of plant derived phytochemicals against DLK. Overall, our strategy establishes the molecular framework of this Ayurvedic formulation in relation to DLK and paves the way for future drug development targeting various neurodegenerative diseases.

Chapter 4

Chapter 4

Identification of Inhibitors for Neurodegenerative Diseases Targeting Dual Leucine Zipper Kinase Through Virtual Screening and Molecular Dynamics Simulations

AIM: Identifying novel inhibitors and repurposing FDA-approved drugs for the treatment of neurodegenerative diseases, along with analysing inhibitor-induced structural plasticity.

This chapter is reproduced from our original article:

Koirala, S., S. Samanta, and P. Kar. "Identification of inhibitors for neurodegenerative diseases targeting dual leucine zipper kinase through virtual screening and molecular dynamics simulations." *SAR and QSAR in Environmental Research* 35.6 (2024): 457-482.

4.1. Introduction

Neurological diseases significantly burden the entire world, affecting millions of people. Factors like an aging population, changing lifestyles, and limited access to healthcare contribute to the rise in neurological disorders. Globally, conditions like stroke, dementia, and neurodegenerative disorders are major public health concerns [277]. Arising from nervous system irregularities, neurological disorders enclose diverse maladies. They range from neurodegenerative conditions like Alzheimer's and Parkinson's to neurodevelopmental conditions such as autism and attention deficit hyperactivity disorder (ADHD), along with debilitating ailments like multiple sclerosis, epilepsy, and migraines [278]. Neurological diseases are associated with numerous signaling pathways that include a diverse range of proteins [279]. Targeting these proteins/signaling represents a promising strategy to address and overcome such diseases.

The dual leucine zipper kinase (DLK) pathway is crucial for mediating neuronal responses to stress and injury, and dysregulation of this pathway has been associated with the onset and progression of several neurodegenerative diseases [67]. DLK is a serine/threonine protein kinase that belongs to the mixed lineage kinase (MLK) subfamily and is a multifunctional kinase that plays a role in various cellular processes, including cell differentiation, apoptosis, and stress response [242]. MLK, as members of the MAP3K family, initiate a phosphorylation cycle that includes phosphorylating MKKs or MAP2Ks, leading to JNK activation. JNKs then phosphorylate transcription factors like c-jun, which can trigger axonal degeneration and apoptosis (see **Figure 4.1**). This signaling pathway is a key mediator of stress response in various cell types, including neurons, and often results in apoptosis [243]. As a key player in the neuronal stress response, DLK plays a vital role in regulating neurodegeneration in both acute neuronal injury and chronic neurodegenerative diseases such as Alzheimer's, Parkinson's, and Amyotrophic lateral sclerosis (ALS), etc [78]. Several studies involving the knockdown of DLK have indicated that it protects against neurodegeneration. The Genentech research team also confirms its role in neurodegeneration [66,126]. There are several FDA-approved drugs available for specific neurological diseases [280], but there are currently no FDA-approved drugs that target DLK. As it is associated with multiple neurological disorders and has an important role in stress-induced neurodegeneration, targeting DLK could be an optimistic strategy for addressing neurological conditions as a whole. As there are no FDA-approved drugs against DLK, we selected sunitinib as a reference compound to compare as it has good potency to inhibit DLK [281].

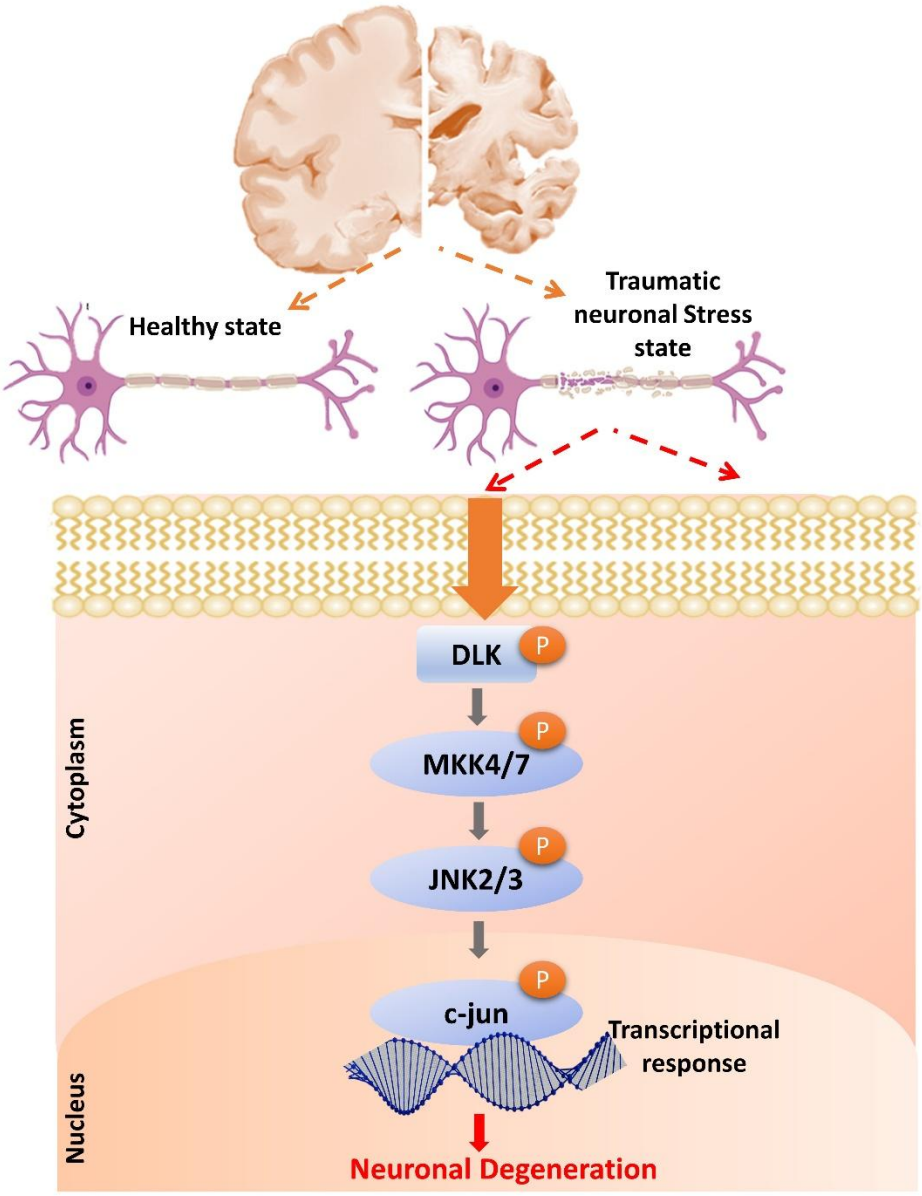


Figure 4.1: DLK signaling overview in neurons. Various cellular stress triggers the activation of DLK, which subsequently phosphorylates and activates downstream proteins, including the transcription factor c-jun. The activation of these transcription factors within neurons triggers stress-induced neuronal degeneration through transcriptional responses.

Several compounds have been explored as inhibitors against DLK in earlier investigations [12,13]. Genentech has recently identified GNE-3511 and GNE-8505 as

potent DLK inhibitors with good central nervous system (CNS) penetration as well as drug metabolism and pharmacokinetics (DMPK) properties [71,78,125]. Additionally, GDC-0134 underwent clinical development for ALS treatment but was ultimately terminated due to safety issues [126]. Recently, Denali Therapeutics reported DN1289 [127], and Kang Le et al. reported IACS-52825 and IACS-8287 as potent DLK inhibitors [128]. The lack of FDA-approved drugs and limited DLK inhibitors identified emphasizes the need to discover novel potent inhibitors along with drugs that can be repurposed against DLK. The present study aims to identify competitive inhibitors against DLK, considering its significance as a potential target against various neurodegenerative diseases. Apart from kinases, there are several other targets that are well-addressed for neuroprotection along with various therapeutic strategies [282–287].

Herein, we used multiple computational approaches like structure-based virtual screening, ADMET analysis, and all-atom molecular dynamics simulations to identify potent novel inhibitors and repurposed drugs against DLK. In this work, we screened ~28000 compounds from the Natural Products Atlas database and 2697 drugs from the MedChemExpress database (FDA-approved) via the structure-based virtual screening against DLK. The MedChemExpress FDA-approved drugs database was considered for identifying whether those can be served as potential inhibitors against DLK as a drug-repurposing strategy. After virtual screening, we assessed the ADMET profiles of the top 80 hit compounds. Subsequently, we refined the binding poses and structures of each complex through molecular dynamics (MD) simulations. Finally, we calculated the binding free energy against DLK using the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method.

4.2. Materials and Methods

4.2.1. Ligand Preparation

The current investigation employed a computational screening approach targeting the ATP-binding pocket present in the kinase domain of the DLK protein. We utilized two comprehensive databases as sources of potential compounds: the Natural Products Atlas database [288] (https://www.npatlas.org/download; ~28000 compounds) and the

MedChemExpress database of FDA-approved pharmaceuticals (https://www.medchemexpress.com/screening/FDA-approved_Drug_Library.html.; ~2697 compounds) The Schrödinger Glide program [289] was utilized to import the ligands and subject them to the Ligprep module [290], which generated all possible conformations and tautomeric states of the ligands under pH 7.0 ±2.0. Following the incorporation of hydrogen atoms, the ligands underwent a procedure of optimization and minimization using the OPLS3e [155] force field.

4.2.2. Protein Preparation and Virtual Screening

Our study utilized the crystal structure of the DLK kinase domain (PDB ID 5CEO) as the protein target, with a resolution of 2.28 Å [250]. Missing residues were added via the Modeller plugin of UCSF Chimera, and the Schrödinger suite protein preparation wizard was employed to prepare the protein for docking [291]. To optimize our molecular docking process, we added missing hydrogen atoms, assigned charges, and generated Het states at a pH range of 7.0 ± 2.0 using Epik. The protein was prepared for molecular docking by deleting all molecules except the co-crystallized ligand and essential water molecules and subsequently minimized using the OPLS3e force field [155]. The co-crystallized ligand present in the active site of the protein target was utilized as a reference point to generate a grid for molecular docking. We centered the grid box for molecular docking on the selected ligand's centroid, with protein atoms' van der Waals radii scaled to 1.00 Å and a partial atomic charge cut-off of 0.25 set for polarity. For virtual screening, the two datasets were first filtered through QikProp module and Lipinski's rule of five, followed by screening in high throughput virtual screening (HTVS; 70% compounds), standard precision (SP; 50% compounds) and extra precision (XP; 30% compounds) docking protocols using Schrödinger Maestro. The screening was done separately for two databases and the ligands with the top docking scores were selected for subsequent ADMET prediction analysis.

4.2.3. Estimation of ADMET Properties

ADMET studies examine a drug candidate's ability for effective absorption, distribution, metabolization, excretion, and toxicity inside the body, ensuring their viability for the development of therapeutics. SwissADME [292] and ADMETlab 2.0

[270] were utilized to predict the ADME properties of the selected ligands. pkCSM [293] was used to predict the toxicity of the selected ligands. Finally, ligands with favorable ADMET features were chosen for MD simulations and subsequent analyses.

4.2.4. Molecular Dynamics Simulation Protocol

To explore the dynamics of the selected five complexes along with apo, we conducted long replica molecular dynamics (MD) simulations using AMBER18 software under periodic boundary conditions [148]. For each complex, two independent simulations with different initial velocities were run for 250 ns. The characterization of proteins in our study was accomplished using the Amber ff14SB force field [153], while the small molecules were modelled using the updated generalized Amber force field (GAFF2) [163]. The solvent for our simulations was comprised of water molecules positioned at a buffer distance of 10.0 Å in all directions from the protein's surface. We described the water molecule parameters using the TIP3P force field to ensure accurate representation [172]. In order to neutralize the system, we added the appropriate number of ions. Further, we used the SHAKE algorithm to fix all covalent bond lengths, including those involving hydrogen atoms [294]. To compute the long-range electrostatic interaction, we used the particle-mesh Ewald (PME) method [263]. We set the non-bonded cut-off distance to 10.0 Å for this calculation. We employed a 2.0 fs time step in our MD simulations and used both the steepest descent and conjugate gradient algorithms for a two-stage energy minimization. In the first minimization stage, the steepest descent algorithm was applied for 5000 cycles, followed by the conjugate gradient algorithm for another 5000 cycles, with a harmonic restraint of 2 $kcal.mol^{-1}$ $Å^{-2}$ to the whole system. The second unconstrained minimization was performed through 100 steepest descent cycles followed by 900 conjugant gradient cycles. The systems were heated from 0 to 300 K in the NVT ensemble. Throughout the simulation, we utilized the Langevin thermostat [295] and Berendsen Barostat [180] to ensure a constant temperature of 300 K and pressure of 1 bar in the system. We proceeded to production simulation run of 2x250 ns for each system under the isothermal-isobaric ensemble (NPT) after completing the prior simulation steps. We collected MD trajectories every 10 ps to use for later analyses, producing a total of

25,000 snapshots for each replica run. The Cpptraj module [264] of Amber18 was employed for all post-simulation trajectory analyses.

4.2.5. Binding Free Energy Using the MM-PBSA Scheme

The binding free energy (ΔG_{bind}) between protein and ligands was calculated using the well-known molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method. The calculation of ΔG_{bind} was performed using the MMPBSA.py script. The MM-PBSA scheme has been extensively explained in previous research articles from our lab [296–301]. In the current investigation, the same procedure was used. The binding free energy (ΔG_{bind}) can be calculated by summing up three distinct energy contributions, namely the molecular mechanics energy (ΔE_{MM}), solvation energy (ΔG_{solv}), and entropy term ($-T\Delta S$), as outlined in equation (1).

$$\Delta G_{\mathrm{bind}} = \Delta H - T\Delta S \approx \Delta E_{\mathrm{MM}} + \Delta G_{\mathrm{solv}} - T\Delta S \quad (1)$$

ΔE_{MM} was computed by adding up the covalent (ΔE_{cov}) electrostatic (ΔE_{elec}) and van der Waals interaction energy (ΔE_{vdW}) terms (equation (2)). ΔG_{solv}, which includes both polar and nonpolar solvation free energy, was determined according to equation (3), respectively.

$$\Delta E_{\mathrm{MM}} = \Delta E_{\mathrm{cov}} + \Delta E_{\mathrm{elec}} + \Delta E_{\mathrm{vdW}} \quad (2)$$

$$\Delta G_{\mathrm{solv}} = \Delta G_{\mathrm{pol}} + \Delta G_{\mathrm{np}} \quad (3)$$

The entropic contribution was estimated by means of normal-mode analysis. In addition, the MM-PBSA scheme was used to perform per-residue-based energy decomposition.

4.2.6. Inter-Residual Correlated Motion and Principal Component Analysis

To better understand the internal dynamics of apo and lead-bound DLK residues, we generated a dynamic cross-correlation matrix DCCM [302] using the Bio3D [303] package of R in RStudio, which allowed us to explore the correlations and anti-correlations between C_α atom-pairs of DLK residues. Correlated motion between residues is indicated by positive values on the map, while negative values correspond to anti-correlated movements. PyMOL software was utilized to visualize the motion.

To obtain a deeper understanding of the most likely conformations of DLK in various lead bound systems and in the apo system, we performed principal component analysis (PCA) [304], a dimension reduction technique taking C_α atom coordinates of the protein's residues from the simulation trajectories. Additionally, we created a 2-dimensional free energy landscape (FEL) for the DLK-inhibitor systems and the apo DLK utilizing the first two principal components (PC1 and PC2) as reaction coordinates.

4.2.7. Protein Structure Network Analysis

To investigate the complex residual network within a protein's structure, we utilized the webPSN v2.0 [221] webserver, which incorporates the protein structure network (PSN) [265] and elastic network model-normal mode analysis (ENM-NMA) into its methodology. The identification of the shortest communication path and network properties, including nodes, hubs, edges, linkages, and communities, was accomplished using PSN analysis on stable states of MD trajectories. The identification and visualization of the network hubs and communities were accomplished through the utilization of VMD software [266]. Prior publications from our lab have extensively detailed the PSN methodology utilized in the study [267].

4.2.8. Optimization of CID156581477 Through Structural Analogues

Using the DeLA-Drug [305] webserver, we employed a sampling with substitutions method and deep learning algorithm to create 100 structural analogues of CID156581477 to improve its binding affinity. After generating 100 analogues of CID156581477, we conducted virtual screening against DLK using the Schrödinger suite using various parameters as described earlier. We then selected the top two leads based on their docking score to further analyze their potential as DLK inhibitors. We then conducted replica runs of 250 ns atomistic molecular dynamics simulations on the DLK complexes of these analogues to evaluate their structural stability. Moreover, we performed MM-PBSA-based binding free energy calculations using 2000 conformations extracted from the simulation trajectory at regular intervals during the last 100 ns.

4.2.9. Molecular Target Prediction

Natural compounds and drugs interact with a large number of proteins and enzymes. These interactions between small molecules and proteins/enzymes can aid in unraveling the molecular mechanisms underlying their inhibitory or modulatory effects [306]. By accessing the Swiss Target Prediction site (http://www.swisstargetprediction.ch/index.php), we input the canonical SMILES of studied ligands in the search bar and subsequently analysed the obtained results.

4.3. Results and Discussion

4.3.1. Virtual Screening of Natural Products and FDA-approved Compounds

Initially, we retrieved the crystal structure of DLK (PDB ID-5CEO) from the Protein Data Bank and prepared it for virtual screening. **Figure 4.2(A)** illustrates the kinase domain of DLK, highlighting the binding pocket and different regions of interest. Secondly, we selected the ligand database for screening. The virtual screening process involved two databases: the Natural Products Atlas database (~28,000 compounds) and the MedChemExpress database (2,697 drugs) consisting of FDA-approved drugs. The two datasets were first filtered through QikProp module of Schrödinger and the Lipinski's rule of five, followed by HTVS (70%), SP (50%), and XP (30%) docking protocols. The top compound in the natural products dataset exhibited the lowest docking score of -12.57 kcal/mol, while the best FDA-approved drug gave a docking score of -10.97 kcal/mol. We selected compounds from the natural products dataset with docking scores lower than -10 kcal/mol and compounds from the FDA-approved dataset with docking scores lower than -9.00 kcal/mol for further analysis. The complete workflow of the virtual screening process, followed by MD simulation and other analyses, leading to the finding of potent inhibitors against DLK, is depicted in **Figure 4.2(B)**.

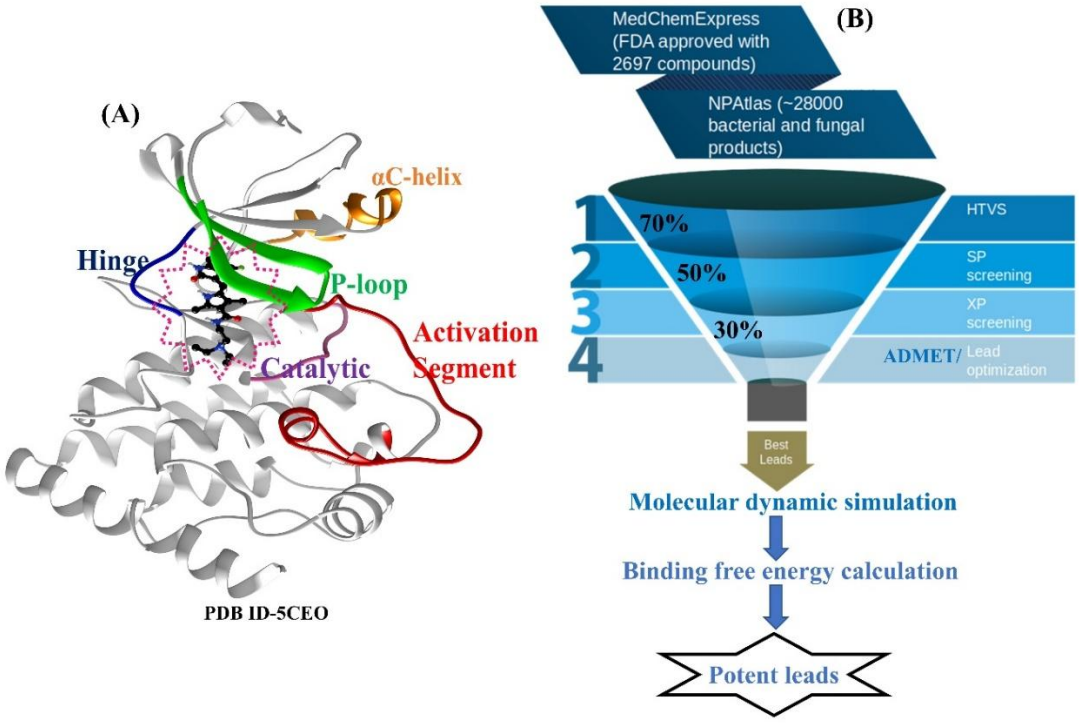


Figure 4.2: **(A)** The crystal structure of DLK, highlighting the binding pocket and various other crucial regions. **(B)** A visual representation of the methodology, which includes virtual screening, ADMET prediction, molecular dynamics simulations, to lead molecule identification.

4.3.2. Drug-like Properties of the Compounds

ADMET analyses are crucial for identifying compounds with advantageous features that make them suitable for further trials in drug development. Here, we evaluated the ADMET properties of 80 compounds from the natural product dataset, as well as the reference compounds sunitinib (control), and the other previously reported DLK-inhibitors (**GNE-3511, CDC-0134, DN1289, and IACS-52825)**. A primary focus on initial selection for the novel compounds were driven by blood-brain barrier (BBB) permeability prediction which is crucial for targeting neurodegeneration. Among all, only two newly identified compounds exhibited favourable overall characteristics (see **Figure 4.3**). As indicated in **Table B1**, these two molecules exhibit a molecular weight well below 500 g/mol similar to the reference compounds, signifying enhanced potential for absorption and permeability across the blood-brain barrier (BBB). These

molecules conform to all established rules, including Lipinski's, Ghose's, Veber's, Egan's, and Muegge's, further supporting their suitability for pharmacological and therapeutic development (see **Table 4.1**).

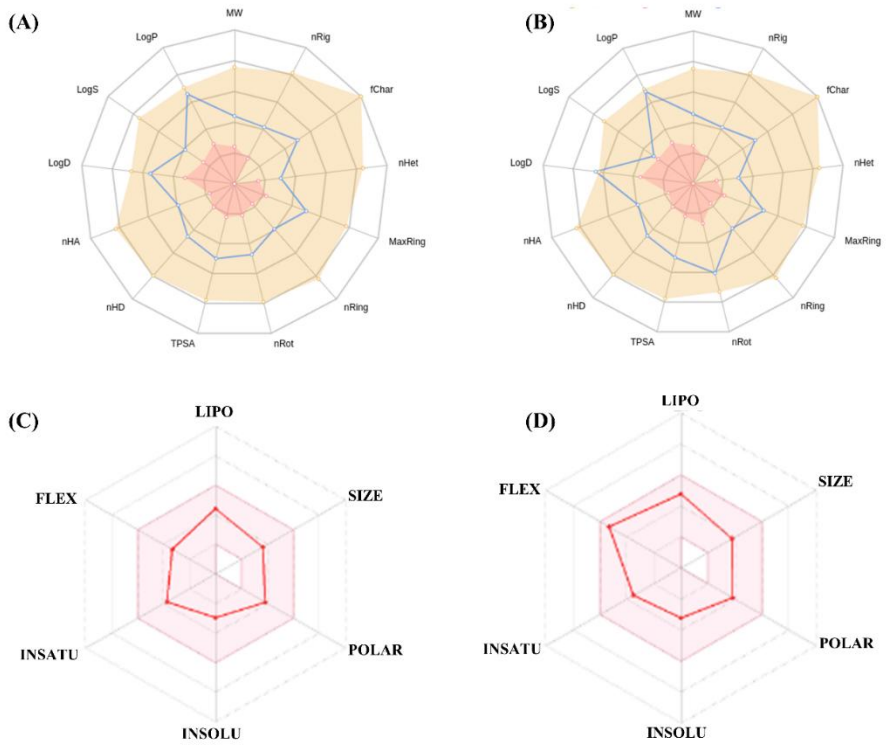


Figure 4.3: Overall ADMET properties of two studies compounds. **(A)** and **(B)** are predicted through ADMETlab 2.0 webserver, while **(C)** and **(D)** are predicted through the SwissADME webserver. **(A)** and **(C)** represents CID156581477, and **(B)** and **(D)** represents CID139591660.

However, the reference compound IACS-52825 violated several of the standard rules. The results shown in **Figure B1** indicate that the two investigated compounds, CID156581477 and CID139591660, are part of yellow Egan eggs, indicating a higher possibility of blood-brain barrier (BBB) permeability. Among the reference compounds, only GNE-3511 showed BBB-permeability (see **Table B2**). Both of the novel compounds were predicted as non-substrates of p-gp, indicating that these will have low probability to be expelled out of the cell by this ATP binding cassette (ABC) efflux membrane transporter. Except for GDC-0134, the other reference compounds

were predicted to be substrates of p-gp. Further, these two novel compounds were non-inhibitors of the cytochrome P450 (CYP) isozymes CYP1A2, CYP2C19, CYP2C9, and CYP3A4, except CYP2D6 according to pharmacokinetics analysis (see **Table B2**).

Table 4.1: The determination of whether a significant regulatory rule governing the classification of compounds as drug-like is approved or denied. The properties are predicted using the SwissADME webserver. The properties is calculated for two new identified compounds, control, and few reported DLK inhibitors.

Compounds CID	Lipinski	Ghose	Veber	Egan	Muegge	Bioavailability Score
156581477	Yes (0 violation)	Yes	Yes	Yes	Yes	0.55
139591660	Yes (0 violation)	Yes	Yes	Yes	Yes	0.55
Sunitinib	Yes (0 violation)	Yes	Yes	Yes	Yes	0.55
GNE-3511	Yes (0 violation)	Yes	Yes	Yes	Yes	0.55
GDC-0134	Yes (0 violation)	Yes	Yes	Yes	Yes	0.55
DN1289	Yes (0 violation)	Yes	Yes	Yes	Yes	0.55
IACS-52825	Yes (0 violation)	No	Yes	No	No	0.55

In addition, we used the pkCSM webserver to anticipate the toxicity profiles of the two hit compounds along with other reference compounds. To assess the molecular toxicity of these compounds, we predicted a range of toxicological properties, including hepatotoxicity, AMES toxicity, acute and chronic oral rat toxicity, hERG inhibition, skin sensitization, among others (see **Table B3**). **Table B3** clearly shows that both

compounds successfully clear the toxicity assessment, exhibiting minimal or negligible toxicity for all toxicological properties. All reference compounds were estimated to be hepatotoxic in the toxicity prediction. For the FDA-approved database, we analyzed BBB-permeability for the top nine drugs, and only two (Dithranol and Danthron) were selected for further study because of their capacity to cross the blood-brain barrier (BBB) (see **Table 4.2**). Overall, these four compounds (two FDA-approved and two natural compounds) possess favorable overall ADMET properties and may act as potent DLK inhibitors for neurodegenerative disease treatment.

Table 4.2: Blood-brain barrier (BBB) penetration of approved drugs using SwissADME webserver. In the context of repurposing approved drugs for neurodegenerative diseases, predicting BBB penetration is essential as these drugs must effectively cross the BBB to reach the central nervous system (CNS).

Sl no	Drug Name	Ml. weight	Docking score	BBB penetration	Associated disease
1	Mitoxantrone	444.50	-10.97	No	Prostate cancer
2	Trifluridine	296.20	-10.55	No	Antiviral medicine
3	Ipragliflozin	404.50	-10.43	No	SGLT-2 inhibitor (diabetes)
4	Dithranol	226.23	-9.98	Yes	Psoriasis
5	Brivudine	333.13	-9.72	No	Antiviral
6	Maribavir	376.20	-9.67	No	Cytomegalovirus (CMV) infection
7	Catechin	290.27	-9.42	No	Anticancer
8	Olaparib	434.50	-9.38	No	Ovarian and fallopian tube cancer
9	Danthron	240.21	-9.38	Yes	Inflammation of the oral cavity and oropharynx
10	Idoxuridine	354.10	-9.37	No	Eye infections

11	Ozenoxacin	363.40	-9.23	No	Impetigo (skin infection)
12	Ruxolitinib	306.40	-9.11	No	Myelofibrosis
13	Clevudine	260.22	-9.00	No	Treatment of hepatitis B

4.3.3. Exploring Inhibitor-Binding and Conformational Stability Through Molecular Dynamics Simulation

Molecular dynamics simulation is a reliable approach to investigate the interactions and structural stability between small molecules and target proteins within a defined physiological environment. In this study, we conducted 2x250 ns MD simulations of the top four hit molecules. The stability of each simulation was evaluated by examining the root-mean-square deviation (RMSD) of the backbone atoms over the 250 ns simulation period. Each independent replica runs showed similar deviations in all complexes along with apo and converged within the simulation period (see **Figure B2**). Furthermore, we also computed the RMSD distribution for the combined trajectory of the two replica runs (see **Figure 4.4(A)**). It is clear from **Figure 4.4(A)** that the backbone RMSD distribution of CID156581477/DLK showed restricted fluctuations with the most probable peak at ~2.00 Å, while apo and CID139591660/DLK exhibited a single peak with broad distribution. Two probable peaks of RMSD are shown in the case of sunitinib/DLK (~2.3 Å, ~2.8 Å) and danthron/DLK (~1.8 Å, ~2.3 Å). The dithranol/DLK complex displayed a slightly broader distribution around two probable peaks (~1.70 Å, ~2.40 Å).

The root-mean-square fluctuation (RMSF) of each residue was computed to explore the residual fluctuations in each complex structure as well as the apo DLK (see **Figure 4.4(B)**). We observed a similar trend of residual fluctuations in all complexes except CID139591660/DLK in the case of the αC-helix region. In all cases, the loop regions consistently displayed high flexibility, with the activation segment region exhibiting the highest degree of fluctuation (see **Figure 4.4(B)**).

Subsequently, the solvent-accessible surface area (SASA) and the radius of gyration (R_g) of each complex were measured, and their probability distribution plot is illustrated in **Figure B3**. The most probable peaks for SASA are - apo (485 Å), sunitinib (445 Å), CID156581477 (465 Å), CID139591660 (440 Å), danthron (460 Å), and dithranol (470 Å) (see **Figure B3(A)**). A slightly higher SASA value was observed for apo as compared to the studied ligands, suggesting the displacement of water from the binding pocket upon ligand binding. Additionally, after ligand binding, we observed slight changes in the compactness of DLK as indicated in the probability distribution of R_g (see **Figure B3(B)**). These findings indicate that ligand binding results in a marginal alteration in protein compactness and a decrease in solvent accessibility.

To track the movement of ligands in the binding pocket, we measured the distance between the center of mass (CoM) of the ATP-binding pocket residues (residues within 5 Å distance from the ligands in the initial conformation) and the heavy atoms of the ligands throughout the simulation course. Supplementary information **Figure B4** shows the temporal changes in each ligand-protein distance, while **Figure 4.4(C)** exhibits the probability density of the combined runs. All the ligands maintained a stable interaction with the proteins, and the most probable peaks were observed at - CID156581477 (1.30 Å), CID139591660 (0.60 Å), dithranol (1.80 Å), danthron (2.20 Å), and the control sunitinib also showed a single peak with wide distribution (See **Figure 4.4(C)**). Interestingly, all ligands exhibited a notable peak at distances below 5.0 Å from the initial binding region (ATP-binding pocket), indicating their strong affinity towards DLK. Furthermore, to examine the dynamics of the ligands within the binding pocket of DLK, we computed the potential of mean force (PMF) for each one based on the corresponding ligand RMSD values (see **Figure 4.4(D)**). As shown in **Figure 4.4(D)**, a single narrow peak was obtained for CID156581477 (~1.20 Å), dithranol (~0.40 Å), and danthron (~0.20 Å), while a peak with a broad base was observed for CID139591660 (~1.10 Å). The control sunitinib showed two probable peaks, one narrow peak (~2.20 Å) and one broad peak (~0.70 Å). This ligand PMF study indicates that all four studied ligands exhibit greater stability within the binding pocket of DLK as compared to the control sunitinib. Additionally, we have also

calculated the torsion angle profile of new identified compounds CID156581477, and CID139591660, which is reported in supplementary information **Figure B5 and B6**. The rotation around a single bond allows a ligand to adopt an optimal conformation, which can be analysed using the torsion angle profile. Distinct torsion profiles were observed for the two ligands around several studied bonds. For CID156581477, the C4-C5 (brown) and C14-C15 (yellow) bonds exhibited one prominent peak with one secondary hump each, indicating one major and one minor conformation based on the rotation around these bonds (see **Figure B5**). The torsion angle profiles of the other studied bonds showed a distribution around a single minimum. In contrast, for CID139591660, a larger degree of variation in the studied rotatable bonds was observed (see **Figure B6**). This indicates a higher degree of change in orientation of the rotatable bonds within the binding pocket for CID139591660. These calculations provide valuable insights into the ligand's conformational preferences and binding interactions.

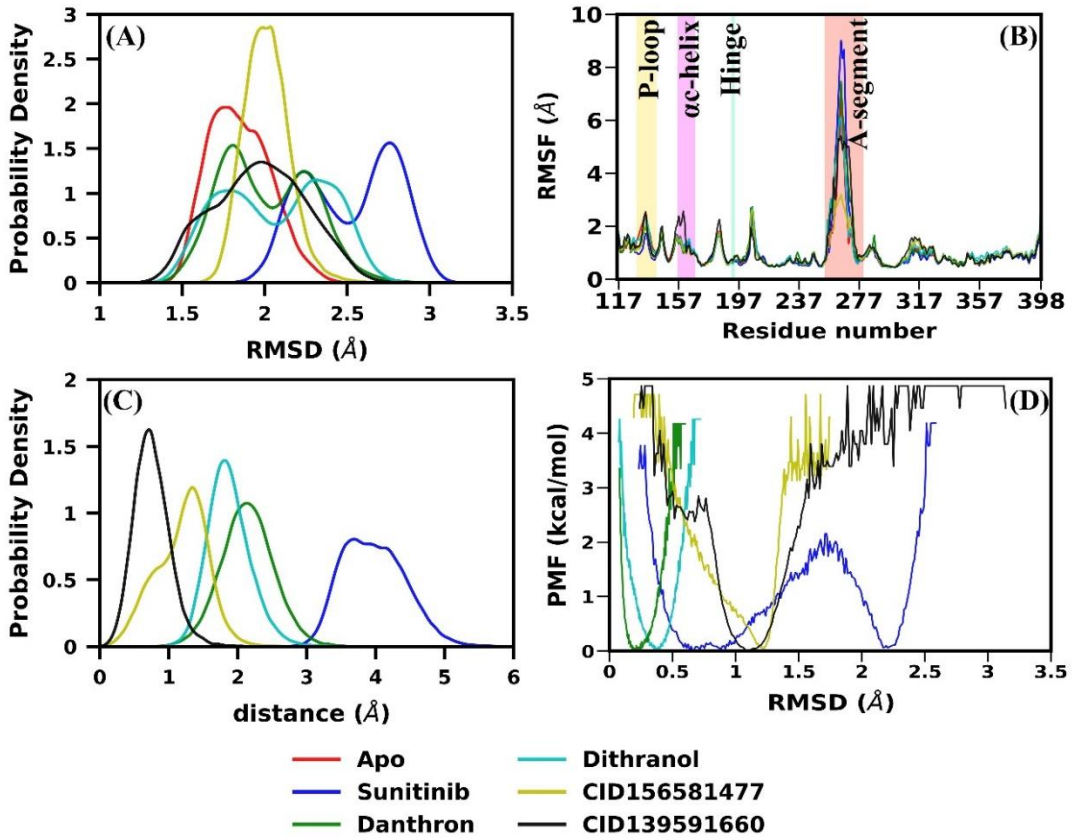


Figure 4.4: (A) The probability distribution of root-mean-square deviation (RMSD) of the entire protein backbone of the DLK kinase domain. **(B)** The individual residue-wise root-mean-square fluctuation (RMSF) of the C_α-atoms in DLK. **(C)** The probability distribution of the ligands' distance from the center of mass of residues within 5 Å of the ligands in the initial docked conformation. **(D)** The potential of mean force (PMF) for each of the five inhibitors was computed using ligand RMSD as the reaction coordinate for each DLK-bound inhibitor.

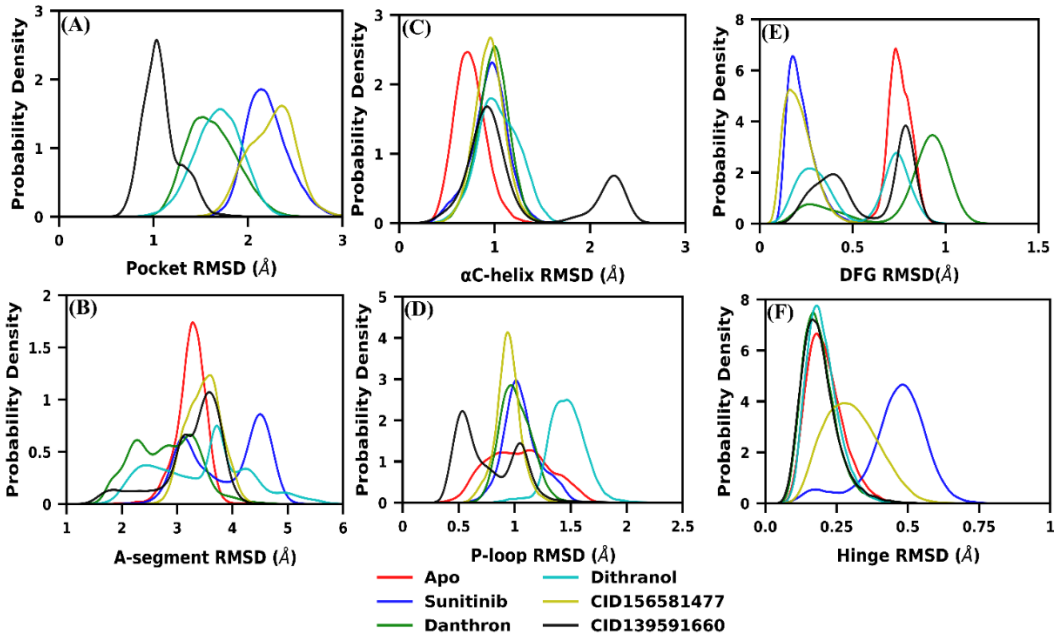


Figure 4.5: The RMSD distribution of backbone atoms of **(A)** binding pocket **(B)** activation segment **(C)** αC-helix **(D)** P-loop **(E)** DFG-motif **(F)** hinge region.

Furthermore, we also investigated the impact of ligand binding on the conformational dynamics of DLK. The probability density distribution of backbone-atoms RMSD was plotted for important regions of DLK proteins, including the binding pocket, P-loop, hinge region, αC-helix, activation segment, and DFG-motif region (see **Figure 4.5**). As depicted in **Figure 4.5(A)**, the binding pocket residues of DLK undergo some conformational rearrangement after the binding of different ligands. Overall, the binding region remained stable and rigid, as indicated by a single peak in each case with RMSD values below 3 Å. After ligand binding, the activation segment of DLK displayed multiple conformations with diverse structural rearrangements, indicating

the highly dynamic nature of the activation segment (see **Figure 4.5(B)**), which is also consistent with our previous RMSF analysis (see **Figure 4.4(B)**). After ligand binding, the αC-helix of DLK underwent minimal conformational alteration where all systems had rigid and stable αC-helix forming a single peak. At the same time, CID139591660/DLK gave a bimodal distribution representing two prominent structural conformations (see **Figure 4.5(C)**). After complexation, the P-loop demonstrated enhanced rigidity and stability, as seen from the narrow peak in its RMSD distribution in the complexed structures (see **Figure 4.5(D)**). The DFG motif showed bimodal distributions in the case of danthron/DLK, dithranol/DLK, and CID139591660/DLK, while a single conformation was observed for apo, sunitinib/DLK, and CID156581477/DLK (see **Figure 4.5(E)**). The hinge region, which is structurally conserved, remained stable and rigid for all systems throughout the simulation, as indicated by a single peak in each case (see **Figure 4.5(F)**).

4.3.4. Exploring DLK Inhibitors Binding: Energetics and Key Residue Identification

MD simulations along with free energy calculations, represent a suitable approach to investigate the stability of ligand binding with proteins and the strength of their interactions. To understand the binding mechanism of compounds against DLK, we calculated specific contributions of free energy for each compound, as displayed in **Table 4.3**. The formation of the complex is favored by intermolecular electrostatic (ΔE_{elec}) and van der Waals (ΔE_{vdW}) interactions, along with the non-polar solvation free energy (ΔG_{np}). However, complexation is disfavoured by the polar solvation free energy (ΔG_{pol}) and configurational entropy ($-T\Delta S$). The average binding affinity was calculated from the two replica runs in all four complexes (DLK/danthron, DLK/dithranol, DLK/CID156581477, DLK/CID139591660) and compared with the control complex DLK/sunitinib. The obtained binding energy of all complexes are as follows - DLK/CID139591660 (-12.80 kcal/mol), DLK/dithranol (-12.55 kcal/mol), DLK/danthron (-11.10 kcal/mol), DLK/sunitinib (-9.45 kcal/mol), and DLK/CID156581477 (-6.09 kcal/mol). Since significantly lower binding affinity was observed for DLK/CID156581477 (-6.09 kcal/mol) as compared to the control

DLK/sunitinib (-9.45 kcal/mol), DLK/CID156581477 was excluded from further analyses. As listed in **Table 4.3**, for all complexes, the range of van der Waals interactions' contribution to the binding free energy varied from -42.90 kcal/mol to -26.58 kcal/mol. The favorable electrostatic interaction (ΔE_{elec}) component was overcompensated by the polar solvation free energy (ΔG_{pol}), leading to an overall binding-unfavorability from the net-polar ($\Delta E_{elec}+\Delta G_{pol}$) component of the binding free energy (see **Table 4.3**). The non-polar solvation energy (ΔG_{np}) was relatively consistent across all complexes, while the configurational entropy term ($-T\Delta S$) ranged from 14.14 kcal/mol to 22.31 kcal/mol. Van der Waals interactions served as the main force for the binding in the complexes. Overall, CID139591660, danthron, and dithranol exhibited greater binding affinity as compared to the control drug, sunitinib, and were considered for further analyses.

Table 4.3: Binding free energy components (kcal/mol) calculated for DLK-inhibitors complexes from the MM-PBSA scheme. the standard errors of the mean are included in parentheses.

DLK/ligand complex	ΔE_{vdw}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	$-T\Delta S$	ΔG_{bind}
DLK/Sunitinib	-42.90 (0.07)	-47.53 (0.36)	66.19 (0.35)	-4.14 (0.00)	18.93 (0.78)	**-9.45 (0.88)**
DLK/Danthron	-27.08 (0.05)	-18.16 (0.07)	22.30 (0.29)	-2.81 (0.00)	14.65 (0.75)	**-11.10 (0.81)**
DLK/Dithranol	-26.58 (0.05)	-13.39 (0.06)	16.03 (0.04)	-2.75 (0.00)	14.14 (0.97)	**-12.55 (1.02)**
DLK/CID156581477	-33.09 (0.05)	-30.86 (0.12)	40.09 (0.08)	-3.61 (0.00)	21.38 (0.80)	**-6.09 (0.88)**
DLK/CID139591660	-37.53 (0.08)	-38.43 (0.17)	44.75 (0.10)	-3.90 (0.00)	22.31 (0.67)	**-12.80 (0.76)**

Furthermore, to explore the impact of crucial residues on the binding free energy, the MM-PBSA method was used to calculate the per-residue decomposition of free energy

(see **Figure 4.6(A-D)** and supplementary **Table B4**). This method predicts energy contributions of residues involved in ligand binding and identify hot-spot residues. Glu191, Phe192, Leu243, Val139, and Leu141 are some of the critical residues within DLK's binding pocket that interacted with the selected inhibitors (see **Figure 4.6(A-D)**). The binding patterns of dithranol and danthron are notably similar, as evidenced by the similarity of the residues involved in their interactions (see **Figure 4.6(B-C)**). The identification of these key residues could aid in the development of novel ATP-competitive inhibitors targeting DLK for the treatment of neurological disorders.

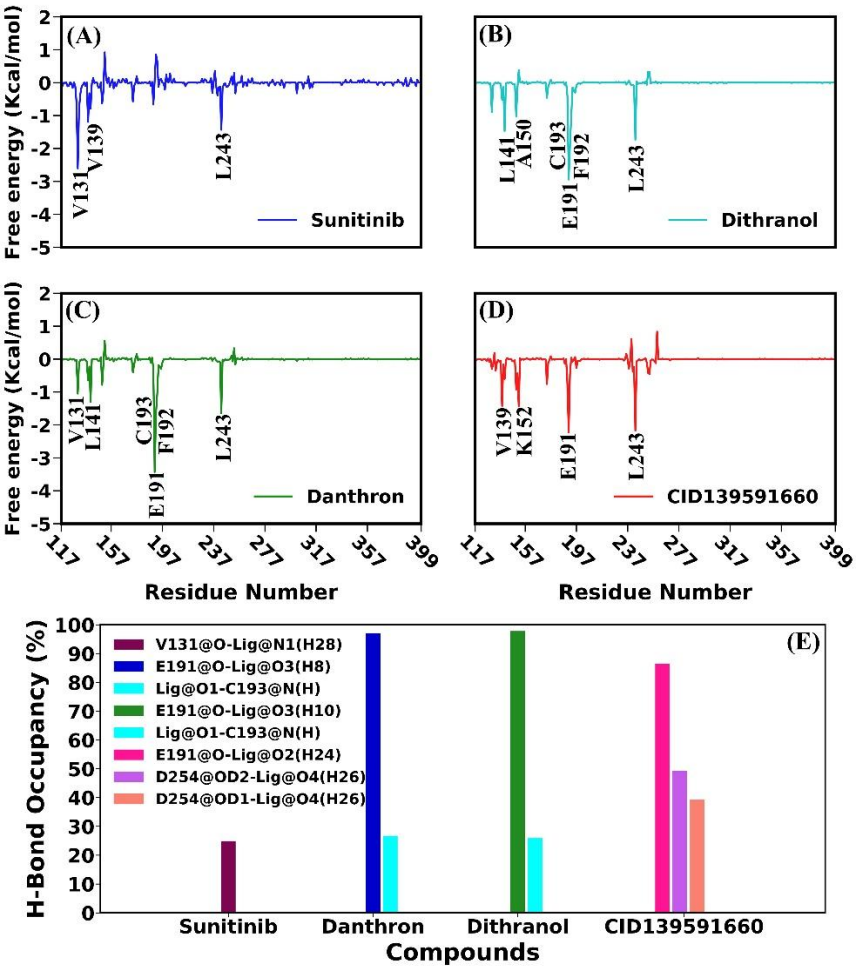


Figure 4.6: ((A)-(D)) The individual residue contributions to the binding of selected inhibitors to DLK. **(E)** The DLK-inhibitor hydrogen bonds with higher than 20% occupancy in the 2 x 250 ns simulations.

4.3.5. Hydrogen Bond and Interaction Profile Analysis

The ligand-protein hydrogen bonds are crucial components of complex stability, and their exploration can aid in understanding molecular interactions [307]. Hence, we calculated the intermolecular H-bonds formed between the ligands and DLK (see **Figure 4.7**). Our analysis reveals that CID139591660 formed the highest number of H-bonds with DLK (≥ 3), as compared to the other inhibitors. Both danthron and dithranol exhibited a similar H-bond formation pattern, with an average of ≥ 1.5 H-bonds formed (see **Figure 4.7**). Sunitinib, the control ligand, formed ≤ 1 H-bond compared to our studied ligands. Overall, our studied inhibitors formed a greater number of H-bond as compared to control sunitinib.

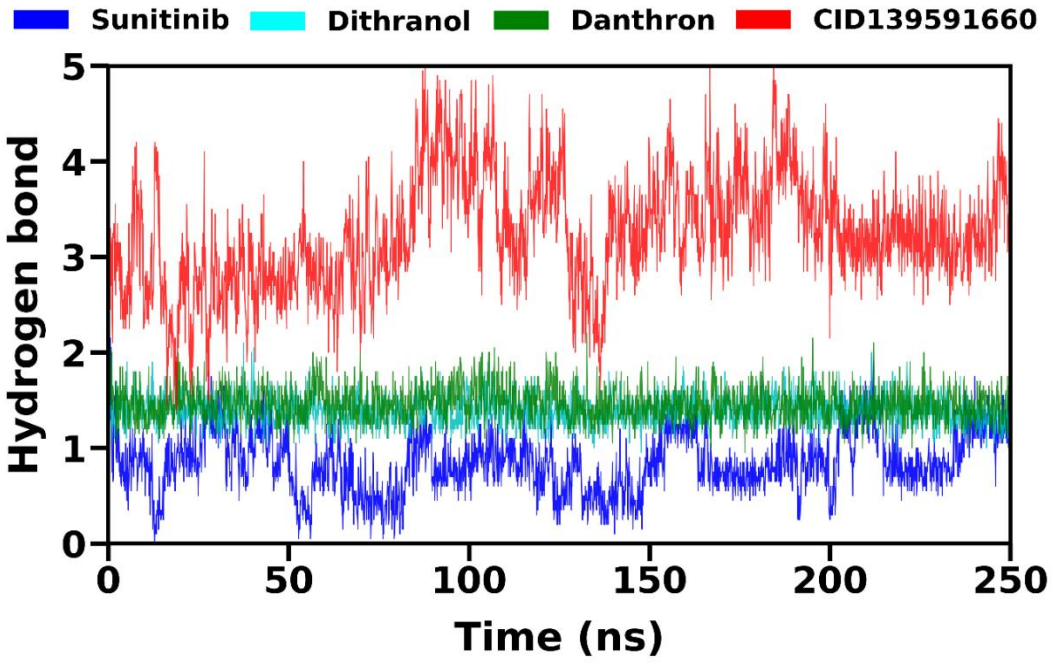


Figure 4.7: Time evolution of block average of the hydrogen bonds formed between DLK and inhibitors.

Furthermore, we also analyzed the hydrogen bond occupancy between the ligand and DLK throughout the simulation trajectories. **Figure 4.6(E)** illustrates the crucial DLK-inhibitor H-bonds with > 20% occupancy which played the most significant role in binding of those ligands. Further, detailed parameters of the H-bonds with > 10% occupancy are listed in **Table 4.4**. Sunitinib interacts mainly with three residues, Val131 (24.88% occupancy), Gln197 (18.38% occupancy), and Ser133 (10.67%

occupancy), forming hydrogen bonds. Danthron and dithranol formed hydrogen bonds with the same residues, Cys193 (26.65% and 25.93% occupancy) and Glu191 (97.02% and 97.81% occupancy), resulting in an equal average number of hydrogen bonds (see **Table 4.4**). CID139591660 formed hydrogen bonds with residues Ser253 (16.60% occupancy), Cys193 (11.82% occupancy), and Glu191 (86.49% occupancy). Additionally, Asp254 forms hydrogen bonds with multiple oxygen atoms of the CID139591660, with occupancies of 49.23%, 39.22%, 17.28%, and 16.91% (see **Table 4.4**).

Table 4.4: Presents the hydrogen bond interactions established by DLK with inhibitors and the corresponding average distance and occupancy percentage. Hydrogen bonds greater than 10% occupancy are presented here.

Binding couples		**Molecular dynamics**	
Acceptor	**Donor**	**Distance(Å)**	**Occupancy[a](%)**
Sunitinib			
Lig@O1	Ser133@N...H	2.90	10.67
Val131@O	Lig@N1...H28	2.77	24.88
Gln197@OE1	Lig@N1...H28	2.79	18.38
Danthron			
Lig@O1	Cys193@N...H	2.91	26.65
Glu191@O	Lig@O3...H8	2.65	97.02
Dithranol			
Lig@O1	Cys193@N...H	2.91	25.93
Glu191@O	Lig@O3...H10	2.64	97.81
CID139591660			
Lig@O3	Ser253@OG...H9	2.78	16.60
Lig@O2	Cys193@N...H	2.87	11.82
Glu191@O	Lig@O2...H24	2.71	86.49
Asp254@OD2	Lig@O4...H26	2.69	49.23
Asp254@OD1	Lig@O4...H26	2.71	39.22
Asp254@OD2	Lig@O3...H25	2.72	17.28

Asp254@OD1	Lig@O3...H25	2.72	16.91

We also investigated the hydrophobic interactions and hydrogen bond formation in the DLK-inhibitor complexes using the structures from the final simulation trajectory. **Figure 4.8 ((A), (C), (E), (G)** displays the 3D orientation of the inhibitors within DLK, highlighting the binding residues and hydrogen bonds. Additionally, **Figure 4.8 ((B), (D), (F), (H))** illustrates the 2D ligand-protein interaction maps, revealing the prominent residues involved in hydrophobic interactions and hydrogen bonding. The H-bond forming residues and their occupancy are already discussed. Herein, we mainly focused on hydrophobic interaction and the residues involved therein. Val131, Val139, Leu141, Ala150, Phe192, Cys193, and Gln197 are some of the critical residues that participated in hydrophobic interactions with a majority of the inhibitors studied (see **Figure 4.8**).

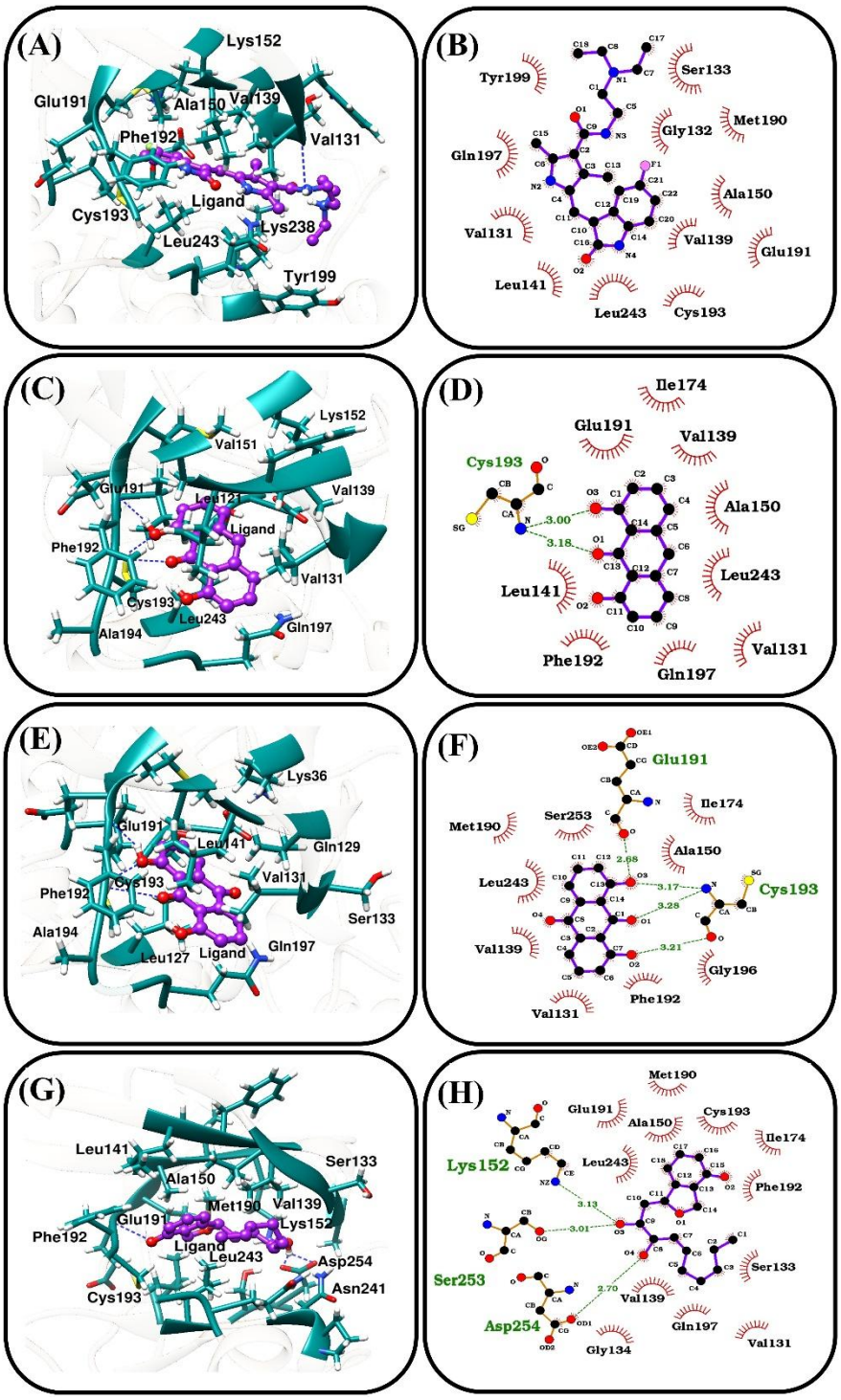


Figure 4.8: The DLK-inhibitor interaction profiles. The 3D interaction profiles were generated using UCSF Chimera, where we display critical residues in the ball and stick model. **(A)** DLK/Sunitinib, **(C)** DLK/Dithranol, **(E)** DLK/Danthron, **(G)** DLK/CID139591660. The 2D interaction profiles were generated using LigPlot+. **(B)**

DLK/Sunitinib, **(D)** DLK/Dithranol, **(F)** DLK/Danthron, **(H)** DLK/CID139591660. The pink semicircles with spikes represent residues engaged in hydrophobic interactions, while green dotted lines depict hydrogen bonds.

4.3.6. Protein Structure Network (PSN) Analysis

To gain more insight into the structural changes induced by ligand binding, we conducted a protein structure network (PSN) analysis (see **Figure 4.9**). PSN represents protein structure as a network with nodes representing amino acids and links representing their interactions, where nodes with higher degrees are called hubs, and communities consist of nodes with stronger connections among themselves. **Table 4.5** provides an overview of the network characteristics, while **Figure 4.9** ((a)-hubs (b)-communities) visually represents the communities, represented by various colors, and the hubs, sorted based on their strength of interaction. The complexes form a greater number of links: DLK/sunitinib (284), DLK/dithranol (285), DLK/danthron (285), and DLK/CID139591660 (283), as compared to apo (277).

Table 4.5: The network properties of DLK (apo) and its interactions with various inhibitors generated using web-PSN.

Parameters	Apo	DLK/Sunitinib	DLK/Dithranol	DLK/Danthron	DLK/CID139591660
No. of linked nodes	251	255	253	251	251
No. of links	277	284	285	285	283
No. of hubs	35	37	36	38	39
No. of linked mediated hubs	128	128	127	131	137
No. of Communities	6	7	7	8	8

No. of nodes involved in communities	43	45	49	48	46
No. of links involved in communities	60	62	71	69	63

A similar trend was observed for communities: DLK/sunitinib (7), DLK/dithranol (7), DLK/danthron (8), and DLK/CID139591660 (8), as compared to apo (6). The number of links and nodes involved in community formation was higher in the case of complexes (see **Figure 4.9** and **Table 4.5**), indicating higher stability of protein backbone post ligand binding, which also agrees with our previous analyses. A similar type of observation was found in our earlier studies with DLK [308].

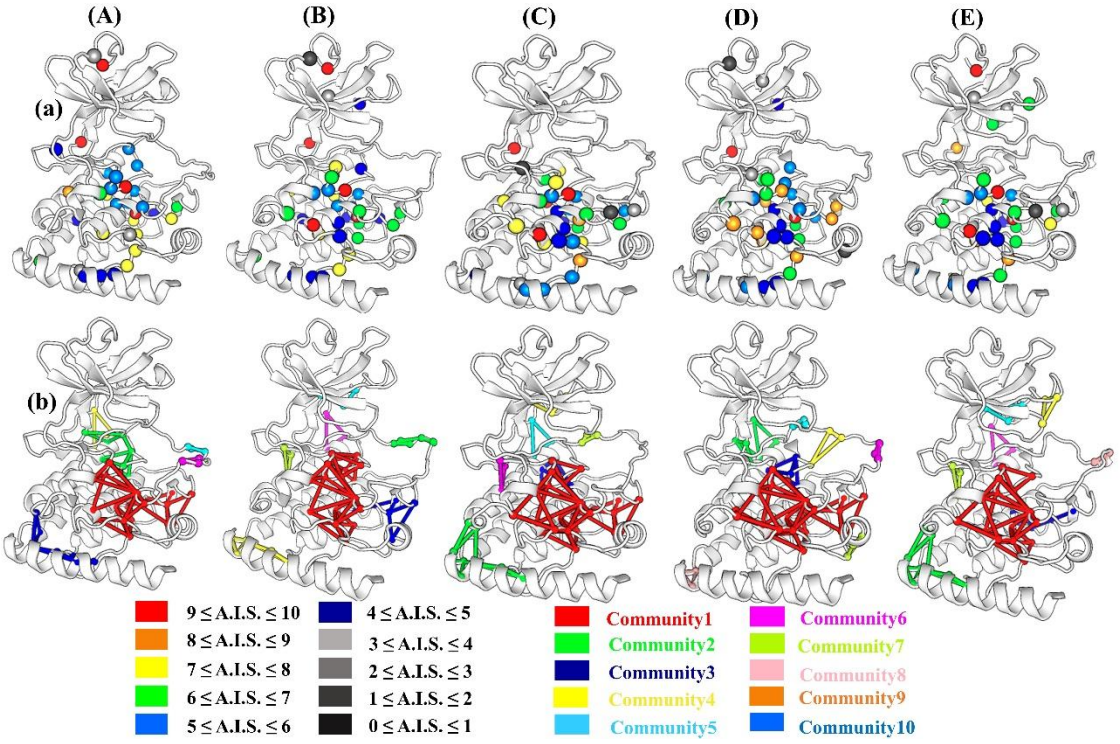


Figure 4.9: The hubs and communities present in apo, and inhibitor bound DLK. The hubs are depicted using colour codes that indicate the level of their interaction strength and the communities are arranged from highest to lowest ranking (C1 to C10). **(A)** apo,

(B) sunitinib **(C)** dithranol, **(D)** danthron, and **(E)** CID139591660. The structures in the two rows are indicative of **(a)** hubs and **(b)** communities.

4.3.7. Principal Component Analysis (PCA) and Correlation and Anticorrelation Motions

Additionally, to gain a further understanding of the structural variation and dynamics of the protein, we conducted the principal component analysis (PCA). Due to the first two principal components (PC1 and PC2) describing a substantial portion of the variance, we used those two eigenvectors as reaction coordinates to create free energy landscapes for the studied systems (see **Figure 4.10**). **Figure 4.10** clearly shows that the ligand-bound DLK adopts distinct conformations, as revealed by the presence of multiple probable minima separated by moderate to high energy barriers, indicating structural rearrangement in DLK post ligand binding. In contrast, the apo structure shows only two probable conformations with a moderate energy barrier. The extracted free energy minima structures revealed slight changes in the αC-helix and significant variability in activation segment conformation (see **Figure 4.10**), which is also consistent with our previous analyses. Overall, the distinct conformational landscape explored by the different systems indicates structural adaptations in DLK induced by ligand binding.

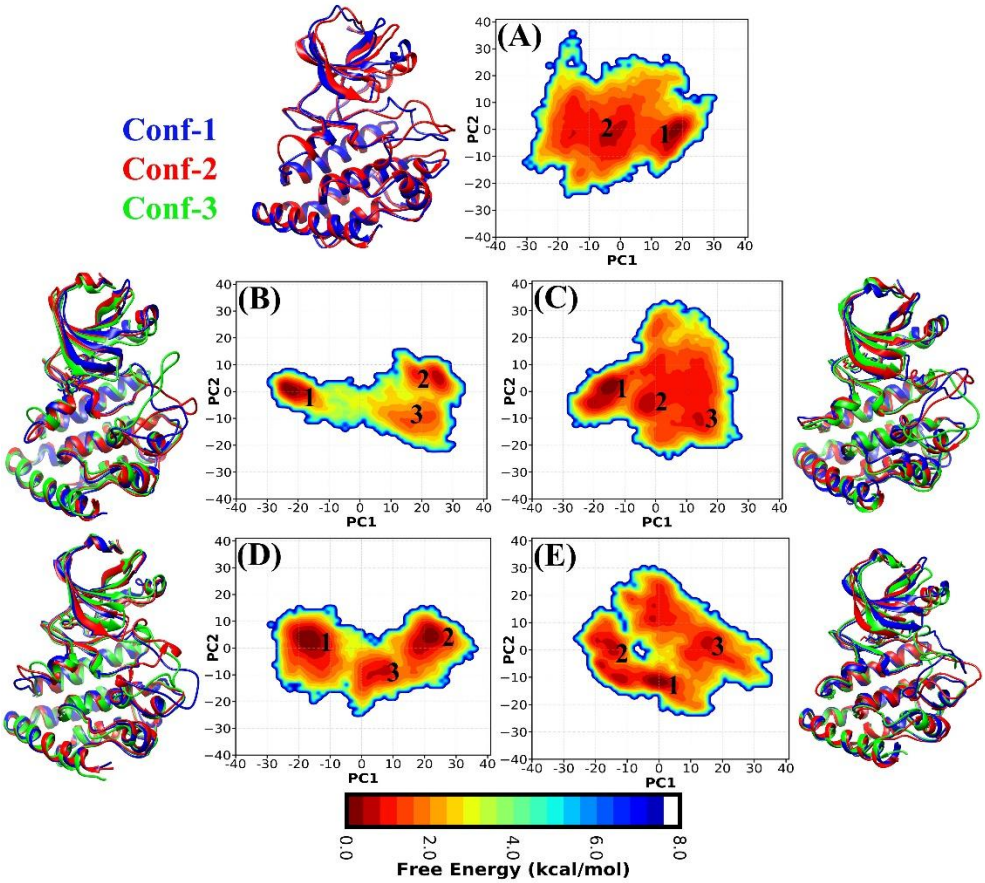


Figure 4.10: The free energy landscape (FEL) for **(A)** apo, **(B)** Sunitinib, **(C)** Dithranol, **(D)** Danthron, and **(E)** CID139591660. We used the atom position vectors projected onto the first two principal components (PC1 and PC2) as reaction coordinates to generate the landscape. The color bar depicts the free energy values in kcal/mol, which we calculated based on the distribution density.

Additionally, to distinguish differences between the apo and complexes in the dynamic movements of residual backbone C_α-atoms, we used dynamic cross-correlation matrix (DCCM) analysis. We used the Bio3D package to create 3D protein structure maps, helping us identify the residue pairs responsible for protein conformational changes**.** **Figure 4.11(A-E)** represents the DCCM plot along with the corresponding 3D structure map. The regions characterized by highly correlated movements are indicated by red-orange colors, while the black-to-whitish-pink regions represent highly negatively correlated movements or anticorrelation. We have highlighted four

significant regions, namely R1, R2, R3, and R4, located in various parts of the protein structure. **Figure 4.11(A-E)** clearly shows that region R1 at the protein's terminal end displays various patches of negative correlation in apo and DLK/CID139591660 complexes and a relatively lower negative correlation motion for DLK/dithranol and DLK/danthron, whereas negligible anticorrelation and positive correlation were observed in the case of the control DLK/sunitinib complex. In region R2, which includes the activation segment, several moderate to strong anticorrelation patches were observed in the apo state in different locations. In other cases, low anticorrelation motion was observed for DLK/dithranol, while high correlations were seen in other DLK/inhibitors complexes in the R2 region (see **Figure 4.11**). In region R3, which includes the correlation motion of the hinge region, αC helix, and P-loop, we observed moderate to high levels of anticorrelation, along with varying degrees of correlations across various segments of the proteins (see **Figure 4.11**). Region R4 further illustrates the inter-segmental motion correlation between several key motifs of DLK. Here, the hinge region exhibited a strong correlation with the catalytic loop in all DLK complexes. Certain sections of the activation segment (residues ~268-275) displayed anti-correlated motion relative to the P-loop region. This was more prominently observed in the CID139591660-bound DLK, as evident from the increased presence of blue patches in that complex. Residues ~156-164 of the αC helix were also observed to be anti-correlated with the residues ~267-286 of the activation segment, most strongly seen in DLK/CID139591660. In all systems, diagonal movements exhibited a strong correlation, whereas off-diagonal movements showed distinct directional changes upon inhibitor binding. It is clear from the study that various inhibitors result in different patterns of protein motion, which was also observed by Sk et al. [309]. Overall, the DLK complex with sunitinib, dithranol, and danthron displayed a definite reduction in anticorrelation motion in comparison to apo, while some areas in CID139591660-bound DLK had moderated patches of anti-correlation.

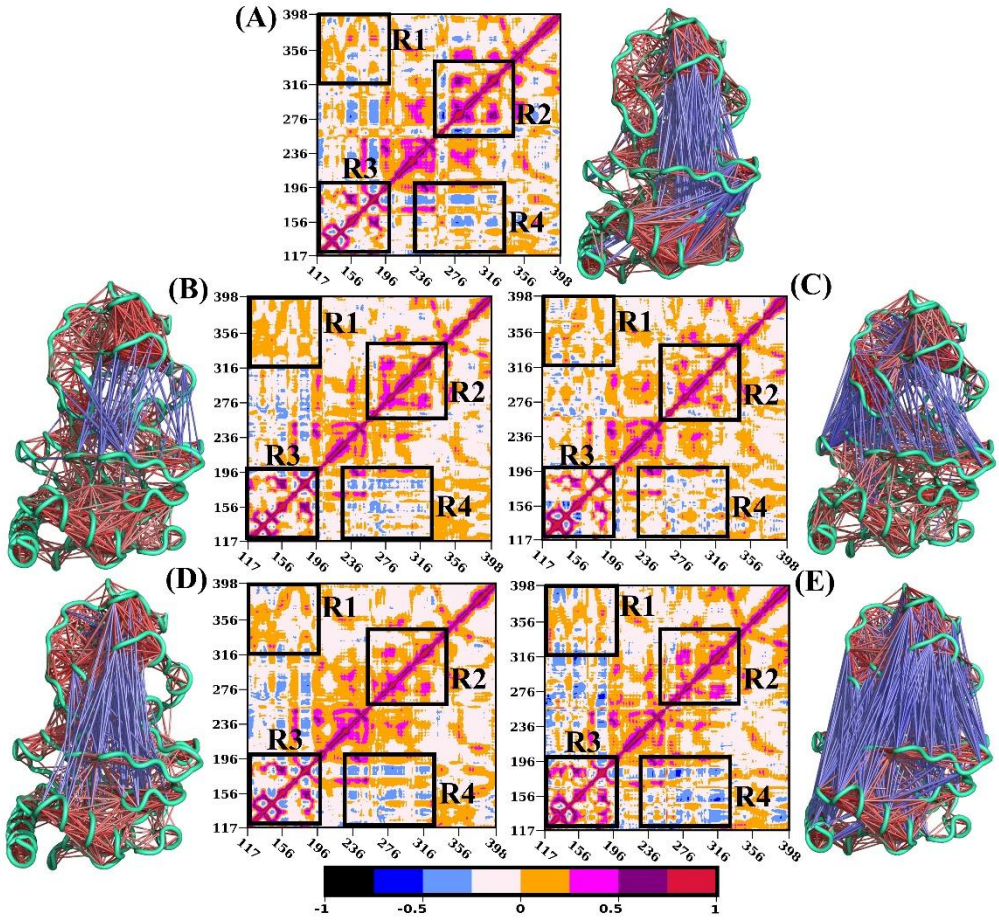


Figure 4.11: The dynamic cross-correlation matrices (DCCM) of the DLK in **(A)** apo, **(B)** Sunitinib, **(C)** Dithranol, **(D)** Danthron, and **(E)** CID139591660 systems. The black to whitish pink (-1 to 0) and red to orange (1 to 0) represents to the anti-correlation and correlation motion. In the 3D structure, the red lines represent correlation motion, while the blue lines represent anti-correlation motion.

4.3.8. Structural Analogues of the Two Natural compounds: A strategy for Lead Optimization

To further optimize the two natural compounds for targeting DLK, we utilized the deep learning based DeLA Drug to generate their analogues, with the objective of investigating whether these analogues would show better binding free energy and stable binding interactions with DLK. Initially, we created 100 analogues for each of CID156581477 and CID139591660 using the DeLA Drug webserver and subjected

them to virtual screening against DLK. As the CID139591660-analogues did not demonstrate improved binding free energy, we have not included those in our analyses and discussion here. We selected the top two analogues of CID156581477 following virtual screening and subjected those to blood-brain barrier penetration (BBB) prediction using SwissADME. After BBB analysis, we found that only one analogue, named CID156581477-ANLG, showed BBB permeability (see **Table B5**). Hence, we considered this analogue for MD simulation and further analyses. The structural modifications of CID156581477 leading to the formation of CID156581477-ANLG, along with the structures of the associated ligands, are shown in **Figure 4.12 (A-B)**. The complex was then subjected to 2x250 ns MD simulation. **Figure 4.12(C)** displays the convergence of the DLK/CID156581477-ANLG complex within the simulation range, as indicated by the time evolution of backbone RMSD. The ligand-protein distance analysis revealed that the ligand remained stable within the binding pocket, with a distance within ~2 Å (see **Figure 4.12(D)**). To further investigate the binding potential, we employed the MM-PBSA method to calculate the binding free energy of the complex. **Table B6** lists a comparison of free energy components between the CID156581477 and CID156581477-ANLG complexes, revealing that CID156581477-ANLG (-11.62 kcal/mol) exhibited significantly higher binding affinity as compared to the parent molecule CID156581477 (-6.09 kcal/mol). CID156581477-ANLG exhibited a higher binding affinity due to a higher favorable van der Waals interaction (ΔE_{vdW}) compared to CID156581477. The number of H-bond formations in CID156581477-ANLG and its parent molecules CID156581477 is similar (see **Figure 4.12(E)**). Overall, CID156581477-ANLG exhibited a greater binding potential compared to the control sunitinib, indicating its prospect as a potent DLK inhibitor.

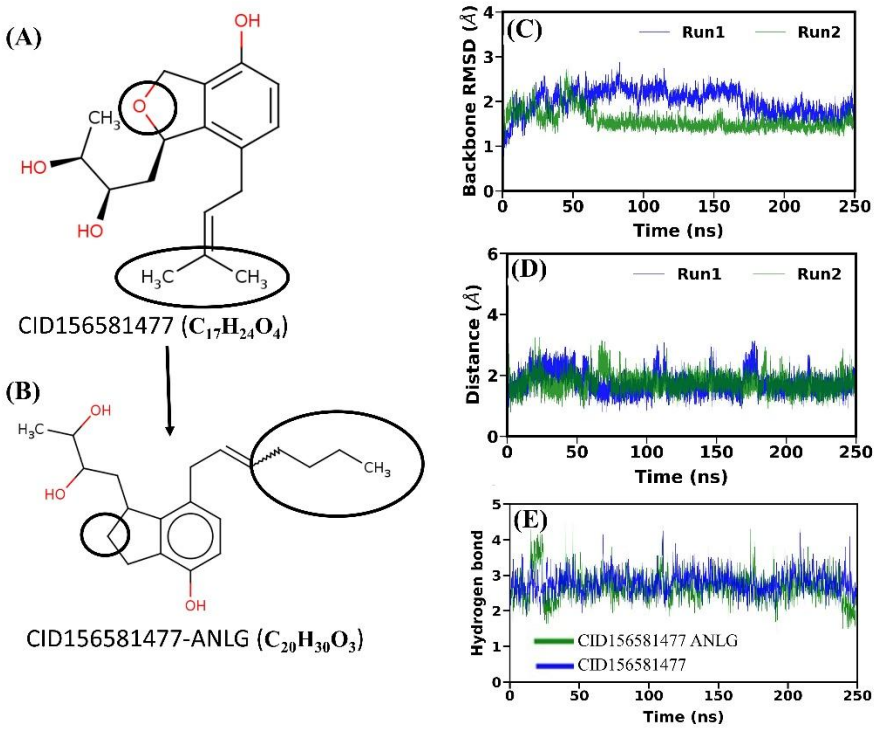


Figure 4.12: (A) and **(B)** represent the modification from the parent molecule CID156581477 to its analogue CID156581477-ANLG. The time evolution of **(C)** RMSD backbone atoms of DLK bound with CID156581477-ANLG, **(D)** distance of CID156581477-ANLG from the center of mass of residues within 5 Å of the ligand in the initial docked conformation **(E)** Number of hydrogen bonds formed between DLK and the inhibitor CID156581477 as well as its analogue CID156581477-ANLG.

4.3.9. Identification of Target Class for Inhibitors via Target Prediction Studies

Molecular target prediction studies provide valuable insights into predicting the therapeutic protein targets that a particular small molecule may interact with. Here, we examined the predicted protein/enzyme targets only for CID156581477-ANLG, danthron, and dithranol, as there was no record for CID139591660 in the webserver. This study has significant relevance as we are targeting protein kinase. The molecular target analysis indicates that all three inhibitors possess favorable drugability properties and interact with a diverse range of proteins/enzymes (see **Figure B7**). The inhibitors exhibit significant kinase target percentages - danthron (33.3%), CID156581477-ANLG (20.0%), and dithranol (33.3%) (see **Figure B7**).

4.3.10. Similarity Index Analysis of the Compounds

The four proposed hits, namely CID139591660, CID156581477-ANLG, danthron, and dithranol, show better binding affinity compared to the control (sunitinib). Here, we investigated the degree of similarity among the four proposed hits and with the control sunitinib. We utilized the ChemMine Tools web server [310], where the "Similarity Workbench" interface calculates similarity scores between compounds using the Tanimoto coefficient. The value ranges from 0 to 1, where a higher score closer to 1.0 indicates a high level of similarity. We calculated the similarity index of these four inhibitors, as shown in **Table 4.6**. All inhibitors showed a high level of dissimilarity among each other and with sunitinib (see **Table 4.6**).

Table 4.6: The Tanimoto Coefficient as a metric for assessing the similarity of molecular structures.

Inhibitor	Sunitinib	Dithranol	Danthron	CID139591660	CID156581477-ANLG
Sunitinib	1.00	0.17	0.20	0.13	0.15
Dithranol	0.17	1.00	0.25	0.21	0.33
Danthron	0.20	0.25	1.00	0.34	0.20
CID139591660	0.13	0.21	0.34	1.00	0.50
CID156581477-ANLG	0.15	0.33	0.20	0.50	1.00

Furthermore, we utilized the Multidimensional Scaling (MDS) approach in ChemMine Tools to assess the structural and physicochemical similarity index between the four inhibitors and sunitinib. The obtained scatter plot is shown in **Figure B8**. All the compounds are dissimilar from each other as they do not overlap with each other and form diverse clusters (see **Figure B8**).

4.4. Conclusion

DLK signaling plays a significant role in assisting neurodegeneration by mediating stress responses and neuronal death pathways, making it a crucial therapeutic target for neurodegenerative diseases. Inhibiting DLK provides neuroprotection by blocking downstream signaling pathways responsible for neuronal cell death. Here, we used natural compounds and FDA-approved drugs (drug repurposing) to inhibit DLK, which can provide neuroprotection during stressed and disease states. In the current study, we found two compounds from the natural products database NPAtlas with suitable ADMET properties and two FDA-approved drugs with BBB permeability after virtual screening/docking studies. Of the two natural compounds, CID156581477 is derived from the fungus *Pestalotiopsis vaccinii*, [311] whereas CID139591660 is sourced from the marine fungus *Zopfiella marina* [312]. Danthron, an FDA-approved drug, is prescribed for the management of acute and chronic inflammation in the oral cavity and oropharynx. Extensive literature supports the well-established neuroprotective properties of danthron and its various derivatives in treating neurological diseases [313], which also supports our findings. The pharmacological research indicates that anthraquinones, including danthron have diverse properties, including antioxidative, antitumoral, antiviral, and anti-inflammatory effects [314–316]. On the other hand, dithranol is an approved medication used in the management of sub-acute or chronic psoriasis [317]. These four hit compounds were then subjected to MD simulations (2x250 ns) and binding free energy calculations. To ensure a thorough analysis of protein-ligand binding, we analyzed multiple factors to guarantee the effective targeting of the protein. This included a detailed study of the protein's binding pocket, focusing on its accessibility, shape, and capacity to form strong and specific interactions of the ATP-binding pocket with potential ligands. We also evaluated the structural stability to ensure that the ligand-binding will not significantly modulate the overall conformational variability of the protein. On the ligand side, we assessed properties such as molecular weight and hydrogen bonding potency to ensure the optimal binding affinity. In the initial analysis, the RMSD results suggest that throughout the 250 ns production simulations, all systems maintained a stable equilibrium. The binding pocket's stability was enhanced upon the inhibitor's binding to the DLK protein. The inhibitor binding had an impact on the overall flexibility of

residual motion and an increase in correlated motion was observed in the complexes compared to the apo form. The PCA findings indicate that structural rearrangements occurred following the binding of inhibitors to DLK, which in turn influenced distance-dependent interactions. Additionally, the PSN analysis also suggests post-ligand binding structural and network alterations. To increase the binding potency of the ligand CID156581477 which originally did not show high binding potency as compared to our control, we have generated structural analogues using the deep learning-based web server DeLA Drug, and found one analogue (CID156581477-ANLG) with higher binding potency as indicated in our binding free energy calculations. The results from MM-PBSA and hydrogen bond analysis suggest a decreasing order of inhibitor affinity as CID139591660 > dithranol > CID156581477-ANLG > danthron > sunitinib > CID156581477. Hence, we identified four compounds with better affinity then control sunitinib. Crucial and potent residues, such as Glu191, Phe192, Leu243, Val139, and Leu141, actively participate in interactions with the inhibitors. The four identified compounds may protect neurons from stress-induced degeneration by inhibiting DLK and its downstream signaling pathway, as the mechanism depicted in **Figure 4.13**.

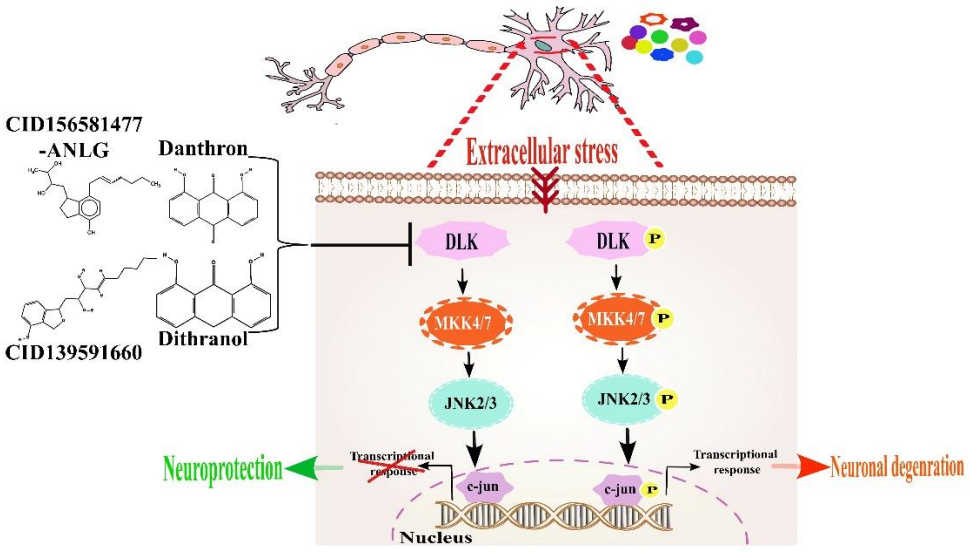


Figure 4.13: The DLK signaling pathway in response to external stress. Under external stress, DLK undergoes phosphorylation, leading to the activation of downstream targets, including the transcription factor c-jun, which ultimately results in neuronal

degeneration through transcriptional response. When inhibitors bind to DLK, it prevents DLK from undergoing phosphorylation, resulting in the inactivation of downstream proteins and transcription factors, including c-jun, which ultimately protect neurons under various stressed and diseased conditions. CID1139591660, CID156581477-ANLG, dithranol, and danthron have the potential to inhibit DLK, leading to neuroprotection through the described pathway.

Overall, the compounds we are proposing represent novel inhibitors for targeting DLK, which offer a fresh structural framework along with high-binding potency, potentially providing new avenues for therapeutic intervention. Further, our findings provide significant theoretical guidance and understanding that can be applied to developing and designing potent novel DLK inhibitors. It is imperative to emphasize that additional *in vitro* experiments and further clinical trials are essential to solidify the efficacy of these promising inhibitors.

Chapter 5

Chapter 5

Characterizing the Conformational Dynamics and Inhibition Mechanism of MKK7 by Different Inhibitor Types via Gaussian Accelerated Molecular Dynamics Simulations

AIM: To investigate the conformational dynamics of active and inactive MKK7 induced by different bound inhibitors and understand their molecular mechanisms of inhibition.

This chapter is reproduced from our original article:

Koirala, S., Samanta, S., & Kar, P. "Characterizing the Conformational Dynamics and Inhibition Mechanism of MKK7 by Different Inhibitor Types via Gaussian Accelerated Molecular Dynamics Simulations". *International journal of Biological Macromolecules* (Submitted)

5.1. Introduction

Mitogen-activated protein kinase kinase 7 (MKK7) is a dual-specificity kinase that plays a critical role in the MAPK pathway by acting as an upstream activator of c-Jun N-terminal kinase (JNK), promoting its phosphorylation and activation [47,318]. JNK activation is triggered by stress stimuli, including inflammatory cytokines, UV radiation, and environmental stress, leading to cellular responses like apoptosis and inflammation. Upon activation, JNK moves into the nucleus and phosphorylates key transcription factors, such as c-Jun, regulating the expression of genes involved in apoptosis, inflammation, and the cellular stress response [107,115,319,320]. MKK7, in conjunction with MKK4, enhances JNK activation, though it is more specific in JNK phosphorylation under stress conditions [318]. The kinase domain of MKK7 consists of a conserved bilobed structure. The smaller N-terminal lobe is mainly involved in

ATP binding, while the larger C-terminal lobe contains the substrate-binding site and key catalytic residues. The activation segment (A-loop) of MKK7 includes serine and threonine residues within the SKAKT motif that must be phosphorylated to activate the kinase. This phosphorylation induces conformational changes, positioning the catalytic residues for effective substrate binding and ATP hydrolysis [83,106,321]. MKK7 is the only human MKK characterized by three distinct docking sites, D1, D2, and D3, which mediate interactions with JNK and enable phosphorylation of the TxY motif threonine [107,321]. Upon stimulation by extracellular stress, cytokines, or growth factors, MAP3Ks such as DLK, ASK1, and MLK3 are phosphorylated and activated [318]. These activated MAP3Ks subsequently bind to the DVD domain of MKK7, leading to its activation. Once activated, MKK7 interacts with JNK through its docking domain, initiating a downstream physiological response. **Figure 5.1** presents an overview of its signalling pathway along with the crystal structure of its kinase domain highlighting its important regions.

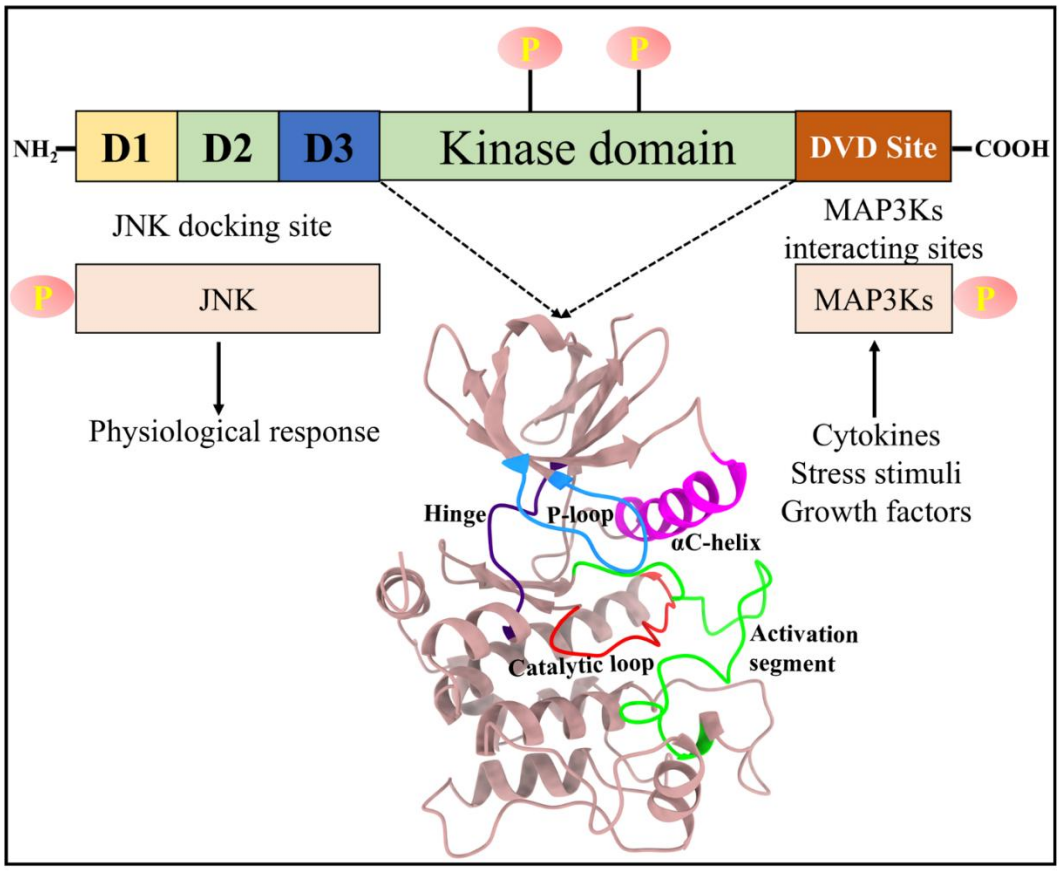


Figure 5.1: Schematic representation of the MKK7 kinase domain along with its docking domain which interact with other upstream and downstream proteins. The diagram highlights key structural elements, including the DVD site which interacts with MAP3Ks, JNK docking site, and kinase domain along with its crystal structure highlighting important regions. Stress stimuli, cytokines, and growth factors activate MAP3Ks, which then phosphorylate MKK7. Phosphorylated MKK7 activates JNK, which translocate to the nucleus to initiate physiological responses

The JNK signaling pathway is triggered by cellular stress and has been linked to numerous diseases[322], including Alzheimer's disease [323], Parkinson's disease [324], inflammatory disorders [123,325] and cancer [318]. Several studies have reported that elevated levels of MKK7 are significantly linked to the advancement of cancer and the promotion of metastatic processes [326,327]. However, several studies also suggest a tumor-suppressive role for MKK7 [328–330], indicating its function may differ significantly based on the disease context, cellular phenotype, and the nature of various stress stimuli [331]. Moreover, the JNK signaling pathway has been linked to infections caused by Hepatitis B [332] and the Zika virus [333]. Consequently, extensive research is focused on designing potent and selective MKK7 inhibitors to address the related pathological conditions.

MKK7 inhibitors may be classified into Type I, Type II, and allosteric inhibitors based on their binding sites and the protein's intrinsic conformations. Type I, Type II, and allosteric inhibitors exert distinct regulatory effects on MKK7 through different modes of action. Type I inhibitors bind competitively to the ATP-binding site of the active kinase (DFG-*in*), directly inhibiting its catalytic function. In contrast, Type II inhibitors target the inactive conformation of MKK7 (DFG-*out*), extending their interaction beyond the ATP-binding pocket to occupy an adjacent hydrophobic region, thereby stabilizing the kinase in its inactive state and preventing activation. Allosteric inhibitors bind to sites distant from the ATP pocket, inducing conformational changes that potentially impacts the ability of MKK7 to phosphorylate downstream targets[334]. Schröder et al. successfully crystallized the complex structures of the Type II inhibitor SB1-G-23, which binds in the Type II

pocket in a DFG-*out* conformation, along with the ibrutinib, which binds to the N-terminal lobe (allosteric pocket) of MKK7 [321]. Type II and N-terminal-bound allosteric inhibitors exhibit greater specificity and reduced off-target effects, making them attractive therapeutic options compared to Type I or ATP-competitive inhibitors. **Figure 5.2** shows the MKK7 structures with different types of inhibitors bound, as well as apo, highlighting key regions essential for its kinase activity.

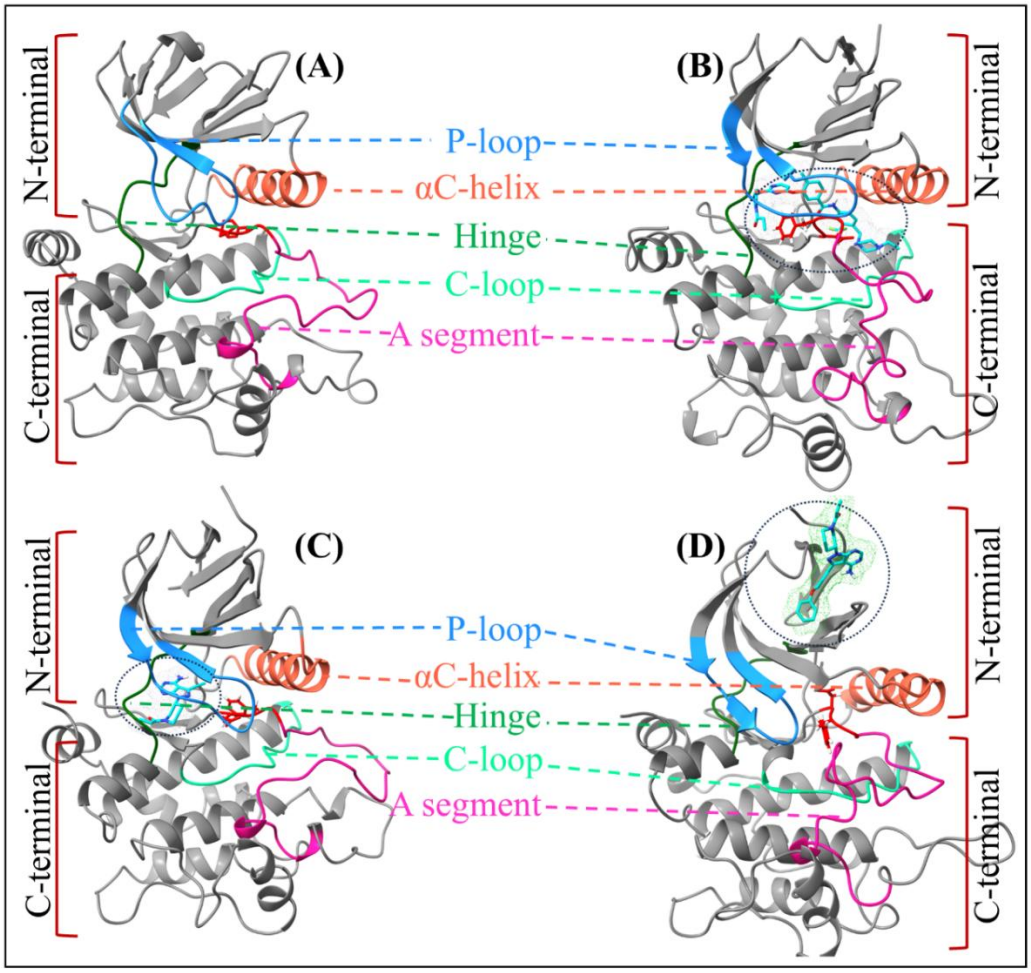


Figure 5.2: The 3D structure of MKK7 in apo and different inhibitor bound system (A) apo, (B) Type II inhibitor bound, (C) Type I inhibitor bound, (D) N-terminal allosteric inhibitor-bound system. Several key regions are highlighted with different colour codes.

The present study investigates the structural features of MKK7 in both its active and inactive states, influenced by Type I, Type II, and allosteric (N-terminal-bound)

inhibitors, elucidating the molecular mechanisms and modes of action through extensive microsecond-long Gaussian accelerated molecular dynamics (GaMD) simulations [199,335]. Additionally, this study aims to understand the intrinsic dynamics of the native MKK7 in apo form along with three types of inhibitors using free energy surface analyses using different collective variables. Furthermore, the protein dynamics of both the apo and complex forms are calculated through principal component analysis (PCA) [304], the free energy of inhibitor-binding are examined utilizing the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) [336–338], along with normal-mode analysis based conformational entropy calculations [237].

5.2. Materials and Methods

5.2.1. System Preparation of Apo and Different Inhibitor-Bound MKK7

The X-ray crystallographic structures of MKK7 in its apo form (PDB ID: 6QFL) [339] and in complex with three types of inhibitors, Type I (PDB ID: 6IB0) [340], Type II (PDB ID: 6YG7) [321], and N-terminal allosteric inhibitor (PDB ID: 6YZ4) [321] were utilized for the current studies. Missing residues were modelled utilizing the MODELLER plugin of UCSF Chimera [341]. The co-crystallized water and irrelevant molecules were removed, keeping only the corresponding inhibitors for each complex system. The LEaP module [342] of AmberTools19 [148] was employed to incorporate all missing hydrogen atoms. A double-phosphorylated state (serine and threonine of SKAKT motif of A-loop) was manually constructed using the experimental crystal structures of MKK7. Finally, four distinct systems were prepared for study the apo and three inhibitor-bound systems: the Type I ATP-competitive inhibitor bound system (T1-I), the Type II inhibitor bound system (T2-I), and the N-terminal allosteric inhibitor-bound system (A-I). Overall, four distinct systems were constructed for subsequent molecular dynamics (MD) simulations studies to examine the influence of various inhibitor types on the overall conformational dynamics of MKK7 from a structure-function relations perspective.

5.2.2. Parameterization and Force Field Assignment

The AMBER ff14SB force field [153] was used to characterize the protein parts. The structures of the inhibitors were optimized using the semi-empirical quantum mechanical Austin Model 1 (AM1) approach, and bond-charge correction (BCC) charges were computed utilizing the Antechamber module [260] of AMBER18. The atomic charges obtained from this model approximate the HF/6-31G* electrostatic potential on the surface of the molecules. The updated General Amber Force Field (GAFF2) [148] was used to assign parameters for electrostatic and non-covalent interactions for inhibitor molecules. The phosaa10 force field [343] was employed to parameterize the phosphorylated residues. All systems were set up by modeling in an explicit TIP3P water model [172]. The systems were solubilized in a TIP3P water box with a truncated octahedral setup, ranging a 10 Å distance from the protein surface in all directions. Appropriate ions were introduced in the required amount to neutralize the systems' overall charges.

5.2.3. Conventional Molecular Dynamics (cMD) Simulations

All of the systems underwent energy minimization utilising both the steepest descent and conjugate gradient algorithms. Subsequently, all systems were systematically heated from 0 to 300 K for a duration of 50 ps in the canonical (NVT) ensemble, with a restraint force constant of 2.0 kcal $mol^{-1}Å^{-2}$ imposed on the solute. The temperature was regulated through a Langevin thermostat [295]. All covalent bonds including hydrogen atoms were constrained using the SHAKE algorithm [169]. Long-range electrostatic interactions were computed via the particle mesh Ewald summation (PME) [344] method with a 10 Å nonbonded cutoff. Thereafter, the systems were equilibrated at a constant pressure of 1 atm for 20.0 ns, monitored by a Berendsen barostat [180]. A time step of 2.0 fs was employed for all simulations, with coordinates recorded every 10.0 ps for subsequent analyses. A 2.0 ns brief production NVT simulation was conducted without constraints to obtain early potential statistics. The MD simulations were carried out using the pmemd.cuda module from AMBER18 software package.

5.2.4. Gaussian Accelerated Molecular Dynamics (GaMD) Simulations

GaMD[199] is an enhanced sampling method that smoothens the potential energy surface by using a harmonic boost potential, therefore lowering potential energy barriers. The key advantages of GaMD[345] is that it does not require predefined reaction coordinates for simulations, Highlighting its uniqueness then other enhanced sampling techniques. The boost potential follows to a Gaussian distribution, which can be reliably retrieved by reweighting methods[201]. Thus, this enhanced sampling approach presents a valuable alternative for studying the dynamics of complex biological systems.

A harmonic boost potential (ΔV) is added when the system potential (V) falls below the threshold energy (E), as described in equations (1) and (2).

$$\Delta V = \frac{1}{2}K(E - V)^2, \qquad \text{if } V < E \tag{1}$$

Here, K represents the harmonic force constant, and the modified potential ($V*$) is defined as:

$$V^* = V + \frac{1}{2}K(E - V)^2, \qquad \text{if } V < E \tag{2}$$

When the system potential (V) exceeds the threshold energy (E), the boost potential (ΔV) is set to zero.

$$V^* = V, \qquad \text{if } V \geq E \tag{3}$$

In GaMD, a total potential boost (ΔV_P), a dihedral potential boost (V_D), or a combination of both (dual potential boost: ΔV_P and V_D) can be applied. Generally, dual-boost simulations achieve higher acceleration compared to the other approaches[346]. The parameters V_{max} , V_{min} , V_{avg}, and σV are derived from the initial equilibrated simulations. In this study, we utilized the dual potential boost for GaMD simulations. The boost parameters were calculated from the initial 20 ns of conventional MD simulations. Following this, 50 ns of GaMD simulations were carried out with updated boost potentials. Finally, GaMD simulations were run for 1 μs x 4 replica per system in the NVT ensemble, saving coordinates every 10 ps, resulting in a total of 400000 snapshots.

5.2.5. Dynamics and Free Energy Surface

The MD trajectories analyses were done through the Cpptraj module in AmberTools19[264]. Stability and dynamic analyses were investigated on production simulation trajectories analyzing protein backbone root-mean-square deviation (RMSD), radius of gyration (RoG), solvent-accessible surface area (SASA), and intermolecular hydrogen bond occupancy. A reference structure for comparisons was selected from the conformation obtained after the energy minimization for all of these analyses. Moreover, structural flexibility was analyzed using root-mean-square fluctuations (RMSF), while intrinsic residual motions were further studied through dynamic cross-correlation matrices (DCCM)[347]. Principal motions of each system were determined by principal component analysis (PCA)[348] in order to reduce the dimension of the data sets obtained from the GaMD trajectories. Additionally, the structural rearrangements in the loop area were investigated with dihedral principal component analysis (dPCA)[216].

The free energy surface (FES) serves as a powerful tool for visualizing the possible conformational landscape of a molecular system, elucidating the spatial organization of interacting molecules along with their corresponding energy profiles[349,350]. Hence, to investigate the structural dynamics of MKK7 systems, FESs was constructed and analysed.

FESs can be represented by the equation provided below:

$$\Delta G = -k_B T \ln(\rho *) - k - \left(\sum_{k=1}^{\infty} \frac{\beta^k}{k!} C_k - const.\right) \quad (4)$$

Here, k_B is the Boltzmann constant, T is the system's temperature, and ρ * represents the modified probability for a chosen reaction coordinate (RC). The parameter is set so that the global minimum corresponds to 0 kcal/mol. β is the Boltzmann factor, and C_k is the expansion coefficient. Next, the potential boost, combined with different RC derived from the MD trajectories, was used to reconstruct the FES using reweighting techniques. We used the "PyReweighting" python scripts for reweighting, which was developed by the McCammon group[201]. All FES analyses in this study were reweighted. However, we did not apply the reweighting procedure to the kernel density violin plots and other probability density plots.

5.2.6. Protein Structure Network (PSN) Analysis

To investigate the structural interactions among amino acid residues in the MKK7 kinase, we employed the webPSN v2.0[221] server, which integrates the Protein Structure Network (PSN) approach with the Elastic Network Model-Normal Mode Analysis (ENM-NMA). In a protein structure network, residues are represented as nodes, while their non-covalent interactions form the edges connecting these nodes. The strength of interactions between two residues is evaluated using a threshold parameter referred to as "I_{min}", which determines whether the residues are classified as interacting. The interaction strength between residues "*i*" and "*j*" is expressed as a percentage and is computed using the following equation.

$$I_{ij} = \frac{n_{ij}}{\sqrt{N_i N_j}} \times 100 \quad (5)$$

In the equation, n_{ij} denotes the number of atom-atom pairs between the side chains of residues "*i*" and "*j*" within a cutoff distance of 4.5 Å. Normalization factors N_i and N_j are applied to standardize interaction assessments by accounting for variations in amino acid sizes and their propensity to form maximum contacts in the protein structure. Residues "*i*" and "*j*" are considered interacting if I_{ij} exceeds the threshold I_{min}, establishing an edge between the corresponding nodes. At a given I_{min}, residues forming at least four edges are identified as "hubs" in the network. Communities are clusters of nodes with dense internal connections and fewer links to other parts of the network. Key network properties were analysed to elucidate structural and interaction differences across the systems.

5.2.7. Binding Free Energy Calculations

To determine the binding free energy (ΔG_{bind}) of the MKK7/inhibitors complex, the MM/PBSA method was used with the MMPBSA.py script from the AMBER18 software package. This energy is influenced by three main factors, internal energy ($\Delta E_{internal}$), desolvation free energy (ΔG_{solv}), and configurational entropy ($T\Delta S$). These components work together and can be expressed through a mathematical equation.

$$\Delta G_{bind} = G_{complex} - G_{protein} - G_{ligand} \quad (6)$$

The free energy for each component in equation 6 is calculated using the corresponding free energy expression given in equation 7.

$$G = < E_{internal} > + < G_{solv} > - T < S_{conf} > \quad (7)$$

The terms E_{internal} and S_{conf} correspond to the free energy components associated with the internal molecular energy in the gas phase and the conformational entropy, respectively. Meanwhile, (G_{solv}) represents the solvation free energy, which is the sum of the polar (G_{pol}) and non-polar (G_{np}) contributions. The internal energy E_{internal} can be further defined as.

$$E_{internal} = E_{cov} + E_{elec} + E_{vdW} \quad (8)$$

E_{cov}, E_{elec}, and E_{vdW} correspond to the covalent, electrostatic, and van der Waals interactions between the interacting molecules. The covalent interaction term, E_{cov}, can be further decomposed as follows.

$$E_{cov} = E_{bond} + E_{angle} + E_{dihedral} \quad (9)$$

The calculation of the non-polar solvation term (G_{np}) is based on an empirical formula that includes adjustable parameters.

$$G_{np} = \gamma \times \text{SASA} + \beta \quad (10)$$

The solvent-accessible surface area (SASA) was calculated using the quick linear combination of pairwise overlap (LCPO) method[351] with a probe radius of 1.4 Å. The parameters for surface tension (γ) and cavity offset (β) were set to 0.0378 kcal/mol $Å^{-2}$ and 0.5692 kcal/mol, respectively.

The binding free energy (enthalpy, *H*) of protein-ligand complexes was computed using the MMPBSA.py script in the AMBER18 software package. The solvation free energy (G_{PB}) was determined with the Adaptive Poisson–Boltzmann Solver (APBS), where the linear Poisson–Boltzmann equation was solved with the default maximum iteration settings in AMBER18.

The term S_{conf}, representing configurational entropy, can be described as,

$$S_{conf} = S_{trans} + S_{rot} + S_{vib} \quad (11)$$

Entropy contributions from translational (S_{trans}), rotational (S_{rot}), and vibrational (S_{vib}) motions were derived using normal mode analysis[352]. The MM/GBSA method was employed to decompose the binding free energy at the amino acid level[170,353,354].

5.2.8. Construction and Validation of Markov State Model (MSM)

We used MSM to analyse simulation trajectories and examine the impact of various inhibitors on the transitions between different states in MKK7. MSM represents a kinetic network model that partitions conformational space into distinct states, where conformations act as independent variables, and kinetic data is derived from transfer probability calculations between states[355,356]. MSM partitions conformational space into sub stable conformations to track state transitions. The transition process is considered Markovian when its probability is determined only by the current state and independent of prior states. The transitions among various sub-stable states can be represented using the following equation.

$$P(n\tau) = [T(\tau)]^n \, P(O) \tag{12}$$

Here, P represents the vector of state totals, τ denotes the lag time, and T signifies the transfer probability within the model.

In this study, the MSM was built utilizing the PyEMMA software[357]. Time-lagged independent component analysis (TICA) was used to project the high-dimensional space onto a lower-dimensional space[358]. Subsequently, the K-means clustering algorithm was employed to categorize conformations into 100 microstates. A 10/20 step lag was determined using the implied time scale. To confirm that the model follows Markovian behaviour, the Chapman-Kolmogorov (C-K) test[359] was conducted, and the results were examined through graphical analysis.

$$T(n\tau) \approx T(\tau)^n \tag{13}$$

T(τ) is the estimated transfer probability matrix associated with the lag time T, where n signifies the count of integer steps. Following this, PCCA+ [360] was applied to cluster the kinetically relevant microstates into 5 sub-stable states (macrostates). Lastly, transition path analysis (TPT) was carried out to determine the transition pathways and quantify their probabilities. The detailed methods were mention by various groups where they had primarily used MSM to model various proteins[361,362].

5.3. Results

In this study, we aim to explore the dynamic behavior of the MKK7 structure and to investigate how binding of various inhibitors influences the protein structure and dynamics. Here we chose four systems for study, the apo system (without inhibitors), and three inhibitor bound systems: the Type I ATP-competitive inhibitor-bound system (T1-I), the Type II inhibitor-bound system (T2-I), and the N-terminal allosteric inhibitor-bound system (A-I). To attain these goals, we conducted Gaussian accelerated molecular dynamics (GaMD) simulations on the kinase domains (KDs) of MKK7. For each system, four independent simulations were performed (1 μs per simulation, totaling 4 μs per system) to enhance sampling and minimize the risk of false positive results. For the analysis, we utilized the final 800 ns trajectories from each simulation.

5.3.1. Structural Stability and Flexibility of Active and Inactive MKK7

To evaluate the structural stability of whole protein for the apo and different inhibitor-bound systems, the root-mean-squared deviations (RMSD) of backbone atoms were calculated with reference to their initial conformations. The convergence of the simulations is indicated by the time evolution of the RMSD of the protein backbone atoms, as shown in **Figure C1** of the supporting information. To further assess the conformational dynamics of the apo and inhibitor-bound MKK7, the combined distribution and variability of the whole protein's RMSD are represented using a violin plot, as shown in **Figure 5.3(A)**. The Apo state, exhibited a broader RMSD distribution and a higher interquartile range, indicating higher RMSD deviation. In contrast, the T1-I, T2-I, and A-I systems displayed narrow RMSD distributions with tighter interquartile range, suggesting that ligand binding enhances the structural stability of the protein. In their converged states, the three inhibitor bound complexes stabilized at different RMSD values with respect to each of their crystal form. **Table C1** of the supporting information represents the average RMSD values calculated across all individual four runs for each system. Overall, the analysis of RMSD trajectories indicates that the inhibitor binding to MKK7 enhances the stability of the protein.

Next, the solvent-accessible surface area (SASA) of all four systems was calculated, and the results are presented in **Figure C2**. The SASA was calculated for the entire protein as well as for the specific ligand-binding pockets to evaluate the correlation between ligand binding and solvent accessibility (see **Figure C2(A)-(D)**). The average SASA varied between ~135-143 nm^2. The distributions were centered around comparable SASA values, with only minimal differences observed across the systems. This indicates that ligand binding does not notably impact the overall solvent accessibility of the protein surface (see **Figure C2(A)**). Upon calculating the SASA for the specific binding pockets, the observed decrease in SASA upon ligand binding, suggesting displacement of water molecules from binding sites post ligand binding (see **Figure C2(B)-(D)**).

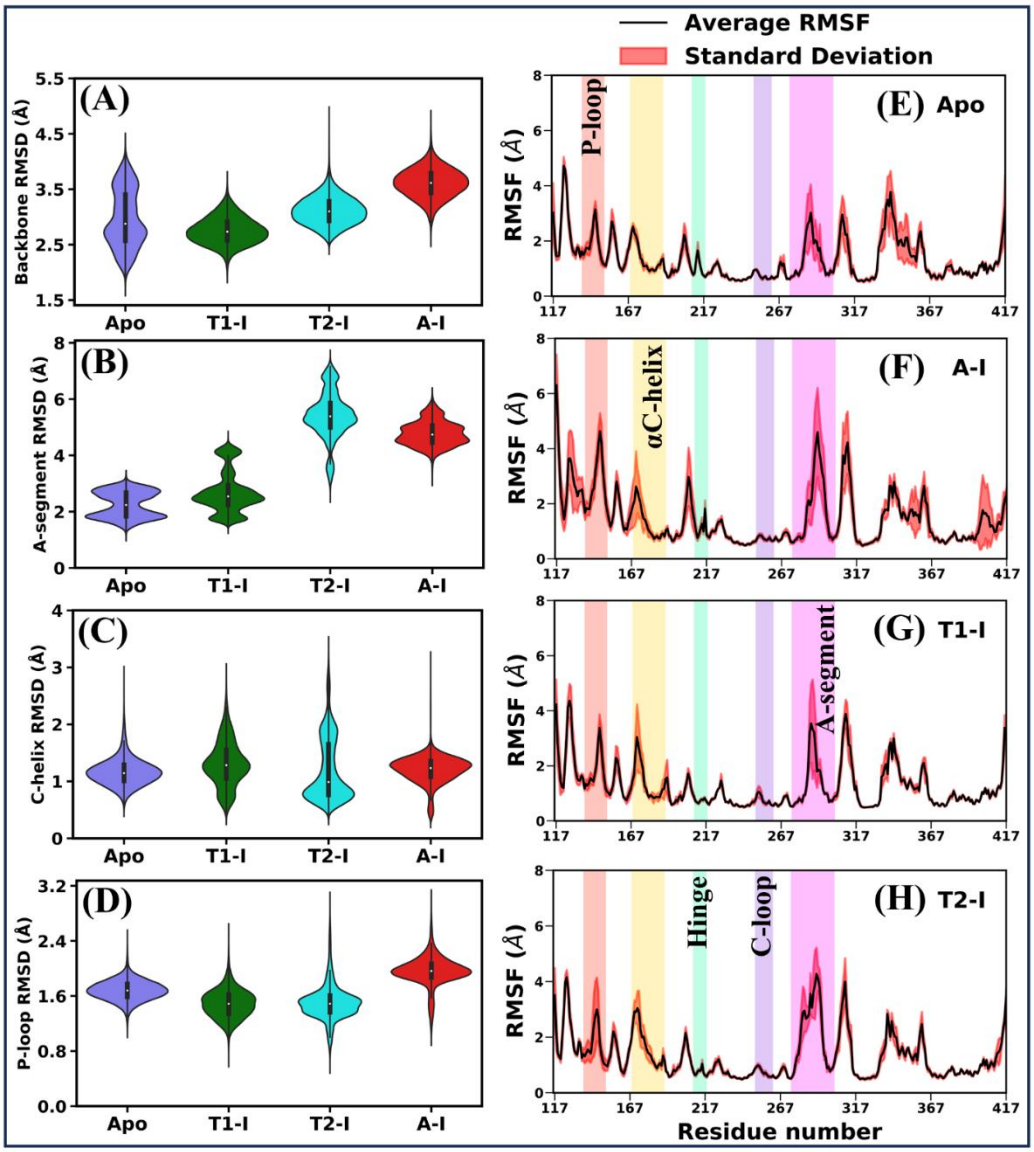


Figure 5.3: The backbone atoms RMSD depicted through Violin plot of four system. **(A)-(D)** is the violin plot for overall protein, A-segment, αC-helix, P-loop of MKK7. **(E)-(H)** The average RMSF of backbone Cα-atoms from four runs (black) and the standard deviation for each run (red).

We further calculated the radius of gyration (RoG) to measure the compactness of all systems as shown in **Figure C3(A)-(D)**. As depicted in **Figure C3(A)**, the overall radius of gyration values had overlaps, indicating structural similarity across the systems. However, the variations in peak positions suggest differences in the predominant conformational compactness among the systems. We also analyzed the

RoG of the binding pocket and observed an increase in compactness for the A-I and T1-I systems following ligand binding. However, the T2-I system was less compact as compared to apo, which may indicate that the binding of the Type II inhibitor induces conformational changes that result in reduced binding pocket compactness (see **Figure C3(B)-(D)**).

To assess the dynamic behaviour through atomic fluctuations in apo and inhibitor-bound systems, the root mean square fluctuations (RMSF) of C_α atoms were calculated using combined trajectory data. The results, presented with standard deviation, are depicted in **Figure 5.3(E)-(H)**. As shown in **Figure 5.3(E)-(H)**, the overall fluctuation patterns were similar across all systems. However, a closer examination in the key regions like, αC-helix and activation segment (A-segment) exhibited slightly higher fluctuations for inhibitor-bound systems as compared to the apo system. In contrast, the hinge region in the T1-I and T2-I systems showed reduced fluctuations relative to the apo and A-I systems, likely due to the binding interactions near the hinge region in case T1-I and T2-I systems. Additionally, we observed increased fluctuations in the terminal region of the A-I system compared to the other systems.

Additionally, we also calculated the backbone RMSD for key dynamic regions, including the A-segment (**Figure 5.3(B)**), αC-helix (**Figure 5.3(C)**), and P-loop (**Figure 5.3(D)**). The violin plot analysis of the A-segment shows differences in structural deviations among different systems. The Apo system had a bimodal RMSD distribution (range 1.5-3 Å), indicating stable conformations with limited variation. The T1-I system (range 1.6-4.5 Å) showed one sharp peak and two smaller peaks, suggesting wider range of conformations. The A-I system (range 3.8-6 Å) exhibited a broad peak with higher median value, indicating greater conformational diversity in A-loop upon allosteric modulation. The T2-I system shows the broadest RMSD range (3.0-7.3 Å) and slightly higher median than A-I, reflecting potentially more conformational sampling in A-segment. Overall, the apo and T1-I systems displayed similar distribution patterns, whereas the T2-I and A-I systems showed a similar distribution trend. These suggest that the A-loop dynamics in the T2-I and A-I systems may undergo allosteric modulation which influence the A-loop dynamics (**Figure 5.3(B)**). We generated the separate probability density plot for the A-segment and

identified the structural conformation corresponding to the peak. Analysis revealed that the A-I and T2-I systems displayed a closed-like A-loop conformation, whereas the apo and T1-I systems had open-like conformations (see **Figure C4**). To gain detailed insights, principal component analysis focusing on the dihedral angles (phi/psi) of the residues (dPCA) of the A-loop were performed, with results discussed in the following section. For the αC-helix region, the apo and A-I systems exhibit a single peak with a similar distribution pattern, while the T1-I and T2-I systems shows a broader distribution with a comparable trend (see **Figure 5.3(C)**). The P-loop region showed a consistent single peak across all systems, suggesting the binding of the Type I, II or the N-terminal allosteric inhibitor induced negligible deviation in the P-loop dynamics of the protein (**Figure 5.3(D)**). Furthermore, we computed the backbone RMSD for several other critical regions, including the hinge region, the catalytic loop, and the DFG segment, to evaluate their significance in ligand binding and the protein's catalytic functions (see **Figure 5.4**).

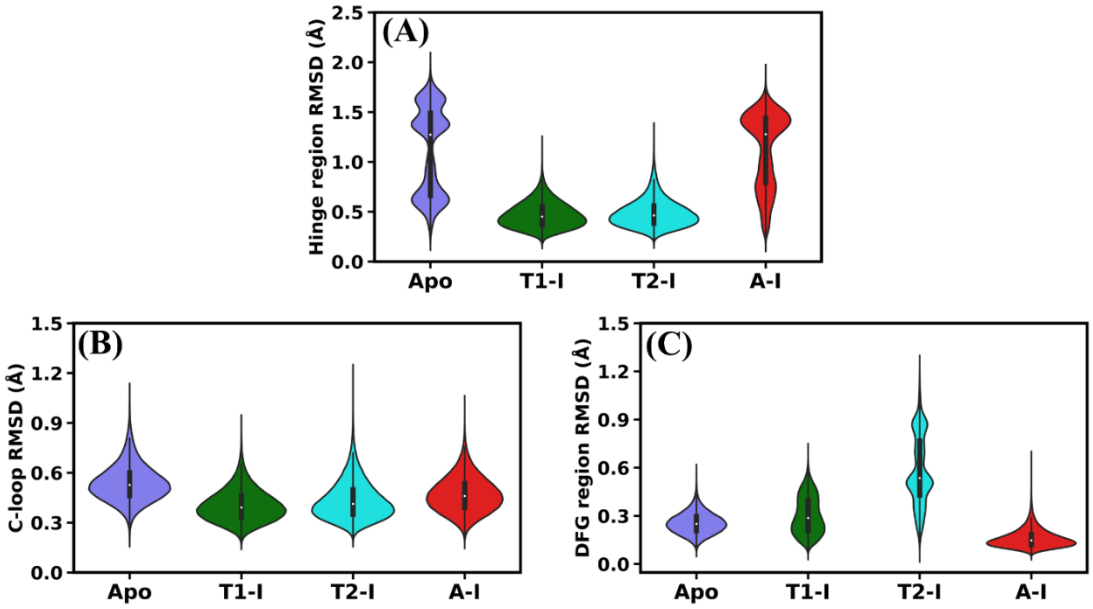


Figure 5.4: Violin plots illustrating the RMSD distribution for **(A)** the hinge region, **(B)** the catalytic loop (C-loop), and **(C)** the DFG motif.

As shown in **Figure 5.4(A)**, the broader bimodal distribution of RMSD in the apo state was due to the hinge region's flexibility without a stabilizing ligand or ATP. For the

A-I system, the allosteric inhibitor binds away from the active site, and may have caused conformational changes that increased flexibility and RMSD. In contrast, T1-I and T2-I inhibitors binds directly/close to the active site, stabilizing the hinge region, as unimodal distribution was observed in both cases at lower RMSD values. In **Figure 5.4(B)**, the catalytic loop (C-loop) displayed unimodal distribution with relatively low RMSD values across all states, indicating it was less dynamic, which is typical for most kinases. This stability is crucial for maintaining the structural integrity of kinases. In **Figure 5.4(C)**, the apo and A-I systems showed a single, unimodal DFG conformation. T1-I exhibited broader RMSD, indicating moderate deviation, while T2-I had the broadest distribution, reflecting higher conformational dynamics.

5.3.2. Global Dynamics, Residual Correlation, and Network Mapping in MKK7

Principal Component Analysis (PCA) is a widely recognized dimensionality reduction technique used to characterize the overall motions within a protein. We applied the *K*-means clustering method using the first two principal components (PC1 and PC2) to classify the sampled conformations into distinct clusters, as shown in **Figure C5**. Cluster analysis identified three optimal clusters across all systems, with the highest cluster comprising ~46.02–52.28% of the conformations, while the lowest contained ~13.93–25.05% (see **Figure C5**). Additionally, we generated a free energy landscape using PC1 and PC2 to explore the conformational space and visualized the primary modes of movement through porcupine plots as shown in **Figure 5.5**. We also extracted structures from the minima region to analyse the motion and probable conformational states of MKK7 (see **Figure 5.5**). Among the four systems, T1-I showed the most restricted conformational sampling, while the others had a wider sampling space with multiple energy minima / regions of high sampling, suggesting higher conformational variability. The apo system exhibited two minima separated by a moderate energy barrier and the extracted structures revealed distinct motions in the β3-αC loop, P-loop and initial part of αC-helix regions (see **Figure 5.5(A)**). For the A-I system showing broader conformational sampling space, two structures were extracted from the high sampled regions. Structures revealed significant dynamic behaviour in the P-loop and A-loop, with the A-loop adopting a closed-like conformation (Conf1), suggesting a likely inactive

state (see **Figure 5.5(D)**). Similarly, the T2-I system exhibited three minima separated by a significant energy barrier. The extracted structures indicate high dynamicity in the β3-αC loop and αC-helix regions, characterized by outward movement in conf1 and conf3. Additionally, the A-loop adopted a closed, inactive-like conformation in conf1, while conf2 and conf3 exhibit a semi-closed state (see **Figure 5.5(C)**). The movements in the different regions resulting in various conformations can also be observed in the porcupine plots mapping the PC1 and PC2 on the protein structure.

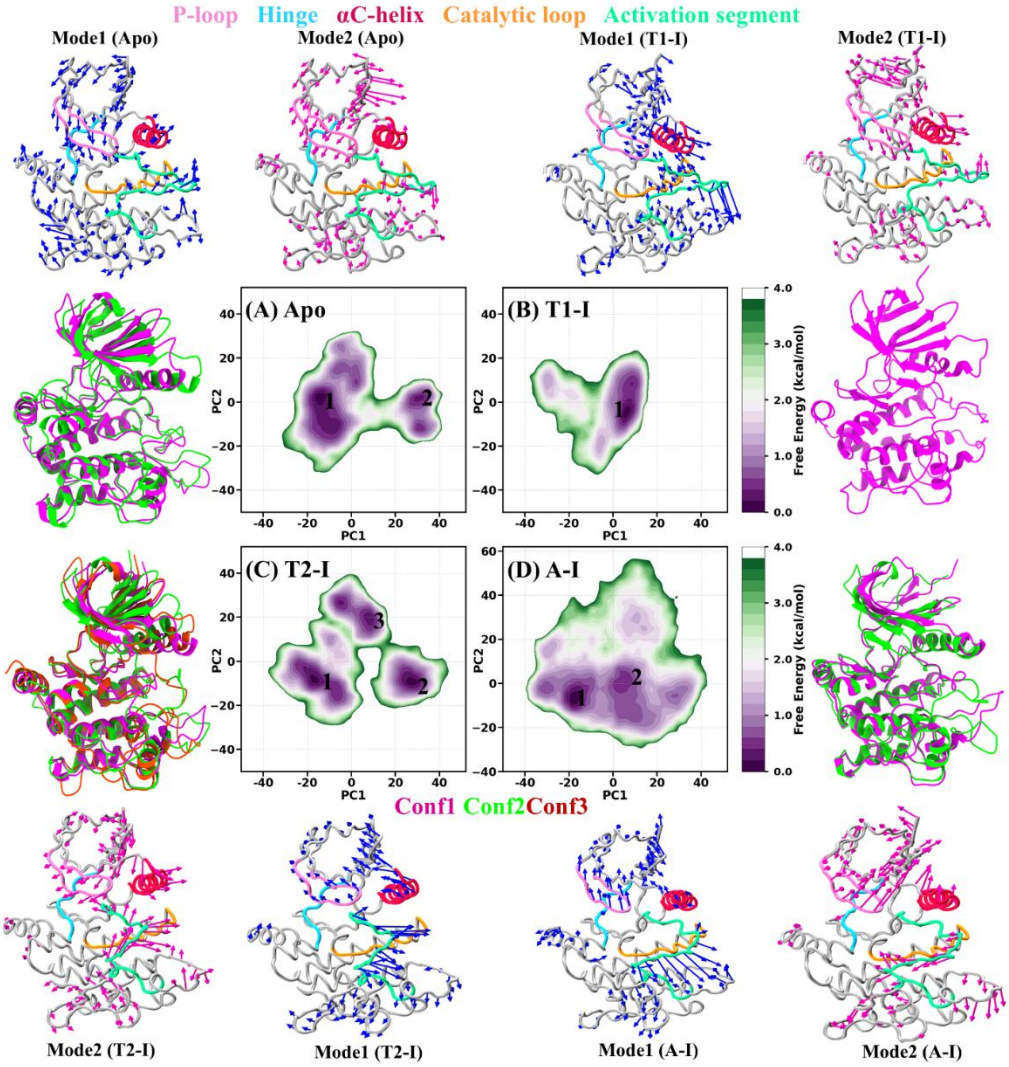


Figure 5.5: Principal Component Analysis (PCA) was performed for four systems. **(A)** apo, **(B)** T1-I, **(C)** T2-I, and **(D)** A-I. The figure illustrates the PCA results, highlighting

the extracted structures from the minima regions of the free energy landscape. Additionally, a porcupine plot is shown to represent the dominant motions of the protein in each system.

Additionally, to examine how different regions of MKK7 move relative to each other, we performed a dynamic cross-correlation matrix (DCCM) analysis, which assesses the correlation of motion based on C_α atoms. **Figure 5.6** presents a color-coded map illustrating the extent of motion correlation, where positive values represent movement in the same direction, while negative values indicate movement in opposite directions. Key regions within the overall architecture of the kinase domain are highlighted in dotted areas, providing a clear visualization of correlation patterns between these segments and serving as positional references for the entire kinase domain. Several correlated motions were observed between spatially close regions, such as the later part of the αC-helix with the final portion of the αE-helix and the adjacent sections of the catalytic loop, where the catalytic loop exhibited a stronger positive correlation motion in the apo and T1-I systems. The hinge region showed a high degree of correlation with some portion of the catalytic loop and the subsequent regions leading up to the activation segment. The final residues of the αE-helix and the catalytic loop showed correlated movement with the activation segment, most notably in T1-I, followed by apo. This correlated motion was significantly diminished in A-I and T2-I, with T2-I displaying a slight anticorrelated motion. Interestingly, the activation segment had a negative correlation with β5 and the later residues of β4 in the N-lobe, particularly in apo and T1-I. A strong correlation between the motion of the P-loop and the β3-αC loop, located just before the αC helix, was prominently observed in the T2-I system, which was also explored in our previous work[348]. Overall, the apo and T1-I systems shared similar correlation patterns across multiple regions, whereas the T2-I and A-I inhibitor-bound systems displayed a comparable correlation pattern distinct from the former.

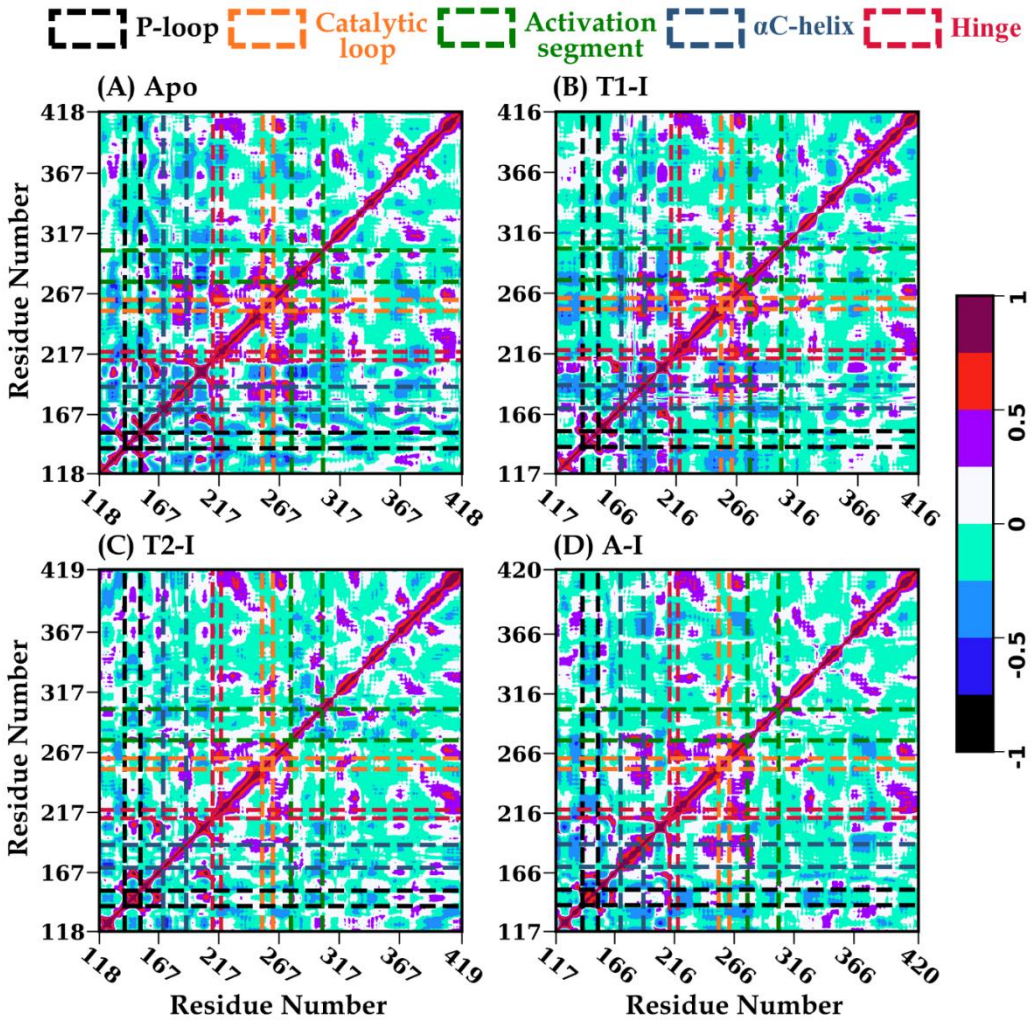


Figure 5.6: Dynamic cross-correlation matrices (DCCM) illustrating the relationship between different regions of MKK7 using the C_α-atom motion relative to their average position. The Pearson's correlation coefficients are indicated using a color-coded scale.

Further, to investigate the structural alterations resulting from inhibitor binding at different pockets, we carried out a protein structure network (PSN) analysis on the apo and three inhibitor-bound conformations of MKK7. PSN represents the protein structure as a network, where amino acids serve as nodes and their interactions create the links between these nodes. The degree of a node is defined by the number of edges or direct connections it establishes, and nodes with higher degrees are identified as hubs. A community in a network is defined as a group of nodes that are more densely connected to each other than to the rest of the network. The webPSN v2.0 webserver was used to perform network analysis on the four studied systems.

For each system, a wide range of network properties were evaluated and summarized in **Table 5.1**. **Figure 5.7** presents the hubs and communities identified across all systems, along with comparison of hub variations between the T1-I against the T2-I and A-I systems. It is worth noting that the various systems exhibited significant diversity in the total number of parameters measured. Following inhibitor binding, there was a notable rise in the number of connected nodes, total links, hubs, and links facilitated by hubs (see **Table 5.1**). Despite having a higher number of hubs in the T1-I (35) and T2-I (39), they formed a lower number communities as compared to the apo state. Although the T1-I system had fewer communities, its strongest community (Community-1) consisted of 30 nodes, 44 links, and 17 hubs formed near A-loop and catalytic-loop in C-terminal region, significantly surpassing those of other systems (see **Table C2** and **Figure 5.7**). As a result, the T1-I system is primarily stabilized by this dominant community, which may play a role in maintaining its stable and active conformation. Refer to **Table C2** for detailed information on node, links, and hubs involved in forming communities from highest to lowest rank.

Table 5.1: The network analysis of the four studied systems highlighting network properties.

Network Properties	Apo	T1-I	T2-I	A-I
No. of linked nodes	270	273	274	280
No. of links	294	301	308	316
No. of hubs	29	35	39	43
No. of links mediated by hubs	111	140	146	154
No. of communities	11	7	10	13
No. of nodes involved in communities	53	55	63	70
No. of links involved in communities	67	72	84	92

Additionally, a comprehensive comparison was conducted, analysing the differences and similarities in network properties between T1-I against T2-I and A-I (see **Table**

C3-C4 and **Figure 5.7(I)-(L)**). The comparison between T1-I and T2-I networks (**Table C3**) shows a significant level of similarity, with approximately 96% of nodes and 75% of links shared between them. However, T2-I exhibited a slightly larger number of nodes, links, and hubs, indicating a denser and more interconnected network. The increased presence of specific hubs and links in T2-I suggests functional differences that may contribute to allosteric regulation. Similarity metrics, including the Jaccard index (62.71%), Otsuka index (69.82%), and Overlap Coefficient (78.79%), demonstrate moderate-to-high overlap between the networks. Similarly, the comparison between T1-I and A-I networks (**Table C4**) reveals a strong core similarity, with 94–96% of nodes and approximately 67–70% of links shared between them. However, A-I exhibited a greater number of nodes (280 and 273), links (316 and 301), and hubs (43 and 35), reflecting a more interconnected network. Additionally, A-I had a higher proportion of specific nodes (5.71%) and specific hubs (58.14%), revealing differences which may have arisen due to allosteric regulations. The Jaccard index (55.23%), Otsuka index (63.97%), and Overlap Coefficient (73.64%) indicated a moderate level of overlap between these two networks. Overall, both T2-I and A-I networks show greater complexity and connectivity than T1-I, with increased nodes, links, and hubs, indicating functional diversity due to allosteric regulation in kinase domain.

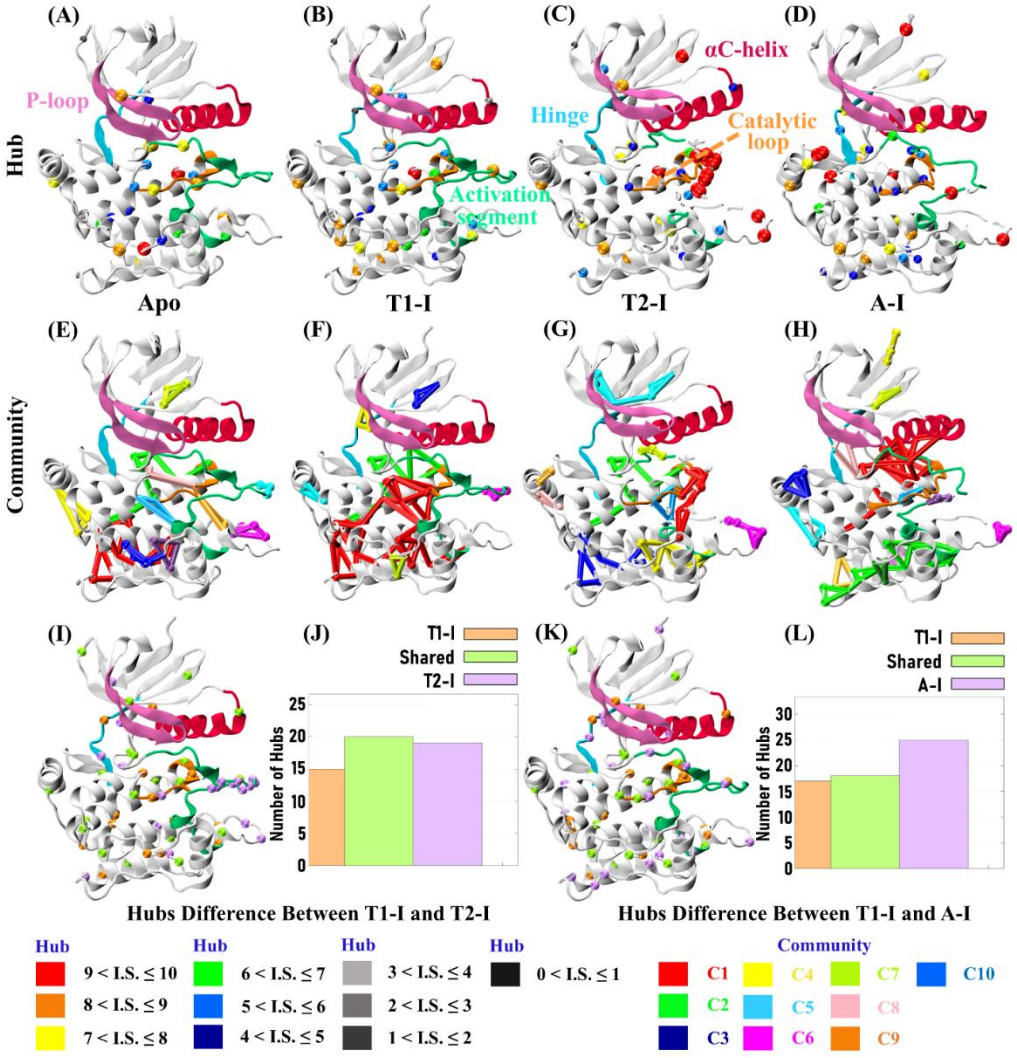


Figure 5.7: (A)-(D) Hubs, and (E)-(H) community maps derived from network analysis. (I)-(L) the differences in hubs between T1-I system with T2-I and A-I system. Color-coding is used to distinguish link interaction strength, average hub interactions, and the hierarchical organization of communities.

5.3.3. Conformational Transitions in The Activation Segment

The activation loop or A-loop, is a regulatory region in protein kinases, located near the active site and one of the critical regions in controlling the activity of the kinase. In kinases, the conformation of this loop region, and its extension the activation segment plays a key role in controlling the activity of kinase and its ability to interact with other proteins [363]. Based on the analysis of the A-segment RMSD and PCA of the entire kinase domain of MKK7, we observed distinct dynamics and identified

several probable conformational states, particularly highlighting the closed A-loop (inactive like) conformation in the T2-I and A-I systems. To further investigate the conformational changes within the activation segment, we performed a principal component analysis focusing on the dihedral angles (phi/psi) of the residues comprising the activation segment. **Figure C6** in the supporting information illustrates the percentage distribution of conformational populations across different clusters for all four systems. This distribution is analysed using the *K*-means clustering method in relation to the first two modes, dPC1 and dPC2. The T2-I system was divided into four clusters, whereas the other three systems were divided into three clusters, with their percentage distributions detailed in **Figure C6**. To examine the free energy basins associated with dPC1 and dPC2, we constructed a free energy landscape that highlights the various prominent conformational states adopted by the activation segment (**Figure 5.8(A)−(D)**). As show in **Figure 5.8(A)−(D)**, all systems showed multiple minima separated by moderate to high energy barriers. Structural analysis from minima regions shows that in the apo and T1-I systems, the A-loop predominantly adopted an open (active-like) conformation, whereas in the T2-I and A-I (Conf3) systems, the A-loop adopted a semi-closed or closed (inactive-like) conformation. The dPCA analysis of the A-loop aligns with our previous A-segment RMSD and PCA analysis.

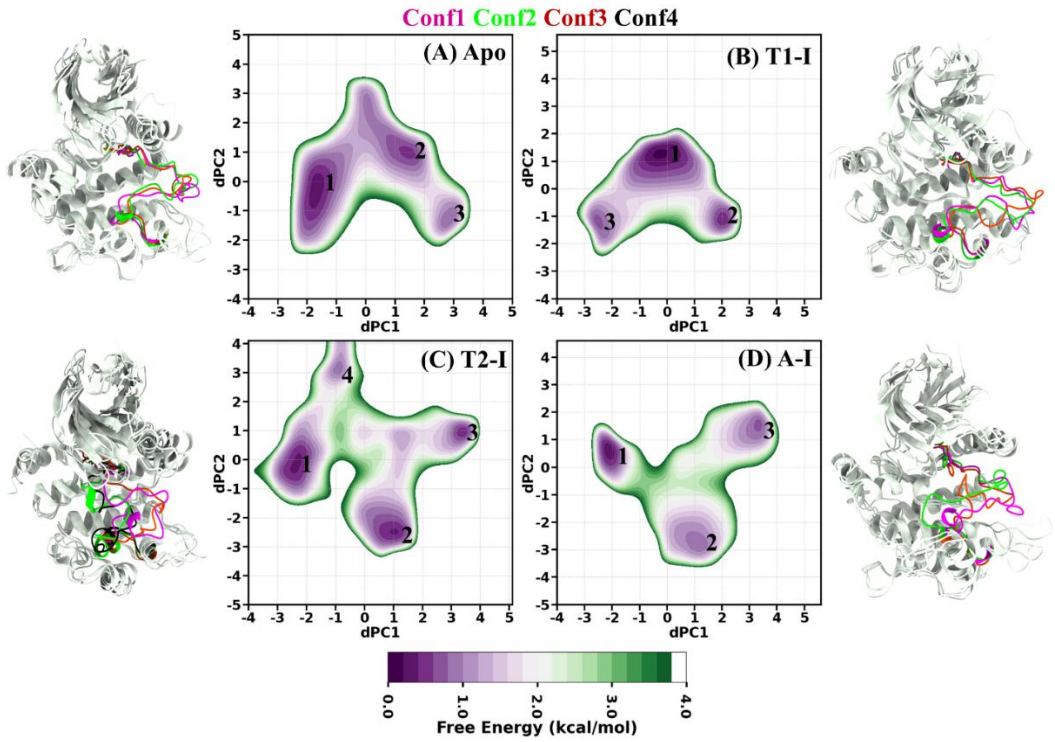


Figure 5.8: dPCA-based Free Energy Landscape (FEL) generated for the four studied systems, **(A)** apo, **(B)** T1-I, **(C)** T2-I, and **(D)** A-I. The extracted structures from minima regions represent the most stable conformations observed during the simulations.

5.3.4. Hydrophobic Spine Dynamics

The Spine model, introduced by Kornev et al. highlights the critical importance of two sets of hydrophobic residues that span from the N-lobe to the C-lobe in the kinase domain, playing a fundamental role in its structural integrity[364]. These two spines, known as the regulatory spine (R-spine) and the catalytic spine (C-spine), play a pivotal role in coordinating the structural organization and functional assembly of the kinase domain of MKK7. The R-spine residues of MKK7 include Val^{186}, Thr^{201}, His^{257}, and Phe^{278} while the C-spine residues comprise of Val^{150}, Ala^{163}, Ala^{219}, Ile^{265}, Leu^{266}, Leu^{267}, Leu^{328}, and Leu^{331}. A common characteristic observed in many catalytically inactive kinase structures is the disruption of the R-spine, often resulting from the inward rotation of Phe^{278} (in MKK7) during the DFG-flip. The probability distribution plot for the R-spine and C-spine RMSD was

generated to analyze their respective distributions, and shown in **Figure C7**. As shown in **Figure C7**, The C-spine in the T2-I system exhibited a unimodal distribution with a narrow peak (~0.5 Å), while the T1-I system showed a single broad peak (~0.3–1.4 Å). In contrast, the apo and A-I systems displayed multimodal distributions, with the apo system having two peaks (~0.7 Å and ~2.0 Å) and the A-I system showing broader variation (peaks at ~0.7 Å and ~1.4 Å). The R-spine RMSD had unimodal distributions for all, but the A-I and T2-I systems show broader distributions (~0.3-2.0 Å) compared to the relatively narrower distributions observed in the apo (~0.3-1.5 Å) and T1-I systems (~0.3-1.3 Å). To gain into deeper insights, we constructed a free energy surface (FES) plot of the R-spine versus C-spine and extracted structures from the minima regions to analyze the alignment and conformational states of the residues as shown in **Figure 5.9**. The apo system exhibited two minima, due to two possible conformational states of the C-spine, while in the other three systems we detected a single minimum, with A-I having minima with larger sampling space. In the apo and the T1-I systems, the structures derived from the minimum region showed an aligned R-spine (active R-spine) while, in the T2-I and A-I systems, the R-spine is misaligned or broken (inactive R-spine). In the T2-I and A-I systems, the R-spine alignment is broken due to the structural displacement of Phe^{278} in the DFG motif.

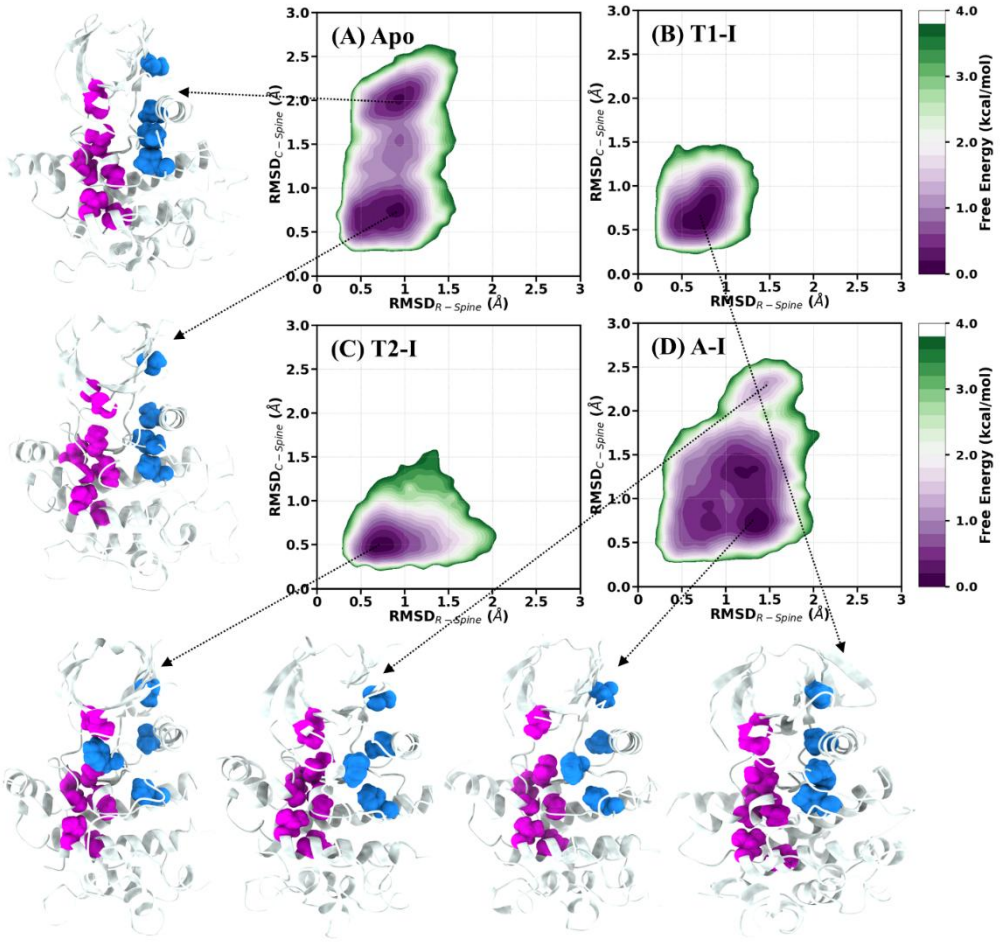


Figure 5.9: The 2D-FEL generated using RMSD of R-spine and C-spine. **(A)** apo, **(B)** T1-I, **(C)** T2-I, and **(D)** A-I. The structure associated with the minima regions are shown.

5.3.5. Diverse Conformational States Observed in The αC-Helix and DFG Motif

In protein kinases, the active state and multiple inactive states are differentiated by distinct structural arrangements of key residues. The positioning of the DFG motif (Asp^{277}, Phe^{278} and Gly^{279}) and the αC helix also regulates both the active and different inactive conformations of kinases [365]. The free energy landscapes of the DFG motif and αC helix can be analysed using two key reaction coordinates (RCs), as described by Tsai et al.[366]. The first RC involves the pseudo torsion angle (ζ(Cys^{276})) calculated using the C$_\alpha$ atoms of Leu^{275}, Cys^{276}, Asp^{277}, and Phe^{278}, which helps differentiate between the DFG-in (ζ < 120°) and DFG-out (ζ ≥ 120°) states [367]. The second RC, proposed by M. Sultan et al., measures the distance (D) between the NZ atom of Lys^{185} and the CG atom of Asp^{182}, representing the hallmark

salt-bridge which tracks the αC helix movement[368]. The conformational space explored using the reaction coordinates (ζ and D), with extracted structure derived from MD simulations displayed in **Figure 5.10(A)-(I)**. As shown in **Figure 5.10(A)-(I)**, the apo and T1-I systems predominantly retained their DFG-*in* conformations throughout the simulations, whereas the T2-I and A-I systems adopted a DFG-*out* state throughout the simulations. In the apo and T1-I systems, the αC-helix sampled both active and inactive conformations, with the inactive (αC-helix-*out*) conformation being less explored in the time-scale of our MD simulations. In the T2-I system, both αC-helix-*in* and αC-helix-*out* conformations were observed, with the αC-helix-*in* conformation being less prevalent. In the A-I system, only the αC-helix-*out* conformation was observed, with the highest sampling at a distance of approximately ~8-10 Å.

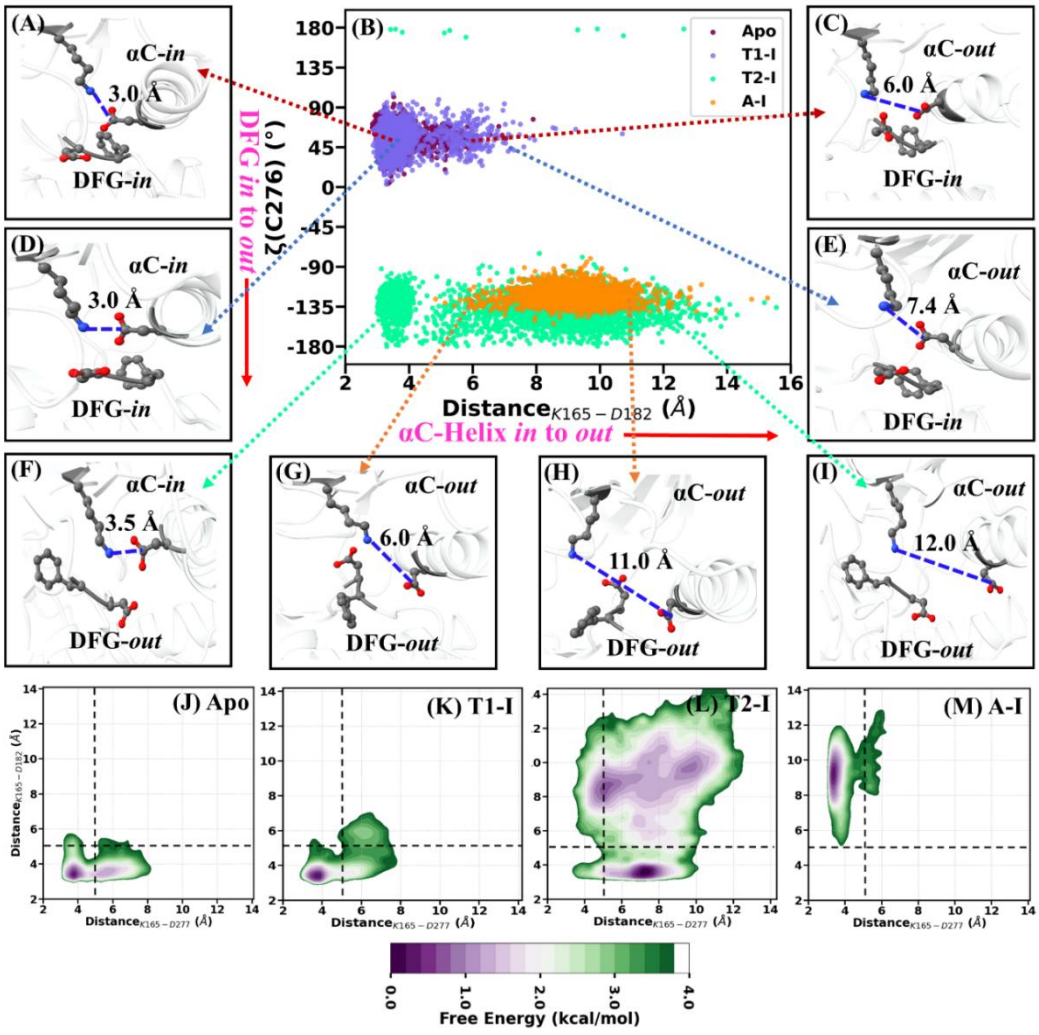


Figure 5.10: (B) The explored conformational space of four systems, mapped with respect to the pseudotorsional angle (ζ) and the distance between Lys^{165}–Asp^{182} residues. **(A)** and **(C)-(I)** represented the conformation of active and inactive MKK7. **(J)-(M)** represent 2D-FEL of two hallmark salt bridge distance between Lys^{165} and Asp^{182} in Y-axis and Lys^{165} and Asp^{277} on X-axis.

In addition to the salt bridge between the β3-lysine (Lys^{165}) and the aspartate (Asp^{182}) of the αC-helix, a prominent salt bridge exists between the β3-lysine (Lys^{165}) and the aspartate (Asp^{277}) of the activation segment, which plays a critical role in stabilizing the active conformation. To gain a more detailed understanding, we constructed the 2D FEL between these two prominent salt bridges, as shown in **Figure 5.10(J)-(M)**. The apo and T1-I systems showed a global-minima within a 5 Å distance range, indicating stable formation of both salt bridges throughout the simulations (see **Figure 5.10(J), (K)**). The T2-I system had a global-minima outside the 5 Å range, with a significantly broader sampling space compared to the apo, T1-I, and A-I systems (see **Figure 5.10(L)**). In contrast, the A-I system displayed a unique distribution as the salt bridge between Lys^{165} and Asp^{182} was disrupted by the outward movement of the αC-helix, while the salt bridge between Lys^{165} and Asp^{277} remains intact and consistent (see **Figure 5.10(M)**). Additionally, we also used *K*-means clustering method to categorize the sampled conformations into distinct clusters, along with their respective occurrence percentages, as shown in **Figure C8**. From **Figure C8**, we can clearly observe the formation and disruption of salt bridges in detail. The above results indicate that the apo and T1-I systems likely represent active kinase conformations, as their salt bridges remain stable within a 5 Å distance range. In contrast, the T2-I and A-I systems exhibit either brokage of both salt bridge (T2-I system), or single salt bridge (A-I system), which are characteristic of inactive kinase states.

5.3.6. P-loop Structural Dynamics

In protein kinases, the phosphate-binding loop (P-loop) is a highly conserved structural motif essential for ATP-binding and catalytic activity. The conformational flexibility of the P-loop is critical for the precise positioning of ATP, enabling

efficient phosphoryl transfer during enzymatic reactions. Recent research on STK17B kinase has shown the importance of P-loop dynamics in modulating inhibitor selectivity and kinase regulatory mechanisms[369]. Similarly, studies on MAP4K4 and some other kinase reveal that unique folded shapes of the P-loop can affect how selective the kinase is, indicating that the P-loop flexibility plays a key role in kinase activity[370]. Additionally, the P-loop's structural flexibility in protein kinases plays a key role in substrate binding, with conformational changes impacting overall kinase activity[371]. Here we created a scatter plot using *K*-means clustering to investigate P-loop dynamics, using the P-loop RMSD and the center of mass distance between the P-loop and the HRD motif, which is conserved in kinases, as reaction coordinates (see **Figure 5.11**). **Figure 5.11** clearly shows that all four systems were divided into clusters with a comparable distribution, consistent with the previously described P-loop RMSD, which shows a unimodal distribution. In the Apo state, the dominant C1 cluster had 46.73% population, with distance in the range of ~18-20 Å. In T1-I, the C1 cluster contained 46.85 of the overall frames, however with reduced distance (range ~17-19 Å), due to Type I inhibitor binding. T2-I showed a higher C1 prominence (54.25%), within a distance range of ~21-23 Å. In A-I, C1 with a distance range of ~23-25 Å, and C2 at a lower distance were close to equally populated (41.27% and 38.02% respectively), with C3 containing 20.70% frames. A-I also showed a slightly greater distance variability with large sampling size in Y-axis (see **Figure 5.11**). Overall, the P-loop showed dynamics with slightly different conformational preference in terms of the P-loop-HRD motif distance across all systems, with an overall distance range of ~17–27 Å in all systems.

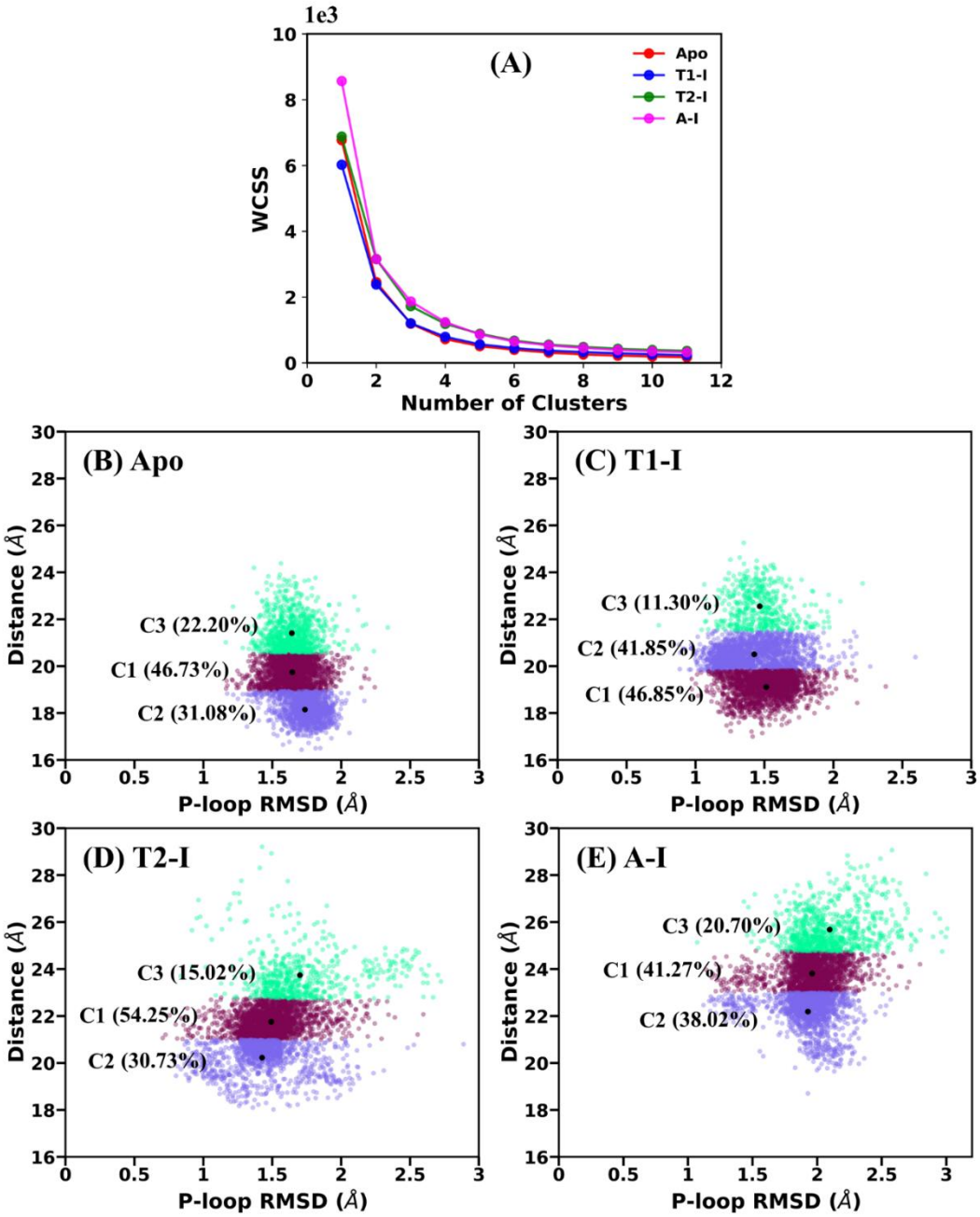


Figure 5.11: (A) A within-cluster sum of squares plot illustrating the four systems based on *K*-means clustering, using P-loop RMSD and P-loop-HRD distance, with elbow points identifying the optimal cluster count. **(B)-(E)** Clustered scatter plots based on the P-loop RMSD and P-loop-HRD distance, displaying population percentages (rounded to two decimal places) for each respective cluster. The black dots represent centroids of the cluster.

5.3.7. Binding Free Energy Analysis

The total binding free energy was calculated, along with highlighting the residues involved in binding and the hydrogen bonds formed between the inhibitors and MKK7. The average binding free energy, along with its various components for the T1-I, T2-I, and A-I systems are listed in **Table 5.2**. As shown in **Table 5.2**, the T2-I inhibitor displayed highest binding affinity or the lowest binding free energy to MKK7 (ΔG_{bind} = -30.62 kcal/mol) than the T1-I (ΔG_{bind} = -6.52 kcal/mol) and A-I (ΔG_{bind} = -7.78 kcal/mol) inhibitors. **Table 5.2** further shows that the T2-I system exhibited more favourable van der Waals interactions (ΔE_{vdw} = -100.58 kcal/mol) and non-polar free energy (ΔG_{np} = -7.25 kcal/mol) compared to the other two systems. The T2-I system exhibited higher polar solvation energy (ΔG_{pol} = 83.07 kcal/mol) than the other inhibitors because of its larger structure, which includes more polar groups. Additionally, due to its higher number of possible rotatable bonds, it also showed higher entropy contribution ($-T\Delta S$ = 38.69 kcal/mol) compared to the other two systems. In all three systems, the favourable intermolecular electrostatic interactions were outweighed by the unfavourable polar solvation free energy. Therefore, van der Waals interactions emerged as the dominant driving force for interactions in case of all the systems.

Table 5.2: The components of the total binding free energy for the T1-I, T2-I, and A-I systems bound to MKK7. All values are in kcal/mol. Enthalpy value represent the average of four simulation runs, whereas entropy is calculated from one run due to high computational cost. the standard errors of the mean are included in parentheses.

MKK7/ligand complex	ΔE_{vdw}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	$-T\Delta S$	ΔG_{bind}
MKK7/T1-I	-42.09 (0.03)	-21.28 (0.05)	40.38 (0.05)	-3.72 (0.00)	20.19 (1.06)	-6.52 (1.06)
MKK7/T2-I	-100.58 (0.03)	-43.96 (0.05)	83.07 (0.06)	-7.25 (0.00)	38.69 (0.48)	-30.62 (0.49)

MKK7/A-I	-46.20 (0.04)	-10.97 (0.04)	30.80 (0.05)	-4.18 (0.00)	22.77 (0.60)	-7.78 (0.61)

Furthermore, we decomposed the total binding free energy at the residue level for all three complexes. We highlighted only the residues whose contribution to the total binding free energy was lower than -1.0 kcal/mol and shown in **Figure C11(D)**. Residues Met^{142} and Val^{150} from the P-loop contribute in the binding of the T1-I system, whereas no P-loop residues contribute to the binding in the T2-I and A-I systems. Interestingly, the αC-helix residues Val^{185} and Val^{186} were involved in the interaction with the T2-I system but not with the T1-I and A-I systems, as suggested by the per residue analysis and their binding poses (**Figure 5.12(A)-(D)**). Interestingly, in the T1-I system, the hinge region residues Glu^{213} and Leu^{214} were involved in interactions, whereas the hinge region did not contribute to the interaction profile in the other systems. The αC-helix region significantly contributed to the complexation process in the T2-I system, while the hinge region played an important role in stabilizing the complex in the T1-I system, revealing the distinct functional roles of these important regions in the binding mechanisms of their respective systems. Furthermore, the LigPlot+ software was used to plot the interaction profiles for all three systems, highlighting both hydrophobic interactions and hydrogen bonds (see **Figure 5.12(A)-(C)**). The interaction profile reveals that the A-I system was characterized by the absence of hydrogen bonds, with hydrophobic interactions as the primary driving force. In the T1-I system, two hydrogen bonds were observed, yet hydrophobic interactions were still dominant. However, the T2-I system displayed six hydrogen bonds, along with a high number of hydrophobic interactions.

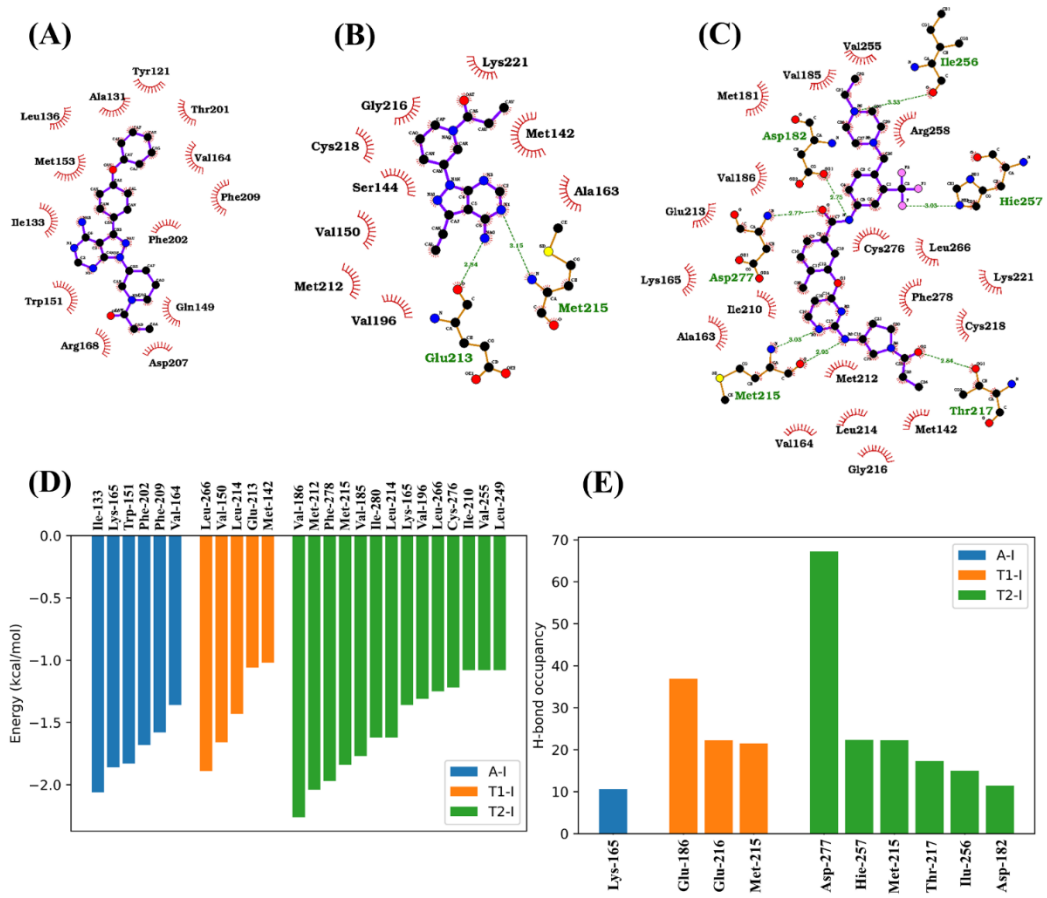


Figure 5.12: Protein-ligand interaction diagram for **(A)** A-I, **(B)** T1-I, and **(C)** T2-I. **(D)** The per-residue energy contribution from the MM/PBSA calculations for each system. **(E)** The hydrogen bonds formed with ≥ 10% occupancy throughout the collective GaMD trajectories.

Furthermore, we also calculated the H-bond occupancy using MD trajectories which determine the stability and strength of the interactions and shown in **Figure 5.12(E)**. Hydrogen bonds with an occupancy greater than 10% are shown here. As shown in **Figure 5.12(E)**, the A-I system displayed only one hydrogen bond (Lys^{165}) with an occupancy of 10.63%, which is consistent with the interaction profile, where no stable hydrogen bond was observed. In the T1-I system, we observed three hydrogen bonds (Glu^{186}, 36.88%), (Glu^{216}, 22.21%), and (Met^{215}, 21.45%), which primarily stabilized the interaction through hydrogen bonding. Similarly, for the T2-I system, we noted six stable hydrogen bonds with occupancies greater than 10% (Asp^{277}, 67.23%), (His^{257},

22.31%), (Met^{215}, 22.23%), (Thr^{217}, 17.31%), (Ile^{256}, 14.96%), and (Asp^{182}, 11.44%), all of which contributed in stabilizing the interaction through hydrogen bonding. Among all, Asp^{277} formed the stable H-bond with higher occupancy (see **Figure 5.12(E)**). All the hydrogen bonds play a crucial role in favourably stabilizing the interactions between the inhibitors and MKK7.

5.3.8. Conformation Transition Elucidated via MSM Analysis

The combined trajectory was utilized for MSM analysis, which models the transitions between distinct conformational states as Markov processes. In this approach, the conformational space is divided into numerous microstates, and the transition probabilities between them are estimated. The MSM was built using the backbone dihedral angles (ϕ and Ψ) of the entire protein, which has been shown to effectively capture its large-scale conformational changes [372]. To validate the constructed MSM models, an implied timescale analysis was performed, followed by the C-K test (**Figure C9-C12**), confirming that the models exhibit Markovian behavior. A five-macrostate model was constructed to examine the structural differences within each system. The MSM network, as shown in **Figure 5.13**, represents the conformational states and their transitions across various studied systems. Each macrostate is represented by a colored circular node, and the percentages indicate the distribution of states. Transitions between states are represented by arrows, with thicker arrows signifying greater transition possibilities and the numbers denoting timescales. Representative structural ensembles for each macrostate are displayed in the surrounding structural figures, showing conformational variations across different macrostate conditions. Macrostate 5 was the most sampled in the apo and T1-I systems, with occupancy of 30.72% and 58.38%, respectively, and its structural representation featured an extended open type A-loop with active-like visual characteristics. In the T2-I system, macrostate 2 (31.53%) and 3 (29.52%) were prominently explored, with structural features indicating a closed A-loop conformation. In the A-I system, the most dominant macrostate 5 had a population of 51.71%, exhibiting an open-like A-loop. However, the lesser sampled macrostates 2 (6.46%) and 4 (15.83%) displayed a closed A-loop conformation in the A-I system. This observation aligns with our previous

analysis, indicating that the inactive-like structure is less frequently sampled in the A-I system, as compared to T2-I. A free energy landscape, derived from the MD trajectory and projected onto the two slowest independent components (ICs), is depicted in **Figure C13**. Overall, the distinct dynamics along with the population percentage and transition pathways calculated in the MSM suggest that different inhibitors stabilize specific conformational states, potentially influencing the functional behavior and regulatory mechanisms of specific systems.

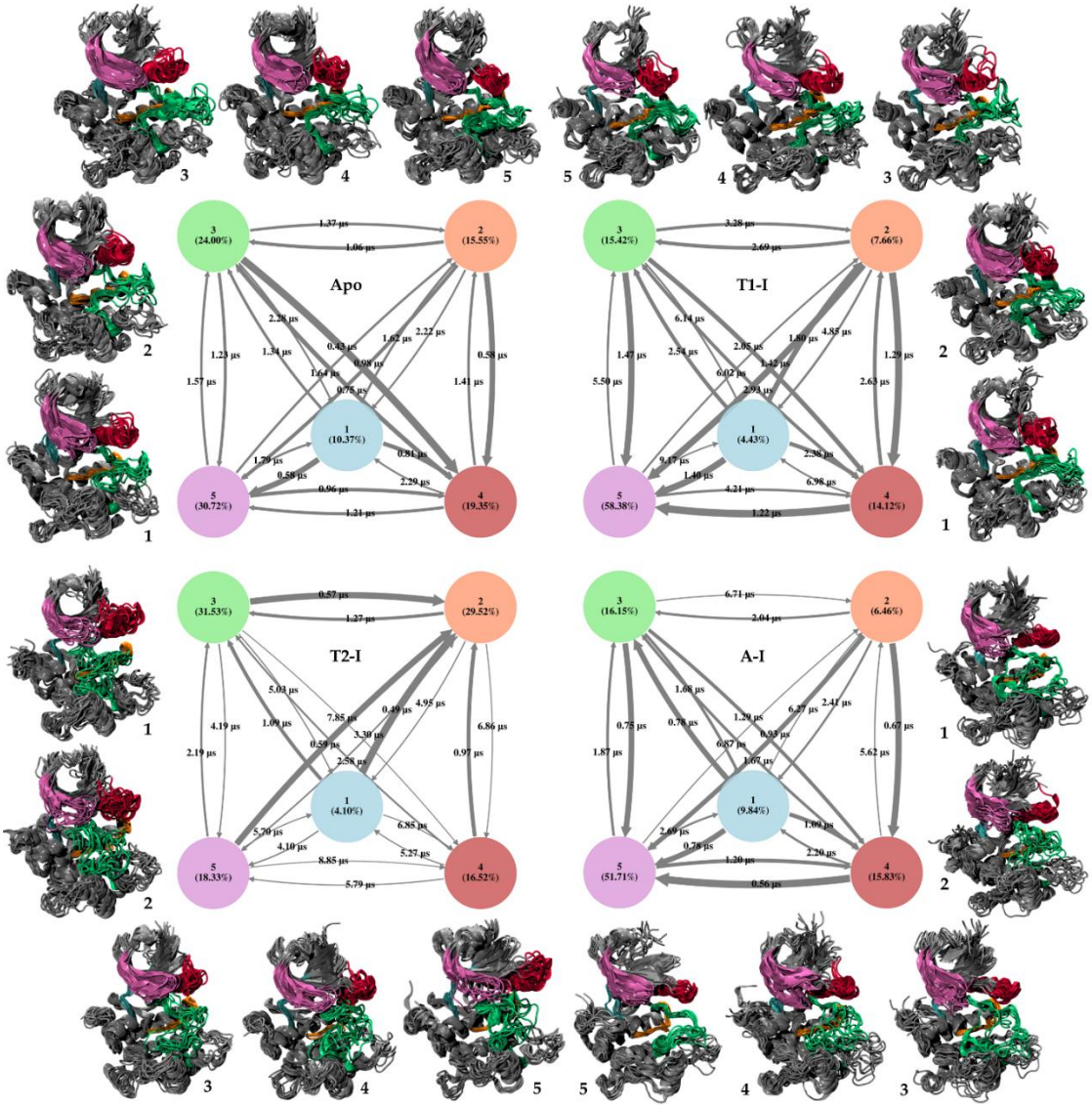


Figure 5.13: A five-state macrostate model built for the four studied systems based on Markov State Model (MSM) analysis, with percentage distributions indicated for each state. The thickness of the arrows represents transition probabilities, and the

accompanying values denote the expected transition times. Ten representative structures from each macrostate are superimposed on the sides. Key structural regions are highlighted using the following color scheme: pink for the p-loop, orange for the catalytic loop, green for the activation segment, red for the αC-helix, and cyan for the hinge region.

5.4. Discussion

Understanding how different inhibitors influence the conformational dynamics of kinases, such as in MKK7, is essential for drug design, as these dynamics play a central role in kinase function and regulation. In this study, we investigated the effects of Type I, Type II, and N-terminal allosteric inhibitors on the conformational dynamics of MKK7. Our findings provide valuable insights into the binding modes of these inhibitors by revealing distinct mechanisms through which they modulate the structure and function of the kinase. By combining structural, dynamic, and energetic analyses, we have identified key differences in how these inhibitors influence MKK7's stability and conformation, offering valuable insights for targeted therapy development.

In this study, we employed GaMD simulations to explore the dynamic behaviour of MKK7 in its apo form and different Type I, Type II, and N-terminal allosteric ligand-bound forms. Analyses such as RMSD of various regions, RMSF, SASA, RoG, PCA, dPCA, DCCM, PSN, and binding free energy calculations provide a broader understanding of MKK7 dynamics in its different forms. The RMSD analysis of the entire protein indicates that inhibitor binding enhances the structural stability of MKK7, as revealed by the single unimodal distribution of RMSD observed in Type I, Type II, and N-terminal allosteric ligand-bound systems compared to broad RMSD in the apo form. This observation aligns with previous studies on kinases like IRAK4 [55], FGFR-1 [267], FAK[373] where ligand binding often stabilizes specific conformations. Variations in the RMSD values of inhibitor-bound complexes indicate that each inhibitor induces distinct conformational changes, affecting the overall stability of MKK7. Next, the SASA analysis reveals that ligand binding has minimal impact on the overall protein. However, a decrease in the SASA of the binding pocket was observed, indicating the displacement of water molecules upon ligand binding also

observed in other kinases like CaMKIV [374]. The T2-I system was less compact than the apo form, suggesting that Type II inhibitor binding induces conformational changes, leading to reduced pocket compactness and has distinct binding mechanisms from Type I inhibitors. RMSF analysis showed similar overall fluctuation patterns, but catalytic regions (αC-helix and A-segment) showed slightly higher fluctuations in inhibitor-bound systems. Similar RMSF pattern was also seen in case of WNK [297] and JAK2 [375] kinase. The hinge region in T1-I and T2-I systems exhibited reduced fluctuations, likely due to binding pocket near hinge region, supporting the idea that inhibitors modulate kinase dynamics, impacting the specific regulatory regions.

Further, RMSD analysis of important regions like A-loop, αC-helix, and P-loop, highlights the distinct effects of different inhibitor types. The apo and T1-I systems followed similar trends in most of the observations, while the T2-I and A-I systems followed a similar trend distinct from former two. This suggests that A-loop dynamics in T2-I and A-I systems may be influenced by allosteric modulation further supported by studies into kinases like MEK1/2 [376] and some other kinase groups [377]. Notably, the A-I and T2-I systems sampled a closed or semi-closed A-loop conformation, whereas the apo and T1-I systems maintained an open-like conformation. Since A-loop conformation is crucial for kinase activation and substrate binding, these findings provide important insights into inhibitor-induced structural and conformational changes. As parts of the A-loop was missing in the crystal structure, we modelled it before simulation. Still, the T2-I and A-I inhibitors allosterically induced an inactive, closed-like conformation of A-loop through allosteric modulations. PCA with *K*-means clustering revealed that the T1-I systems showed smaller sampling space, whereas the T2-I and A-I systems explore a broader sampling space, indicating higher conformational variability. The free energy landscape further supports these findings, with extracted structures from minima regions highlighting the dynamic behaviour of important regions, consistent with previous studies on allosteric inhibitor-induced kinase conformational changes [378]. Salt-bridge disruption and outward αC-helix movement were observed in the T2-I and A-I systems, consistent with similar findings in kinases such as JAK2 [375], EGFR [379], and WNK [297]. DCCM analysis showed similar correlation patterns in apo and T1-I systems, while

T2-I and A-I systems exhibited a distinct but comparable pattern agrees with other analysis. PSN analysis showed that ligand binding increases connected nodes, links, and hubs, thus reorganizing the communication network for specific inhibitor types also reported in DLK kinase[308]. The T1-I system, despite fewer communities, formed a dominant cluster near the A-loop and catalytic loop, supporting its stable active conformation. Binding free energy analysis shows that the Type II system has the lowest binding energy, followed by the N-terminal allosteric and Type I inhibitors. The higher affinity of Type II inhibitors over Type I is consistent with findings in other kinases like JAK2 [375] and WNK [297]. Further, through MSM analysis, we characterized the protein dynamics and visualized in terms of five macrostates, providing insights into transition pathways, probable timescales, and detailed conformational sampling across each system.

Overall, our results reveal that allosteric ligand binding induces conformational shifts in important regions on the kinase domain of MKK7, which are crucial for its catalytic activity. These results are in agreement with earlier studies on kinases such as BRAF [380], EGFR [379], JAK2 [375], Cdk2 [381,382], where allosteric ligand binding was found to affect the structural stability and dynamics of the kinase. Our simulations specifically reveal that Type I inhibitors, stabilize the kinase in an active-like conformation, while Type II and N-terminal allosteric inhibitors, sample inactive like, closed conformations. Similar to the active and inactive dynamics of MKK7, comparable observations have been reported in several kinases, including Akt [383] and Src family kinases [384], JAK2 [375] , WNK [297] and AGC [385]. Studies on receptor tyrosine kinase (RTKs) like FGFR1 [386] and VEGFR2 [387] and non-receptor kinases like Src [388] show that the A-loop conformation controls kinase activation, with ligands stabilizing either active or inactive states based on their binding site agrees with our studies.

The different types of ligands induced conformational dynamics of MKK7 are highly relevant for drug discovery effort, especially in the context of cancer and neurological diseases. The strategy of allosteric inhibition is emerging strategy in protein kinase, involving the targeting of specific inactive kinase conformations to modulate their

function. Allosteric inhibitors that target these conformations can be more selective and effective than traditional ATP-competitive inhibitors. Our findings further enrich the broader understanding of MKK7 structural dynamics and how they are regulated by small molecules, laying a strong foundation for the development of next-generation MKK7 inhibitors that offer improved efficacy, specificity, and therapeutic potential.

5.5. Conclusion

In this study, we run multiple GaMD simulations total 16 μs (with four different initial velocities) to explore the conformational dynamics of MKK7 induced by various types of inhibitors. This study examined the conformational dynamics and inhibition mechanisms of ATP-competitive inhibitors (Type I) and allosteric inhibitors (Type II and N-terminal bound) and comparing them to the native apo structure. Our results reveal distinct conformations of MKK7 in the allosteric inhibitor-bound system, which are unique compared to their crystal conformations. Thus, all the conformations can be explored to develop unique and selective allosteric inhibitors. Our findings highlight significant conformational changes in the A-loop, αC-helix, B3-αC loop, and hydrophobic spines following inhibitor binding. We also observed the breaking of hallmark salt bridge between Lys^{165} and Asp^{182} in the inactive conformations due to the outward movement of αC-helix. Finally, we calculated the binding free energy using the MM/PBSA approach and found that Type II inhibitors have the highest affinity, followed by N-terminal allosteric inhibitors and Type I inhibitors. We have also identified several key residues from all three binding pockets that could be useful for future drug discovery approaches. Thus, these inactive state of allosteric binding MKK7 and their binding patterns offer a crucial foundation for developing and designing potent and specific allosteric inhibitors in the near future.

Chapter 6

Chapter 6

Unravelling the Molecular Choreography of JNK Isoforms to Understand the Structural Dynamics and Phosphorylation Driven Conformational Transitions

AIM: To investigate how phosphorylation and ATP-competitive inhibitor (SP600125) binding affects the conformational dynamics of JNK isoforms.

This chapter is reproduced from our original article:

Koirala, S., Samanta, S., Ursal, K.D, & Kar, P. "Unravelling the Molecular Choreography of JNK Isoforms to Understand the Structural Dynamics and Phosphorylation Driven Conformational Transitions". *Archives of Biochemistry and Biophysics* (Submitted)

6.1. Introduction

The c-Jun N-terminal kinase (JNK) signalling pathway is an essential part of the mitogen-activated protein kinase (MAPK) signalling pathway, functioning as an important link between external signals and cellular responses like extracellular stress, gene expression, and cell growth [100]. External stress like oxidative stress, UV radiation, and inflammatory cytokines, in addition to growth factors and other cellular signals, play a key role in activating this signalling pathway. Upon activation, JNK phosphorylates numerous transcription factors like c-Jun, ATF2, and p53, which play an important role in regulating gene expression that influences apoptosis, cell differentiation, stress response, and cell cycle progression [115]. JNK activation is mediated by two MAPK kinases MKK4 and MKK7, which selectively phosphorylate tyrosine and threonine residues within the activation loop [389]. Since both tyrosine and threonine phosphorylation are required for JNK activation, MKK4 and MKK7 may

function cooperatively to achieve this [390]. The activation of MKK4 and MKK7 is mediated by multiple MAPKK kinases, including MLK, ASK, and DLK [391]. The JNK signalling pathway is evolutionarily conserved among different species, underscoring its critical function in regulating cellular homeostasis. **Figure 6.1** shows the context-dependent roles of JNK isoforms in neuronal growth, differentiation, and stress response, highlighting their regulatory interactions. However, dysregulation of JNK signalling has been linked to neurological diseases like neurodegeneration [117,392], chronic inflammatory conditions such as rheumatoid arthritis (RA) [325,393], inflammatory bowel disease (IBD) [123,394], infectious diseases [395], and cancer [113,396]. Additionally, its dysregulation is also linked to various pathological conditions, including metabolic disorders such as obesity [397], type 2 diabetes (T2D) [398,399], and non-alcoholic fatty liver disease (NAFLD) [400]. As the JNK pathway contributes to both physiological and pathological events, the it is a crucial component for understanding how cells respond to stress and developing targeted therapies.

JNKs are structurally related to other MAP kinases, such as ERK2 and p38, as their kinase domains exhibit a conserved fold and a similar mode of activation through phosphorylation[48]. The structure includes a smaller N-terminal lobe, predominantly made up of β-strands, and a larger C-terminal lobe, primarily composed of α-helices, connected by a flexible hinge-like region. The ATP-binding site, located within a deep cleft at the interface of the N and C lobes, has been a key target in JNK drug discovery, along with other strategies to disrupt its interaction with scaffold protein JIP. Most known inhibitors of JNK are type I kinase inhibitors that bind to the highly conserved ATP-binding pocket. The key challenge with these inhibitors is their lack of selectivity, as the ATP-binding site shows up to 98% homology across JNK isoforms [401]. Several comprehensive reviews on JNK inhibitor development have been reviewed by Siddiqui and Reddy [402], Graczyk [403], and Koch et al. [404], who detailed the progression and the discovery of JNK inhibitors along with challenges.

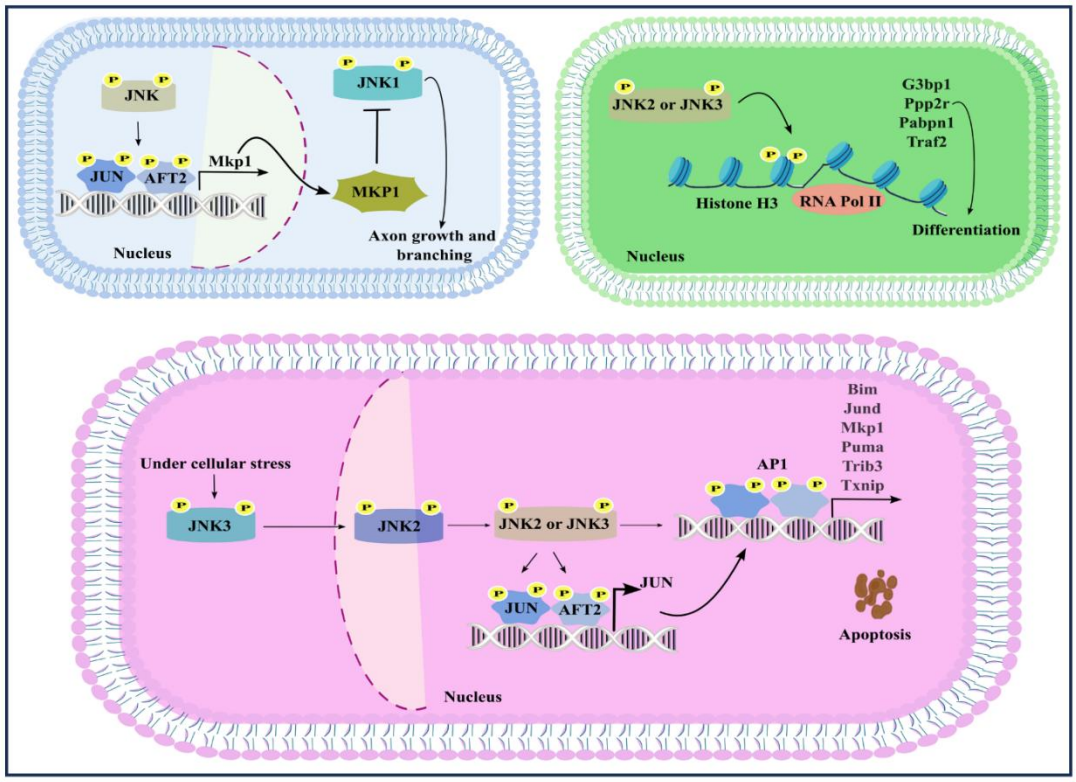


Figure 6.1: JNK signalling modulates cellular processes in response to environmental conditions. In neurons, JNK1 contributes to axon elongation via JUN and AFT2, whereas MKP1 negatively regulates JNK1. During differentiation, histone H3 phosphorylation by JNK2 and JNK3 facilitates RNA Pol II transcription. Under stress, JNK3 activates JNK2, which subsequently phosphorylates JUN and AFT2, forming the AP1 complex and initiating apoptosis. These mechanisms highlight the diverse functional roles of JNK isoforms.

The three isoforms of the JNK family, namely JNK1, JNK2, and JNK3, are encoded by different genes, and they vary in their expression patterns, substrate specificity, and individual functions within the cell [90]. JNK1 and JNK2 are widely expressed across different tissues, where they contribute to various cellular processes, including metabolic regulation, cell cycle control, and immune reactions. Unlike JNK1 and JNK2, JNK3 is primarily expressed in the brain, where it plays a key role in neuronal functions and stress-related processes in the central nervous system along with low level expression in heart, and testes [405,406]. Although these isoforms share similar

structures, they produce different biological effects because of variations in their activation kinetics, substrate specificity, and nature of interactions with other proteins[407]. Each JNK isoform has a unique role in disease progression, which underscores the need for depth studies on their signalling pathways and their mechanisms of action. Understanding how they contribute to cellular processes and disease is important for developing isoform-specific therapy with more specificity.

Molecular dynamics simulations were performed in this study to explore the comparative structural dynamics and energetic characteristics of the three JNK isoforms JNK1, JNK2, and JNK3 bound with the inhibitor SP600125 in both phosphorylated and unphosphorylated forms. SP600125, an anthrapyrazolone inhibitor, is recognized as one of the most potent ATP-competitive inhibitors of JNK reported in 2001[408]. Therefore, we used this inhibitor for all three JNK isoforms included in our study. A thorough analysis of the structural dynamics of all three JNK isoforms in both their phosphorylated and unphosphorylated forms, with also focus on the impact of phosphorylation and inhibitor selectivity, has not been previously reported in the literature. Investigating the structural behaviour of these isoforms in phosphorylation states is crucial for understanding their functional variety and regulatory mechanism. Phosphorylation serves as a key mechanism for controlling JNK activity, triggering structural rearrangements and influencing substrate affinity. This study explores the binding interactions of SP600125 with JNK isoforms under phosphorylated and unphosphorylated conditions, offering insights into how post-translational modifications influence kinase-inhibitor binding and structural integrity.

6.2. Materials and Methods

6.2.1. System Preparations

The structures of all three JNK isoforms were retrieved from the PDB database. The structures for JNK1 (PDB ID: 3ELJ)[409], JNK2 (PDB ID: 7N8T)[109], and JNK3 (PDB ID: 3TTI)[410] were obtained from the database, and unnecessary components were deleted while keeping ligand and missing residues were modelled using MODELLER[411] of UCSF Chimera[412]. Moreover, since the sequence similarity among JNK isoforms is more than 80%, the coordinates of the inhibitor SP600125

from the complex of JNK1/SP600125 (PDB ID:1UKI) were taken and fitted to other JNK isoforms with the relative orientation to generate the respective protein/ligand complexes. In this way, all the complexes of SP600125 with JNK1/2/3 isoforms were prepared. Previously, the same methodology was followed for the WNK1 kinase isoform[413]. After the formation of complexes, we phosphorylated the Thr and Tyr residues in the TPY motif of A-loop, creating six systems, three with phosphorylation and three without phosphorylation and subsequently subjected them to MD simulations. The phosphorylated system are denoted as JNK1-p, JNK2-p, JNK3-p and the unphosphorylated system are denoted as JNK1-u, JNK2-u, JNK3-u. Additionally, pairwise sequence alignment was carried out using BLAST[414], and multiple sequence alignment was performed using the Clustal Omega[415] server for each of the four isoforms. Further, we used the DALI webserver[416] to generate isoform similarity heat maps.

6.2.2. Molecular Dynamic Simulations

To explore the structural and conformational dynamics of JNK1, JNK2, and JNK3 in their phosphorylated and unphosphorylated states while interacting with the inhibitor SP600125, we carried out triplicate run of 1μs all-atom molecular dynamics simulations. Amber ff14SB [153] force field was used to model the protein, and the updated general Amber force field (GAFF2) [163] force field was used to model to the ligand SP600125. To represent the phosphorylated tyrosine and threonine residues in the activation loop of JNK, the phosaa10 [343] force field was used. The TIP3P [172] water model was used in a truncated octahedron box with a 10 Å distance from the walls. To ensure neutrality of overall system, appropriate amounts of Na+ and Cl- ions were added. MD simulations were conducted using pmemd.cuda module of AMBER18[417] with GPU acceleration. The Particle Mesh Ewald (PME) [418] method was used to compute long-range electrostatic interactions, providing an representation of non-bonded forces with a cut-off of 10 Å. The SHAKE[169] algorithm was used to constrain all covalent bonds involving hydrogen atoms. The energy minimization process was carried out in two sequential steps. The first minimization with force constrains of 2.0 kcal mol^{-1} $Å^{-2}$ involved 5000 steps using

steepest descent followed by 5000 steps using conjugant gradient algorithm. The second minimization without any constrains involved 100 steps using steepest descent followed by 900 steps using conjugant gradient algorithm. Following energy minimization, the system was heated from 0 K to 300 K in the NVT ensemble with harmonic constraint of 2.0 kcal mol^{-1} $Å^{-2}$. This was followed by equilibration of 50 ps in the NVT with restraint constraint of 2.0 kcal mol^{-1} $Å^{-2}$, followed by equilibration of 1.0 ns with NPT without any restraint. The next step involved running 3 x 1 μs per system long MD simulations, using a 2.0 fs time step in the NPT ensemble, with constant environment of 1.0 atm pressure and 300 K temperature. The berendsen barostat [180] and langevin thermostat [295] were used to maintain constant temperature and pressure throughout the simulation period. Overall, 300,000 frames were generated for each system.

6.2.3. Trajectories Analysis

All the analyses, including the calculation of root-mean-square deviation (RMSD), root-mean-square fluctuations (RMSF), radius of gyration (R_g), and solvent accessible surface area (SASA), were performed using the Cpptraj [264] module from AmberTools19. Scientific plots were generated with the use of Python, and for the visualizing the 3D structures of the proteins, the ChimeraX [291] software was used.

6.2.4. Essential Dynamics Analysis and Free Energy Surface

We used Principal Component Analysis (PCA)[304] techniques to identify structures with low energy (probable conformation) from the high dimensional phase space trajectories using the Cpptraj module in AmberTools19. PCA calculations involve diagonalizing the covariance matrix of positional deviations within a structural ensemble, where the eigenvectors and eigenvalues represent the direction and amplitude of the motion, respectively. Detailed methods on PCA was discussed in our previous work [348,354]. In addition, we study the structural and conformational selections in the loop region using dihedral PCA (dPCA) [216] for further insights.

Free Energy Surface (FES) provides a comprehensive map of all feasible conformations of a molecular entity, aiding in the exploration of spatial arrangements

of interacting molecules and their respective energy states [349]. To investigate the structural reorganization of JNK isoforms, FES plots were generated using the following equation.

$$G_i = -k_B T \ln \left(\frac{N_i}{N_{max}}\right) \quad (1)$$

Here, k_B represents the Boltzmann constant, and T denotes the absolute temperature. N_i refers to the number of molecules in the ith bin, while N_{max} is the highest populated bin. Any bin without a population was assigned an artificial barrier scaled according to the minimum probability value.

6.2.5. Protein Structure Network (PSN)

The structural network analysis was conducted using the web PSN v2.0 [221] server, which applies graph theory-based protein structure networks (PSN) and Elastic Network Model-Normal Mode Analysis (ENM-NMA). In the PSN framework, amino acid residues are represented as nodes, and the percentage strength of their noncovalent interaction between two nodes, '*i*' and '*j*', is determined using the following formula.

$$I_{ij} = \frac{n_{ij}}{\sqrt{N_i N_j}} \times 100 \quad (2)$$

The number of side chain atom-atom pairs between nodes '*i*' and '*j*' within a 4.5 Å cutoff distance is represented as 'n_{ij}'. 'N*i*' and 'N*j*' represent normalization factors. When 'n_{ij}' exceeds the set threshold 'I_{min}', it signifies the presence of a connection or edge between the nodes. A node is considered a hub if it maintains a minimum of four connections (edges) with other nodes at a given I_{min}. Within a network, a community is made of highly interconnected nodes that share stronger interactions internally than with external nodes.

6.2.6. Dynamic Cross-Correlation Matrix (DCCM)

To investigate the motion correlations across various regions of the protein, we performed a dynamic cross-correlation matrix (DCCM) [419] analysis on different phosphorylated and unphosphorylated JNK isoforms. The analysis was conducted using Bio3D [303] module of R in RStudio, where we used the combined trajectories

of triplicate run from last 800 ns for each system. Using the Cα atoms of the protein, the correlation coefficients were determined by applying the time-averaged deviation of the *i*th residue from its mean position (Δr*i*) as follows.

$$C_{ij} = \frac{<\Delta r_i \Delta r_j>}{\sqrt{<\Delta r_i^2><\Delta r_j^2>}} \quad (3)$$

Here, positive C_{ij} value signifies correlated motions between the *i*th and *j*th residues, whereas a negative C_{ij} value suggests anticorrelated motions. When the C_{ij} value equals 0, it shows no correlation between the residues. The correlation data was then mapped onto the protein structure and viewed using PyMOL.

6.2.7. Binding Free Energy Calculations

The binding free energy (ΔG_{bind}) of JNK isoforms with SP600125 was calculated using the MM/PBSA method and the MMPBSA.py script from the AMBER18 software package. This energy is determined by three key factors, internal energy ($\Delta E_{internal}$), desolvation free energy (ΔG_{solv}), and configurational entropy ($T\Delta S$), which are explained through the mathematical equation as follows.

$$\Delta G_{bind} = G_{complex} - G_{protein} - G_{ligand} \quad (4)$$

The free energy for each component in equation 4 is calculated using the corresponding free energy expression given in equation 5.

$$G = < E_{internal} > + < G_{solv} > - T < S_{conf} > \quad (5)$$

The terms $E_{internal}$ and S_{conf} refer to the internal molecular energy in the gas phase and the conformational entropy respectively, while (G_{solv}) represents the solvation free energy, including both polar (G_{pol}) and non-polar (G_{np}) components. the $E_{internal}$ is further described as.

$$E_{internal} = E_{cov} + E_{elec} + E_{vdW} \quad (6)$$

E_{cov}, E_{elec}, and E_{vdW} correspond to the covalent, electrostatic, and van der Waals interactions. The covalent interaction term, E_{cov}, can be further decomposed as follows.

$$E_{cov} = E_{bond} + E_{angle} + E_{dihedral} \quad (7)$$

The calculation of the non-polar solvation term (G_{np}) is based on an empirical formula that includes adjustable parameters.

$$G_{np} = \gamma \times \mathrm{SASA} + \beta \quad (8)$$

The LCPO method[351], with a 1.4 Å probe radius, was used to compute the solvent-accessible surface area (SASA). The surface tension (γ) and cavity offset (β) parameters were assigned values of 0.0378 kcal $mol^{-1}Å^{-2}$ and 0.5692 kcal/mol, respectively. The MMPBSA.py script in the AMBER18 package was employed to calculate the binding free energy (enthalpy, H) of the protein-ligand complexes. Using the Adaptive Poisson–Boltzmann Solver (APBS), the solvation free energy (GPB) was computed, where the linear Poisson–Boltzmann equation solved using the default maximum iteration settings in AMBER18. The term S_{conf}, representing configurational entropy, can be described as,

$$S_{conf} = S_{trans} + S_{rot} + S_{vib} \quad (9)$$

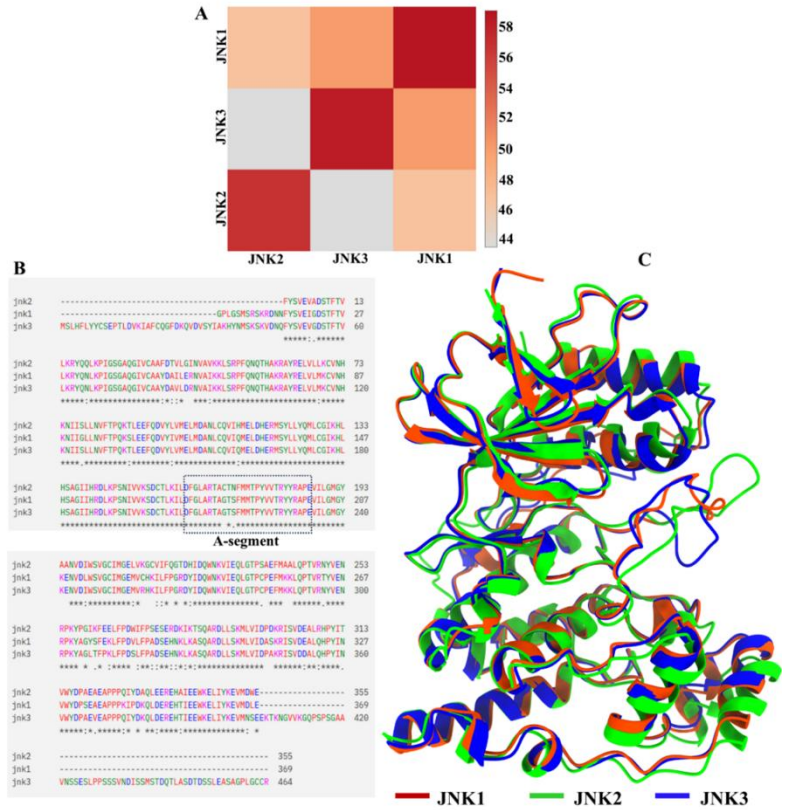
Entropy contributions from translational (S_{trans}), rotational (S_{rot}), and vibrational (S_{vib}) motions were calculated from normal mode analysis [352]. Further, the MM/PBSA method was used to decompose the binding free energy at the amino acid level [308,353,420].

6.3. Results

6.3.1. Structural Comparison Between JNK Isoforms

A systematic structural analysis of JNK isoforms is essential for identifying common structural features while accurately mapping the locations of amino acid modifications, which could have significant implications for their functional regulation in isoforms. The structural similarities between JNK1, JNK2, and JNK3 isoforms were analysed using the DALI web server, and the results are visualized in the heatmap presented in **Figure 6.2A**. As shown in **Figure 6.2A**, the Z-scores between JNK3 and JNK2 (43.7), JNK3 and JNK1 (50.3), and JNK1 and JNK2 (47) suggest a high degree of structural similarity among the isoforms, with JNK3 and JNK1 showing the closest similarity. Additionally, we used the clustal omega server to align the sequences of JNK1, JNK2, and JNK3. The observation highlights the identification of conserved regions and sequence variations, offering insights into their structural and functional similarities

and the results is shown in **Figure 6.2B**. The multiple sequence alignment also shows that JNK1 and JNK3 exhibit greater similarity. The activation loop, highlighted in the **Figure 6.2B**, demonstrates that JNK1 and JNK3 have a conserved A-loop, while JNK2 differs by two residues. For a detailed comparison of the point-to-point residue differences and similarities between the JNK isoforms, please refer to **Figure 6.2B.** To assess the structural similarities and differences among JNK1, JNK2, and JNK3, we performed a structural alignment using the MatchMaker tool of UCSF Chimera and visualized their 3D conformations and shown in **Figure 6.2C** and different RMSD values in **figure 6.2D**.



D

	JNK1	JNK2	JNK3
JNK1	0 Å; 0 Å	0.959 Å; 2.448 Å	0.632 Å; 0.674 Å
JNK2	0.959 Å; 2.448 Å	0 Å; 0 Å	0.972 Å; 2.474 Å
JNK3	0.632 Å; 0.674 Å	0.972 Å; 2.474 Å	0 Å; 0Å

Figure 6.2: (A) Heat map representing sequence similarity among JNK1/2/3 isoforms, generated using the DALI webserver. **(B)** Residue level sequence similarity analysis

based on multiple sequence alignment, performed using Clustal Omega. **(C)** Superimposed 3D structures of JNK1/2/3 isoforms visualized using Chimera, displaying structural similarities and differences in 3D visuals. **(D)** The RMSD differences between JNK1, JNK2, and JNK3 kinase domain structures. The first value shows the RMSD calculated using pruned atom pairs, while the second value is the RMSD calculated using all atom pairs.

6.3.2. Molecular Dynamics Simulations

We performed triplicate run (3 x 1 μs) all-atom molecular dynamics (MD) simulations to investigate the conformational dynamics and stability of the JNK isoforms. MD simulations provide important information about protein dynamics over time, such as their structural flexibility, structural stability, ligand interaction mechanism, and conformational changes under applied conditions. These results provide deeper insights into the overall dynamic behaviour of proteins, along with important key regions such as active and catalytic sites and binding pockets. Our main aim through simulations was to compare the structural characteristics of the JNK isoforms and identify any significant differences or similarities that may contribute to their distinct functional roles. From the simulation trajectories, we have performed various analysis such as root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), solvent-accessible surface area (SASA), radius of gyration (RoG), principal component analysis (PCA), dihydral principal component analysis (dPCA), salt bridge analysis, network analysis, among others, to investigate the dynamics and conformational plasticity of JNK isoforms in both phosphorylated and unphosphorylated form, aiming to gain insights into their structural and functional characteristics. Except for the time-series plots, the last 800 ns trajectories out of the 1 μs independent runs were considered to carry out the various analyses.

6.3.3. Structural Stability and Flexibility Analysis of JNK Isoform

We initially assessed the stability of the simulation trajectories for each system by calculating the RMSD of the backbone atoms. The time evolution plot of RMSD is shown in provided in supplementary information as **Figure D1**. The RMSD time evolution indicates that each system reaches stability after a ~400 ns of simulation time, and all systems show stable and convergence in the later stages of the simulations.

To investigate further, we performed kernel density estimation (KDE) on the RMSD data of kinase domain backbone atoms from triplicate simulations, and the resulting probability distribution is shown in **Figure 6.3A**. The JNK1 system exhibits a unimodal distribution with single peak, with phosphorylated JNK1 having an RMSD of ~2.7 Å and unphosphorylated JNK1 showing an RMSD of ~1.9 Å. For JNK2, the phosphorylated system exhibits a small single peak at ~2.8 Å with a broad distribution, whereas the unphosphorylated system shows a narrow single peak around ~2.8 Å. For JNK3, both the phosphorylated and unphosphorylated systems exhibit a broad distribution with lower density, and the RMSD ranges between ~1.5 Å and ~3.2 Å. Further, we have also calculated the average RMSD for all three runs per system, and the values are provided in supporting information **Table 6.1**. The RMSD values for all systems fall within the range of 1.87 Å to 2.77 Å. For more detailed RMSD values of each individual run, please refer to **Table 6.1**. Prior to conducting further analyses, we first verified whether the inhibitor SP600125 remained within the binding pocket. To achieve this, we calculated the center-of-mass (COM) distance between the ligand and amino acid residues of binding pocket within a 5 Å radius in the initial conformations (see **Figure D2A**). The results confirmed that in all six systems, the ligand remains within the binding pocket throughout the simulations. Additionally, ligand RMSD was calculated, revealing a unimodal distribution, indicating that all ligands remained stable within the pocket (see **Figure D2B**).

Table 6.1: Average Root-mean-square deviation (RMSD) for each production simulation. The standard deviation values are reported in parentheses.

Systems	**Run1 (RMSD (Å))**	**Run2 (RMSD (Å))**	**Run3 (RMSD (Å))**
JNK1-p	2.62 (0.16)	2.57 (0.26)	2.48 (0.24)
JNK1-u	1.87 (0.15)	1.96 (0.18)	1.99 (0.18)
JNK2-p	2.72 (0.19)	2.77 (0.26)	2.23 (0.20)

JNK2-u	2.60 (0.27)	2.70 (0.35)	2.72 (0.20)
JNK3-p	2.57 (0.32)	2.55 (0.24)	2.11 (0.21)
JNK3-u	1.98 (0.23)	2.37 (0.21)	2.17 (0.27)

Moreover, we have evaluated the RMSD of backbone atoms of several critical regions in the kinase domain that are essential for the catalytic activity of JNK kinase. **Figure 6.3B** represents the probability density distribution of the RMSD for the activation segment (A-segment) across different JNK isoforms in both phosphorylated and unphosphorylated states. As shown in **Figure 6.3B**, JNK1 systems shows unimodal distribution with different RMSD values. In case of JNK2 and JNK3 we have observed more than one peak with broad distributions (RMSD ranges from 1.0-5.8 Å). The presence of a multimodal distribution suggests that the system samples multiple conformational states throughout the simulation. The broad RMSD range signifies notable flexibility in this region, with distinct peaks corresponding to different conformational substates. The stability of the phosphorylated and unphosphorylated systems varies, as reflected by the differences in peak positions and their distributions. This suggests that the A-segment undergoes dynamic structural rearrangements that may impact the functional properties of JNK isoforms. For a more comprehensive understanding of A-segment dynamics, we employed dPCA on the activation segment and also calculated the residue-wise RMSD considering the A-segment residues. The interpretation of these findings will be discussed in the subsequent sections. As depicted in **Figure 6.3C**, we also calculated the backbone RMSD of the P-loop and observed a single distribution across all systems, with some exhibiting a narrow peak while others show a broader peak. A comprehensive analysis of the P-loop will be presented in a later section, where we have dPCA and extracted representative structures to display the actual dynamics of the P-loop across different isoforms. We also computed the RMSD for the αC-helix, catalytic loop, hinge region, and DFG motif. The RMSD of the αC-helix exhibited a unimodal distribution across all systems,

with minimal variability and values ranging from ~0.25-1.0 Å (see **Figure 6.4C**). Similarly, the catalytic loop showed a unimodal distribution with minimal variation, with RMSD values between ~0.2-1.0 Å (see **Figure 6.4D**). The hinge region displayed nearly identical RMSD patterns with peak at ~0.5 Å (see **Figure 6.4E**). Similarly, the RMSD of the DFG motif exhibits a unimodal distribution across all systems, however in the case of JNK2-p the distribution is broader (see **Figure 6.4F**). Overall, these observations indicate that the activation loop exhibits a highly dynamic nature.

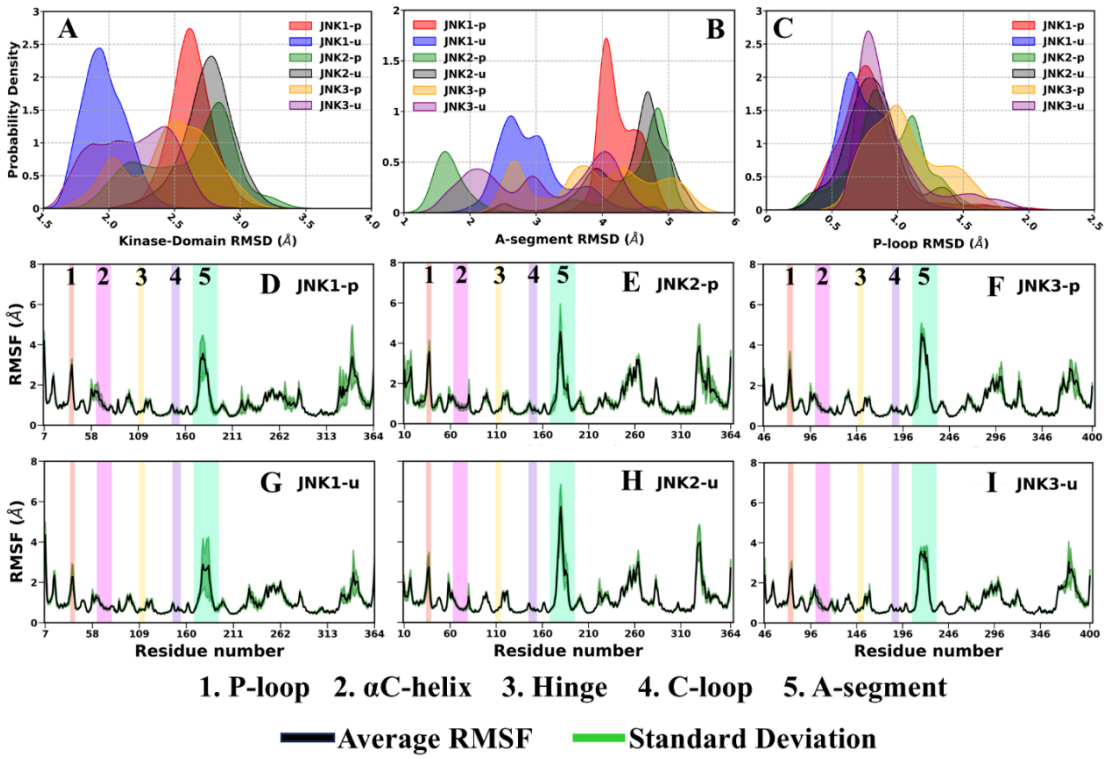


Figure 6.3: Probability distribution of backbone RMSD for **(A)** kinase domain of JNK1/2/3, **(B)** activation segment, **(C)** P-loop. (D-I) The average RMSF of each three runs (black) and the standard deviation of each run (green).

Next, to assess the flexibility of individual residues of the proteins, we computed the root-mean-square fluctuations (RMSF) for all isoform systems, and the results are shown in **Figure 6.3 D-I**. The several important regions are highlighted in the figure. All systems exhibit a similar fluctuation pattern, with the loop regions showing higher flexibility and following a consistent trend across all isoform systems. The activation segment of the JNK2 system exhibits greater fluctuations in both its phosphorylated

and unphosphorylated states compared to the other systems (see **Figure 6.3 E** and **H**). The A-segment of both JNK1 and JNK3 exhibits similar fluctuations in their unphosphorylated and phosphorylated states, likely due to the sequence identity in the A-segment region. Overall, across all isoform systems the RMSF exhibits a consistent fluctuation pattern.

Additionally, we assessed the protein compactness and solvent accessibility in various isoform systems by calculating the probability distributions of the radius of gyration (R_g) and solvent accessible surface area (SASA), respectively, as depicted in **Figure 6.4 A-B** of supporting information. As depicted in **Figure 6.4-A**, the SASA value for JNK1 increases upon phosphorylation, with the peak for JNK1-p being ~172 nm², compared to around 165 nm² for JNK1-u. Similar trend is observed for JNK3, where the JNK3-p system shows peak around ~170 nm², while in case of JNK3-u it is around ~165 nm². In the case of JNK2, there is also slightly increase in SASA value after phosphorylation. Overall, phosphorylation induced activation of the kinase may correlate with the outward movement of Activation loop (open conformation), leading to increased SASA values due to the creation of more solvent-accessible area. The radius of gyration (R_g) for all isoform systems was also evaluated and presented in **Figure 6.4-B**, showing values range between 21 and 22.30 Å. The unphosphorylated systems exhibit a slight decrease in R_g values, which may result from the A-loop not being fully extended, leading to increased compactness relative to the phosphorylated systems.

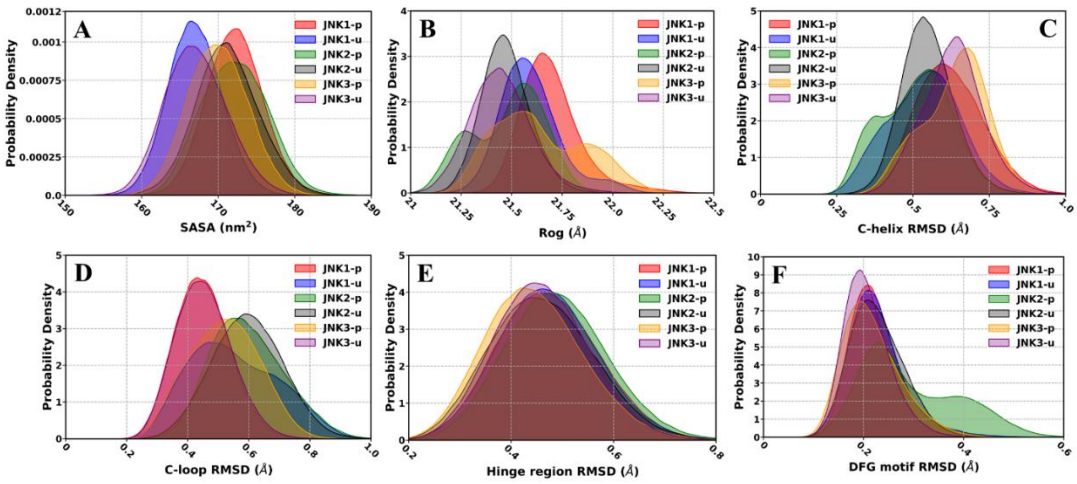


Figure 6.4: The probability distribution of **(A)** Solvent-accessibility surface area, **(B)** Radius of gyration, **(C)** αC-helix RMSD, **(D)** Catalytic loop RMSD, **(E)** Hinge region RMSD, **(F)** DFG motif RMSD.

6.3.4. PCA-based Analysis of Protein Global Dynamics

Principal Component Analysis (PCA) is a method used to simplify complex datasets by transforming correlated variables into independent principal components. It is widely used in structural biology to interpret MD trajectory data, enabling the visualization of prominent motions within biomolecular systems. PCA was performed on the Cα atoms of amino acid residues to characterize the broad conformational motions in the kinase domain of JNK isoforms. We have taken the first two principal components, PC1 and PC2 for our analysis as it accounts maximum global motion exhibited by JNK isoforms. Using K-means clustering, we classified the sampled conformations within the systems into distinct clusters based on PC1 and PC2, as shown in **Figure D3**. Based on our observations, three clusters optimally represent the data distribution, where the most populated cluster comprises around ~33.70-60.20% of frames, and the least populated cluster includes ~15.87-33.87% of frames (see **Figure D3**). We also examined the conformational space in terms of free energy basins by generating free energy landscapes based on PC1 and PC2 as shown in **Figure 6.5**. Additionally, we retrieved the structures and highlights various conformations corresponding to the minima regions with high-population (see **Figure 6.5**). The multiple minima with low to moderate energy barriers observed in case of all systems with broad sampling space suggesting high degree of conformational variability. The PCA plot indicates that the phosphorylated system occupies a wider conformational space than the unphosphorylated system for all isoforms, indicating structural rearrangement after phosphorylation. Structures extracted from the minima regions indicate that the loop regions, mainly the P-loop and A-loop along with others, display high dynamic behaviour across different conformations (see **Figure 6.5**). We observed a slight movement of the αC-helix outward in conformation 3 (conf3) in the case of JNK1-p and JNK2-p systems.

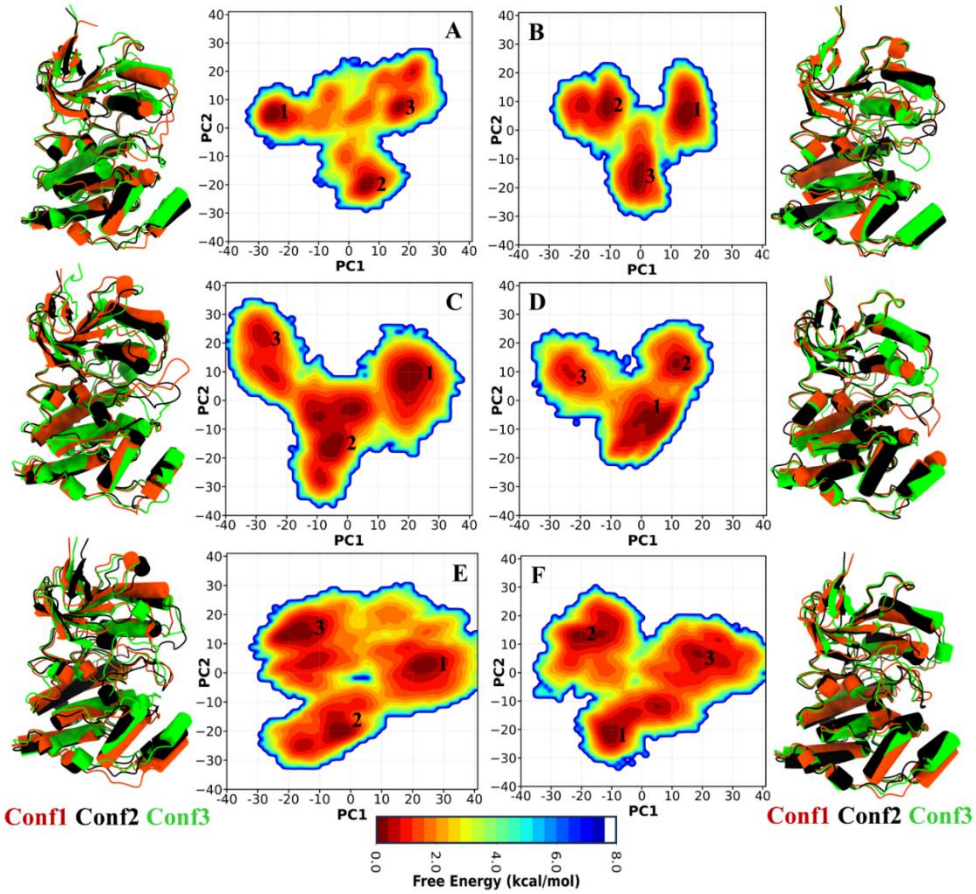


Figure 6.5: PCA analysis was carried out for six systems, including **(A)** JNK1-p, **(B)** JNK1-u, **(C)** JNK2-p, **(D)** JNK2-u, **(E)** JNK3-p, and **(F)** JNK3-u. The figure illustrates the key structural conformations obtained from the low-energy regions of the free energy landscape.

6.3.5. Structural Plasticity of The Activation Segment

The activation segment of protein kinase is essential for regulating its activity, with phosphorylation acting as a key mechanism to control this function and cellular homeostasis[421]. The conformation of the activation segment impacts kinase catalysis, and its phosphorylation status determines its active form, hence crucial for kinase regulation and therapeutic targeting[422]. As mentioned earlier, the RMSD of the A-segment exhibits a wide range of distribution, except for JNK1. To better understand the conformational shifts in the activation segment, we performed principal component analysis using the phi/psi dihedral angles (dPCA) of the activation segment

residues. The percentage distribution of conformation populations in various clusters for each system is shown in **Figure D4** of the Supporting Information, with respect to the first two modes (dPC1 and dPC2) via K-means clustering. A free energy landscape was created to explore the free energy basins using dPC1 and dPC2, showing the different prominent conformation states of the activation segment as shown in **Figure 6.6**. As observed from **Figure 6.6**, the unphosphorylated system shows a broader conformational space, whereas upon phosphorylation it narrows the space and restricts conformational selection. As seen in **Figure 6.6A-B**, the JNK1 phosphorylated system (JNK1-p) shows three energy minima, with conformations that are all extended and open (active conformations). In contrast, the JNK1 unphosphorylated system (JNK1-u) shows a broader sampling space with three minima highlighted, where conformation 1 and 2 (conf1 and conf2) are close or semi-extended (inactive-like). Additionally, we observed the formation of a short helix in the later part of the A-loop in conf2 and conf3 of the JNK1-p system (see **Figure 6.6A**). For JNK2 (see **Figure 6.6C-D**), both the phosphorylated and unphosphorylated forms exhibit similar distribution patterns with multiple minima. The structures extracted from these minima show that all adopt an open A-loop conformation. In the JNK2-p system, conf1 forms a short helix in the distal part, whereas in the JNK2-u system, conf3 forms helix in both the middle and distal regions of the A-loop. For JNK3 (see **Figure 6.6E-F**), both the phosphorylated and unphosphorylated forms show similar distribution patterns with multiple minima. In the JNK3-p system, conformation 3 is semi-extended, while the other two conformations are extended with an open A-loop. In the JNK3-u system, conformations 1 and 2 are semi-extended and move towards the αC-helix (inactive like), whereas conformation 3 is extended with a helix formation in the distal part. JNK1 and JNK3 exhibit similar conformations (conf1 and conf2) in their unphosphorylated forms, while JNK2 displays similar open conformations in both its phosphorylated and unphosphorylated states. Overall, phosphorylation has minimal effect on the A-loop dynamics of JNK2, while in JNK1 and JNK3, phosphorylation triggers a conformational shift towards an active-like A-loop state.

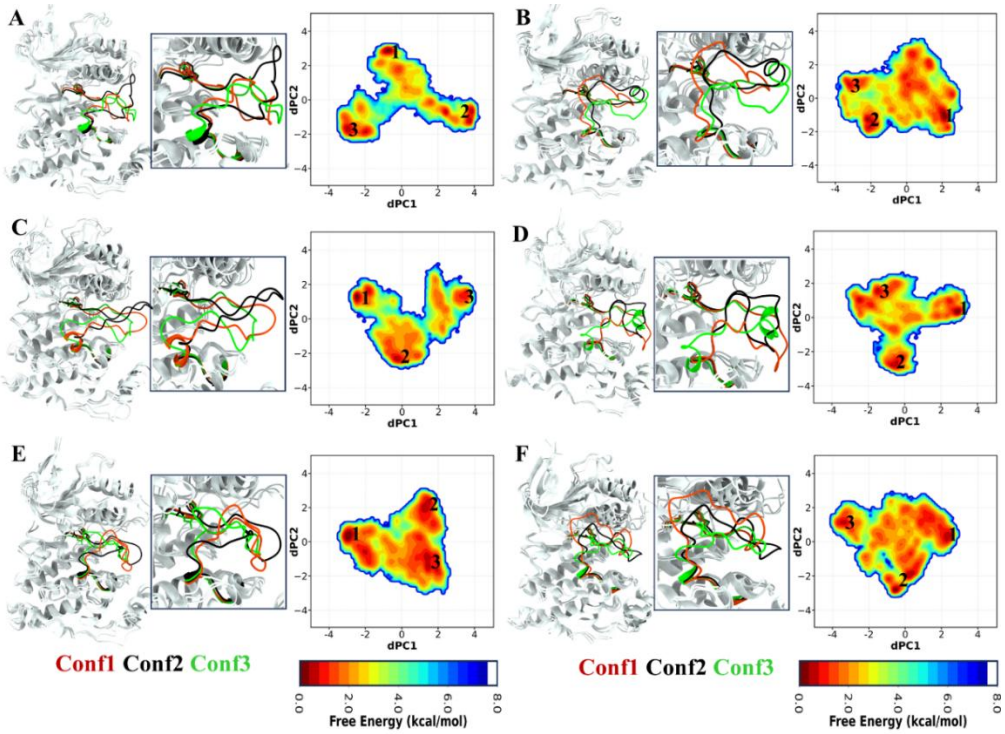


Figure 6.6: dPCA based free energy landscape analysis was done for the six studied systems, **(A)** JNK1-p, **(B)** JNK1-u, **(C)** JNK2-p, **(D)** JNK2-u, **(E)** JNK3-p, and **(F)** JNK3-u. The structure highlights the key conformations extracted from the lowest energy regions, representing the most stable A-loop states observed during the simulations.

Furthermore, the heavy-atom RMSD time evolution for each residue in the activation segment was calculated, as depicted in **Figure 6.7**. Analysis of residual RMSD heatmaps for the activation segment of JNK1/2/3, in both phosphorylated and unphosphorylated states, shows distinct conformational changes over the simulation period. Across the three isoforms, phosphorylation enhances the dynamic nature, as indicated by increased RMSD values, particularly around important residues that may stabilize the activation loop, in comparison to the unphosphorylated states. The phosphorylating residues, Tyr and Thr exhibit higher RMSD values in the phosphorylated systems of JNK1 and JNK3 compared to JNK2 (see **Figure 6.7**). Broadly, as phosphorylation is introduced from the unphosphorylated crystal structure, we observe an increase in RMSD values due to structural rearrangement to achieve A-loop stabilization, mainly in case of JNK1/3. These results indicate that

phosphorylation increases the structural flexibility of the activation segment, potentially contribute in substrate recognition and enzymatic activity. The varying degrees of flexibility may play a role in the functional differences observed among JNK isoforms.

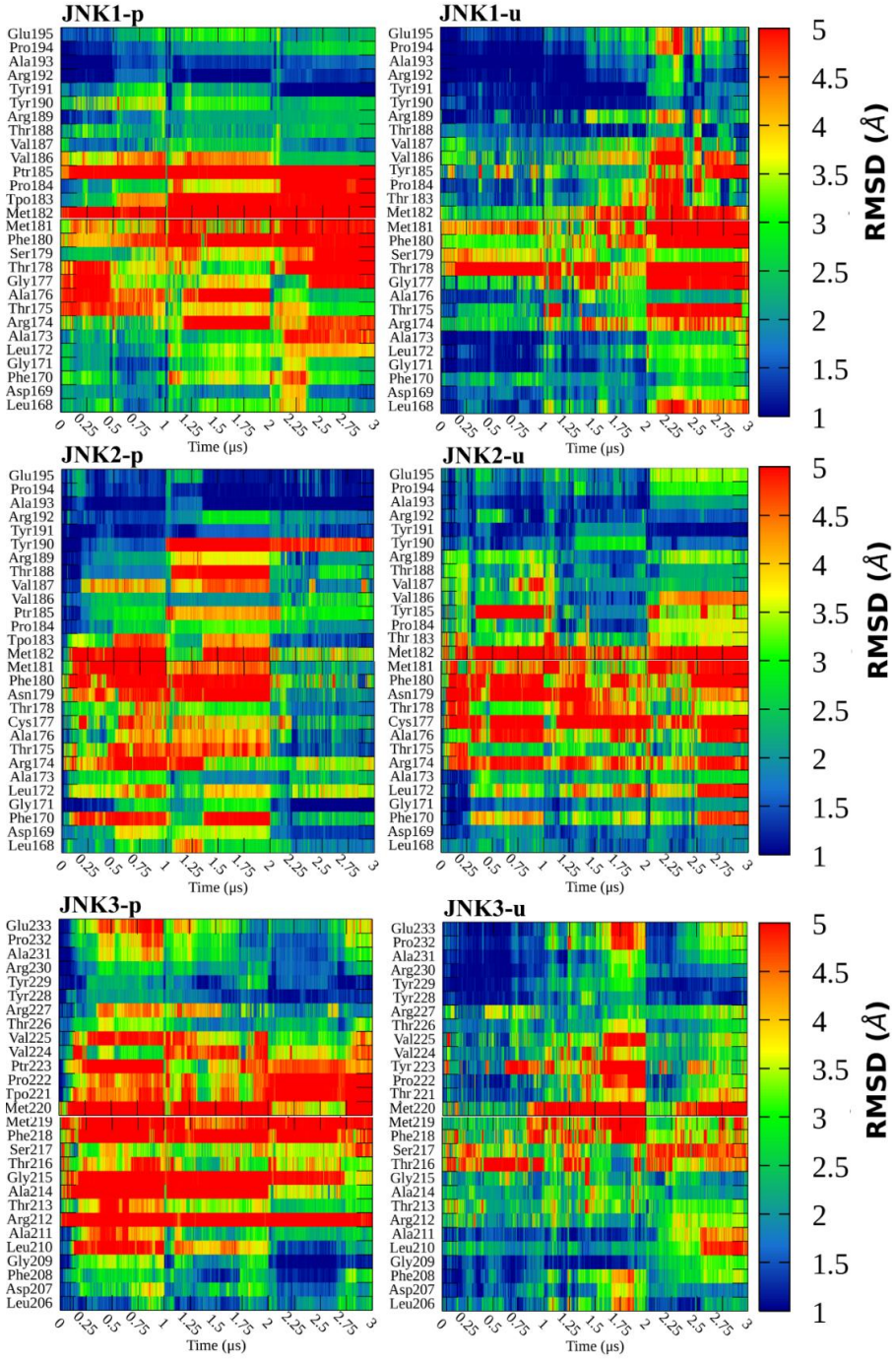


Figure 6.7: Time evolution of residue specific RMSD of activation segment region. The time evolution from 0-3 μs represent to production simulation of three runs (run1-run3) respectively.

6.3.6. Salt Bridge Triad and Hydrophobic Spines Analysis

The hallmark salt bridges in kinases are important as they play important roles in stabilizing both active and inactive conformations, thereby influencing kinase activity and substrate recognition[423]. The highly conserved regulatory triad, the salt bridge between the β3-Lys with αC-helix-Glu, and DFG-motif-Asp residues in JNK stabilizes its active conformation, crucial for ATP binding and catalysis. We measured the distance between the NZ atom of β3-Lys and two specific residues, the CD atom of αC-helix Glu and the CG atom of the Asp residue in the DFG motif. **Figures 6.8A–C** display the measured distance between hallmark salt bridge β3-Lys and αC-helix Glu in different JNK isoforms in both phosphorylated and unphosphorylated states. All the JNK isoform shows distinct characteristics regarding these salt bridge formation. In the case of JNK1 phosphorylated state (JNK1-p), it leads to reduction in the probablity density of salt bridge formation as compared to unphosphorylated JNK1 (JNK1-u). This is evident from the narrow, high-density peak observed at ~3.5 Å in the unphosphorylated (JNK1-u) system, indicating a modulation of this interaction (see **Figure 6.8A**). In the case of JNK2, the β3-Lys and αC-helix Glu salt bridge maintains a consistent distance distribution, with a dominant peak around ~3.5 Å and a small peak at ~5.2 Å in case of both phosphorylated (JNK2-p) and unphosphorylated (JNK2-u) systems, indicating phosphorylation has minimal impact on this interaction (see **Figure 6.8B**). In the JNK3 system, an opposite trend is observed as compared to JNK1. Upon phosphorylation (JNK3-p), the formation of the β3-Lys and αC-helix Glu salt bridge is consistent, as indicated by a sharp peak at ~3.5 Å. In contrast, the unphosphorylated state (JNK3-u) exhibits a weaker interaction, represented by a smaller peak at ~3.5 Å with an overall broad distribution (see **Figure 6.8C**).

Additionally, the distance between β3-Lys and DFG-motif-Asp was calculated for all JNK isoforms to determine the presence and stability of this salt bridge formation, and shown in **Figure 6.8D-F**. Across all JNK isoforms, no stable salt bridge between β3-

Lys and DFG-motif-Asp formation was observed. For JNK1, the unphosphorylated system (JNK1-u) exhibits a sharp peak around 6.5 Å, whereas the phosphorylated system (JNK1-p) shows a broad and low density peak (see **Figure 6.8D**). For JNK2, both the phosphorylated (JNK2-p) and unphosphorylated (JNK2-u) systems shows sharp peaks at ~5.5 Å and ~6.3 Å, indicating a slight reduction in distance upon phosphorylation (see **Figure 6.8E**). For JNK3, both the phosphorylated (JNK3-p) and unphosphorylated (JNK3-u) systems demonstrates a similar broad distribution pattern, with distances ranging from ~3.5 Å to ~10.8 Å, indicating no significant effect of phosphorylation (see **Figure 6.8F**).

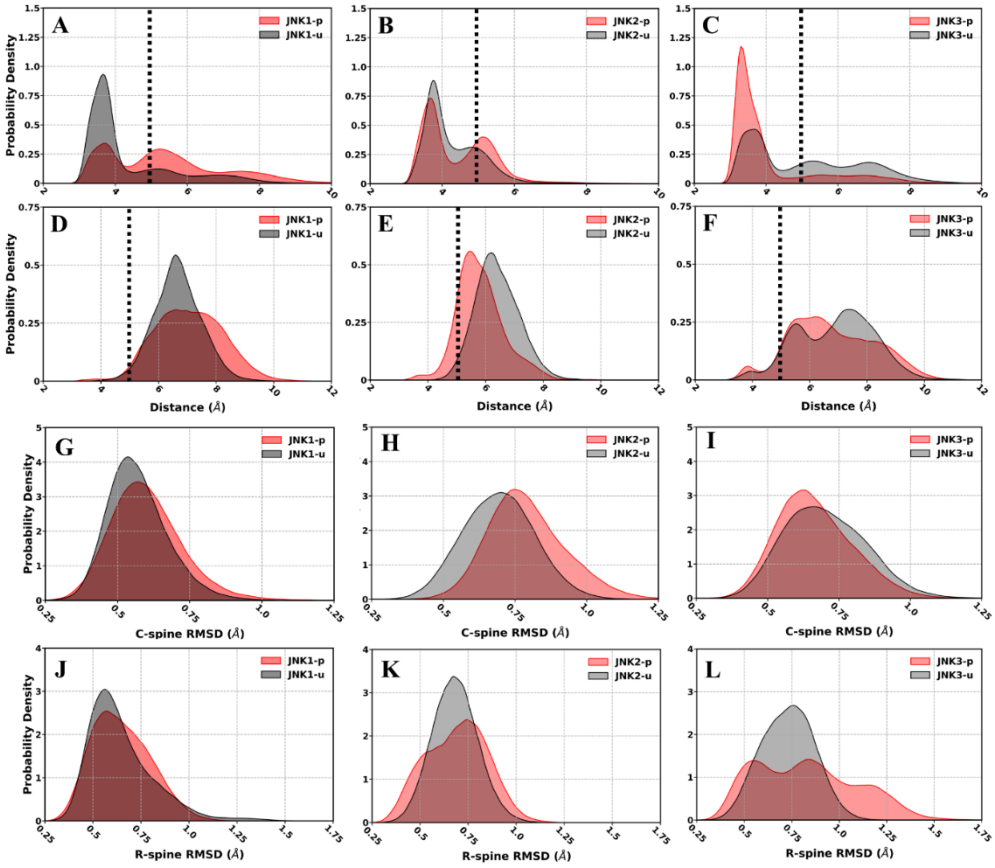


Figure 6.8: (A-C) Probability density distribution of the salt bridge distance between β3-Lys and αC-helix Glu during the simulation period. **(D-F)** Probability density distribution of the salt bridge distance between β3-Lys and DFG-motif-Asp during the simulation period. **(G-I)** Probability distribution of backbone RMSD for C-spine, and **(J-L)** for R-spine.

Studies by Kornev et al. has identified important hydrophobic residues responsible for assembling the catalytic (C-spine) and regulatory spines (R-spines) in protein kinases, ensuring structural integrity and proper ATP positioning [424]. The active form of JNK kinase maintains aligned R-spine, whereas inactivation leads to disrupt the alignment. We calculated RMSD for both R-spine and C-spine considering all heavy atoms, and their KDE plots presented in **Figure 6.8G-L**. The C-spine and R-spine RMSD plots indicate minimal variation among different studied systems. In the C-spine (**Figure 6.8G-I**), JNK1-u peaks at ~0.5 Å, with JNK1-p showing a slightly broader distribution. JNK2-p and JNK2-u exhibit overlapping peaks around ~0.7-0.75 Å. JNK3-p and JNK3-u shows peak around ~0.6 Å with similar distribution. In the case of R-spine (**Figure 6.8J-L**), JNK1-u and JNK1-u system shows peaks at ~0.6 Å with similar type of distribution. JNK2-p and JNK2-u shows nearly identical peaks ~0.7 Å. In the case of JNK3-p broader distribution (distance ranges ~0.25-1.6 Å) compared to JNK3-u which shows sharp peak around ~0.75 Å. The structures extracted from the stable regions of the R-spine RMSD indicate that, in all cases, the R-spine remains aligned, with no R-spine breakage observed in any system (see **Figure 6.9**). Overall, these findings indicate that phosphorylation along with different isoform JNK protein does not significantly alter the hydrophobic spine organization in JNK.

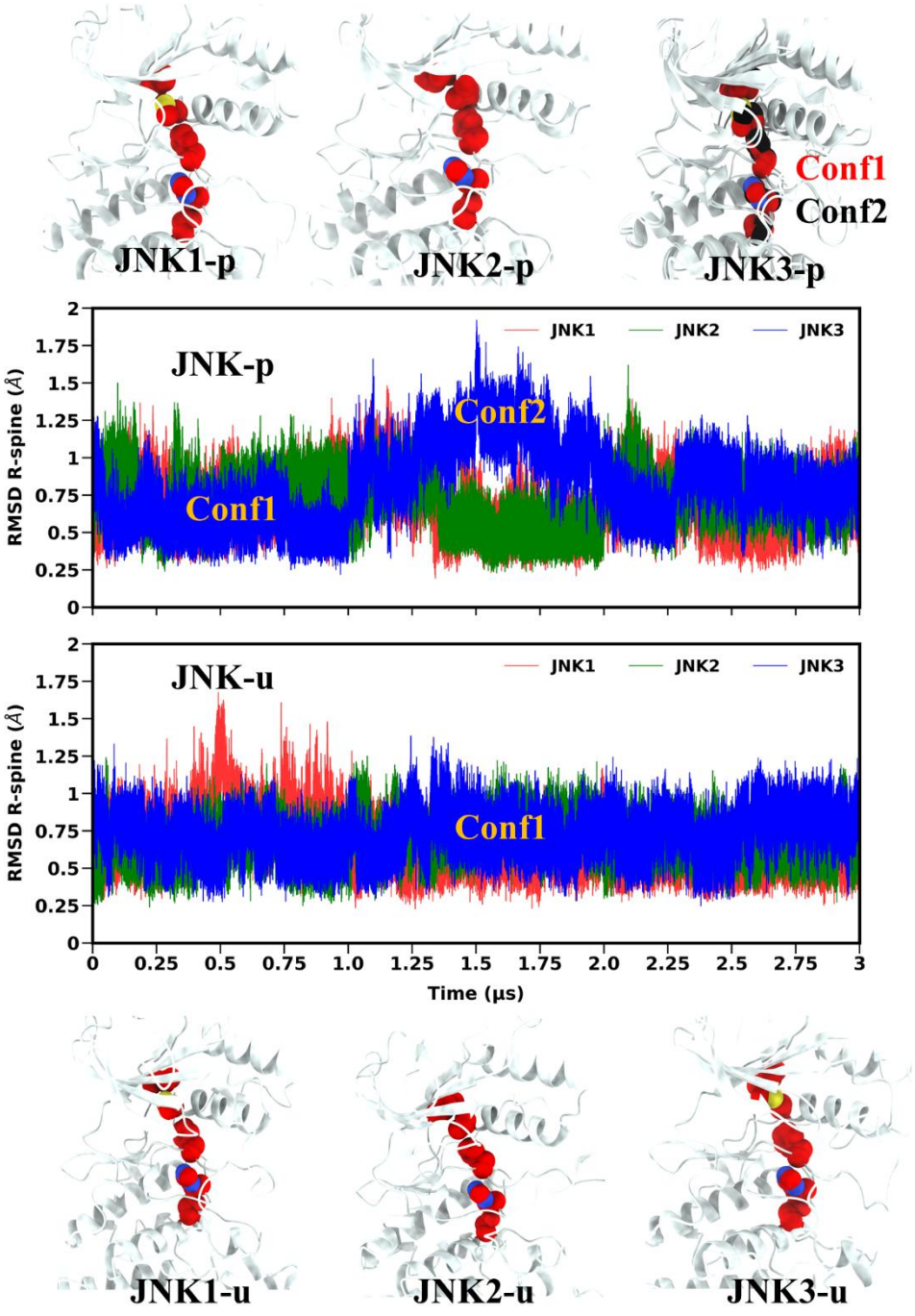


Figure 6.9: Time evolution of the regulatory spine (R-spine) RMSD in JNK isoforms in phosphorylated (JNK-p) and unphosphorylated (JNK-u) states. Structural snapshots show the R-spine residues (red and black spheres) in different conformations, with phosphorylation inducing a shift from Conf1 to Conf2 in JNK3.

6.3.7. Protein Structure Network (PSN) and DCCM Analysis

Protein structure network (PSN) [221] methods were used to study how proteins communicate over long distances and to examine phosphorylation influences and isoform differences through network-based methods. Our objective through PSN analysis was to gain a more detailed insight into the residue level connections within the system and compare with other systems. **Table 6.2** highlights the average number of links, hubs, and community parameters, and **Figure 6.10** represent communities along with shared and unique hubs between phosphorylated and unphosphorylated systems of JNK isoforms. Here, amino acid residues serve as nodes, and their non-covalent interactions with other nodes are represented as edges with a cutoff distance of 4.5 Å. Interestingly, the average network properties varied significantly among the different JNK systems indicating structural and functional variations among them. The number of nodes remains relatively constant for all the systems (323-330), but the number of links among different systems differs, with JNK3-u having the highest (368) and JNK2-p the lowest (350). Hub formation varies, with JNK1-p, JNK2-u, and JNK3-u having the highest number of hubs (44), while JNK1-u and JNK2-p have lower number (37 and 36). The number of links mediated by hubs follows the same pattern, with JNK2-u at the highest (169) and JNK1-u at the lowest (140). The number of community ranges from 6 (JNK1-u) to 13 (JNK2-p and JNK2-u), with JNK2-u exhibiting the high interconnected network (65 nodes, 85 links). **Table D1** provides detailed information on the nodes, links, and hubs contributing to the formation of each community. The overall results indicates that phosphorylation affects residue connectivity by modifying hub distribution and community structure in different JNK isoforms.

Table 6.2: The network analysis of the four studied systems highlighting network properties.

Network Properties	JNK1-p	JNK1-u	JNK2-p	JNK2-u	JNK3-p	JNK3-u
No. of linked nodes	329	330	327	323	324	327

No. of links	367	354	350	362	358	368
No. of hubs	44	37	36	44	41	44
No. of links mediated by hubs	158	140	142	169	155	160
No. of communities	10	6	13	13	7	7
No. of nodes involved in communities	59	43	55	65	51	57
No. of links involved in communities	80	57	69	85	69	79

Additionally, a comparison was conducted, to analysing the differences and similarities in network properties between phosphorylated and unphosphorylated systems of JNK isoforms (see **Table D1-D4** and **Figure 6.10G-I**). The network analysis between JNK1-p and JNK1-u shows high similarity with 97.58% shared nodes and 83.33% shared links. However, JNK1-u has more links (367) and hubs (44), with a higher proportion of specific hubs (38.64%), suggesting increased network in the JNK1-u system. Similarity indicator Jaccard 72.84%, Otsuka 79.93%, and Overlap Coefficient 87.80% indicate strong overlap in neighbouring interactions, while a graphlet similarity of 60.27 suggests moderate topological differences (see **Table D2**). The network analysis between JNK2-p and JNK2-u shows high similarity with 98.76% shared nodes and 84.25% shared links. However, JNK2-p has more links (362) and hubs (44), with a higher proportion of specific hubs (34.09%), suggesting increased network in the JNK2-p system. Similarity indicator Jaccard 77.40%, Otsuka 83.24%, and Overlap Coefficient 90.08% indicate strong overlap in neighboring interactions, while a graphlet similarity of 66.02 suggests moderate topological differences (see **Table D3**). Similarly, for JNK3-p and JNK3-u which shows high similarity with 97.25% shared nodes and 83.97% shared links. However, JNK3-p has more links (368) and hubs (44), with a higher proportion of specific hubs (22.73%), suggesting increased network in the JNK3-p system. Similarity indicator Jaccard 75.47%, Otsuka 81.52%, and Overlap Coefficient 88.26% indicate strong overlap in neighboring

interactions, while a graphlet similarity of 64.44 suggests moderate topological differences (see **Table D4**).

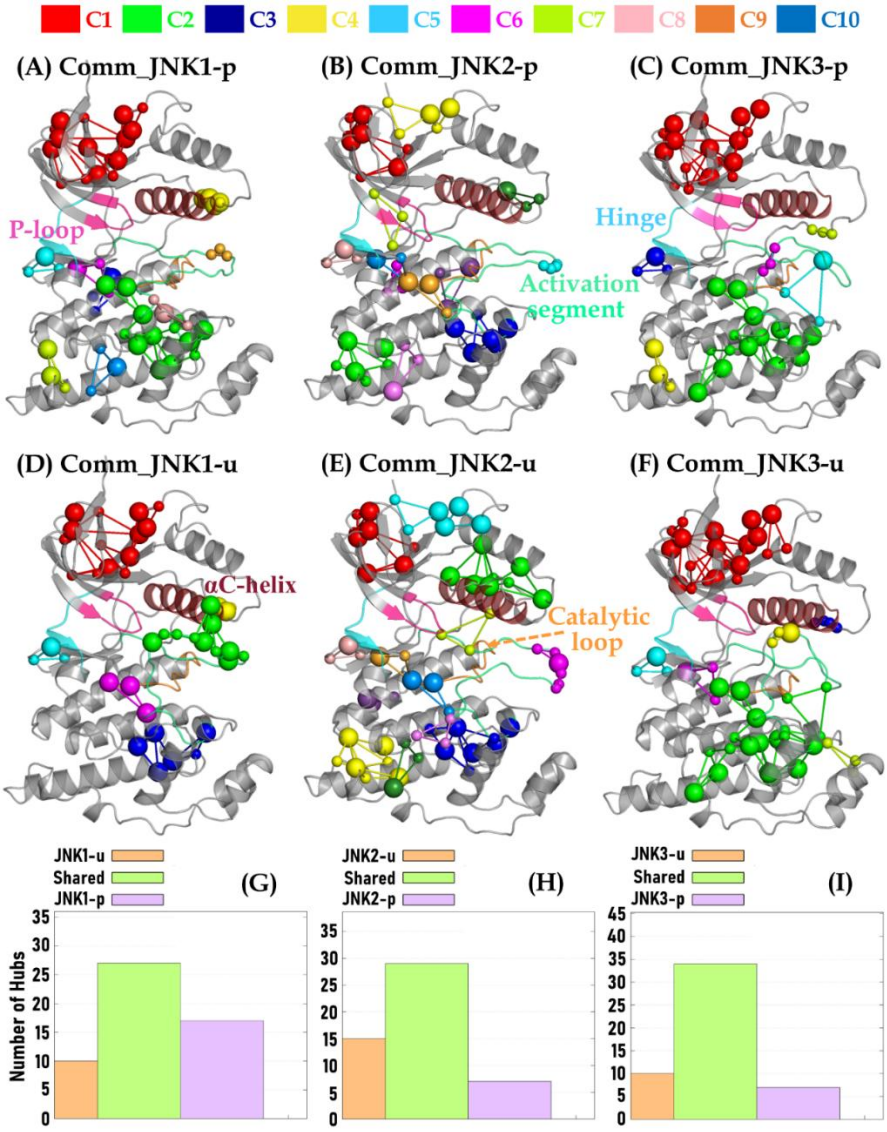


Figure 6.10: (A)-(F) Network derived community maps for the six studied systems. **(G)-(I)** Variations in hub organization between the **(G)** JNK1-p and JNK1-u, **(H)** JNK2-p and JNK2-u, **(I)** JNK3-p and JNK3-u. Color coding differentiates hub interaction averages, and the hierarchical structure of communities from higher to lower level.

Additionally, dynamic cross-correlation matrix (DCCM) analysis was performed to systematically examine how residual movements are correlated across different

regions in JNK isoforms in its phosphorylated and unphosphorylated states. **Figure 6.11** presents color-coded DCCM maps, where positive values indicate correlated residue movement (moving in the same direction), whereas negative values represent anti-correlated motion (moving in opposite directions). An overall increase in anti-correlated motions (more blue patches) is observed as the system shifts from the unphosphorylated to the phosphorylated state, indicating that phosphorylation facilitates structural rearrangement in JNK isoforms. Among the three isoforms, JNK3 systems shows the high degree of anti-correlated motion in both its phosphorylated and unphosphorylated states.

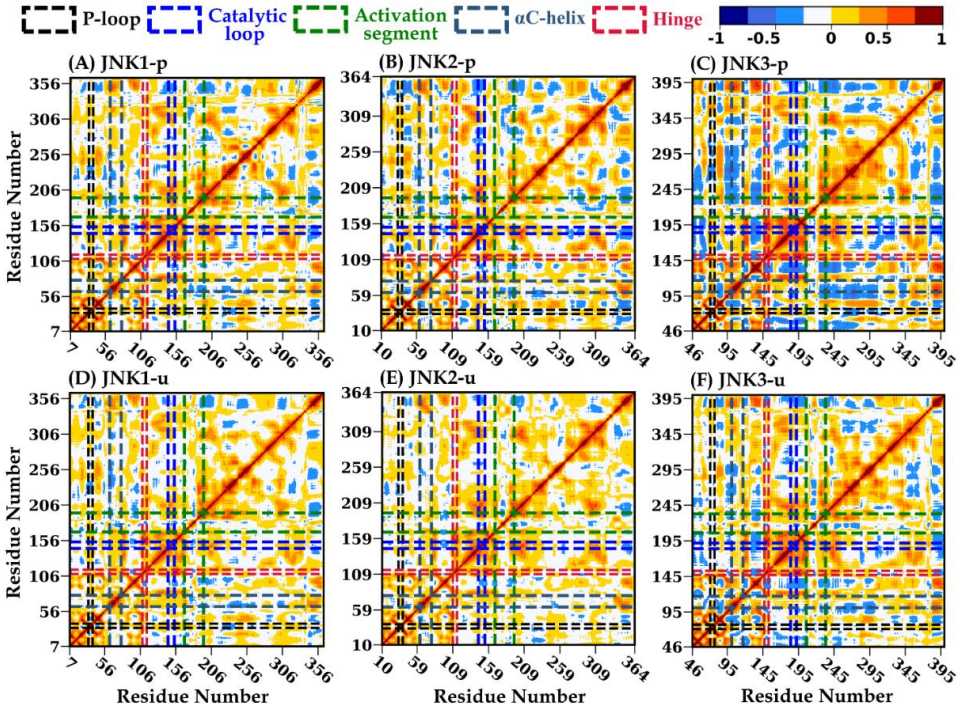


Figure 6.11: Dynamic cross-correlation matrices (DCCM) displaying the interactions between various regions of JNK1/2/3 based on the Cα-atom motion relative to their mean positions. The Pearson's correlation coefficients are represented using a color-coded scale.

6.3.8. Conformational Dynamics of the P-loop

The phosphate binding loop (P-loop) is the important structural element in kinases, playing a crucial role in ATP and competitive inhibitor binding along with regulating

water accessibility within the binding pocket, thereby influencing enzymatic function and activity [425]. Therefore, gaining insight into the conformational dynamics of the P-loop is essential for the rational design of selective inhibitors targeting JNK. In a recent study, Collie et al. proposed that competitive inhibitors interact with c-Met kinases in two distinct conformations, exhibiting both folded and extended states P-loop conformation [426]. The P-loop in Abl kinase moves closer to the αC-helix in its inactive excited state, triggering a rotational shift that disrupts the β3-Lys and αC-Glu salt bridge interaction and forces the αC-helix into an "out" conformation [427]. In this study, we investigate the conformational dynamics of the P-loop in three JNK isoforms, both in their phosphorylated and unphosphorylated states. Our initial backbone RMSD analysis of the P-loop revealed less conformational variation, except for JNK2-p system, which exhibited a bimodal distribution, and JNK3-p system, which had a broader RMSD distribution, while the others show similar type of distributions (see **Figure 6.3C**). To gain deeper insights, we examined P-loop dynamics across JNK isoforms in both phosphorylated and unphosphorylated states, utilizing dihedral angle-based principal component analysis (dPC1 and dPC2) and shown in **Figure 6.12**. As shown in **Figure 6.12**, all systems exhibit similar distributions with one or more minima separated by low to moderate energy barriers. In the JNK1 unphosphorylated system, the P-loop retains a secondary β-sheet structure, which transitions into a loop upon phosphorylation (see **Figure 6.12A-B**). For JNK2, both phosphorylated and unphosphorylated systems shows similar type of extended conformations of P-loop, with minimal secondary β-sheet formation observed in the JNK2-p system (see **Figure 6.12C-D**). For JNK3, the unphosphorylated (JNK3-u) system display a single-global minima, whereas in the case of phosphorylated (JNK3-p) system, it shows two minima separated by a moderate energy barrier. Additionally, phosphorylation in JNK3 also induces secondary β-sheet formation in the P-loop (see **Figure 6.12E-F**).

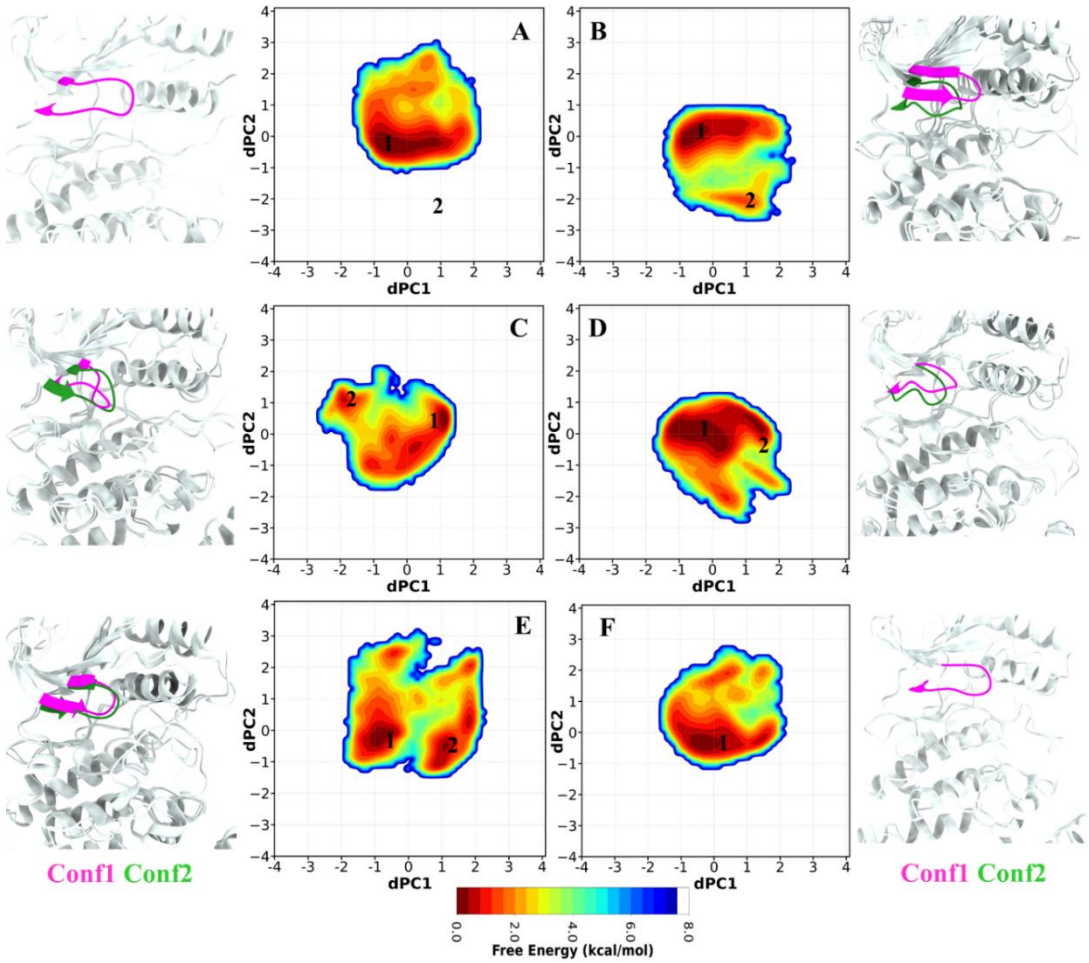


Figure 6.12: Free energy landscape (FEL) analysis along the first two dominant principal components (dPC1 and dPC2) for different JNK isoforms, highlighting major conformational states. Structural snapshots on the left and right illustrate the transition between **Conf1** (magenta) and **Conf2** (green), with energy minima corresponding to dominant conformations observed in the FEL plots. **(A)** JNK1-p, **(B)** JNK1-u, **(C)** JNK2-p, **(D)** JNK2-u, **(E)** JNK3-p, **(F)** JNK3-u.

6.3.9. Binding Affinity and Energetics of JNK Isoform Towards SP600125

To investigate how inhibitor SP600125 interacts with JNK isoforms in terms of binding mechanism, recognition, and specificity, we calculated the total binding energy and its individual free energy components (ΔG_{bind}). The total binding energy consists of multiple components, including van der Waals forces (ΔE_{vdW}), electrostatic

interactions (ΔE_{ele}), polar solvation free energy (ΔG_{pol}), non-polar solvation free energy (ΔG_{np}), and entropy, as detailed in **Table 6.3**. Tables 2 display that (ΔE_{vdW}) and (ΔE_{ele}) contribute to JNK/SP126000 binding, whereas (ΔG_{pol}) and ($T\Delta S$) unfavor the complex formation. Among all interactions, ΔE_{vdW} contributes the most to the binding in each system, with values ranging from -32.01 to -33.41 kcal/mol. Among the isoforms, JNK2-p has exhibited the highest favorable electrostatic interaction (-25.08 kcal/mol), whereas JNK3-p displayed the weakest (-18.20 kcal/mol). The contribution of polar solvation energy (ΔG_{pol}) is unfavorable for binding in each system, with values ranging from 25.96 to 30.66 kcal/mol. A minor stabilizing energy is observed from the non-polar solvation term (ΔG_{np}) (-2.70 to -2.80 kcal/mol), whereas variations in the entropic penalty (-TΔS) influenced the overall binding affinity (ΔG_{bind}). JNK3-u showed the highest binding affinity with an energy of -15.55 kcal/mol, whereas JNK3-p shows the lowest affinity with an energy of -10.24 kcal/mol among all studied systems. JNK3 binding affinity is reduced after phosphorylation, whereas JNK1 and JNK2 show minimal impact. Since all systems shows significant binding affinity of SP600125 toward JNK1/2/3, designing selective inhibitors for individual isoforms may be challenging in drug discovery against JNK.

Table 6.3: The components of the total binding free energy for the JNK isoforms in phosphorylated and unphosphorylated form. All values are in kcal/mol. Enthalpy value represent the average of three simulation runs, whereas entropy is calculated from one run due to high computational cost. the standard errors of the mean are included in parentheses.

JNK/ SP600125 complex	ΔE_{vdw}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	$-T\Delta S$	ΔG_{bind}
JNK1-p/ SP600125	-33.41 (0.03)	-21.45 (0.04)	27.60 (0.03)	-2.78 (0.00)	18.18 (1.12)	**-11.86** (1.12)
JNK1-u/ SP600125	-32.01 (0.03)	-21.06 (0.04)	28.31 (0.03)	-2.80 (0.00)	16.13 (1.28)	**-11.43** (1.28)
JNK2-p/ SP600125	-33.29 (0.03)	-25.08 (0.04)	30.66 (0.03)	-2.70 (0.00)	17.90 (1.11)	**-12.51** (1.11)

JNK2-u/ SP600125	-32.61 (0.03)	-22.60 (0.05)	29.58 (0.04)	-2.75 (0.00)	17.67 (0.68)	**-10.71** (0.68)
JNK3-p/ SP600125	-32.30 (0.03)	-18.20 (0.05)	25.96 (0.04)	-2.79 (0.00)	17.09 (0.82)	**-10.24** (0.82)
JNK3-u/ SP600125	-32.74 (0.03)	-22.55 (0.06)	28.68 (0.03)	-2.75 (0.00)	13.81 (1.19)	**-15.55** (1.19)

To explore the binding mechanism of SP600125 across JNK1/2/3, we computed the individual residue contributions to the overall binding energy using MM/PBSA scheme. **Figure 6.13** highlights residues that significantly contribute to the overall binding energy, while **Table D5** details energy components for residues with contributions less than -1.0 kcal/mol. Significant electrostatic interactions (~ -6 kcal/mol) along with highest affinity was observed for Glu109 in JNK1/JNK2 and Glu147 in JNK3 from hinge region. Glu109 and Glu147 refer to the same residue, with the difference in numbering as per PDB in case of JNK3. Favorable van der Waals interactions (~ -1.5 to -2 kcal/mol) are observed from hydrophobic residues such as Val158, Leu110, and Ile32 in JNK1/2, as well as Val196, Leu148, and Ile70 in JNK3, which had a potent contribution in overall affinity. Overall, analysis of **Table D5** and **Figure 6.13** suggests that the inhibitor binds with same top residues across all isoforms in case of both phosphorylated and unphosphorylated systems, highlighting difficulties in achieving isoform selectivity in inhibitor design.

Table 6.4: Hydrogen bond interactions formed by JNK1/2/3 with SP600125, and the corresponding average distance, and percentage occupancy determined using simulation trajectories. H-bond with more then 10% occupancy are listed here.

Binding couples		**Molecular dynamics**	
Acceptor	**Donor**	**Distance(Å)**	**Occupancy (%)**
JNK1-p			
GLU_109@O	Lig@H24	2.81	90.97
Lig@ N23	MET_111@H	2.93	13.32

JNK1-u			
GLU_109@O	Lig@H24	2.81	91.30
Lig@ N23	MET_111@H	2.93	13.32
JNK2-p			
GLU_109@O	Lig@H24	2.80	92.42
Lig@O22	GLN_37@HE22	2.85	22.05
Lig@N23	MET_111@H	2.93	15.97
JNK2-u			
GLU_109@O	Lig@H24	2.81	92.07
Lig@N23	MET_111@H	2.93	14.87
JNK3-p			
GLU_147@O	Lig@H24	2.81	90.52
Lig@N23	MET_149@H	2.93	13.78
JNK3-u			
GLU_147@O	Lig@H24	2.81	91.00
Lig@N23	MET_149@H	2.93	15.20

Furthermore, we have calculated the H-bond occupancy for all systems, as presented in **Table 6.4**. A highly stable with more than 90% occupancy of H-bond exists between Glu109 (Glu147 in case of JNK3) and SP600125, while other H-bond with Met111 or Met149 exhibit significantly lower occupancy 13-22%. We have presented the 2D interaction patterns (see **Figure D5**) of all studied complex systems and observed a high similarity in the interacting residues across all systems.

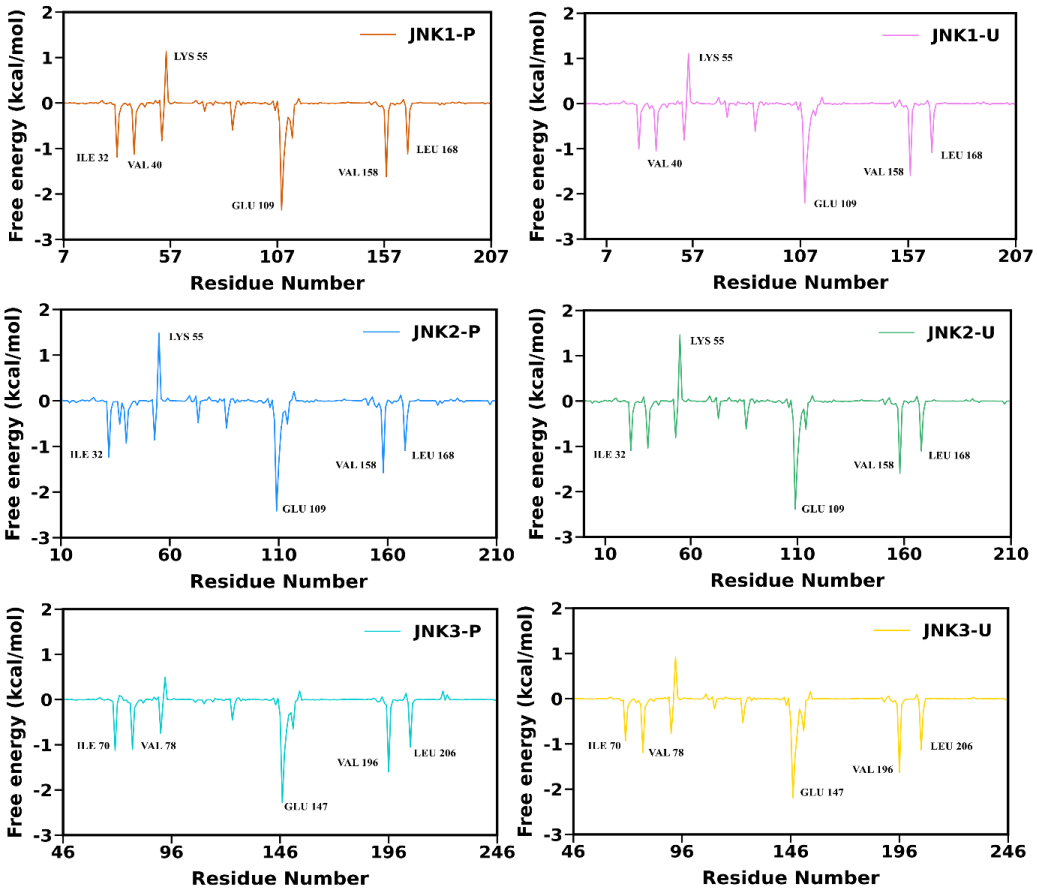


Figure 6.13: The residue level free energy decomposition of JNK1/2/3 in both phosphorylated and unphosphorylated state bind with the inhibitor SP600125. The residues contributing less than -1.0 kcal/mol are highlighted here.

6.4. Discussions

The c-Jun N-terminal kinases (JNK) are important members of the mitogen activated protein kinase (MAPK) family, playing a crucial role in eukaryotic cellular responses to both biotic and abiotic stress. Apart from stress, JNKs also regulate important physiological processes such as neuronal signaling, immune system modulation, embryonic development, cytoskeletal protein dynamics, and pathways involved in cell survival or apoptosis [428]. In this study, we conducted MD simulations of JNK1, JNK2, and JNK3 isoforms bound with the inhibitor SP600125 in both phosphorylated and unphosphorylated states to investigate the structural and conformational dynamics along with binding energetics. The structural dynamics of JNK isoforms are key

determinants of their functional diversity and regulatory roles. Despite their high sequence identity, JNK1, JNK2, and JNK3 adopt different biological functions and show tissue specific expression, mainly regulated by their structural conformations and molecular interactions[429]. MD simulations is an important tool to investigate variations in flexibility, stability, and active site dynamics among these JNK1/2/3 isoforms. Investigating the phosphorylated and unphosphorylated states of JNK1/2/3 allows us to understand the conformational changes responsible for their distinct regulatory mechanisms.

Previously, a study using MD simulations on PKM1 and PKM2 isoforms identified unique conformational states and allosteric sites which is crucial for inhibitor design[430]. Similarly, computational analysis of PAK1 and PAK4 demonstrated how minor active site variations impact inhibitor selectivity [431]. Jonniya et al. have deeply characterized the conformational dynamics of WNK kinase isoforms in complex with an inhibitor, providing insights into their structural transitions and binding mechanisms [432]. In a similar context, the characterization of JNK1, JNK2, and JNK3 isoforms through MD simulations is important to uncover their distinct structural dynamics, phosphorylation induced conformational shift and ligand binding energetics. This investigation helps to elucidate how phosphorylation and inhibitor selection influence the conformational states of different JNK isoforms, revealing crucial dynamics mechanisms.

Following MD simulations of studied systems, analyses including RMSD of different regions, RMSF, SASA, R_g, PCA, dPCA, DCCM, PSN, and binding free energy calculations offer comprehensive insights into the dynamic behavior of JNK1, JNK2, and JNK3 isoforms in their phosphorylated and unphosphorylated states. The DALI webserver based structural comparison suggests that JNK isoforms are well conserved, with JNK1 and JNK3 showing the highest similarity. The whole protein backbone RMSD analysis indicates that all isoforms stabilize after ~400 ns, while phosphorylation triggers structural rearrangements, reflected in varying RMSD ranges with JNK3 shows broader RMSD suggesting higher conformational variability. The multimodal RMSD distribution observed in JNK2 and JNK3 indicates that the

activation segment explores multiple conformation as compared to JNK1. The RMSF analysis indicated that the loop regions exhibit significant flexibility and showed almost similar fluctuation pattern for all isoforms. SASA and R_g analysis revealed phosphorylation induced structural changes, with increased SASA and slight compactness reduction suggesting a more extended activation loop. Phosphorylated JNK isoforms shows little greater conformational variability than their unphosphorylated forms, as shown by PCA analysis. This aligns with findings in other kinases, such as p38, where phosphorylation leads to structural shifts in regulatory elements like the activation loop and catalytic regions [433]. Multiple low-energy states separated by moderate barriers indicate conformational plasticity, consistent with studies on other kinases that associate such structural variability with functional adaptation [434]. Our analysis on A-loop highlights the crucial role of activation segment phosphorylation in regulating JNK isoform dynamics. dPCA analysis shows, JNK1 and JNK3 shift toward an open conformation upon phosphorylation, while JNK2 A-loop remains unaffected. Similar mechanisms in Aurora B/C [435], WNK1 [436], and FGFR1 [267] show phosphorylation rearrange the activation loop, enhancing catalytic activity. As observed from the RMSD of individual residues in the activation segment, phosphorylation introduced from the unphosphorylated crystal structure increases RMSD due to structural rearrangement, A-loop stabilization, substrate recognition, and functional divergence among JNK isoforms. Phosphorylation induces isoform specific effects in JNK, weakening the β3-Lys/αC-Glu salt bridge in JNK1, strengthening it in JNK3, while JNK2 remains largely unaffected. The β3-Lys and DFG-motif Asp does not form stable salt bridge in any JNK isoform. The PSN analysis highlights that phosphorylation induces alterations in residue connectivity, hub distribution, and community dynamics, leading to denser and more interconnected networks in JNK isoforms.

The binding affinity analysis of SP600125 across JNK1/2/3 isoforms reveals important insights into its interaction patterns and selectivity challenges for ATP-competitive inhibitors. A decline in JNK3 binding affinity after phosphorylation implies potential structural or electrostatic modifications and entropy penalty that impact inhibitor binding, whereas JNK1 and JNK2 shows minimal effect. Studies on other kinases,

including WNK isoforms [432] show that phosphorylation does not affect binding affinity in some isoforms, whereas some variations are observed in others as like JNK. The observed good binding affinity across the three isoforms reveals a significant difficulty in achieving isoform selectivity, which is essential for drug development against JNK isoform. Per-residue decomposition analysis indicates that SP600125 interacts with the same key residues across all isoforms, highlighting challenges in developing highly specific isoform-selective inhibitors. To address therapeutic challenges in JNK-related diseases, future studies should focus on designing inhibitors that overcome structural and conformational differences between isoforms or better finding novel distinct allosteric pocket.

6.5. Conclusions

In the present work, MD simulations of JNK1, JNK2, and JNK3 isoforms in phosphorylated and unphosphorylated states highlights structural and conformational changes influencing their functional diversity. Structural comparisons reveal a high degree of similarity among JNK isoforms, with JNK1 and JNK3 being highly similar. The RMSD analysis of overall protein indicates that phosphorylation inducing structural rearrangements and shows shift in RMSD values. The RMSF analysis suggests that loop regions show significant fluctuations, following similar fluctuation pattern in case for systems. Phosphorylation induced dynamics as per SASA and R_g analyses suggest increased solvent exposure and decreased compactness, reflecting extended open-type A-loop and some other structural rearrangement. PCA results indicate that phosphorylation induces structural plasticity in JNK isoforms, with increasing conformational sampling space and promoting dynamic loop motions which may leads to more active like conformations. dPCA results show that phosphorylation promotes a more open type A-loop in JNK1 and JNK3, while JNK2 retains its open like A-loop conformation. Further, increased RMSD of individual residues of activation segment upon phosphorylation indicates structural reorganization of the activation segment, leading to activation loop stability and may the substrate recognition. Phosphorylation affects the β3-Lys/αC-Glu hallmark salt bridge which differently across JNK isoforms, decrease density of formation in JNK1, increase in

JNK3, and unaffected in JNK2. PSN analysis further highlights the effect of phosphorylation in residue connectivity, leading to increase in networks network and shifts in communication pathways. SP600125 exhibits strong binding affinity across all isoforms, with similar interacting residues, hence signifying challenges for designing isoform specific inhibitors. Overall, our study offers a detailed comparative analysis of JNK isoform structural and conformational dynamics with a known competitive inhibitor SP600125, enhancing our understanding of JNK kinase isoforms and the effects of phosphorylation on their structure and dynamics.

Chapter 7

Chapter 7

Dynamic Variability in JNK1: The Role of ATP and MKK7-Mediated Interactions in Structural Regulation and Energy Landscape

AIM: To comprehensively investigate how the individual and combined interactions of ATP and MKK7 influence the conformational plasticity of JNK1.

This chapter is reproduced from our original article:

Koirala, S., Samanta, S., Ursal, K.D, & Kar, P. "Dynamic Variability in JNK1: The Role of ATP and MKK7-Mediated Interactions in Structural Regulation and Energy Landscape". *Archives of Biochemistry and Biophysics* (Submitted)

7.1 Introduction

The c-Jun N-terminal kinases (JNKs) are key players within the MAPK pathway, influencing essential processes such as regulating cell migration, differentiation, proliferation, and apoptosis [437,438]. The JNK family comprises three distinct yet highly homologous isoforms, JNK1, JNK2, and JNK3, which possess over 10 splice variants [405,439]. With 92% sequence similarity, JNK1 and JNK3 differ at only one position within the 23 residues of their ATP-binding sites. With 83% sequence identity, JNK1 and JNK2 are nearly identical except for one residue in the 23-residue ATP-binding pocket [48,440]. The expression of JNK1 and JNK2 is ubiquitous, in contrast to JNK3, which is largely restricted to the brain, heart, and testicular tissue. JNKs are mainly present in an inactive, non-phosphorylated form and become activated in response to various extracellular signals such as UV radiation, oxidative stress, and inflammatory cytokines [441]. JNKs become activated in response to cell stimulation, primarily through the action of upstream kinases MKK4 and MKK7, which phosphorylate the TPY motif in the activation segment. Once activated, JNKs

phosphorylate numerous downstream target proteins [389]. Given their role in determining cell fate through numerous signaling pathways, JNKs are under strict regulatory control. Irregularities in JNK signaling have been associated with a range of diseases, including cancer, neurodegeneration, and autoimmune conditions [442,443].

JNKs possess a relatively simple structure, characterized by a single bilobed kinase domain that contains a central ATP-binding site. Adjacent to the ATP-binding pocket, JNKs interact with their phospho-acceptor substrates in an extended conformation. Although JNKs prefer substrates with specific phospho-acceptor sequences, they also use several docking sites located away from the catalytic center to distinguish among cellular targets [94,444]. These interactions serve to refine substrate selectivity, thereby these docking interactions improve signaling specificity. JNK substrates and various protein partners often contain interaction sequences known as docking domains or D-domains. Typically, D-domains are composed of a small number of basic residues and a hydrophobic motif, separated by a spacer of one to six residues [445,446]. Proteins containing D-domains interact with JNK at D-recruitment sites (DRSs), primarily located on the C-terminal lobe, opposite the ATP-binding surface. Multifunctional in nature, these DRSs mediate interactions with scaffolds, upstream kinases like MKKs, and a range of substrates [82,447].

MKK7 functions as an upstream activator of the JNK pathway, which plays a central role in stress and inflammatory signaling. It is unique among human MKKs in having three putative docking sites D1, D2, and D3, within its 100 amino acid regulatory domain. It has been previously reported that each of the three predicted docking sites can bind to JNK, with the interactions highlighting a degree of synergistic cooperation [94]. Moreover, Kragelj et al. reported the crystal structure of JNK1 in association with the second docking site of MKK7. They provide a detailed understanding of the structure, stoichiometry, and kinetics of the MKK7-JNK signaling complex [107]. Despite structural insights, the role of second docking site of MKK7 in modulating the conformational dynamics of JNK1 has not been studied in detailed.

The present study explores the structural dynamics of JNK1 in its apo form, ATP and Mg^{2+} bound state, MKK7 docking motif bound state, and in the state where both ATP along with Mg^{2+} and the MKK7 docking motif are bound simultaneously, using extensive microsecond-long Gaussian accelerated molecular dynamics (GaMD) simulations. To gain deeper insights into the intrinsic dynamics of JNK1, this study employs free energy surface analysis along with several trajectory-based post-simulation analytical strategies. Furthermore, the free energy of various interactions is calculated using the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) method, combined with conformational entropy calculations based on normal-mode analysis.

7.2. Materials and Methods

7.2.1. Preparation of Apo and ATP/MKK7 Docking Domain Bound Forms of JNK1

All the systems used for these studies were developed from the crystal structure of JNK1 complexed with the MKK7 docking motif (PDB ID: 4UX9) [107]. To create the apo state of JNK1, the MKK7 interaction region was deleted, and other irrelevant structural parts were stripped away to leave the protein in its free form. To model the second system, ATP and Mg^{2+} were docked into the JNK1 active region, resulting in a ligand-bound protein complex. JNK1 in complex with MKK7, utilizing the endogenous MKK7 docking motif from PDB ID 4UX9, defined the third system. Finally, an integrated system was constructed, combining JNK1 with both ATP-Mg^{2+} and the MKK7 docking motif, with ATP and Mg^{2+} positioned via molecular docking.

7.2.2. Conventional Molecular Dynamics Simulation Protocol

All systems were initially subjected to energy minimization using both the steepest descent and conjugate gradient algorithms. Following energy minimization, all systems were gradually heated from 0 K to 300 K over a 50 ps period within the canonical (NVT) ensemble. During this heating phase, positional restraints were applied to the solute atoms with a force constant of 2.0 $kcal \cdot mol^{-1} \cdot Å^{-2}$ to maintain structural integrity and temperature control was achieved using a Langevin thermostat

[295]. To preserve the geometry of covalent bonds involving hydrogen atoms, the SHAKE algorithm [169] was employed throughout the simulations. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) [263] method with a nonbonded cutoff of 10 Å. After heating, the systems were equilibrated at a constant pressure of 1 atm for 20 ns using the Berendsen barostat [180]. All simulations were performed with a 2.0 fs time step, and coordinates were saved every 10 ps for subsequent analyses. Finally, a short 2.0 ns production run in the NVT ensemble was conducted without any restraints to collect initial potential energy statistics. All molecular dynamics simulations were carried out using the pmemd.cuda module from the AMBER18 software package.

7.2.3. Gaussian Accelerated Molecular Dynamics (GaMD) Simulations

Gaussian accelerated molecular dynamics (GaMD) [199] was employed to enhance the sampling of conformational space by smoothing the system's potential energy surface through the application of a harmonic boost potential. This approach reduces energy barriers and enables efficient exploration of biologically relevant states without the need for predefined reaction coordinates, distinguishing GaMD from many other enhanced sampling techniques [345]. The boost potential follows to a Gaussian distribution, which can be reliably retrieved by reweighting methods [201].

For a system with N atoms at positions r = {r_1, r_2, ..., r_n}, when the system potential V(r) falls below a threshold energy E, a harmonic boost potential ΔV(r) is applied and can be expressed as:

A harmonic boost potential (ΔV) is added when the system potential (V) falls below the threshold energy (E), as described in equations (1) and (2).

$$\Delta V(r) = \frac{1}{2}K(E - V(r))^2, \quad \text{if } V(r) < E \tag{1}$$

Here, K represents the harmonic force constant, and the modified potential (V x r) is defined as:

$$V \text{ x } r = V(r) + \frac{1}{2}K(E - V(r))^2, \quad \text{if } V(r) < E \tag{2}$$

If the system's potential energy is greater than or equal to the threshold energy (i.e., V(r)≥E), no boost potential is applied, and the modified potential remains the same as the original (V x (r) = V(r)). To effectively smooth the potential energy landscape, GaMD applies three key principles to the boost potential. First, consider two arbitrary potential values, V1(r) and V2(r), found on the original energy surface; if $V_1(r) < V_2(r)$, then should be a monotonic function and does not change the modified bias potential energy, i.e., $V \times_1(r) < V \times_2(r)$. Second, if $V_1(r) < V_2(r)$, the potential difference observed on the smoothened energy surface should be smaller than that of the original, i.e., $V_2 \times (r) - V \times_1(r) < V_2(r) - V_1(r)$. Combining these two criteria, the threshold energy needs to be set in the following range:

$$V_{max} \leq E \leq V_{min} + \frac{1}{k} \tag{3}$$

Where V_{max} and V_{min} represent the highest and lowest potential energy values of the system, respectively. For Eq. (3) to be valid, the parameter kk must meet the following condition:

$$k \leq \frac{1}{V_{max} - V_{min}}; \qquad \text{let us define} \tag{4}$$

$$k = k_0 \frac{1}{V_{max} - V_{min}} \text{ and } 0 \leq k_0 \leq 1 \tag{5}$$

To ensure that the boost potential can be accurately reweighted, the standard deviation of ΔV must be kept sufficiently low, resulting in a narrow distribution. This requirement is expressed mathematically as: $\sigma\Delta V = k(E - V_{avg})\sigma V \leq \sigma 0$ where V_{avg} and σV are the mean and standard deviation of the system's potential energy, respectively, and σ0 is a user-defined upper limit (such as $10k_BT$) for reliable reweighting. When the threshold energy E is chosen at its lower bound ($E = V_{max}$), the corresponding value for k_0 can be determined accordingly.

$$k_0 = \min\left(1.0, \ \frac{\sigma 0}{\sigma V} X \frac{V_{max} - V_{min}}{V_{max} - V_{avg}}\right) \tag{6}$$

On the other side, if E is set to the upper bound, $V_{min} + 1/k$; k_0 can be calculated as:

$$k_0 = \left(1.0 - \frac{\sigma 0}{\sigma V}\right) . \left(\frac{V_{max} - V_{min}}{V_{max} - V_{avg}}\right) \quad (7)$$

In this study, Gaussian accelerated molecular dynamics (GaMD) simulations were conducted using three types of boost potentials: total potential (ΔV_p), dihedral potential (ΔVD), and a combination of both (dual boost: ΔV_D and ΔV_p). Among these, the dual-boost approach where both dihedral and total potential boosts are applied was selected for its superior ability to enhance sampling efficiency. The threshold energy for the boost potential in all simulations was set to the system's maximum potential energy (E = Vmax). Initial values for the maximum (V_{max}), minimum (V_{min}), average (V_{avg}), and standard deviation (σv) of the potential energy were determined from a 2.0 ns conventional MD (cMD) run performed without any boost. Subsequently, a 50.0 ns GaMD equilibration was carried out, during which the boost parameters were updated every 200 ps to allow the system to reach stable equilibrium values. Following equilibration, each system underwent three independent production simulations, each lasting 1 μs, in the NVT ensemble. Throughout these production runs, atomic coordinates were recorded every 10 ps, resulting in a total of 300,000 frames for downstream analyses.

7.2.4. Analysis of Simulation Trajectories

All molecular dynamics simulation trajectories were processed and analyzed using the Cpptraj module [264] from AmberTools19, which provides a comprehensive suite of tools for trajectory analysis. This software package enables efficient processing of large-scale molecular dynamics data with high computational performance and accuracy.

The conformational stability of each molecular system was evaluated through multiple complementary metrics. Root-mean-square deviation (RMSD) calculations were performed on protein backbone atoms to quantify global structural changes over time relative to the energy-minimized reference structure. In parallel, the radius of gyration (RoG) was calculated to assess the compactness and overall shape of the protein structures. While solvent-accessible surface area (SASA) calculations were performed

to quantify the degree of solvent exposure for each system. Intermolecular contacts were evaluated through hydrogen bond occupancy analysis was conducted to monitor the formation and stability of key intermolecular interactions throughout the simulation trajectories. Root-mean-square fluctuation (RMSF) calculations were performed to characterize the flexibility of individual residues throughout the simulation period. Finally, collective motions were characterized by using Dynamic cross-correlation matrices (DCCM) were constructed to investigate correlated and anti-correlated motions between residues.

Principal component analysis (PCA) [304] was employed to identify the dominant collective motions within the molecular systems. By reducing the dimensionality of the trajectory data, PCA retains the most significant conformational changes, enabling the identification of essential dynamics. This approach is particularly advantageous for characterizing large-scale conformational motions and pinpointing the principal modes of protein flexibility.

To further explore the conformational landscape, free energy surfaces (FES) [349,350] were constructed using multiple reaction coordinates. The integration of these reaction coordinates such as root mean square deviation (RMSD), radius of gyration, inter-residue distances, and principal components ensures a comprehensive coverage of the system's conformational space.

The free energy profiles were calculated as follows:

$$\Delta G = -k_B T \ln(\rho *) - k - \left(\sum_{k=1}^{\infty} \frac{\beta^k}{k!} C_k - const.\right) \quad (8)$$

Here, kB is the Boltzmann constant, T is the system's temperature, and ρ * represents the modified probability for a chosen reaction coordinate (RC). The parameter is set so that the global minimum corresponds to 0 kcal/mol. β is the Boltzmann factor, and Ck is the expansion coefficient. Next, the potential boost, combined with different RC derived from the MD trajectories, was used to reconstruct the FES using reweighting techniques. We used the "PyReweighting" python scripts for reweighting, which was developed by the McCammon group [201]. All FES analyses in this study were

reweighted. However, we did not apply the reweighting procedure to the kernel density violin plots and other probability density plots.

Protein structure network (PSN) analysis was performed using the webPSN v2.0 server [221] to investigate residue-residue communication patterns. In this network representation, amino acid residues serve as nodes, and edges are defined based on the strength of non-covalent interactions between residues.

$$I_{ij} = \frac{n_{ij}}{\sqrt{N_i N_j}} \times 100 \quad (9)$$

In the equation, n_{ij} denotes the number of atom-atom pairs between the side chains of residues "i" and "j" within a cutoff distance of 4.5 Å. Normalization factors N_i and N_j are applied to standardize interaction assessments by accounting for variations in amino acid sizes and their propensity to form maximum contacts in the protein structure. Residues "i" and "j" are considered interacting if I_{ij} exceeds the threshold Imin, establishing an edge between the corresponding nodes. At a given Imin, residues forming at least four edges are identified as "hubs" in the network.

Binding free energy calculations were performed using the MM/PBSA [229] approach as implemented in the MMPBSA.py module of AMBER18. This method provides a computationally efficient approach for estimating protein-ligand binding affinities by combining molecular mechanics energies with continuum solvation models. The details of mathematical formulations are discussed in our earlier studies [308,353]. For binding free energy calculations, 15,000 snapshots were extracted from the equilibrated trajectory (3 independent sets of 5,000 frames each) to ensure adequate statistical sampling. Additionally, MM/PBSA calculations were performed to obtain per-residue decomposition of binding contributions, providing detailed insights into the molecular basis of protein-ligand interactions.

7.3. Results and Discussions

This study systematically examines how binding ATP-Mg^{2+}, the MKK7 docking motif, and their simultaneous interaction modulate the conformational dynamics of the JNK1 kinase. To thoroughly investigate the conformational dynamics of JNK1, we designed

and simulated four specific systems: the ligand-free apo JNK1 (JNK1), the ATP-Mg^{2+} bound state (JNK1-A), the complex with the MKK7 docking motif (JNK1-M), and a system in which both ATP-Mg^{2+} and the MKK7 motif are engaged with JNK1 (JNK1-A-M). These systems were systematically evaluated to characterize the respective and interactive roles of ATP-Mg^{2+} and MKK7 docking motif binding in shaping the structural dynamics and stability of JNK1. Each of the four systems incorporated a doubly phosphorylated activation loop, with both Thr-183 and Tyr-185 residues of JNK1 modified to mimic the active state of JNK1. A visual representation of the system under investigation is shown in **Figure 7.2(A)**.

7.3.1. Analysis of Structural Stability and Flexibility of JNK1

Following the molecular dynamics simulations, backbone root-mean-square deviation (RMSD) was calculated to assess overall structural stability and simulation convergence. As shown in **Figure E1**, the RMSD trajectories remained largely stable across all systems, with only minor fluctuations observed (e.g., in JNK1 run3), indicating well-converged simulations. In addition, solvent accessibility and structural compactness were analyzed to assess the presence of any broad conformational changes. As illustrated in **Figure 7.1(A)**, a minor reduction in the total SASA was noted following ligand binding, suggesting a trend toward reduced solvent exposure. Notably, **Figure 7.1(C)** highlights a marked decrease in SASA within the ATP binding pocket, signifying water displacement after ATP binding. As depicted in Figure 7.1(B), overall Rg values were similar for all systems. However, ATP binding led to a noticeable increase in Rg within the binding pocket, as seen in Figure 7.1(D). The time evolution of SASA and Rg of the ATP-binding pocket are given in Figures E2 and E3, respectively, indicating an overall similar trend throughout the systems. The contrasting increase in Rg at the ATP site, despite overall structural compactness, highlights an adaptive conformational adjustment that stabilizes ATP within the binding pocket.

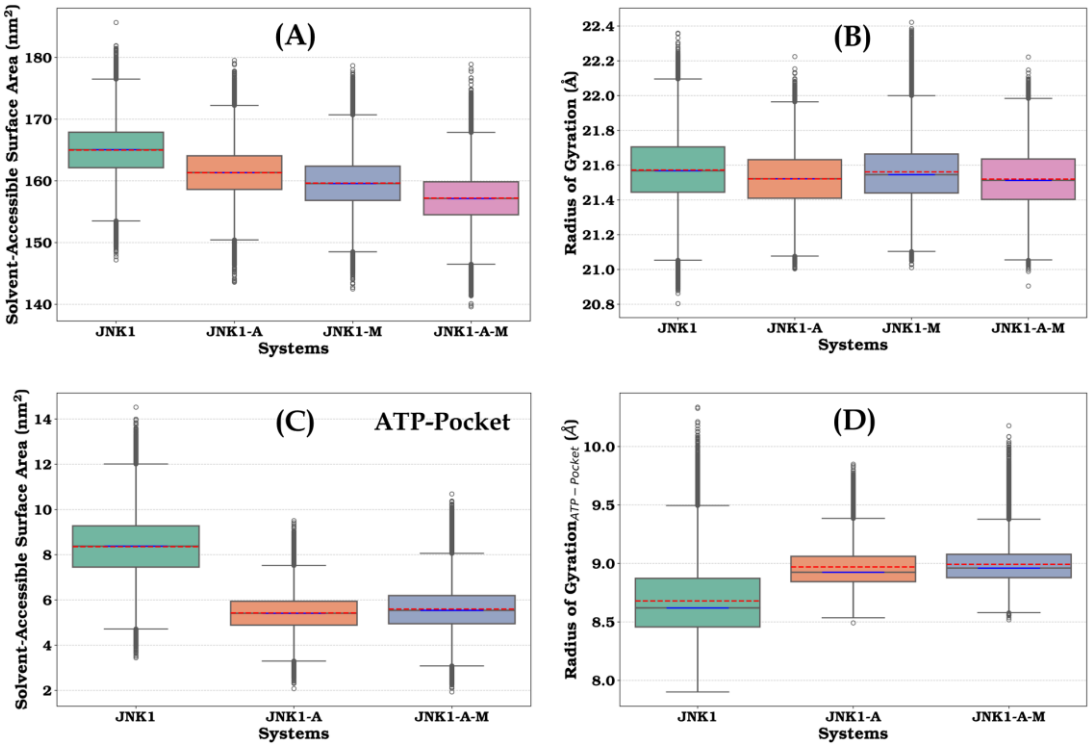


Figure 7.1: Box plot analysis of solvent accessibility and compactness in the studied systems. (A) Solvent-accessible surface area (SASA) of the whole protein, (B) radius of gyration (RoG), (C) SASA of the ATP-binding pocket, and (D) RoG of the ATP-binding pocket.

In order to evaluate the global and local stability, we computed the root-mean-square deviations (RMSD) of the protein backbone atoms along with different important regions relative to their initial structural configuration. **Figures 7.1(B-G)** collectively present a detailed analysis of the conformational dynamics of different important regions of JNK1 under four distinct molecular conditions. The RMSD of different structural elements of JNK1 provides insights into the stability, flexibility, and dynamic behavior of the protein in response to various binding events. Box plots were used to visualize and compare the distributions and variances of RMSD values across all studied systems.

Initially, in order to evaluate the global stability and convergence behavior, we computed the RMSD of the protein backbone as shown in **Figure 7.2 (B)**. The box plot for the Apo system is the longest (1.9-2.3 Å) among the four, reflecting greater

variability and conformational deviation in the absence of binding partners. In contrast, the ATP-Mg^{2+} bound (1.8-2.0 Å), MKK7 motif-bound (1.9-2.2 Å), and dual-bound systems (1.9-2.16 Å) exhibit shorter, more compact boxes. This indicates a reduction in backbone deviation and enhanced structural stability upon ligand or MKK7-motif binding. Notably, the JNK1-A system displays the shortest box, suggesting the most consistent and stable backbone configuration under ATP-Mg^{2+} bound conditions. Additionally, RMSD analysis of the P-loop was performed to investigate its conformational changes in the various systems examined as shown in **Figure 7.2(C)**. The box for the Apo system and MKK7 motif-bound system is relatively long, indicating substantial fluctuations and inherent flexibility in the absence of ATP-Mg^{2+}. In contrast, the ATP-Mg^{2+} bound and dual-bound systems display much smaller boxes, signifying that the P-loop is markedly stabilized and less variable when ATP-Mg^{2+} is present. These findings indicate that the presence of ATP-Mg^{2+} significantly restricts the conformational flexibility of the P-loop, leading to a more stabilized structure. This suggests that ATP binding imposes structural constraints on the P-loop region, thereby minimizing its dynamic fluctuations and contributing to its enhanced stability within the system. **Figure 7.2(D)** and **Figure E4** depict the RMSD for the αC-helix region. All systems show short, compact boxes for the αC-helix RMSD, indicating limited deviations and a generally stable across all studied conditions. The Apo and MKK7 motif-bound systems have slightly large boxes as compared to other systems reflecting some fluctuation. **Figure 7.2(E)** and **Figure E5** show the RMSD of activation segment, a region known for its regulatory flexibility. The boxes for the Apo, ATP-Mg^{2+} bound, and the dual-bound systems shows relatively long boxes as compared to MKK7 motif-bound system which shows compact box. These may be because of substrate binding stabilizes the activation loop (A-loop) in a fixed conformation to ensure proper alignment for phosphorylation, thereby restricting its flexibility. However, when both substrate and ATP are bound, as shown in **Figure 7.2(E)** the A-loop retains some flexibility which may be due to structural adjustments required for efficient catalysis and phosphate transfer.

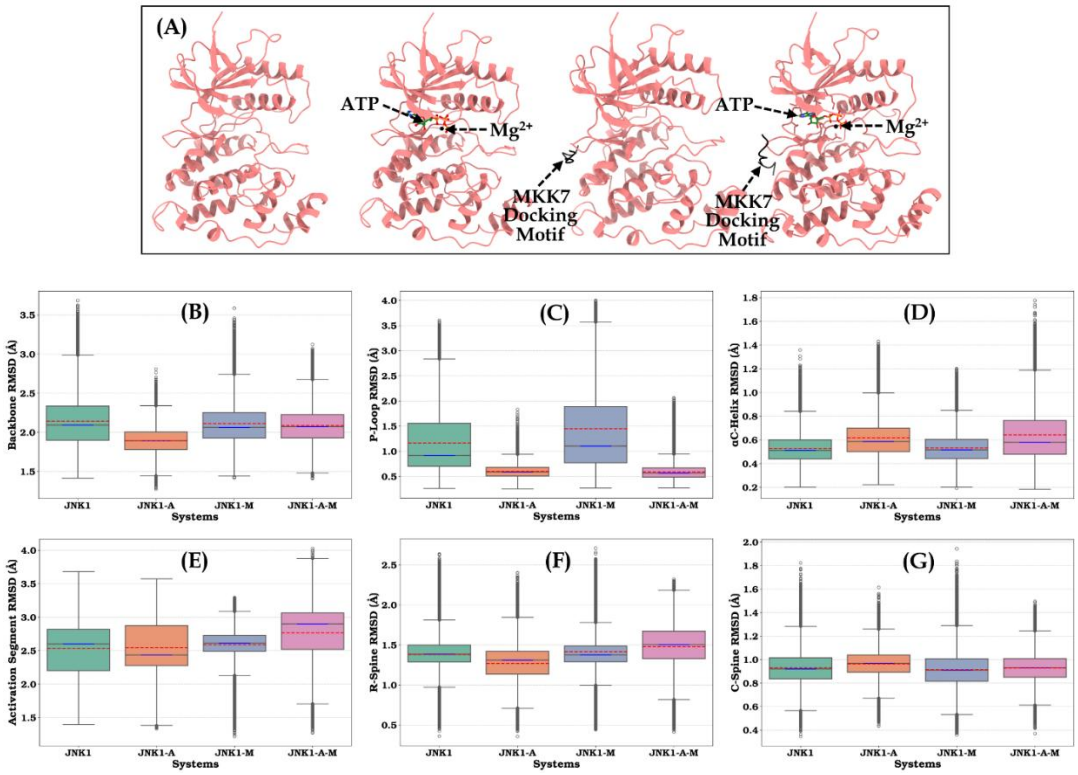


Figure 7.2: (A) Structural representation of the four studied systems - apo JNK1, JNK1 bound to ATP and Mg^{2+}, JNK1 bound to MKK7, and JNK1 bound to both MKK7 and ATP with Mg^{2+}. (B–G) Box plot analysis of the root-mean-square deviation (RMSD) of the whole protein and various important regions.

7.3.2. Hydrophobic Spines Dynamics

In earlier work by Kornev et al. [424,448], two distinct sets of hydrophobic residues were characterized as essential structural elements of kinases, comprising eight residues forming the catalytic spine (C-spine) and four non-sequential residues making up the regulatory spine (R-spine). The proper alignment and stabilization of protein kinase domains are largely governed by the hydrophobic spines, which serve as internal scaffolds to maintain active conformations required for catalytic activity. A misaligned or broken R-spine indicates that the kinase is in an inactive state, as the proper orientation of these hydrophobic residues is essential for catalytic activity. We have calculated the RMSD of both the R-spine and C-spine, as shown in **Figure 7.2 (F-G)**.

The boxplots for the R-spine (**Figure 7.2(F)**) and C-spine (**Figure 7.2(G)**) RMSD values across the four systems display a broadly similar pattern of distribution. Both spines show comparable ranges and medians, indicating that the structural variability of the R-spine and C-spine is consistent regardless of the system. Additionally, we plotted the RMSD of the R-spine over time and extracted the corresponding structures to examine their alignment under different conditions as shown in **Figure 7.3**. In some individual simulation runs, we observe minimal fluctuation in the RMSD values, for each of these regions, representative structures were extracted, and in all cases, the alignment of the R-spine residues (highlighted in black) is consistently maintained. In a free energy landscape (FEL) analysis conducted using the R-spine and C-spine RMSD as reaction coordinates (**Figure E6**), a single prominent minimum was observed in all systems. This demonstrates that, the core alignment of the R-spine is preserved across all studied systems.

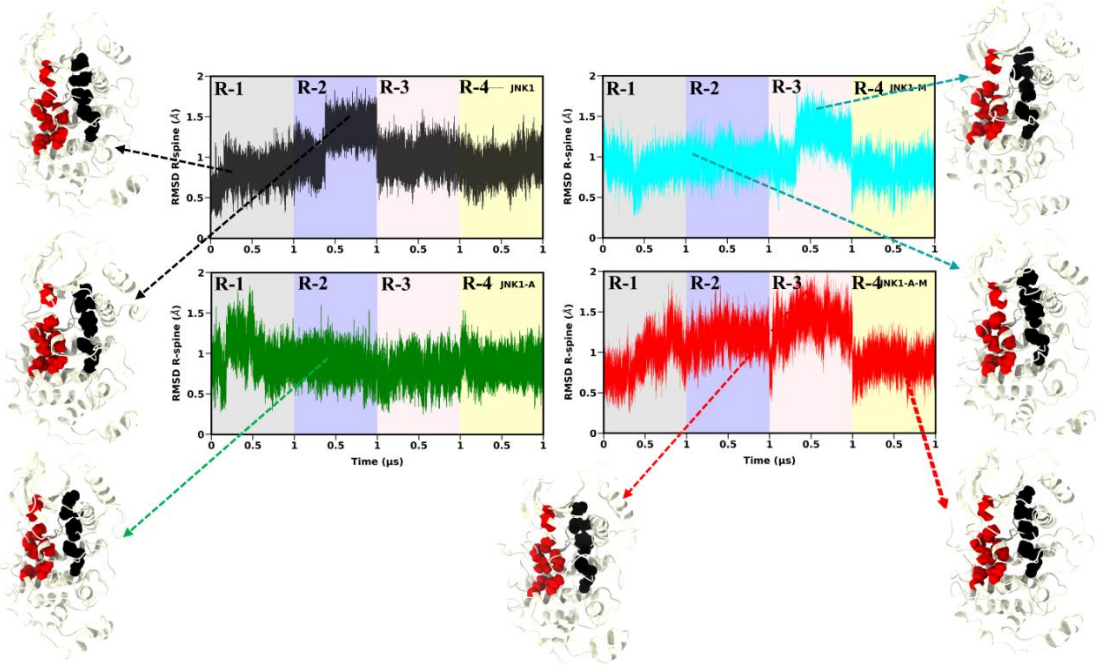


Figure 7.3: The time evolution of RMSD of the R-spine for four systems (JNK1, JNK1-A, JNK1-M, JNK1-A-M), segmented into regions R-1 to R-4 representing individual simulation runs. Representative structures from each region are shown, with R-spine residues highlighted in black and C-spine residues highlighted in red.

7.3.3. Global Motion of the JNK1 Kinase Domain.

Principal Component Analysis (PCA) is a well-established computational method for identifying collective motion in proteins by reducing the complexity of their dynamic trajectories [348]. In this study, PCA was performed using the Cα atoms of amino acid residues to investigate the large-scale motions of the JNK1 kinase domain. To evaluate the conformational space via energy minima, we derived free energy landscapes using the first two principal components (PC1 and PC2, **Figure 7.4A-D)**), with porcupine plots (1 Å cutoff) to display dominant motions as shown in **Figure 7.4(a1-d1 and a2-d2)**. Key structural elements are color-coded: the P-loop (orange), hinge region (magenta), αC-helix (green), catalytic loop (cyan), and activation segment (red). Additionally, key conformations from the energy minima with high occupancy were highlighted in **Figure 7.4(a3-d3)**. The broad distribution of sampled conformations with numerous energy minima across the four states reveals significant structural plasticity. Compared to the other two systems, JNK1 and JNK1-A systems exhibit slightly narrower conformational sampling with multiple minima separated by low-energy barriers, suggesting that ATP binding might enhances system stability. Although the JNK1-M system explores a wide conformational landscape, the existence of a distinct global minimum suggests structural stabilization after protein-protein interactions. The major source of conformational variability appears to be the loop regions, as significant motions were observed in the porcupine plots of the first two principal components.

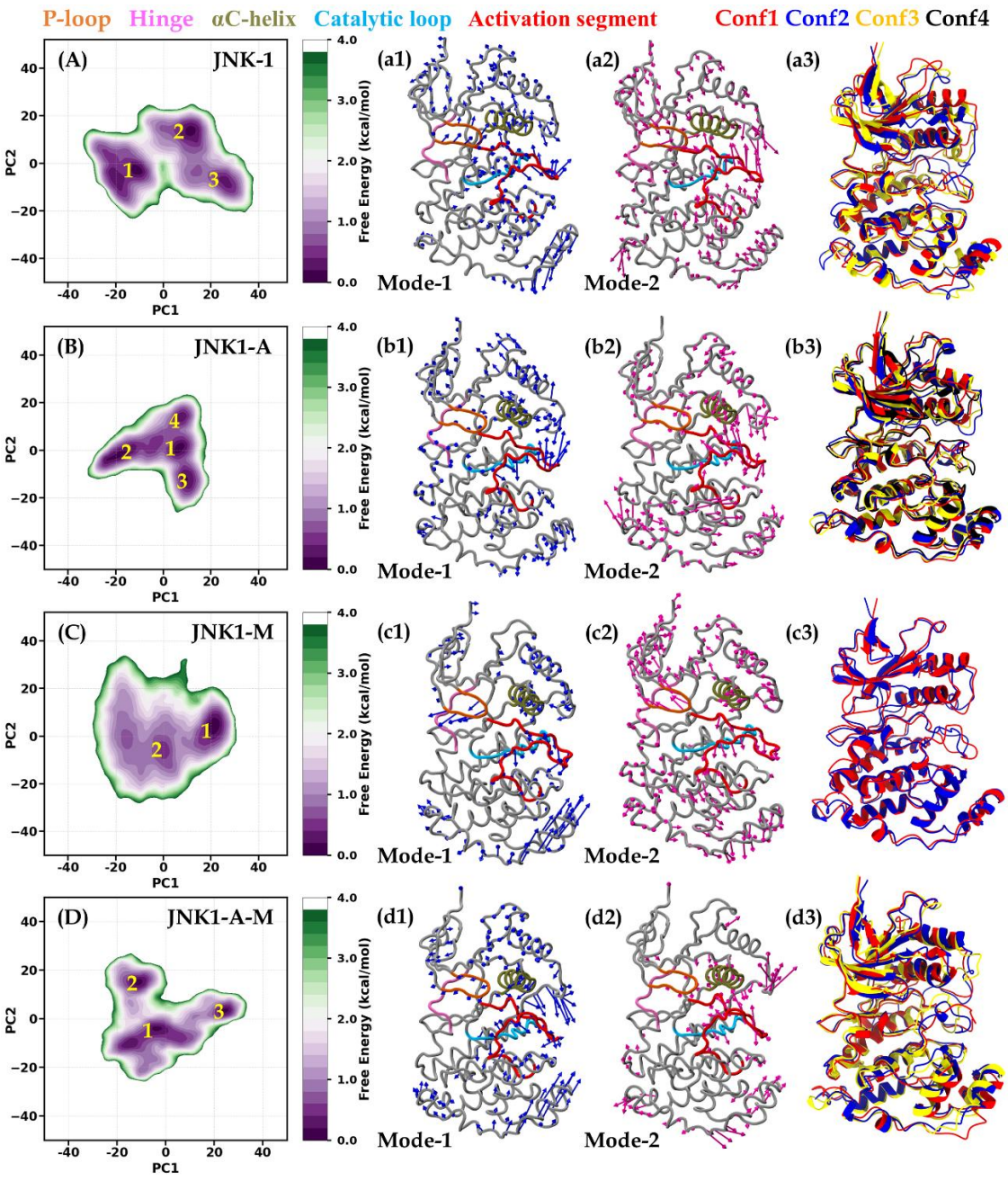


Figure 7.4: Principal component analysis (PCA) was performed for all four systems to characterize their conformational landscapes. (**A–D**) Two-dimensional free energy surfaces are shown, projected onto the first two principal components (PC1 and PC2) for each system. Porcupine plots in panels (**a1–d1**) and (**a2–d2**) illustrate the dominant motions along mode 1 and mode 2, respectively. Panels (**a3–d3**) display representative conformations extracted from the major basins in the free energy landscapes (**A–D**),

each aligned to Conf1 for structural comparison. Key functional regions are color-coded for clarity.

7.3.4. Functional Dynamics of the JNK1 Activation Segment

To elucidate the effects of ATP binding and MKK7 interaction on the conformational dynamics of the JNK1 activation loop (A-loop), we performed dihedral principal component analysis (dPCA) on four studied systems. Unlike Cartesian coordinate-based PCA, dPCA focuses on the protein's internal dihedral angles (φ/ψ), allowing a more detailed examination of local conformational variations relevant to kinase function. The two-dimensional free energy surfaces (FESs) as shown in **Figure 7.5** projected onto the first two dihedral principal components (dPC1 and dPC2) reveal that all systems sample a broad and heterogeneous conformational landscape, with multiple distinct minima separated by varying energy barriers. In the apo system, four major conformational basins are evident, while the addition of ATP, MKK7, or both leads to a more complex landscape with additional populated regions, indicating enhanced conformational diversity. Representative structures extracted from the principal and local minima (conformations 1-4) were superimposed for each system to visualize the structural transitions of the A-loop. Across all systems analyzed, the A-loop is observed in either an extended or semi-extended conformation, with no system exhibiting a fully closed or inactive-like state. Notably, in contrast to other systems where the A-loop remains extended or semi-extended, conformation 4 of the JNK1-A-M system shows a pronounced shift of the A-loop toward the αC-helix, forming a partially closed conformation. Despite this variability, the overall alignment of the kinase core is preserved across all systems, with the most significant structural changes localized to the activation segment and adjacent loops.

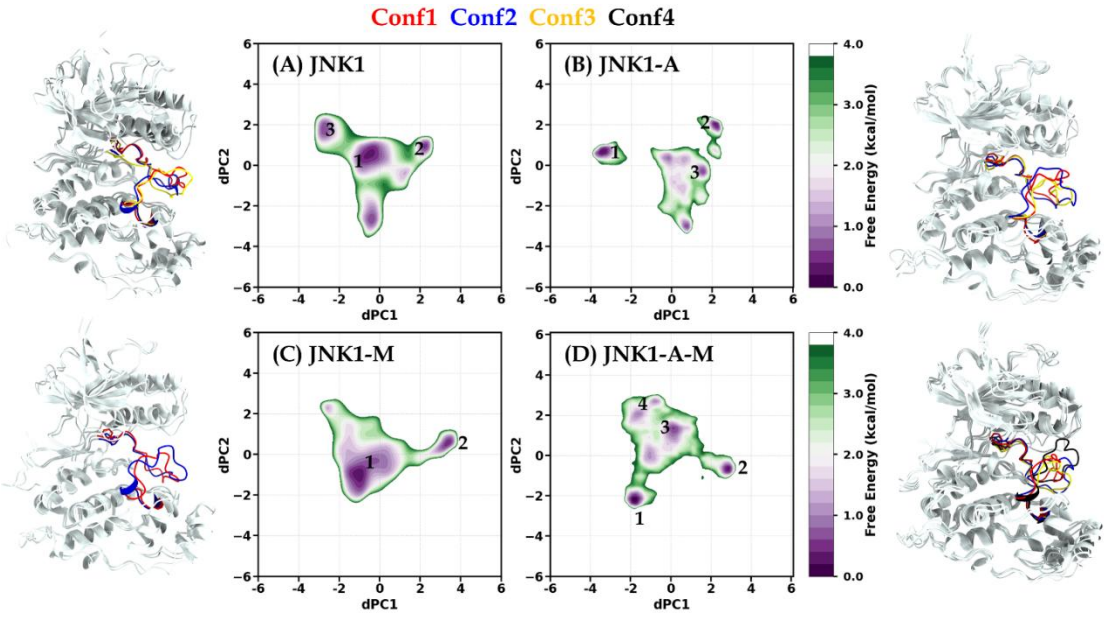


Figure 7.5: Dihedral principal component analysis (dPCA) of the activation segment in the four studied systems. Subplots (A–D) depict the free energy landscapes (FELs) constructed using the first two dihedral principal components (dPC1,2). Representative structures extracted from the highlighted regions in the FELs are shown alongside.

7.3.5. Salt-bridge Triad Formations

The stability of kinase conformations is heavily influenced by salt bridges, especially those formed between the αC-helix and the activation segment with β3-Lysine. These interactions function as regulatory elements that control kinase activity [449]. In **Figure E7(C-D)**, we perfomed a box plot analysis of the Lys55-Asp169 and Lys55-Glu73 distances obtained from the trajectories. For the Lys55-Asp169 distance, the lowest distribution range was observed in JNK1-A, followed by JNK1-A-M, while the other two systems demonstrated a slightly broader distribution. For the Lys55-Glu73 distance in **Figure E7(C)**, the prominent distribution was well within a 4 Å range, with JNK1-A having the smallest box width. We further generated the free energy landscape (FEL) using the distances between these salt-bridge-forming residues on the X and Y axes as shown in **Figure 7.6**. Across the examined systems, the Lys55–Glu73 salt bridge is stably maintained, reflected by a high-density region at short distances in the free energy plots. The consistent presence of this interaction reflects strong electrostatic attraction and precise structural alignment, likely serving as a core

stabilizing element of the kinase, independent of ligand or partner binding. In contrast, the salt bridge involving Asp169 of the DFG motif exhibits greater flexibility and forms less frequently, as reflected by broader free energy distributions and a notable population at longer distances. It suggests that Lys55-Glu73 provides a stable anchoring interaction, while the Asp169 salt bridge is more dynamic and less stable. The stable Lys55–Glu73 salt bridge likely helps maintain the alignment of the αC-helix and β3 strand, key for kinase function, whereas the flexible Asp169 interaction permits structural changes which may be essential for kinase regulation.

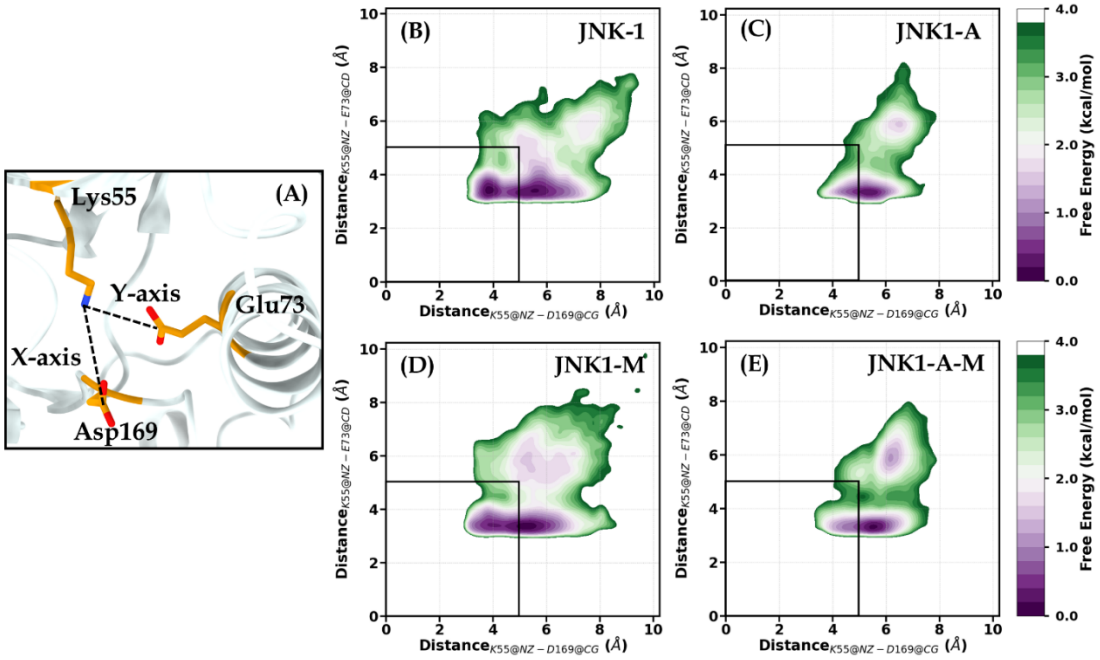


Figure 7.6: (A) Illustration of distances between potential salt-bridge forming residues: β3 strand Lys55, αC-helix Glu73, and DFG-motif Asp169. (B-E) Free energy landscapes (FELs) constructed using the distances between Lys55@NZ–Asp169@CG and Lys55@NZ–Glu73@CD.

7.3.6. P-loop Structural Dynamics

The P-loop is essential for coordinating ATP and inhibitor interactions and governs water movement within the binding cavity [425]. Hence, understanding the conformational dynamics of the P-loop is crucial for systematically designing selective inhibitors targeting JNK1. Collie et al. recently reported that in c-Met kinases,

competitive inhibitors bind to two different conformations of P-loop, displaying either folded or extended structures that are critical for their function [426]. In the inactive excited state of Abl kinase, the P- loop moves towards the αC-helix, leading to a rotational shift that disrupts the β3-Lys αC-Glu ion pair and causes the αC-helix to adopt the "out" conformation [427]. Initially, a box plot was used to analyze P-loop RMSD (**Figure 7.2(C**)), revealing that the JNK1-A and JNK1-A-M systems exhibited narrower boxes, indicating more restricted fluctuations compared to other systems. Further, a distance analysis of the P-loop from the conserved HRD-motif of the catalytic loop (**Figure E7(B)**) indicated that the ATP bound systems JNK1-A and JNK1-A-M, had the smallest box indicating most restricted movement of the P-loop. To gain deeper insights, we constructed the free energy landscape (FEL) using P-loop RMSD plotted against the HRD motif distance as shown in **Figure 7.7**. The energy landscapes of the apo JNK1 and JNK1-M systems span a broad conformational space, with multiple minima indicating significant conformational transitions. Each minimum reflects a unique interplay between the P-loop's flexibility and the HRD motif's positioning, highlighting the conformational diversity between open and closed forms. In contrast, the ATP-bound systems settle into a single dominant conformational state, as indicated by their free energy landscapes. The lack of additional minima in ATP-bound systems suggests that ATP binding enforces a rigid conformational restraint on the P-loop, locking it into a defined orientation that likely stabilizes the active site and ensures proper catalytic alignment.

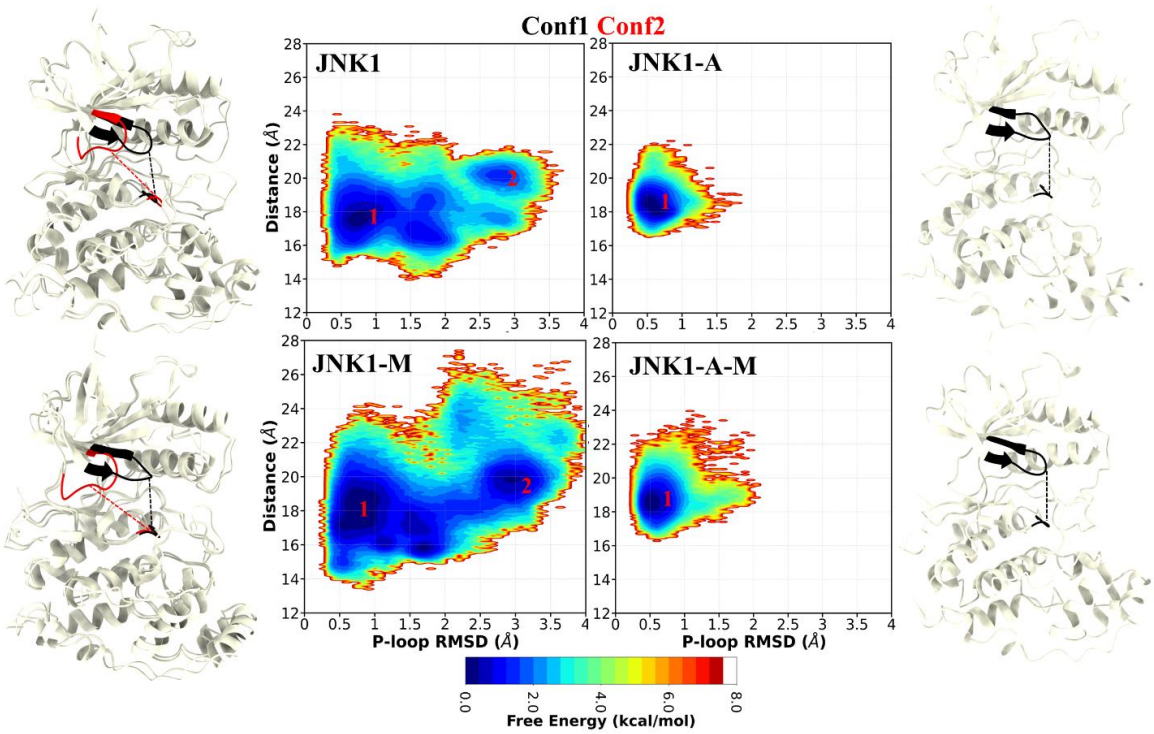


Figure 7.7: Free energy landscape (FEL) analysis using the backbone RMSD of the P-loop and its distance from the conserved HRD motif as reaction coordinates. Representative structures from the highlighted regions of the FEL are shown alongside.

7.3.7. Residue Motion Correlation and Network Formation

Dynamic cross-correlation matrix (DCCM) analysis was used to systematically evaluate the inter-residue movement correlations throughout the JNK1 kinase domain. The DCCM maps in **Figure 7.8** use color coding to represent correlated (positive) and anti-correlated (negative) motions between residues. Highly correlated motions (above 0.4 or below –0.4) are visualized in the matrix. The diagonal and near-diagonal regions of the DCCM map showed highly correlated motions, whereas the off-diagonal areas displayed varying degrees of both correlated and anti-correlated movements. Overall, the JNK1-A-M systems exhibited reduced anti-correlated motions compared to the other systems, as indicated by a reduction in the low and light blue regions. While the correlation pattern for specific residue pairs was broadly maintained, slight changes in the degree of correlation were present.

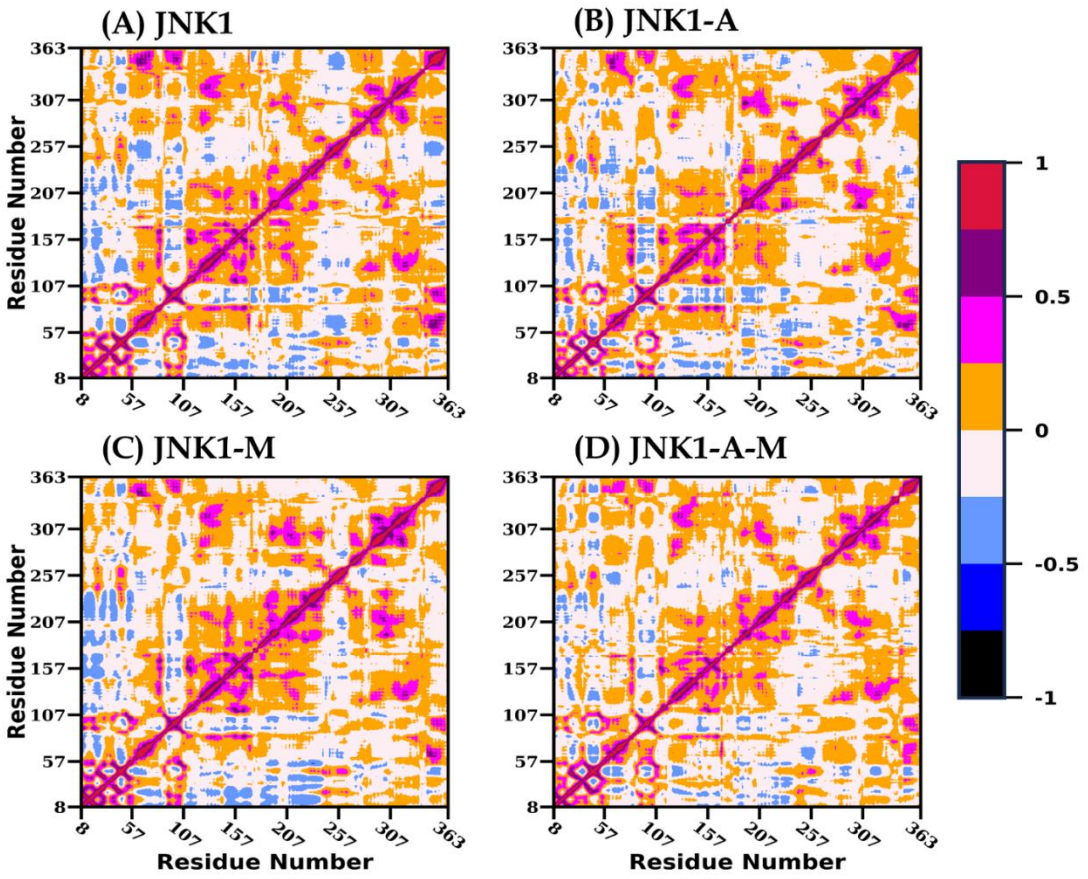


Figure 7.8: Dynamic cross-correlation matrix (DCCM) analysis of the four systems based on Cα atomic coordinates from the simulation trajectories. The degree of correlation is represented using a color scale shown alongside.

To further explore inter-residue network formation across all systems, structure network analysis was performed using webPSN v2.0 [221], which integrates protein structure networks (PSNs) with elastic network model–based normal mode analysis (ENM-NMA). A summary of the network parameters is provided in **Table 7.1**, while **Figure 7.9** illustrates the hubs and communities identified in the protein structure graphs (PSGs). The strongest community C1 was formed in the N-lobe of the JNK1 kinase domain **(Figure 7.9).** The differences in the parameter values reported in **Table 7.1** indicate that structural network alterations occurred across the systems. Interestingly, the JNK1 and JNK1-A systems exhibited the same number of hubs and communities, however, JNK1-A showed a decrease in other parameters, including the number of links, hub-mediated links, and the number of nodes and links involved in

communities. The JNK1-M system, on the other hand, exhibited a higher number of linked nodes, total links, hubs, and hub-mediated links. In contrast, the JNK1-A-M complex, where both ATP and MKK7 were bound, showed a marked reduction in several parameters, including the total number of hubs and communities. Overall, these findings suggest that ligand binding, particularly the simultaneous presence of ATP and MKK7, impacts the protein's structural network, resulting in altered connectivity.

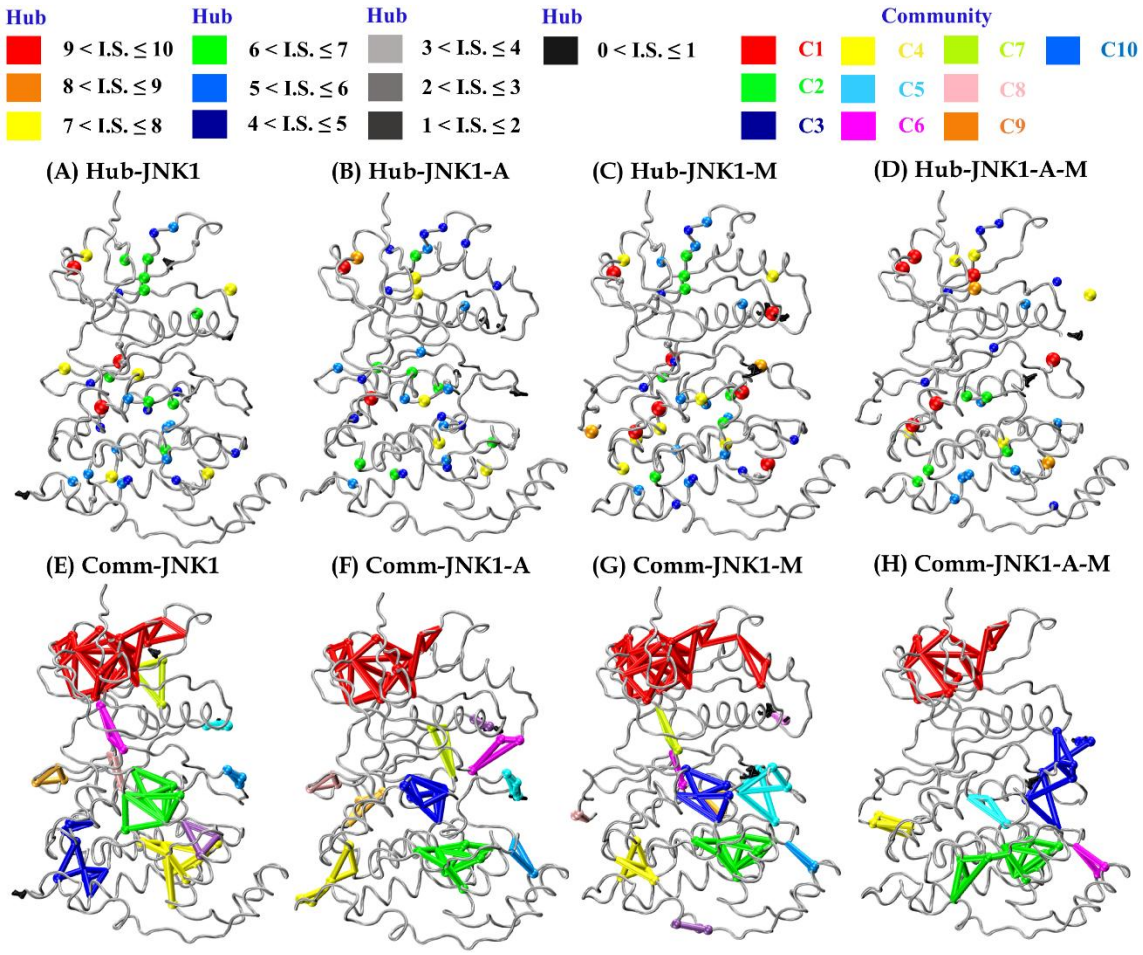


Figure 7.9: Visual representation of hubs and their connections derived from protein structure network (PSN) analysis. Average interaction strength and community rankings are indicated using color codes.

Table 7.1: Overview of network analysis results for the four examined systems.

Network Properties	JNK1	JNK1-A	JNK1-M	JNK1-A-M
No. of linked nodes	318	317	338	332
No. of links	373	366	392	360
No. of hubs	47	47	56	39
No. of links mediated by hubs	189	182	210	151
No. of communities	11	11	12	6
No. of nodes involved in communities	73	63	70	48
No. of links involved in communities	102	84	95	65

7.3.8. Interactions Profile of Complex Systems

We conducted binding free energy calculations, along with residue-level decomposition and hydrogen bond analysis for ATP and the MKK7 docking domain interacting with JNK1. **Table 7.2** presents the average total binding free energy between ATP and JNK1, in JNK1-A and JNK1-A-M systems. Distinct binding energy profiles were observed through MM/PBSA energy calculations for the two systems. Interestingly, while ATP-bound JNK1 exhibited a ΔG_{bind} of –4.31 kcal/mol, the additional binding of MKK7 significantly enhanced ATP affinity towards JNK1, resulting in a ΔG_{bind} of -12.67 kcal/mol for the JNK1-A-M complex. This observation aligns with the known signaling relationship, where JNK1 functions downstream of MKK7. The binding of MKK7 activates JNK1, providing a rational explanation for the increased ATP binding affinity upon MKK7 association. The primary contributor to the increased binding affinity appears to be the interplay between favorable electrostatic interactions and unfavorable polar solvation energy. In the JNK1-A system, the net polar contribution ($\Delta E_{elec} + \Delta G_{pol}$) was slightly unfavorable at 3.51 kcal/mol. In contrast, the corresponding value in the JNK1-A-M complex was –12.67 kcal/mol, indicating a substantially more favorable net polar contribution to ATP

binding. These results demonstrate that MKK7 binding alters the electrostatic environment of JNK1, leading to increased ATP affinity and possibly promoting its activation.

Table 7.2: Binding free energy calculation using MM/PBSA with component-wise energy decomposition (in kcal/mol). Enthalpy and entropy value represent the average of three simulation runs. the standard errors of the mean are included in parentheses.

Systems	**ΔE_{vdw}**	**ΔE_{elec}**	**ΔG_{pol}**	**ΔG_{np}**	**$-T\Delta S$**	**ΔG_{bind}**
JNK1-A	-25.48 (0.09)	-127.43 (0.44)	130.94 (0.39)	-4.26 (0.00)	21.92 (1.61)	**-4.31** (1.72)
JNK-A-M (ATP)	-23.59 (0.09)	-161.64 (0.44)	150.18 (0.37)	-4.31 (0.00)	26.69 (1.30)	**-12.67** (1.42)

Furthermore, to gain deeper insights, residue-wise decomposition of total binding free energy was performed, and those contributing below -1.00 kcal/mol are shown in **Figure 7.10(C)**. Each bar reflects the contribution of individual residues to the stability of the respective complexes. In the JNK1-ATP complex, Asp151, Gln-37, Asn114, Gly35, Ser155, Ala36, Lys55, and Leu168 show the most favorable energy contributions, highlighting their importance in ATP binding. In the JNK1-MKK7 complex, residues like Glu329, Trp324, Cys163, Asp162, Arg127, Ser161, Met151, and Tyr130 show significant contributions, reflecting their involvement in the protein-protein interface. The MKK7 residues involved in interactions include Pro39, Pro41, Arg40, Leu43, Gln44, Leu45, Pro46, and Leu47. In addition, 2D interaction plots were generated to visualize various hydrogen bonds and hydrophobic interactions, as depicted in **Figure 7.10(A-B)**. The frequency of hydrogen bonding by important residues in JNK1-ATP and JNK1-MKK7 complexes is depicted in **Figure 7.10(D)**. The key contributor to hydrogen bonding between JNK and ATP was Ser155, with over 60% occupancy, whereas Ser161 and Arg127 were the most significant in the JNK–MKK7 interaction, showing approximately 50–70% occupancy of hydrogen bonds. Overall, the interaction analysis of JNK1 with ATP and MKK7, along with the

identification of key residues involved, provides valuable functional insights that could inform therapeutic strategies targeting these sites.

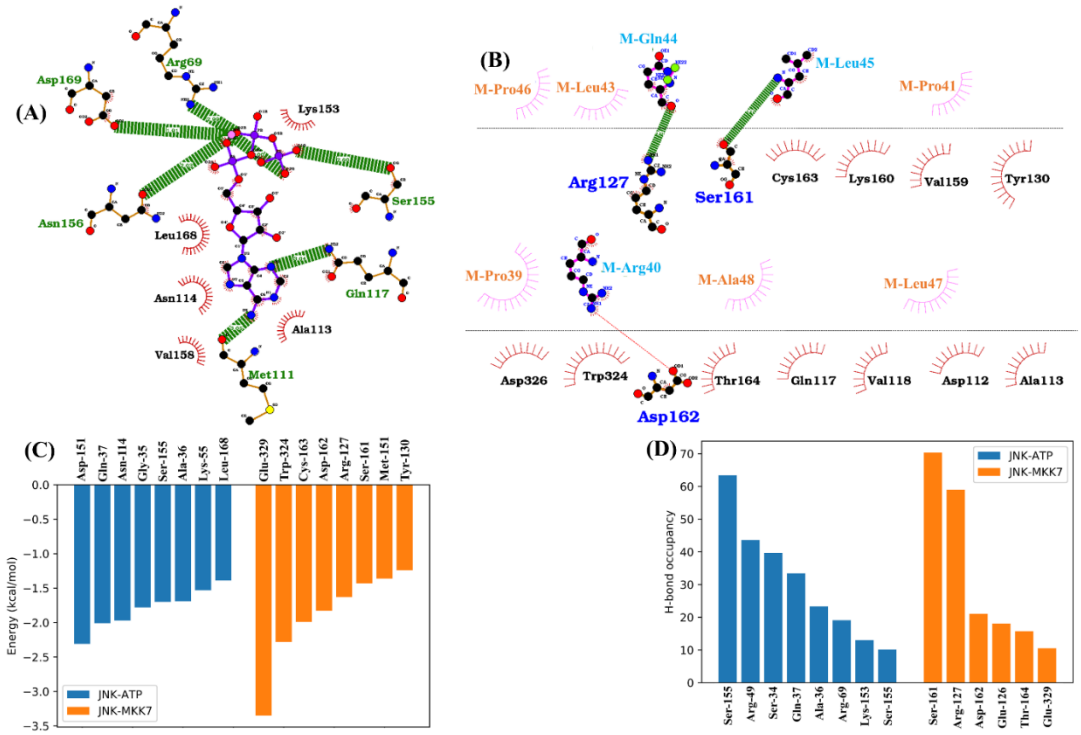


Figure 7.10: Interaction profile, per-residue free energy contribution and formation of hydrogen bonds.

7.4. Conclusions

This study offers a comprehensive investigation of the structural dynamics and conformational plasticity of JNK1 under various molecular conditions using extensive Gaussian accelerated molecular dynamics (GaMD) simulations. Through a detailed structural comparison between the apo form of JNK1 and its ligand-bound counterparts, including ATP-Mg^{2+} alone, MKK7 alone, and the combination of both we have elucidated how each ligand individually and synergistically alters the kinase's conformational landscape and functional dynamics. The RMSD-based analysis revealed that ATP binding confers considerable structural stability, particularly to motifs like the P-loop and αC-helix, while the MKK7 interaction contributes localized

stabilization mainly within the activation segment, highlighting distinct stabilizing effects imparted by each binding partners. Our analysis of the hydrophobic spines revealed that both the regulatory (R-spine) and catalytic (C-spine) residues consistently retain their conserved alignment across all simulated systems, suggesting that the structural integrity of the kinase core is preserved irrespective of ligand binding or docking partner association. Salt-bridge analysis revealed that the interaction between Lys55 and Glu73 remains consistently stable across all systems, serving as a key structural anchor. In contrast, the salt bridge involving Asp169 displayed significant variability, suggesting its dynamic nature may be crucial for modulating activation loop movements during transitions between active and inactive-like conformations. Analysis of the free energy landscapes revealed that ATP binding imposes conformational restriction on the P-loop, supporting its structural role in catalytic site stabilization and its potential impact on the coordination of ligand entry or release. Additionally, analysis of the binding free energy between ATP and JNK1 in both JNK1-A and JNK1-A-M systems demonstrated enhanced ATP affinity in the presence of MKK7, attributed to an improved net polar energy component. Residue-wise energy decomposition further revealed that distinct JNK1 residues are involved in binding, emphasizing the specificity of their interaction interfaces. These findings enhance our understanding of the structure-function relationship of JNK1 and its interaction mechanism with MKK7. We anticipate that the insights gained from this study will support the development of novel inhibitors against JNK1.

Chapter 8

Chapter 8

Conclusions and Future Perspective

8.1. Overall Conclusions

In this doctoral dissertation, I have described an in-depth computational investigation of the conformational dynamics and inhibition mechanisms of key kinases involved in the DLK-JNK signaling pathway, a cascade known to mediate stress responses and neurodegeneration. The primary aim of this thesis has been to elucidate the structural and mechanistic features of DLK, MKK7, and JNK isoforms, and to identify potential therapeutic strategies for modulating their activity through structure-based drug discovery. The study integrates molecular docking, molecular dynamics (MD) simulations, free energy calculations, and enhanced sampling techniques to gain mechanistic insights and guide to study conformational dynamics and inhibitor design.

The first segment of this thesis explored the DLK kinase as a promising therapeutic target for neurodegenerative diseases. Considering the crucial role of DLK in initiating sustained JNK activation after cellular stress and its role in axonal degeneration, the study focused on screening phytochemicals from the SWRT formulation of Ayurvedic medication, typically prescribed for neurological diseases. Through an extensive virtual screening combined with post-docking MD simulations, four phytocompounds were identified that possess favourable ADMET characteristics and high affinity towards the ATP-binding pocket of DLK. The four identified lead compounds are sourced from different plants, specifically *Anethum sowa*, *Acorus calamus*, *Zingiber officinale*, and *Syzygium aromaticum*. By narrowing down the essential phytoconstituents of this ancient formulation, this study may contribute to the design of potent inhibitors through lead optimization against DLK. This study identified key phytochemicals of SWRT formulations and characterized their binding interactions with DLK. To the best of our knowledge, this is one of the first studies to provide molecular insight into the potential of Ayurvedic compounds for DLK-targeted neurodegenerative therapy.

After recognizing the neuroprotective effects of SWRT through DLK inhibition, we directed our efforts toward identifying new ATP-competitive inhibitors targeting DLK. Hence, the second study expanded the drug discovery framework by targeting the ATP-binding site of DLK. The lack of FDA-approved drugs targeting DLK, highlights the necessity for the development of potent inhibitors to address neurological pathologies. Therefore, the design of targeted therapies that are both potent and selective, while avoiding cytotoxic effects, has attracted considerable attention. The second part of the thesis focuses on the conventional systematic approach of high-throughput virtual screening of natural product libraries (NPAtlas) and FDA-approved compounds (MedChemExpress), followed by molecular dynamics simulations integrated with free energy calculations. We carried out virtual screening of natural products and FDA-approved drugs in combination with pharmacokinetic and pharmacological profiling, followed by unbiased molecular dynamics simulations to evaluate the stability of the complexes. We have also evaluated the ADMET properties of the selected compounds. We have also explored the conformational dynamics of the kinase upon inhibitor binding. Finally, we identified two compounds from the NPAtlas database exhibiting favorable ADMET profiles, along with two FDA-approved drugs, danthron and dithranol, showing blood-brain barrier (BBB) permeability based on virtual screening and subjected to MD simulation. Binding free energy estimates using MM-PBSA revealed higher affinity for CID139591660, dithranol, and danthron compared to the control compound sunitinib, whereas CID156581477 exhibited reduced affinity. Furthermore, employing the deep learning-based DeLA Drug server, we identified an analogue of CID156581477 (CID156581477-ANLG) with higher affinity compared to both the parent compound and the control. Based on MM-PBSA and hydrogen bonding results, the inhibitor affinities followed the order: CID139591660 > dithranol > CID156581477-ANLG > danthron > sunitinib > CID156581477. Thus, four compounds were found to exhibit greater affinity than the control compound, sunitinib. However, further molecular and pharmacokinetic studies are required to test the selectivity and affinity *in vitro* and *in vivo*.

The third part of the thesis focused on MKK7 kinase, the immediate downstream effector of DLK and a dual-specificity kinase central to JNK activation. Using both

type I, type II, and allosteric inhibitors, we systematically investigated the conformational transitions and stability of MKK7 through Gaussian accelerated MD (GaMD) simulations. We observed that the inactive state is characterized by a closed A-loop, outward displacement of the αC-helix, disruption of the hallmark salt bridge, and misalignment of the hydrophobic spines. Notably, the type II inhibitors promoted a more stable inactive conformation of MKK7 by enlarging the ATP-binding cleft and disrupting key catalytic residues. Finally, we computed the binding free energy and compared the binding configurations using the MM/PBSA approach. According to the binding free energy calculations, type II inhibitors show the highest affinity, followed by N-terminal allosteric inhibitors and type I inhibitors. These structural and energetic analyses could serve as a foundation for developing more effective type I/II and N-terminal allosteric inhibitors, with significant implications for treating MKK7-related diseases.

In addition, we have also explored the comparative structural dynamics and energetic characteristics of the three JNK isoforms JNK1, JNK2, and JNK3 bound with the inhibitor SP600125 in both phosphorylated and unphosphorylated forms. Investigating the structural behaviour of these isoforms in phosphorylation states is crucial for understanding their functional variety and regulatory mechanisms. We observe that the structural integrity of isoforms remains largely conserved, but phosphorylation leads to conformational changes in the activation loop and some other key regions, affecting accessibility, flexibility, and compactness. The phosphorylation affects the β3-Lys and αC-helix-Glu salt bridge in JNK1 and JNK3, while JNK2 remains unchanged, and no isoform exhibits a stable β3-Lys and DFG-Asp salt bridge interaction. Finally, we computed the binding free energy and compared the binding configurations using the MM-PBSA approach. The results suggest that SP600125 shows significant affinity for all JNK isoforms. Overall, our findings offer valuable insights into the structural and conformational dynamics of JNK isoforms in complex with a competitive inhibitor, revealing fundamental challenges in isoform-specific drug discovery.

Finally, we have investigated the JNK1 dynamics in complex with ATP and MKK7. Here, we primarily focused on the conformational dynamics of JNK1 in its apo form,

ATP-bound state, MKK7-bound complex, and the dual ATP-MKK7 bound state, along with their associated binding free energy calculations. The GaMD studies revealed conformational variability among the different states. The binding of ATP resulted in a restricted mobility of the P-loop, leading to a rigid free energy landscape in terms of its up and down motion towards the conserved HRD-motif. The activation segment adopted a range of conformations, including states where the loop shifted toward the αC-helix. Throughout all systems, the alignment of the hydrophobic spine architecture remained consistently intact. The salt bridge between β3-Lys55 and Asp169 of the DFG-motif alternated between formed and broken states. In contrast, the key interaction between β3-Lys55 and αC-helix-Glu73 remained mostly intact across systems, except in the ATP-MKK7 bound complexes, where an additional prominent broken conformation was also observed. Furthermore, the binding mechanism, residue-specific energy contributions, and interaction profiles were thoroughly analysed. Overall, the study provides critical mechanistic insights into JNK1 in complex with MKK7 and ATP-bound states, emphasizing how binding influences their dynamic behaviour and interaction patterns.

Collectively, this thesis offers a comprehensive computational framework to understand the DLK-MKK7-JNK signaling axis at an atomistic level. Each chapter contributes new insights ranging from natural product-based inhibition, structure-function relationships, protein-protein interactions, and isoform-specific modulation to kinase conformational dynamics, which are critical for guiding future targeted drug development. The strategies and findings reported here hold translational potential and serve as a foundation for future experimental validations and clinical exploration.

8.2. Outlook and Future Directions

In this thesis, I have explored the conformational dynamics and molecular mechanisms of inhibition within the DLK-JNK signaling pathway, which plays a crucial role in neuronal stress response, axonal degeneration, and neuroinflammation. Using multiscale simulation techniques including molecular docking, classical and enhanced sampling molecular dynamics, and free energy calculations, I have investigated the mechanistic landscape of kinase regulation and inhibition, offering new perspectives

for structure-based drug discovery. This work presents the comprehensive computational attempts to investigate the structural and conformational plasticity and inhibition mechanisms of the DLK-MKK7-JNK cascade across multiple levels.

Despite the significance of DLK in mediating pathological signaling in neurological diseases, there is currently no FDA-approved drug targeting DLK. Moreover, its C-helix displays a highly distorted conformation, distinct from canonical kinase structures. This unusual conformation not only challenges conventional drug design approaches but also opens up a novel opportunity to investigate the structural determinants behind this distortion and its functional implications in kinase regulations. Also, a critical gap in our knowledge is the lack of experimentally validated phosphorylation sites in human DLK, which limits our ability to accurately model its active or inactive forms. Future studies, integrating phosphorylation and mass spectrometry data with computational modeling, will be essential to bridge this gap and to explore DLK's regulation at the post-translational level. Additionally, opportunities exist in predicting novel allosteric sites and identifying inhibitors capable of inhibiting them, offering a valuable perspective for drug discovery against DLK.

In the case of JNK isoforms, the high sequence similarity among JNK1, JNK2, and JNK3 presents a significant challenge for selective drug development. We observed that phosphorylation induces subtle but important conformational differences among isoforms, yet creating highly selective inhibitors for each isoform is still a significant challenge. Achieving high selectivity among these isoforms is a key future goal, which will require the integration of long-timescale simulations, isoform-specific structural data, and advanced machine-learning-based binding prediction models. Future studies could focus on predicting allosteric pockets in JNK isoforms and MKK4 along with identifying inhibitors targeting these pockets as a promising strategy. As illustrated in Figure 8.1, we present the future scope of work on the DLK-JNK signaling pathway, detailing the possible analyses for each kinase.

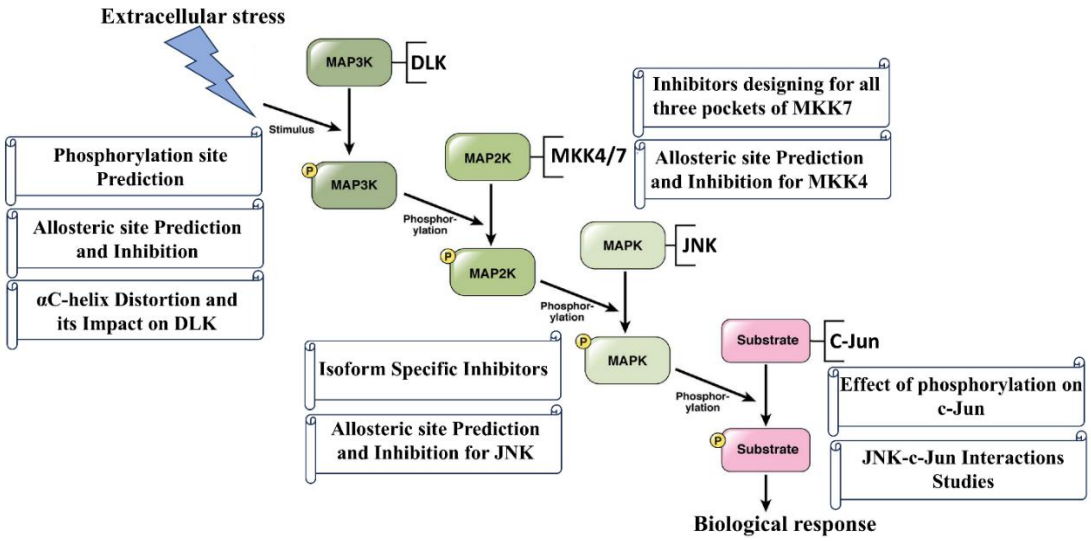


Figure 8.1: Illustration of the DLK–JNK signaling pathway, showing the wide range of research possibilities at each kinase level. These include phosphorylation analysis, mapping of allosteric pockets, studying structural changes, and the search for novel inhibitors.

Additionally, while my work focused on the ATP-binding pockets and allosteric sites of individual kinases, future researchers may benefit from expanding the scope to include protein-protein interactions, such as the DLK-MKK7 and MKK7- with individual JNK isoform complexes, which are central to the functionality of signaling. Targeting these interfaces could offer a new layer of selectivity beyond what is achievable through ATP-competitive and other allosteric inhibitors. Therefore, our study would be a great starting point for further novel drug development and conformational study of the DLK-JNK signaling pathway.

Appendix

Appendix A

Table A1: The ADMET analysis of all four selected leads.

Properties	Lead-1	Lead-2	Lead-3	Lead-4
Physiochemical and other properties				
Molecular weight (Da)	**492.389**	**330.29**	**374.433**	**484.366**
Volume	**445.387**	**323.279**	**388.945**	**430.778**
Density	**1.105**	**1.021**	**0.962**	**1.124**
nHA	**13**	**7**	**6**	**14**
nHD	**7**	**4**	**3**	**9**
nRot	**5**	**6**	**10**	**7**
nRing	**4**	**2**	**2**	**3**
MaxRing	**10**	**6**	**6**	**6**
nHet	**13**	**7**	**6**	**14**
fChar	**0**	**0**	**0**	**0**
nRig	**25**	**15**	**13**	**20**
Flexibility	**0.2**	**0.4**	**0.769**	**0.35**
Stereo Centers	**5**	**0**	**1**	**5**
TPSA	**216.58**	**124.29**	**96.22**	**243.9**
logS	**-4.164**	**-2.821**	**-3.236**	**-1.96**
logP	**0.667**	**2.089**	**1.829**	**-0.295**
logD	**1.001**	**1.813**	**2.263**	**0.119**
$T_{1/2}$	**0.918**	**0.956**	**0.934**	**0.967**
MDCK permeability	**1.8e-05**	**1.4e-05**	**1.5e-05**	**2.6e-05**
Pgp-inhibitor	**0.0**	**0.006**	**0.039**	**0.005**
Pgp-substrate	**0.729**	**0.006**	**0.994**	**0.032**
Toxicity Evaluation				
hERG Blockers	0.006	0.011	0.086	0.029

H-HT	0.207	0.141	0.395	0.037
DILI	0.098	0.936	0.045	0.954
AMES Toxicity	0.366	0.072	0.224	0.076
Rat Oral Acute Toxicity	0.063	0.518	0.018	0.007
FDAMDD	0.009	0.031	0.897	0.003
Skin Sensiti zation	0.029	0.956	0.429	0.919
Carcinogen city	0.041	0.197	0.043	0.008
Eye corrosion	0.003	0.016	0.004	0.003
Eye Irritation	0.01	0.925	0.155	0.258
Respiratory Toxicity	0.029	0.042	0.058	0.011

Volume: Van der Waals volume. **Density:** MW/Volume. **nHA:** Number of hydrogen bond acceptors; optimal: 0~12. **nHD:** Number of hydrogen bond donors; optimal:0~7. **nRot:** Number of rotatable bonds; optimal:0~11. **nRing:** Number of rings; optimal:0~6. **MaxRing:** Number of atoms in the biggest ring; optimal: 0~18. **nHet:** Number of heteroatoms; optimal:1~15. **fChar:** Formal charge; optimal:-4~4. **nRig:** Number of rigid bonds; optimal:0~30. **Flexibility:** nRot/nRig. **Stereo centers:** Optimal ≤ 2. **TPSA:** Topological polar surface area; optimal:0~140. **logS:** Log of the aqueous solubility; optimal:-4~0.5 log mol/L.

logP: Log of the octanol/water partition coefficient; optimal:0~3. **logD:** logP at physiological pH 7.4; optimal: 1~3. **$T_{1/2}$:** The output value is the probability of having a long half-life; long half-life: > 3 h. **MDCK permeability:** low: < 2 x 10^{-6} cm/s; medium: 2-20 x 10^{-6} cm/s; high: > 20 x 10^{-6} cm/s. **Pgp-inhibitor/substrate:** The output value is the probability of being Pgp-inhibitor or Pgp-substrate.

The toxicity results indicate the probability of the ligand being active/toxic in the corresponding analyses: selection of drugs with less toxicity/adverse effects: 0-0.3; excellent; 0.3-0.7: medium; 0.7-1.0: poor.

Table A2: List of leads along with their compound ID and SMILES

Leads	**Compound ID**	**SMILES**
Lead-1	CID5491630	COC1=C(C=CC(=C1)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)OC4C(C(C(C(O4)C(=O)O)O)O)O)O
Lead-2	CID16066851	C1=CC(=C(C=C1C=CC(=O)OCC(=O)C2=CC(=C(C=C2)O)O)O)O
Lead-3	CID5318039	COC1=C(C=CC(=C1)CCC(CC(=O)CCC2=CC(=C(C=C2)O)OC)O)O
Lead-4	CID471118	C1=C(C=C(C(=C1O)O)O)C(=O)OC2C(C(OC(C2OC(=O)C3=CC(=C(C(=C3)O)O)O)O)O)CO)O

Table A3: Tanimoto Coefficient for Molecule Structure Similarity Measurement.

System	Lead-1	Lead-2	Lead-3	Lead-4	sunitinib
Lead-1	1.0	0.18	0.19	0.21	0.14
Lead-2	0.18	1.0	0.24	0.20	0.12
Lead-3	0.19	0.24	1.0	0.17	0.12
Lead-4	0.21	0.20	0.17	1.0	0.10
sunitinib	0.14	0.12	0.12	0.10	1.0

Table A4: Residual decomposition of the binding free energy in kcal/mol between the DLK and Different Ligand Molecules. Binding free energy with values < –1.0 kcal/mol are listed here.

Residue	ΔE_{vdw}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	ΔG_{bind}
Lead-1					
VAL-139	-1.69	-0.40	0.43	0	-1.66
LEU-243	-1.85	0.13	0.20	0	-1.52
GLU-191	-0.06	-3.42	2.33	0	-1.15
PHE-192	-1.62	-0.86	1.49	0	-1.00
Lead-2					
ASP-254	1.07	-13.75	10.46	0.01	-2.21
LEU-243	-1.62	-0.13	0.27	0	-1.48
PHE-192	-1.62	-1.12	1.30	0	-1.44
LEU-141	-1.20	-0.11	0.18	0	-1.13
Lead-3					
PHE-192	-1.74	-2.11	1.46	0	-2.38
CYS-193	-1.0	-3.53	2.81	0	-1.71
VAL-139	-1.36	-0.13	0.14	0	-1.36
LEU-141	-1.43	0.27	-0.01	0	-1.17
ALA-150	-1.11	-0.16	0.20	0	-1.03
Lead-4					
ASP-236	0.82	-8.36	5.17	0	-2.37
VAL-139	-1.87	-0.20	0.20	0	-1.87
LEU-243	-1.35	-0.01	0.25	0	-1.11

Table A5: Main hydrogen bond interactions formed by DLK with leads and the corresponding average distance, angle, and percentage of occupancy determined using the trajectories of production simulations. Hydrogen bonds with more than 10% occupancy are listed here.

System	Acceptor	Donor	Distance	Angle	Occupancy
Lead-1					
Run1	Lig@O5	Ser253@OG	2.84	155.92	18.32
	Lig@O4	Cys193@N	2.92	158.87	13.11
	Glu191@O	Lig@O3	2.75	152.61	70.39
	Asp254@OD1	Lig@O8	2.67	158.33	24.85
	Asp254@OD2	Lig@O8	2.67	157.90	20.74
	Ser133@O	Lig@O13	2.75	152.71	10.00
Run-2	Lig@O5	Ser253@OG	2.83	156.70	24.26
	Lig@O4	Cys193@N	2.92	159.57	10.46
	Glu191@O	Lig@O3	2.74	152.91	81.68
	Asp254@OD2	Lig@O8	2.65	161.52	25.30
	Ser133@O	Lig@O13	2.71	162.93	20.27
	Asp254@OD1	Lig@O8	2.66	159.34	15.98
Lead-2					
Run-1	Lig@O7	Cys193@H	2.87	159.76	59.14
	Lig@O6	Ser253@HG	2.76	156.44	36.33
	Asp254@OD2	Lig@O6	2.65	159.65	39.04
	Asp254@OD1	Lig@O6	2.65	159.52	37.42
	Asp254@OD2	Lig@O5	2.70	161.94	33.35
	Asp254@OD1	Lig@O5	2.70	161.68	32.11
	Lig@O7	Cys193@N	2.87	158.89	52.76
	Lig@O5	Ser253@OG	2.79	152.90	20.68
	Asp254@OD2	Lig@O5	2.68	160.01	47.50

Run-2	Asp254@OD2	Lig@O6	2.66	161.27	46.69
	Asp254@OD1	Lig@O6	2.67	161.34	45.28
	Asp254@OD1	Lig@O5	2.69	160.26	38.12
Lead-3					
Run-1	Lig@O5	Cys193@N	2.89	153.54	51.59
	Glu191@O	Lig@O3	2.71	155.93	86.35
	Cys193@O	Lig@O5	2.84	144.35	28.29
	Ser133@O	Lig@O6	2.79	148.31	12.23
Run-2	Lig@O5	Cys193@N	2.88	151.24	57.24
	Glu191@O	Lig@O3	2.73	154.71	86.63
	Cys193@O	Lig@O5	2.84	145.30	38.27
Lead-4					
Run-1	Asp236@OD1	Lig@O2	2.61	164.48	45.01
	Gly134@O	Lig@O13	2.71	162.06	41.07
	Cys193@O	Lig@O12	2.74	162.71	40.41
	Asp236@OD1	Lig@O1	2.62	165.02	35.60
	Asp236@OD1	Lig@O3	2.61	164.77	16.84
	Cys193@O	Lig@O3	2.70	147.56	15.36
	Asp254@OD2	Lig@O12	2.66	162.74	14.84
Run-2	Lig@O6	Ser253@HG	2.81	153.26	14.48
	Lig@O11	Gln197@N	2.91	156.76	13.81
	Lig@O12	Ser253@OG	2.84	156.52	11.37
	Cys193@O	Lig@O3	2.72	152.19	70.04
	Asp236@OD2	Lig@O12	2.64	164.12	39.71
	Asp254@OD1	Lig@O12	2.64	164.07	35.26
	Asp254@OD1	Lig@O13	2.70	153.83	35.02
	Asp254@OD2	Lig@O13	2.70	154.75	32.44
	Glu200@OE1	Lig@O10	2.61	165.73	14.32
	Glu200@OE1	Lig@O11	2.61	165.43	13.13
	Glu200@OE2	Lig@O10	2.61	165.77	12.39

	Glu200@OE2	Lig@O11	2.61	165.24	11.13

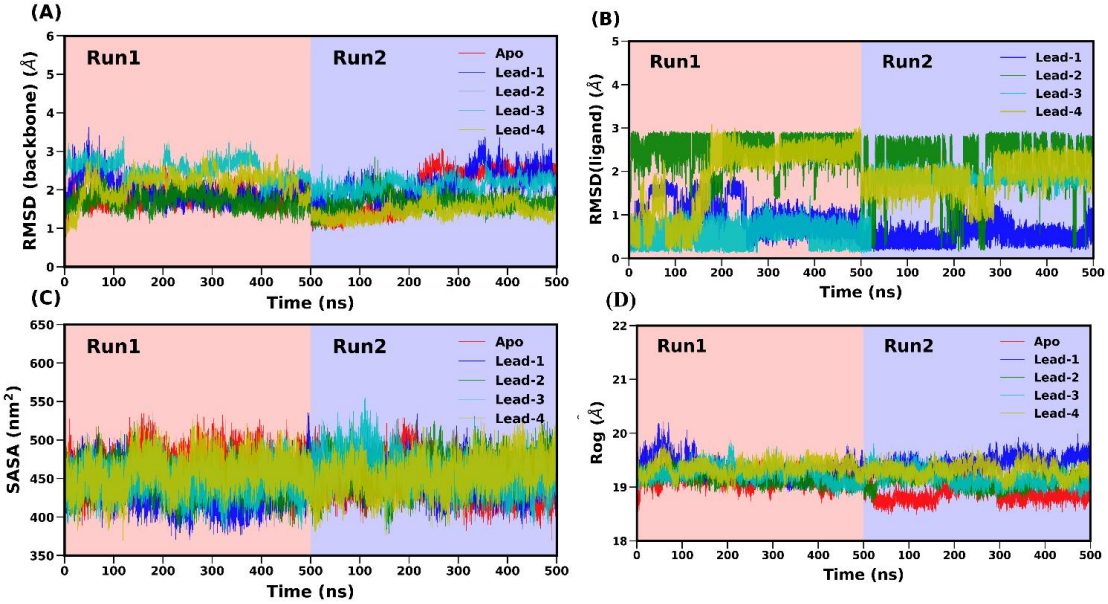


Figure A1: Time evolution of **(A)** root-mean-square deviation (RMSD) of backbone atoms of DLK, **(B)** ligand RMSD, **(C)** Solvent accessible surface area (SASA) of DLK, and **(D)** Radius of gyration of DLK.

Appendix B

Table B1: The drug-likeliness properties calculated using the SwissADME webserver. The properties are calculated for two new identified compounds, control, and few reported DLK inhibitors.

Compounds CID	Molecular formula	Molecular weight (Da)	No. of H-bond acceptors	No. of H-bond donors	Molar Refractivity
156581477	C17H24O4	292.37 g/mol	4	3	82.64
139591660	C18H26O4	306.40 g/mol	4	3	87.29
Sunitinib	C22H27FN4O2	398.47 g/mol	4	3	116.31
GNE-3511	C23H26F2N6O	440.49 g/mol	7	1	123.36
GDC-0134	C19H20F4N6O2	440.39 g/mol	9	1	109.46
DN1289	C18H19F4N7O2	441.38 g/mol	10	1	107.26
IACS-52825	C16H13F7N4O2	426.29 g/mol	11	2	83.91

Molecular weight range: 150-500 g/mol

No. of H-bond donors: ≤5

No. of H-bond acceptor: ≤10

Molecular refractivity: 40-130

Table B2: The Pharmacokinetics properties predicted using SwissADME webserver. The properties are calculated for two new identified compounds, control, and few reported DLK inhibitors.

Compounds CID	GI absorption	BBB permeant	P-gp substrate	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor	Log K_p
156581477	High	Yes	No	No	No	No	Yes	No	-6.49 cm/s
139591660	High	Yes	No	No	No	No	Yes	No	-6.15 cm/s
Sunitinib	High	Yes	Yes	No	Yes	No	Yes	Yes	-6.86 cm/s
GNE-3511	High	Yes	Yes	No	No	Yes	Yes	Yes	-6.69 cm/s
GDC-0134	High	No	No	No	No	No	Yes	Yes	-6.88 cm/s
DN1289	High	No	Yes	No	No	No	Yes	Yes	-7.38 cm/s
IACS-52825	High	No	Yes	No	No	No	Yes	Yes	-7.21 cm/s

Table B3: Toxicity prediction of the compounds using pkCSM webserver. Toxicity prediction is essential to ensure that selected compounds are safe for human use and reduce the risk of failure of compounds in later stages of drug development. The properties are predicted for two new identified compounds, control, and few reported DLK inhibitors.

Model name	**CID156581477**	**CID139591660**	**Sunitinib**	**GNE-3511**	**CDC-0134**	**DN1289**	**IACS-52825**	**Unit**
AMES toxicity	No	No	No	No	No	No	No	Categorical (Yes/No)
Max. tolerated dose (human)	0.402	-0.103	-0.291	-0.935	-0.579	-0.014	0.409	Numeric (log mg/kg/day)
hERG l inhibitor	No	No	No	No	No	No	No	Categorical (Yes/No)
hERG ll inhibitor	No	No	Yes	Yes	No	No	No	Categorical (Yes/No)
Oral Rat Acute Toxicity (LD50)	2.32	2.12	2.327	3.082	2.595	2.69	2.29	Numeric (mol/kg)
Oral Rat Chronic Toxicity (LOAEL)	1.63	1.73	0.926	-0.191	1.14	0.62	0.988	Numeric (log mg/kg_bw/ day)
Hepatotoxicity	No	No	Yes	Yes	Yes	Yes	Yes	Categorical (Yes/No)
Skin Sensitisation	No	No	No	No	No	No	No	Categorical (Yes/No)

T.Pyriformis toxicity	1.027	0.858	0.648	0.325	0.288	0.286	0.285	Numeric (log ug/L)
Minnow toxicity	1.261	1.115	3.811	1.916	2.279	2.752	2.312	Numeric (log mM)

Table B4: Residual decomposition of free energy between the DLK and inhibitors molecules. Residues contributing Energy greater than -1 kcal/mol are listed here. All values are in kcal/mol.

DLK-inhibitor complex	**Residue**	**ΔE_{vdw}**	**ΔE_{elec}**	**ΔG_{pol}**	**ΔG_{np}**	**ΔG_{total}**
DLK/Sunitinib	Val131	-3.09	-4.93	5.42	0	**-2.60**
	Leu243	-1.77	0.54	-0.21	0	**-1.44**
	Val139	-1.42	-0.27	0.49	0	**-1.20**
DLK/Dithranol	Glu191	0.74	-6.32	2.63	0	**-2.95**
	Phe192	-1.26	-2.14	1.46	0	**-1.94**
	Leu243	-1.78	-0.42	0.47	0	**-1.73**
	Leu141	-1.49	-0.28	0.31	0	**-1.46**
	Cys193	-0.92	-2.00	1.86	0	**-1.06**
	Ala150	-1.07	-0.05	0.10	0	**-1.02**
DLK/danthron	Glu191	0.71	-7.32	3.17	0	**-3.44**
	Phe192	-1.14	-1.96	1.44	0	**-1.66**
	Leu243	-1.81	-0.23	0.41	0	**-1.63**
	Leu141	-1.36	-0.12	0.19	0	**-1.29**
	Val131	-1.26	-0.31	0.53	0	**-1.04**
	Cys193	-0.98	-2.21	2.15	0	**-1.04**
DLK/CID139591660	Glu191	0.16	-5.12	2.73	0	**-2.23**
	Leu243	-2.31	-0.17	0.32	0	**-2.16**
	Val139	-1.78	0.03	0.31	0	**-1.44**
	Lys152	-0.59	-6.71	5.87	0	**-1.43**

Table B5: List of CID156581477 analogues with their docking score, blood-brain barrier penetration (BBB) ability, and SMILES. The BBB prediction was done using the SwissADME webserver.

Sl. no	SMILE ID	Docking score	Ml. weight	BBB-penetration
1	N(O)C(C)(O)CCc1ccc(O)c2c1CCCC2O	-11.51	267.15	No
2	CCCCC=CCc1ccc(O)c2c1C(CC(O)C(C)O)CC2	-11.51	318.20	Yes

Table B6: Binding free energy components for DLK with two inhibitors complexes calculated using the MM-PBSA scheme. All values are in kcal/mol.

DLK/ligand complex	ΔE_{vdw}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	$-T\Delta S$	ΔG_{bind}
CID156581477-ANLG	-38.61 (0.06)	-24.57 (0.13)	33.84 (0.08)	-3.86 (0.00)	21.58 (0.72)	**-11.62 (0.74)**
CID156581477	-33.09 (0.05)	-30.86 (0.12)	40.09 (0.08)	-3.61 (0.00)	21.38 (0.80)	**-6.09 (0.88)**

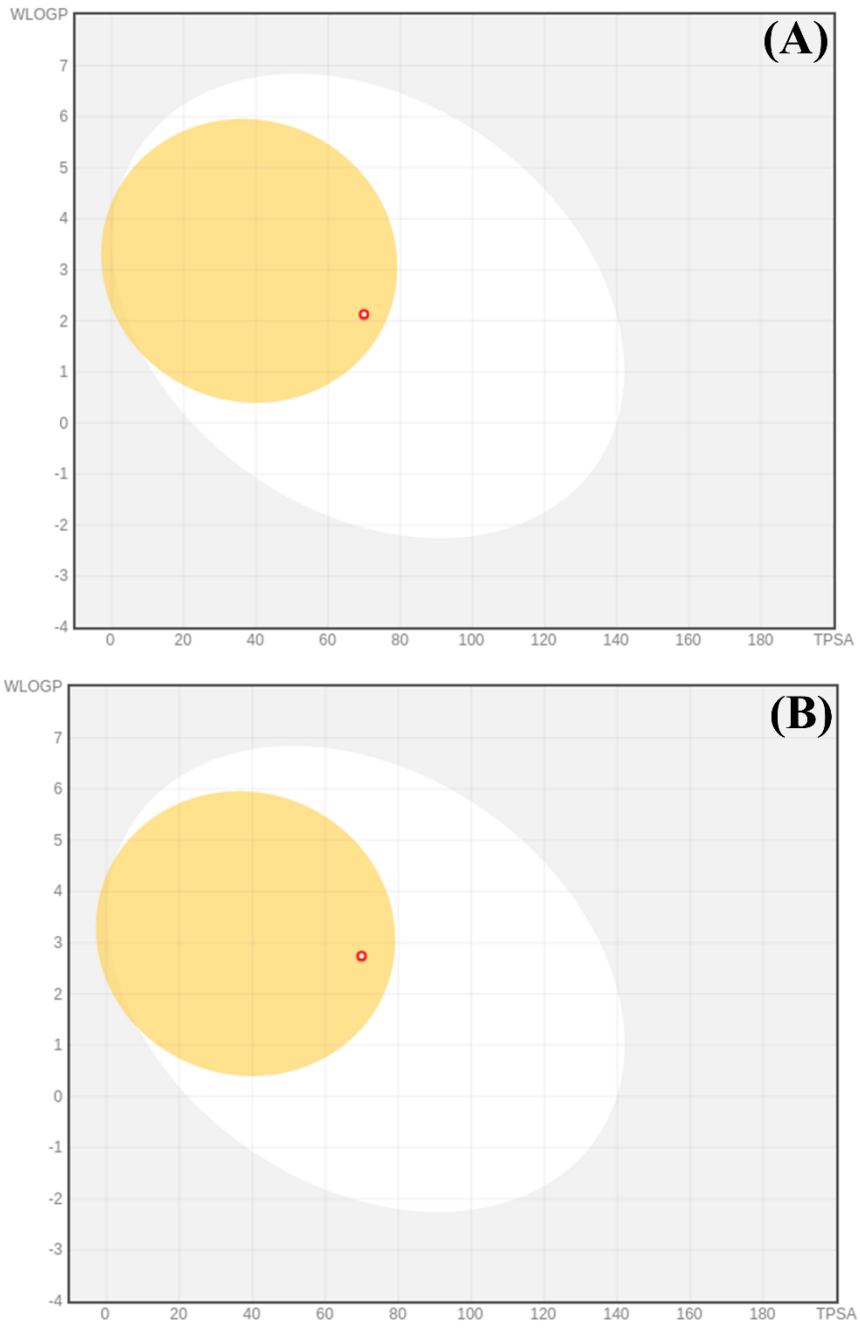


Figure B1: Visual representation of two compounds **(A)** CID156581477 **(B)** CID139591660 using the BOILED-egg model of the SwissADME webserver. Both compounds are a part of yellow Egan eggs, indicating a higher possibility of blood-brain barrier (BBB) permeability.

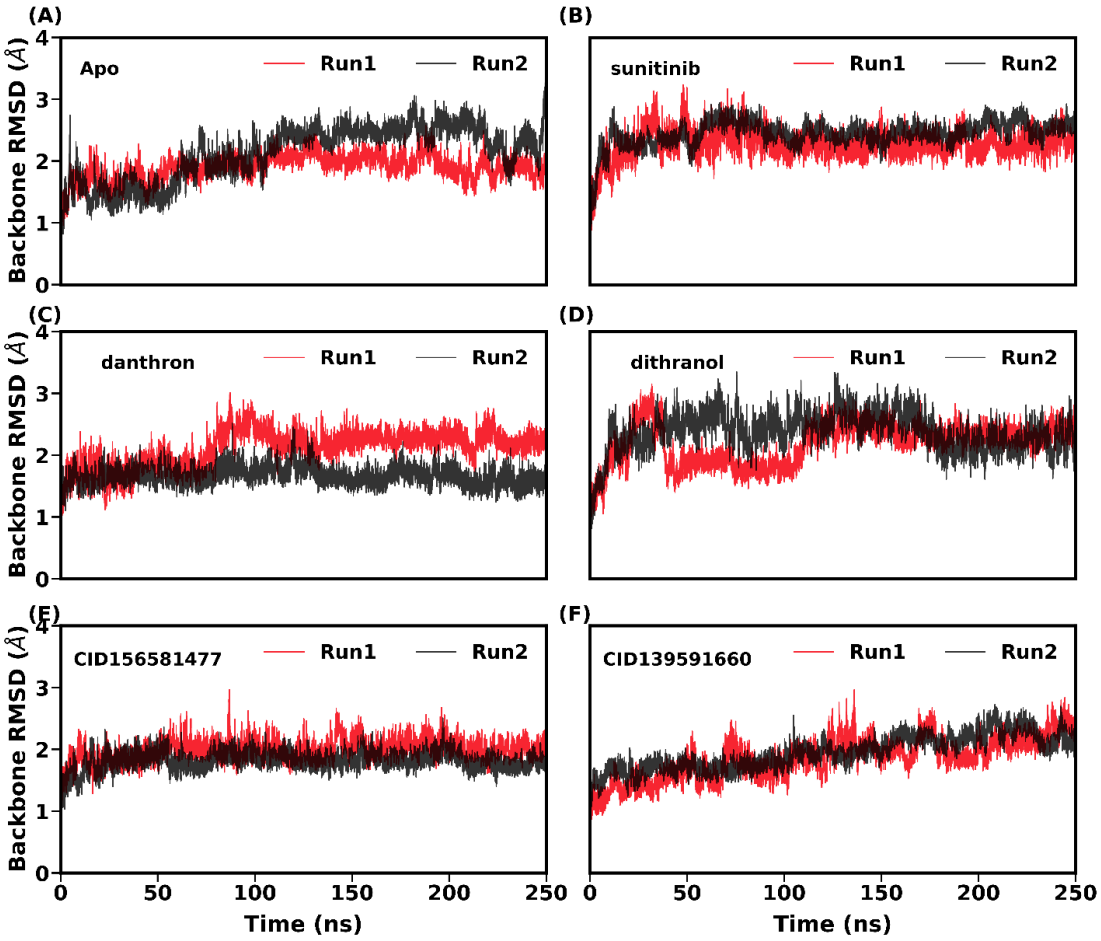


Figure B2: Time evolution of backbone atom RMSD of apo and inhibitor bound DLK.

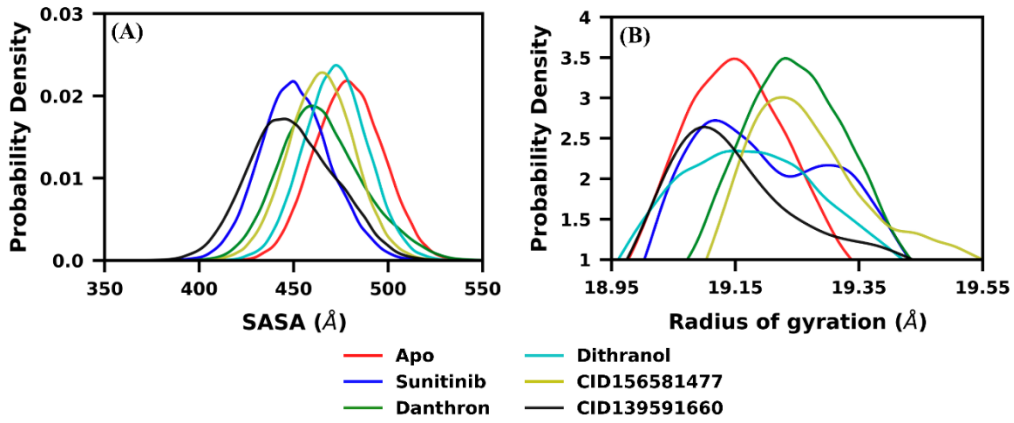


Figure B3: The probability distribution of **(A)** Solvent accessible surface area (SASA), and **(B)** Radius of gyration (R_g) of apo form and inhibitor bound DLK.

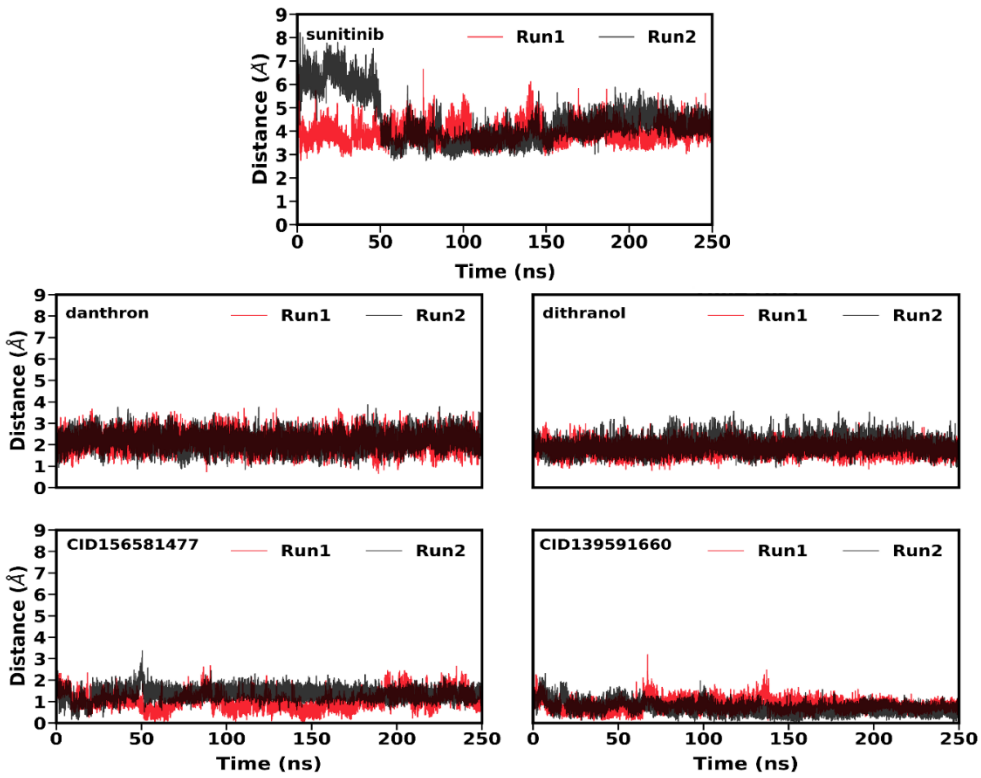


Figure B4: Time evolution of the ligands' distance from the center of mass (CoM) of residues within 5 Å of the ligand in the initial docked conformations.

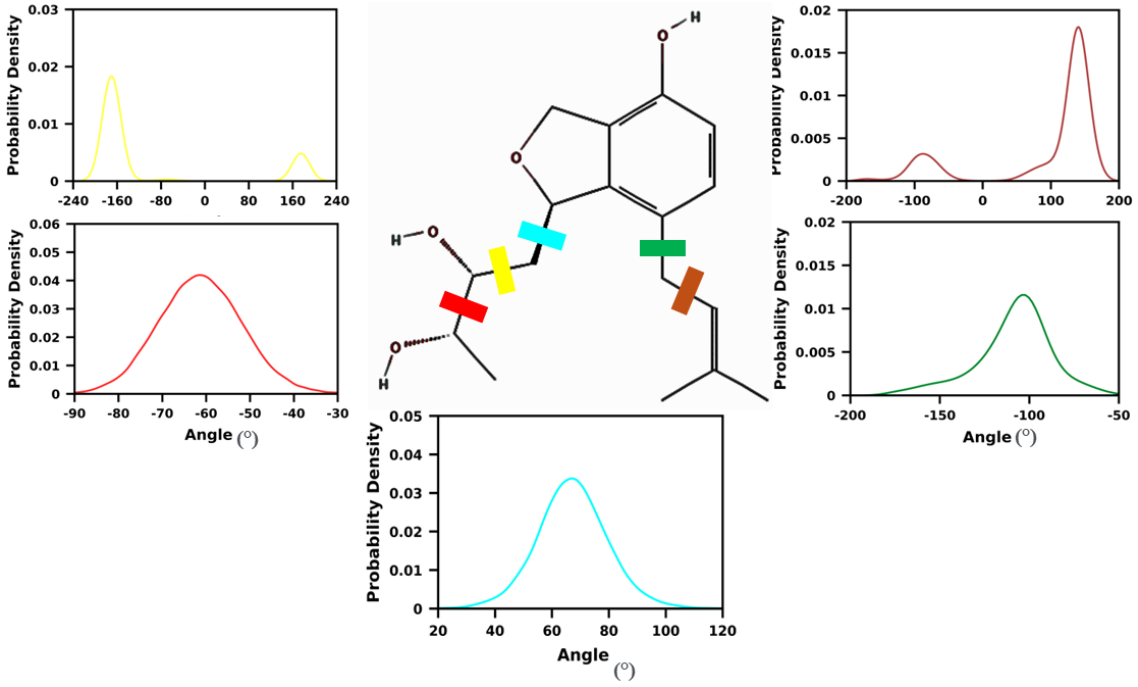


Figure B5: Calculation of various torsion angles' probability density for CID156581477 conducted using the molecular dynamics trajectory.

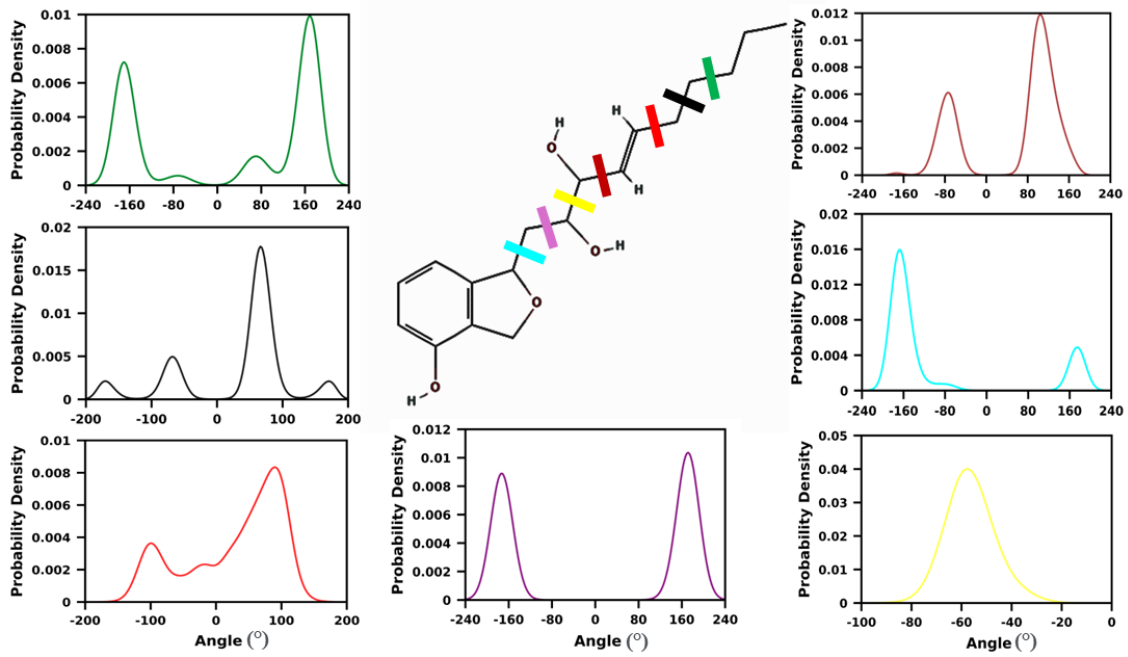


Figure B6: Calculation of various torsion angles' probability density for CID139591660 conducted using the molecular dynamics trajectory.

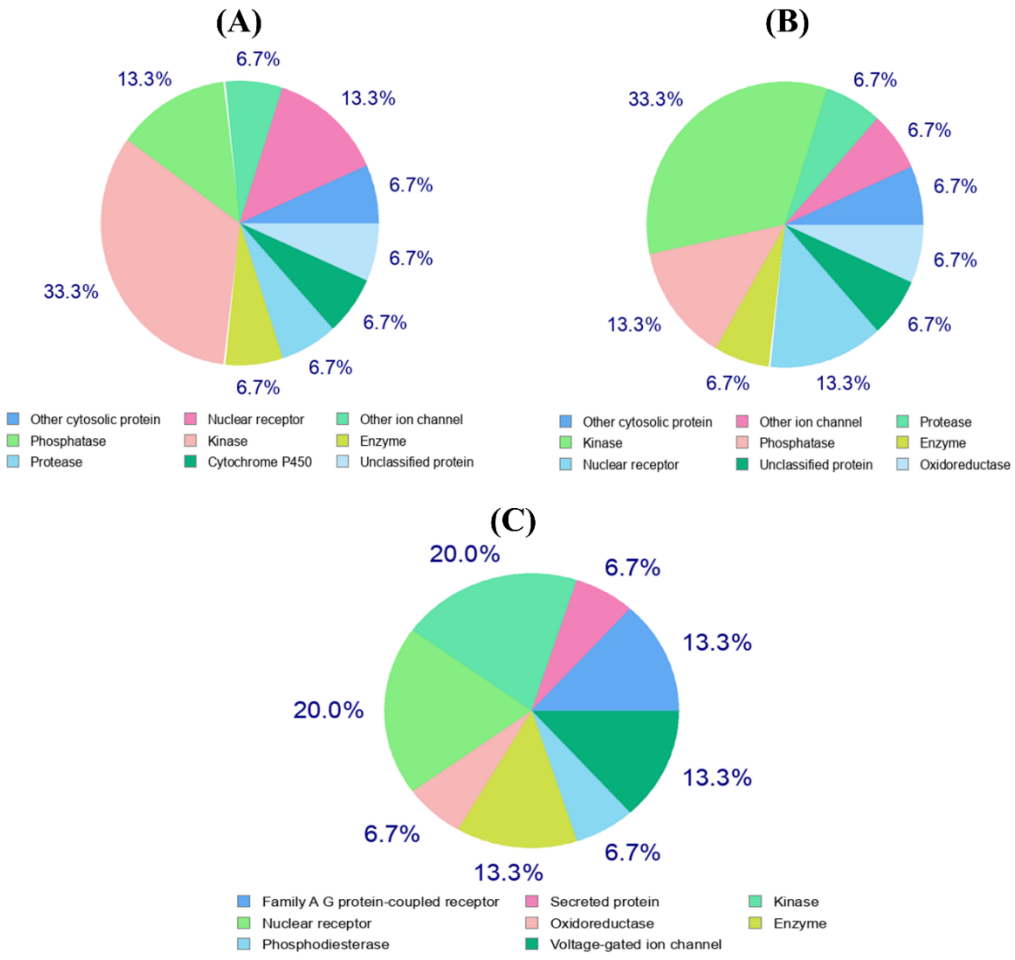


Figure B 7: Molecular target predictions for **(A)** dithranol, **(B)** danthron, and **(C)** CID156581477-ANLG obtained from Swiss target prediction report. The frequency of the target classes is depicted in the pie chart.

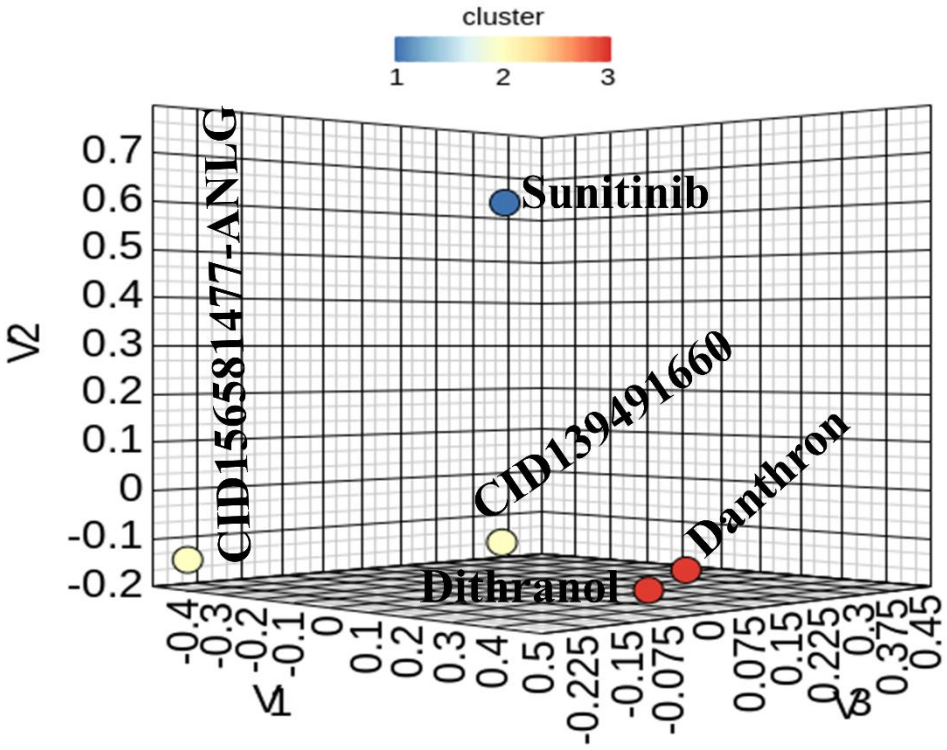


Figure B8: Compound similarity analysis using the ChemMine Tools that utilize Multidimensional Scaling (MDS) to evaluate the structural and physicochemical similarity.

Appendix C

Table C1: Average Root-mean-square deviation (RMSD) for each GaMD production simulation. The standard deviation values are reported in parentheses.

Systems	Run1 (RMSD (Å))	Run2 (RMSD (Å))	Run3 (RMSD (Å))	Run4 (RMSD (Å))
Apo	3.51 (0.41)	2.95 (0.35)	2.60 (0.29)	2.46 (0.25)
T1-I	2.71 (0.31)	2.78 (0.29)	2.69 (0.18)	2.62 (0.19)

T2-I	2.94 (0.25)	3.19 (0.25)	2.98 (0.18)	3.25 (0.25)
A-I	3.55 (0.27)	3.69 (0.27)	3.51 (0.31)	3.52 (0.22)

Table C2: Number of nodes, links, and hubs involved in the formation of various communities, identified through the protein structure network (PSN) analysis.

Communities	**Apo**	**T1-I**	**T2-I**	**A-I**
Community-1	11 nodes 16 links 6 hubs	30 nodes 44 links 17 hubs	13 nodes 24 links 10 hubs	18 nodes 26 links 8 hubs
Community-2	9 nodes 11 links 2 hubs	10 nodes 13 links 3 hubs	9 nodes 11 links 4 hubs	14 nodes 21 links 8 hubs
Community-3	6 nodes 9 links 2 hubs	3 nodes 3 links 1 hubs	9 nodes 11 links 4 hubs	5 nodes 9 links 4 hubs
Community-4	4 nodes 6 links 2 hubs	3 nodes 3 links 1 hubs	7 nodes 9 links 2 hubs	4 nodes 5 links 1 hubs
Community-5	4 nodes 5 links 0 hubs	3 nodes 3 links 1 hubs	6 nodes 7 links 3 hubs	4 nodes 5 links 2 hubs
Community-6	4 nodes 5 links 0 hubs	3 nodes 3 links 0 hubs	6 nodes 8 links 2 hubs	4 nodes 5 links 1 hubs
Community-7	3 nodes 3 links 0 hubs	3 nodes 3 links 2 hubs	4 nodes 5 links 0 hubs	3 nodes 3 links 1 hubs
	3 nodes		3 nodes	3 nodes

Community-8	3 links 2 hubs		3 links 1 hubs	3 links 2 hubs
Community-9	3 nodes 3 links 2 hubs		3 nodes 3 links 0 hubs	3 nodes 3 links 1 hubs
Community-10	3 nodes 3 links 2 hubs		3 nodes 3 links 2 hubs	3 nodes 3 links 2 hubs
Community-11	3 nodes 3 links 1 hubs			3 nodes 3 links 1 hubs
Community-12				3 nodes 3 links 1 hubs
Community-13				3 nodes 3 links 1 hubs

Table C3: Comparison of network properties between the T1-I and T2-I systems.

Network Properties	**T1-I**	**T2-I**
Number of Linked Nodes	273	274
Number of Specific Nodes	10 (3.66%)	11 (4.01%)
Number of Shared Nodes	263 (96.34%)	263 (95.99%)
Number of Links	301	308
Number of Specific Links	73 (24.25%)	80 (25.97%)

Number of Shared Links	228 (75.75%)	228 (74.03%)
Number of Hubs	35	39
Number of Specific Hubs	15 (42.86%)	19 (48.72%)
Number of Shared Hubs	20 (57.14%)	20 (51.28%)
Average % shared Neighbours (Jaccard)	62.71	
Average % shared Neighbours (Otsuka)	69.82	
Average % shared Neighbours (Overlap Coefficient)	78.79	
Average % shared Cliques (K3-6)	56.82	
Graphlets Similarity	66.61	

Table C4: Comparison of network properties between the T1-I and A-I systems.

Network Properties	**T1-I**	**A-I**
Number of Linked Nodes	273	280
Number of Specific Nodes	9 (3.30%)	16 (5.71%)
Number of Shared Nodes	264 (96.70%)	264 (94.29%)
Number of Links	301	316
Number of Specific Links	90 (29.90%)	105 (33.23%)
Number of Shared Links	211 (70.10%)	211 (66.77%)
Number of Hubs	35	43
Number of Specific Hubs	17	25

	(48.57%)	(58.14%)
Number of Shared Hubs	18 (51.43%)	18 (41.86%)
Average % shared Neighbours (Jaccard)	55.23	
Average % shared Neighbours (Otsuka)	63.97	
Average % shared Neighbours (Overlap Coefficient)	73.64	
Average % shared Cliques (K3-6)	55.10	
Graphlets Similarity	65.30	

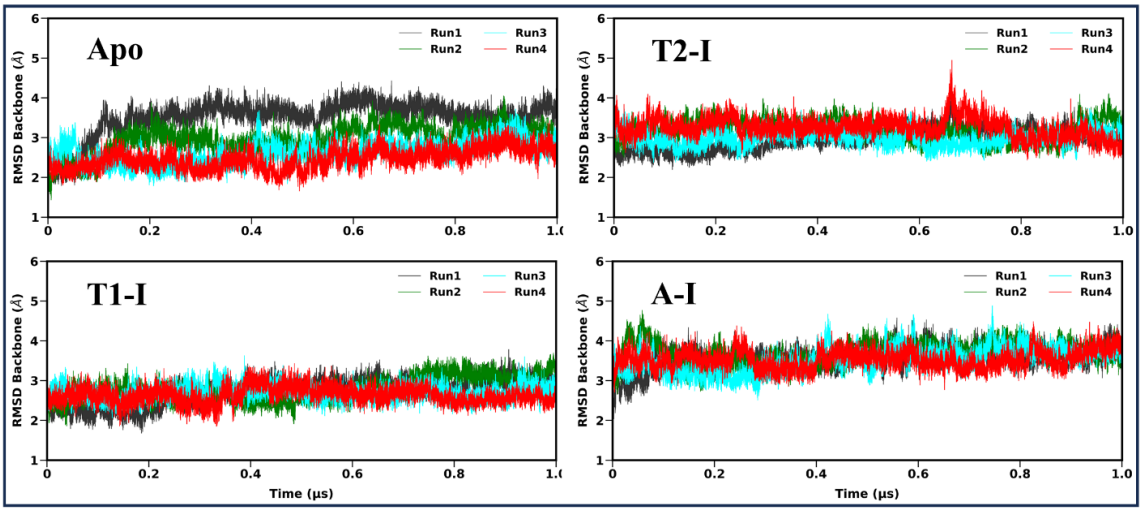


Figure C1: Time evolution of backbone root-mean-square deviation (RMSD) for the MKK7 kinase domain across four systems from four replica 1 μs long GaMD simulations.

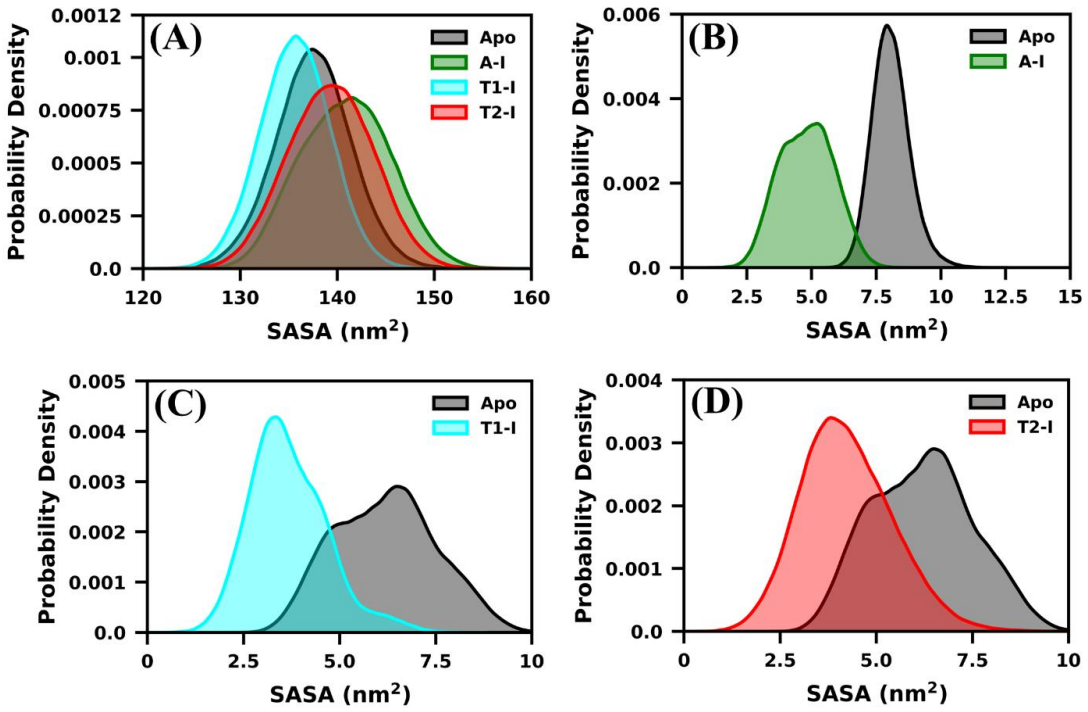


Figure C2: Probability density analysis depicting the solvent accessible surface area (SASA) for **(A)** the entire protein, **(B)** the allosteric binding region of Apo and A-I, **(C)** the Type I inhibitor binding pocket for Apo and T1-I, and **(D)** the Type II inhibitor binding pocket for Apo and T2-I systems.

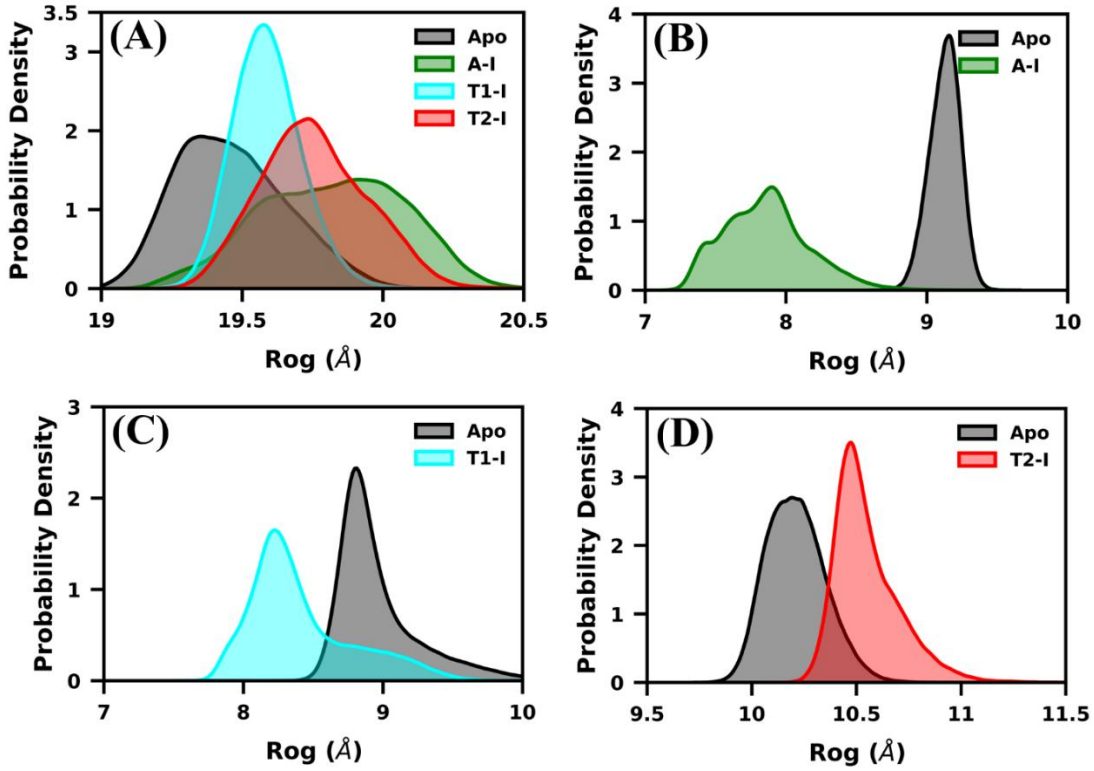


Figure C3: Probability density analysis depicting the radius of gyration (RoG) for **(A)** the entire protein, **(B)** the allosteric binding region of Apo and A-I, **(C)** the Type I inhibitor binding pocket for Apo and T1-I, and **(D)** the Type II inhibitor binding pocket for Apo and T2-I systems.

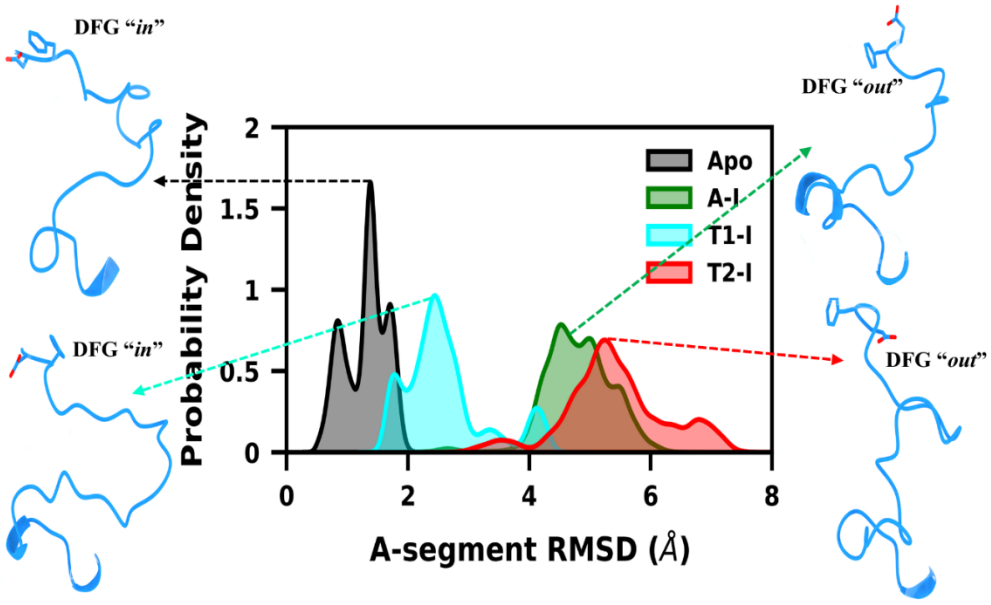


Figure C4: The backbone RMSD of the activation segment represented via a probability density graph. Several structures corresponding to the peaks are highlighted, depicting the backbone and side-chain orientations of the loop and the DFG motif, respectively.

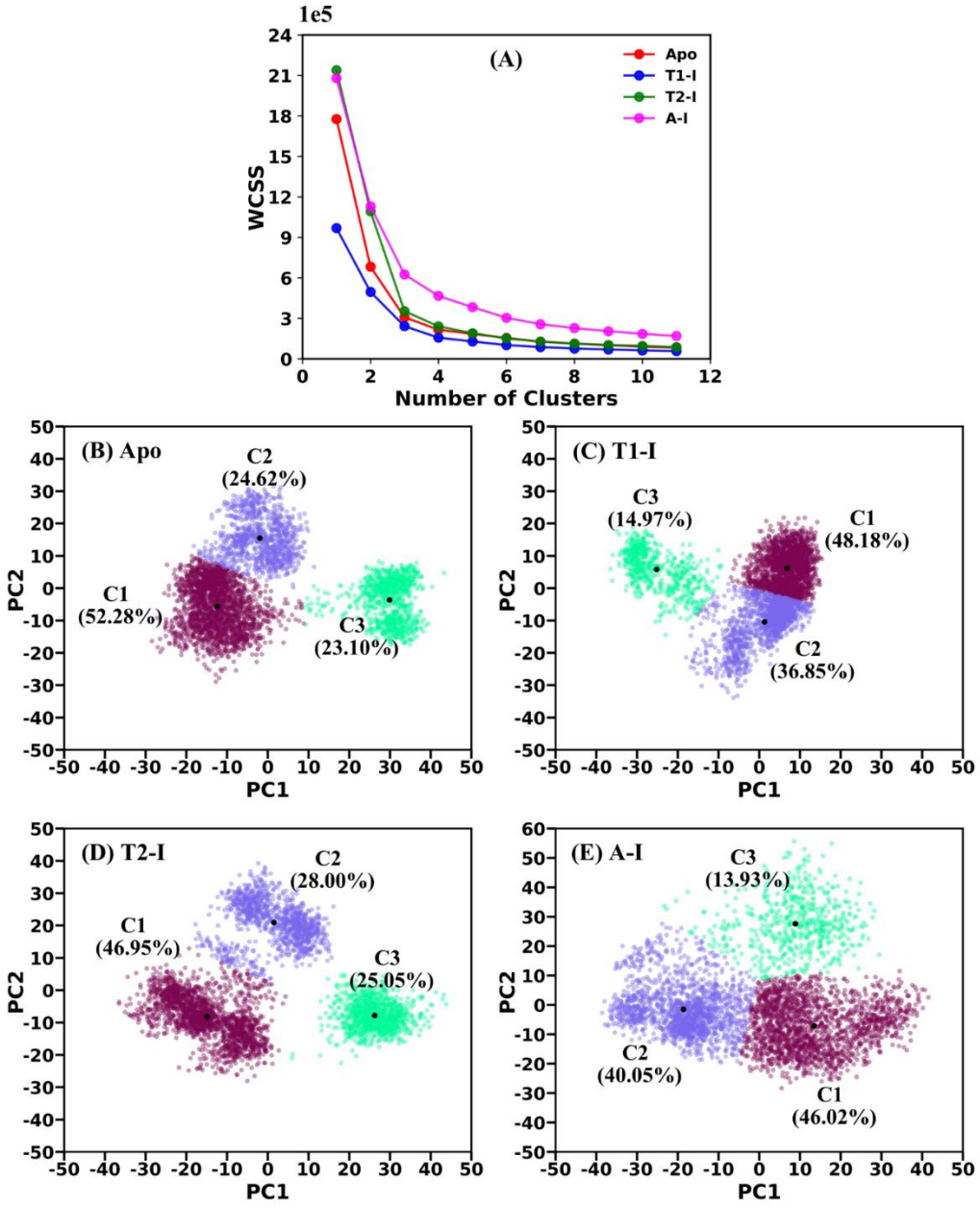


Figure C5: (A) A within-cluster sum of squares plot illustrating the four systems based on *K*-means clustering, using the first and second principal modes (PC1 and PC2) from PCA, with elbow points identifying the optimal cluster count. **(B)-(E)** Clustered scattered plots based on PC1 vs. PC2, displaying population percentages (rounded to two decimal places) for each respective cluster. The black dots represent the centroids of the clusters.

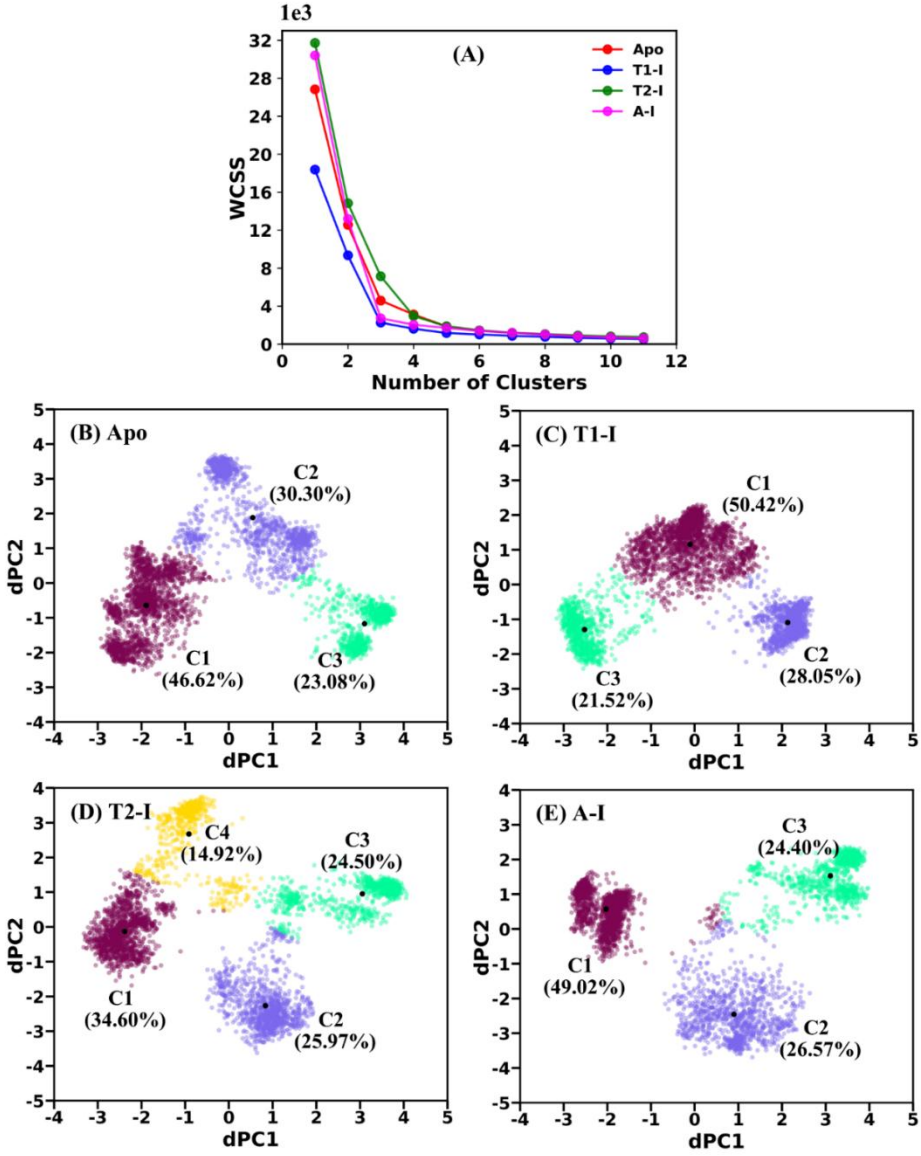


Figure C6: (A) A within-cluster sum of squares plot illustrating the four systems based on *K*-means clustering, using the first and second principal modes (dPC1 and dPC2) from dPCA of the activation segment, with elbow points identifying the optimal cluster count. **(B)-(E)** Clustered scattered plots based on dPC1 vs. dPC2, displaying population percentages (rounded to two decimal places) for each respective cluster. The black dots represent the centroids of the clusters.

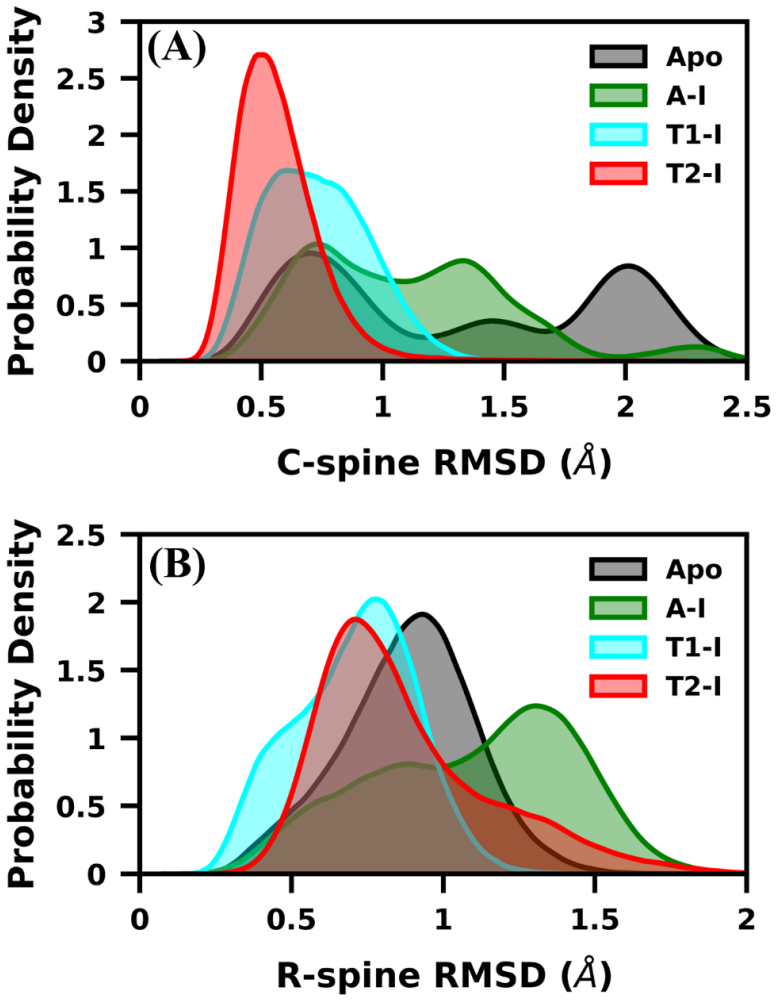


Figure C7: Probability density plots depicting the RMSD distribution for **(A)** C-spine and **(B)** R-spine of the MKK7 kinase domain.

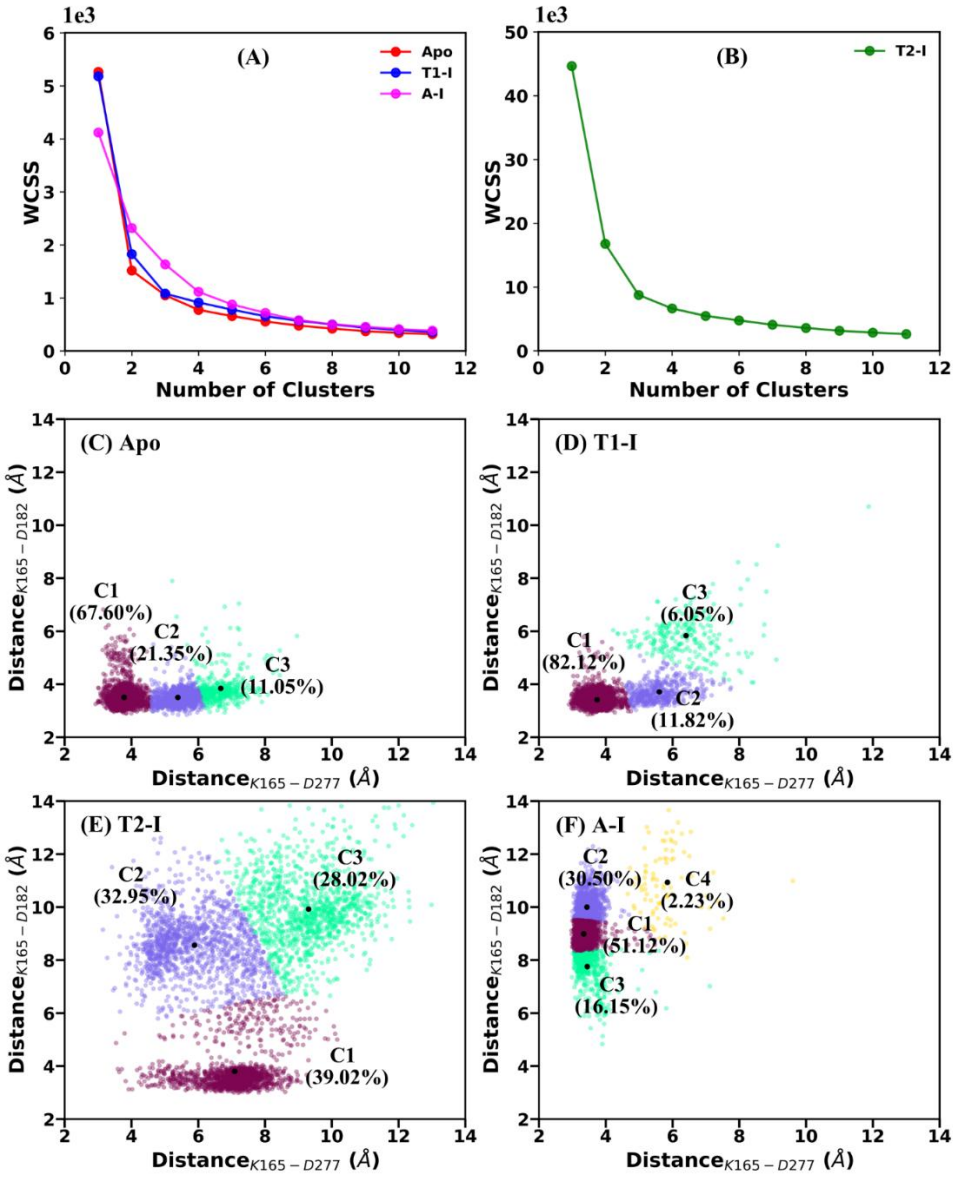


Figure C8: (A)-(B) A within-cluster sum of squares plot illustrating the four systems based on *K*-means clustering, using the Lys^{165}-Asp^{277} and Lys^{165}-Asp^{182} distances, with elbow points identifying the optimal cluster count. **(C)-(F)** Clustered scattered plots based on Lys^{165}-Asp^{277} vs. Lys^{165}-Asp^{182} distances, displaying population percentages (rounded to two decimal places) for each respective cluster. The black dots represent the centroids of the clusters.

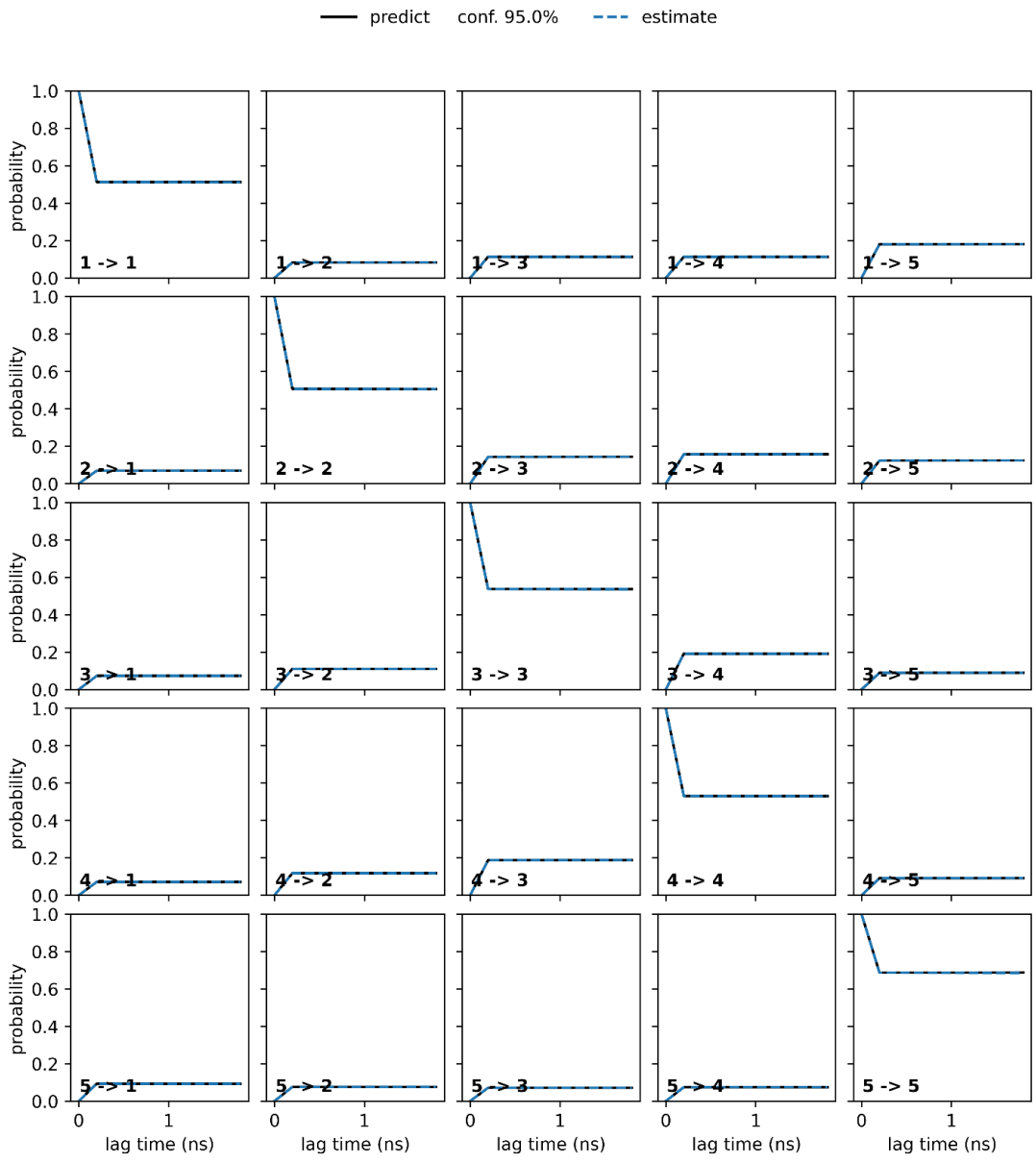


Figure C9: Validation of MSM for apo system through C-K test. Shown are the data from MSM (black line) and the observed trajectory (blue dotted line with estimated error).

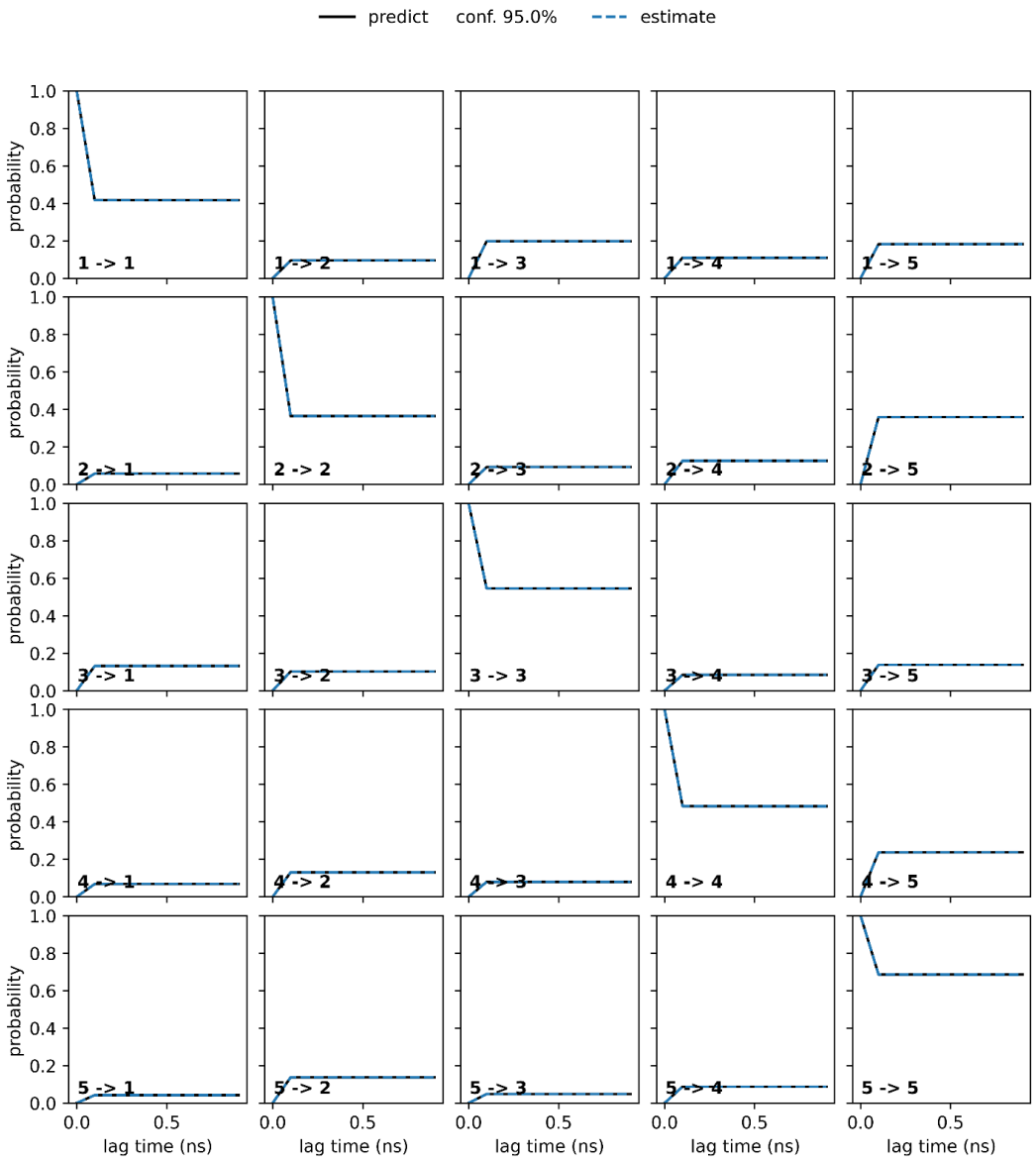


Figure C10: Validation of MSM for A-I system through C-K test. Shown are the data from MSM (black line) and the observed trajectory (blue dotted line with estimated error).

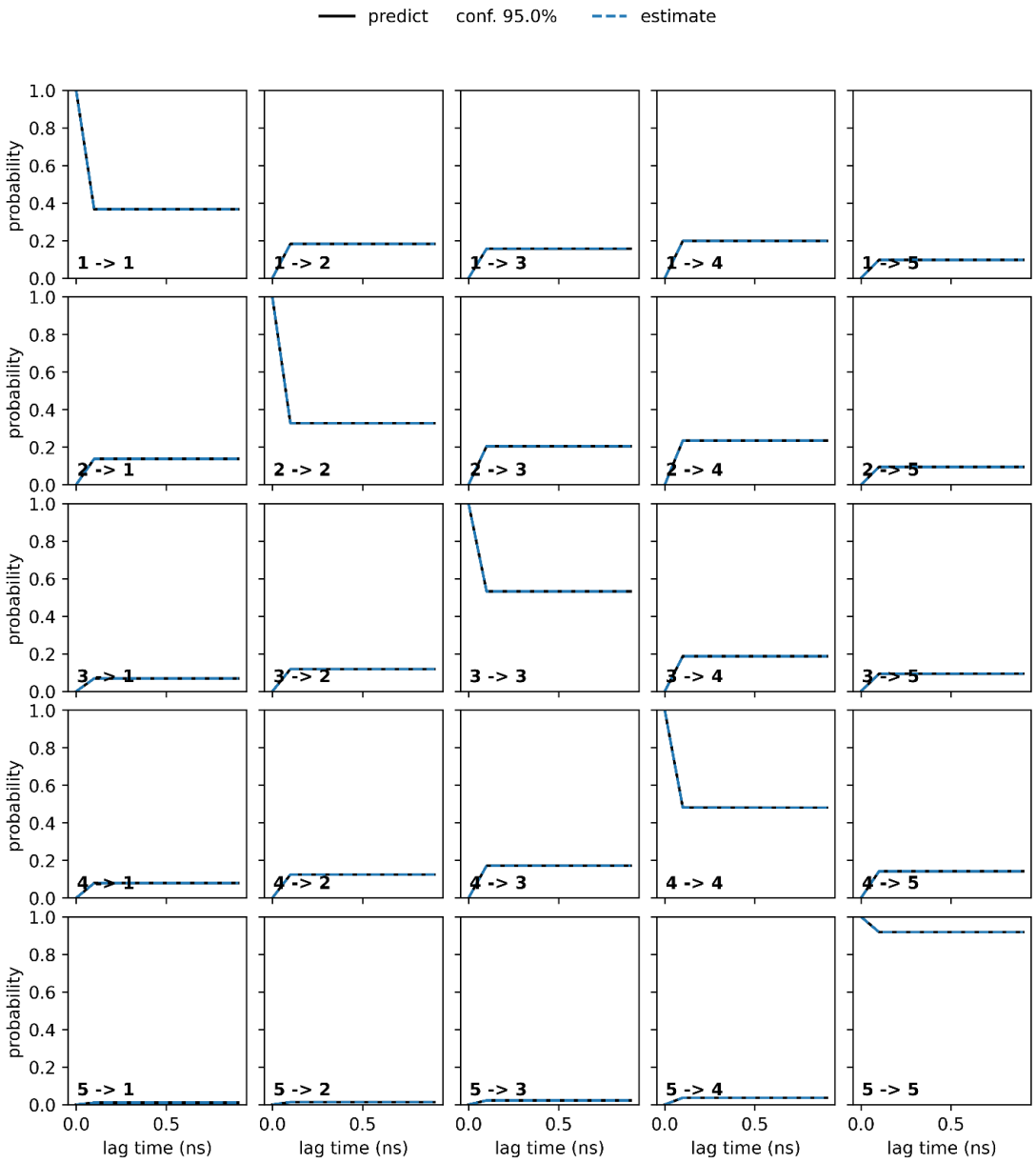


Figure C11: Validation of MSM for T1-I system through C-K test. Shown are the data from MSM (black line) and the observed trajectory (blue dotted line with estimated error).

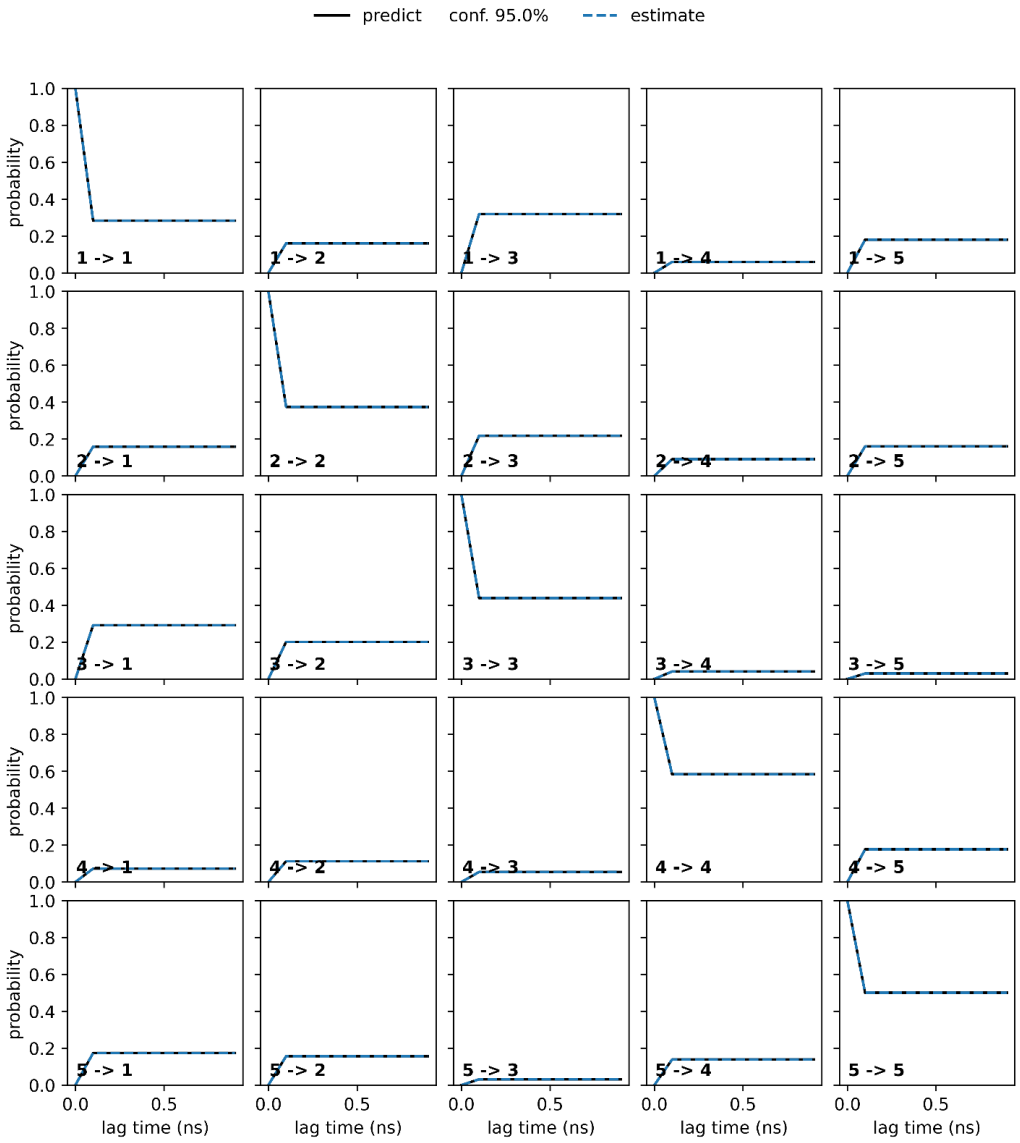


Figure C12: Validation of MSM for T2-I system through C-K test. Shown are the data from MSM (black line) and the observed trajectory (blue dotted line with estimated error).

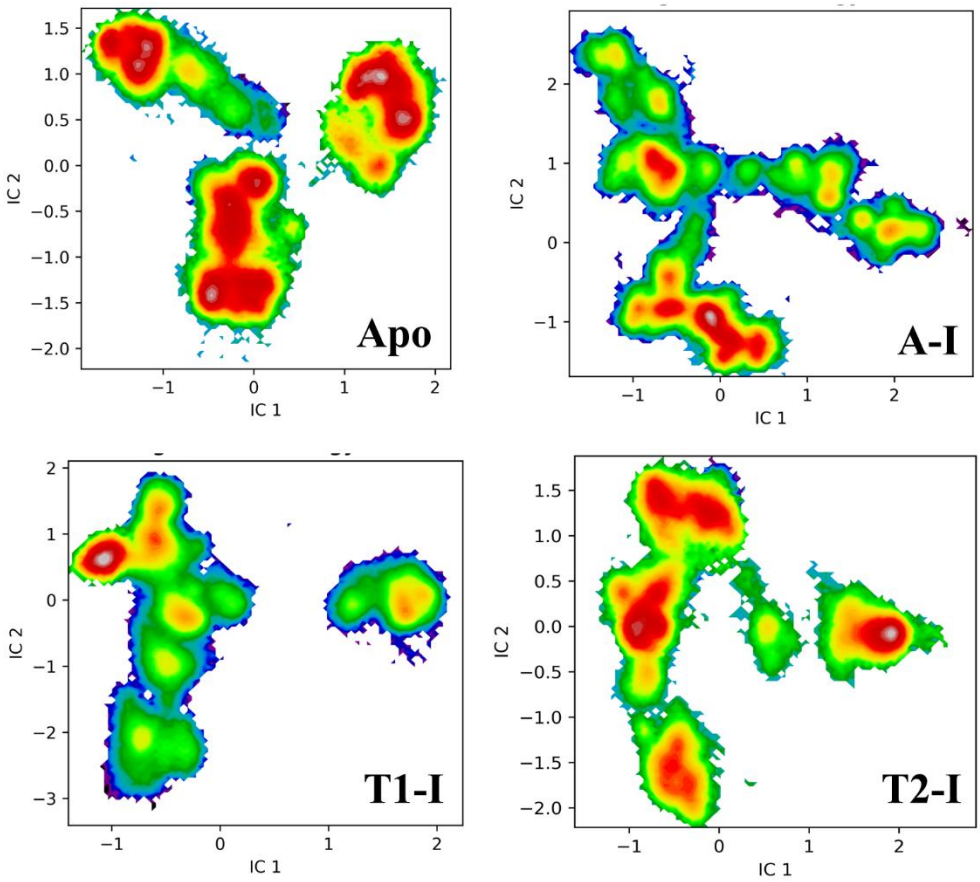


Figure C13: The MD trajectory frames projected onto the two slowest ICs, represented as free energy surface (FES).

Appendix D

Table D1: Number of nodes, links, and hubs involved in the formation of various communities, identified through the protein structure network (PSN) analysis.

Communities	JNK1-p	JNK1-u	JNK2-p	JNK2-u	JNK3-p	JNK3-u
Community-1	18 nodes 27 links 9 hubs	16 nodes 22 links 7 hubs	9 nodes 14 links 4 hubs	9 nodes 14 links 5 hubs	20 nodes 30 links 10 hubs	21 nodes 33 links 11 hubs
Community-2	13 nodes 20 links 9 hubs	11 nodes 16 links 6 hubs	6 nodes 9 links 3 hubs	8 nodes 12 links 6 hubs	16 nodes 24 links 12 hubs	20 nodes 29 links 12 hubs
Community-3	5 nodes 6 links 1 hubs	6 nodes 7 links 5 hubs	6 nodes 9 links 5 hubs	8 nodes 11 links 5 hubs	3 nodes 3 links 1 hubs	4 nodes 5 links 0 hubs
Community-4	5 nodes 9 links 4 hubs	4 nodes 6 links 2 hubs	5 nodes 6 links 2 hubs	7 nodes 11 links 4 hubs	3 nodes 3 links 2 hubs	3 nodes 3 links 1 hubs
Community-5	3 nodes 3 links 1 hubs	3 nodes 3 links 1 hubs	4 nodes 5 links 0 hubs	6 nodes 8 links 4 hubs	3 nodes 3 links 1 hubs	3 nodes 3 links 1 hubs
Community-6	3 nodes 3 links 1 hubs	3 nodes 3 links 3 hubs	4 nodes 5 links 0 hubs	6 nodes 8 links 1 hubs	3 nodes 3 links 0 hubs	3 nodes 3 links 1 hubs
Community-7	3 nodes 3 links 2 hubs		3 nodes 3 links 0 hubs	3 nodes 3 links 0 hubs	3 nodes 3 links 0 hubs	3 nodes 3 links 0 hubs

Community-8	3 nodes 3 links 1 hubs		3 nodes 3 links 1 hubs	3 nodes 3 links 1 hubs		
Community-9	3 nodes 3 links 0 hubs		3 nodes 3 links 2 hubs	3 nodes 3 links 1 hubs		
Community-10	3 nodes 3 links 1 hubs		3 nodes 3 links 1 hubs	3 nodes 3 links 2 hubs		
Community-11			3 nodes 3 links 2 hubs	3 nodes 3 links 2 hubs		
Community-12			3 nodes 3 links 1 hubs	3 nodes 3 links 0 hubs		
Community-13			3 nodes 3 links 1 hubs	3 nodes 3 links 1 hubs		

Table D2: Comparison of network properties between the JNK1-p and JNK1-u systems.

Network Properties	**JNK1-p**	**JNK1-u**
Number of Linked Nodes	330	329
Number of Specific Nodes	8 (2.42%)	7 (2.13%)
Number of Shared Nodes	322 (97.58%)	322 (97.87%)
Number of Links	354	367

Number of Specific Links	59 (16.67%)	72 (19.62%)
Number of Shared Links	295 (83.33%)	295 (80.38%)
Number of Hubs	37	44
Number of Specific Hubs	10 (27.03%)	17 (38.64%)
Number of Shared Hubs	27 (72.97%)	27 (61.36%)
Average % shared Neighbours (Jaccard)	72.84	
Average % shared Neighbours (Otsuka)	79.93	
Average % shared Neighbours (Overlap Coefficient)	87.80	
Average % shared Cliques (K3-6)	72.64	
Graphlets Similarity	60.27	

Table D3: Comparison of network properties between the JNK2-p and JNK2-u systems.

Network Properties	**JNK2-p**	**JNK2-u**
Number of Linked Nodes	323	327
Number of Specific Nodes	4 (1.24%)	8 (2.45%)
Number of Shared Nodes	319 (98.76%)	319 (97.55%)
Number of Links	362	350
Number of Specific Links	57 (15.75%)	45 (12.86%)
Number of Shared Links	305 (84.25%)	305 (87.14%)

Number of Hubs	44	36
Number of Specific Hubs	15 (34.09%)	7 (19.44%)
Number of Shared Hubs	29 (65.91%)	29 (80.56%)
Average % shared Neighbours (Jaccard)	77.40	
Average % shared Neighbours (Otsuka)	83.24	
Average % shared Neighbours (Overlap Coefficient)	90.08	
Average % shared Cliques (K3-6)	87.18	
Graphlets Similarity	66.02	

Table D4: Comparison of network properties between the JNK3-p and JNK3-u systems.

Network Properties	**JNK3-p**	**JNK3-u**
Number of Linked Nodes	327	324
Number of Specific Nodes	9 (2.75%)	6 (1.85%)
Number of Shared Nodes	318 (97.25%)	318 (98.15%)
Number of Links	368	358
Number of Specific Links	59 (16.03%)	49 (13.69%)
Number of Shared Links	309 (83.97%)	309 (86.31%)
Number of Hubs	44	41
Number of Specific Hubs	10 (22.73%)	7 (17.07%)

Number of Shared Hubs	34 (77.27%)	34 (82.93%)
Average % shared Neighbours (Jaccard)	75.47	
Average % shared Neighbours (Otsuka)	81.52	
Average % shared Neighbours (Overlap Coefficient)	88.26	
Average % shared Cliques (K3-6)	87.14	
Graphlets Similarity	64.44	

Table D5: Residual decomposition of the binding free energy in kcal/mol between JNK1/2/3 and SP600125. Binding free energy with values less then -1.0 kcal/mol are listed here.

Residue	E_{vdW} **(kcal/mol)**	E_{elec} **(kcal/mol)**	G_{pol} **(kcal/mol)**	G_{np} **(kcal/mol)**	G_{total} **(kcal/mol)**
JNK1-p					
GLU 109	0.11	-6.01	3.54	0	-2.35
VAL 158	-1.63	-0.27	0.28	0	-1.61
LEU 110	-1.10	-1.05	0.89	0	-1.26
ILE 32	-1.90	-0.63	1.34	0	-1.19
LEU 168	-1.25	-0.21	0.34	0	-1.12
VAL 40	-1.32	-0.09	0.29	0	-1.12
LYS 55	-0.39	0.53	0.99	0	1.13
JNK1-u					
GLU 109	0.11	-6.03	3.72	0	-2.19
VAL 158	-1.59	-0.25	0.24	0	-1.59
LEU 110	-1.11	-1.02	0.88	0	-1.25
LEU 168	-1.23	-0.22	0.37	0	-1.08
VAL 40	-1.24	-0.09	0.28	0	-1.05

LYS 55	-0.34	0.59	0.86	0	1.10
JNK2-p					
GLU 109	0.16	-6.17	3.59	0	-2.41
VAL 158	-1.58	-0.26	0.26	0	-1.57
LEU 110	-1.17	-1.09	0.87	0	-1.38
ILE 32	-1.98	-0.52	1.26	0	-1.23
LEU 168	-1.22	-0.27	0.41	0	-1.08
LYS 55	-0.39	0.40	1.46	0	1.48
JNK2-u					
GLU 109	0.10	-6.09	3.60	0	-2.39
VAL 158	-1.61	-0.25	0.27	0	-1.59
LEU 110	-1.13	-1.04	0.84	0	-1.33
LEU 168	-1.22	-0.22	0.34	0	-1.10
ILE 32	-1.75	-0.34	1.00	0	-1.09
VAL 40	-1.23	-0.082	0.27	0	-1.03
LYS 55	-0.44	0.50	1.40	0	1.46
JNK3-p					
GLU 147	0.08	-5.93	3.57	0	-2.27
VAL 196	-1.62	-0.27	0.29	0	-1.60
LEU 148	-1.08	-1.01	0.84	0	-1.25
ILE 70	-1.93	-0.46	1.27	0	-1.12
VAL 78	-1.26	-0.06	0.22	0	-1.10
LEU 206	-1.14	-0.19	0.29	0	-1.05
JNK3-u					
GLU 147	0.14	-6.08	3.75	0	-2.18
VAL 196	-1.62	-0.26	0.25	0	-1.63
LEU 148	-1.06	-0.98	0.78	0	-1.26
LEU 206	-1.20	-0.19	0.27	0	-1.12
VAL 78	-1.39	-0.13	0.33	0	-1.19

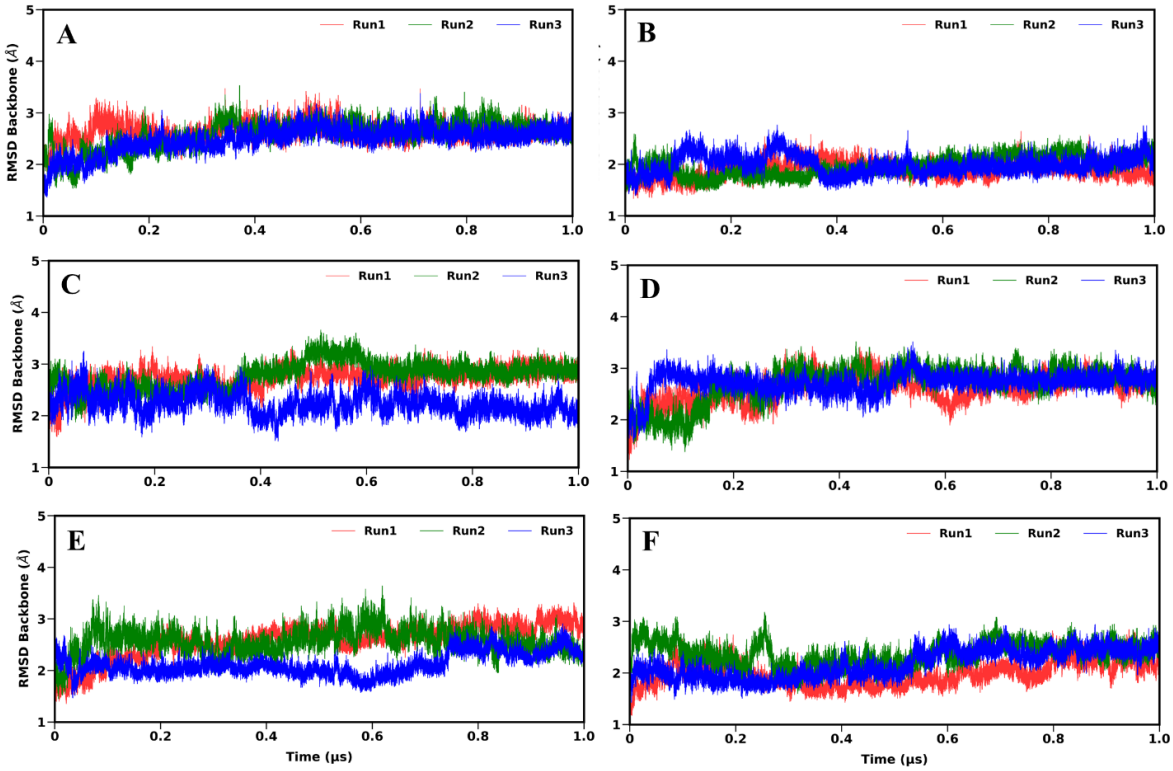


Figure D1: Time evolution of root-mean-square deviation (RMSD) of backbone atoms of whole protein. **(A)** JNK1-p, **(B)** JNK1-u, **(C)** JNK2-p, **(D)** JNK2-u, **(E)** JNK3-p, **(F)** JNK3-u.

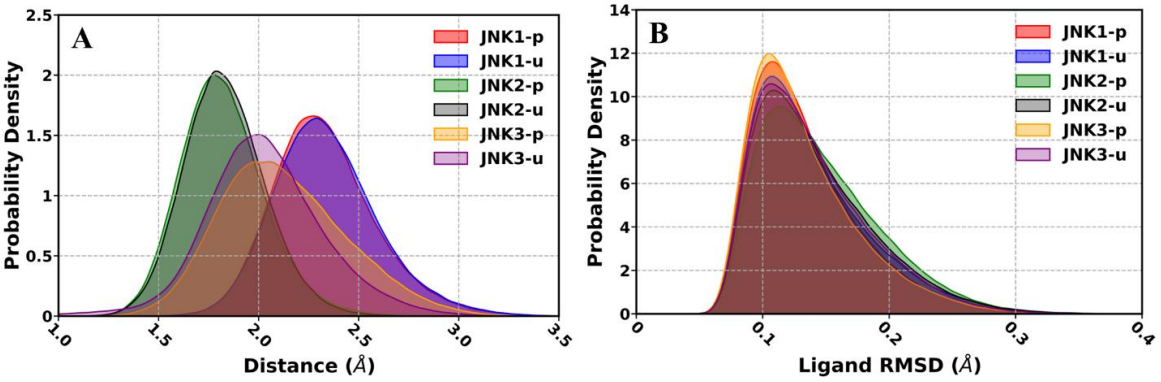


Figure D2: (A) Measurement of the COM distance between the ligand and amino acid residues located within 5Å. **(B)** The probability distribution of ligand RMSD.

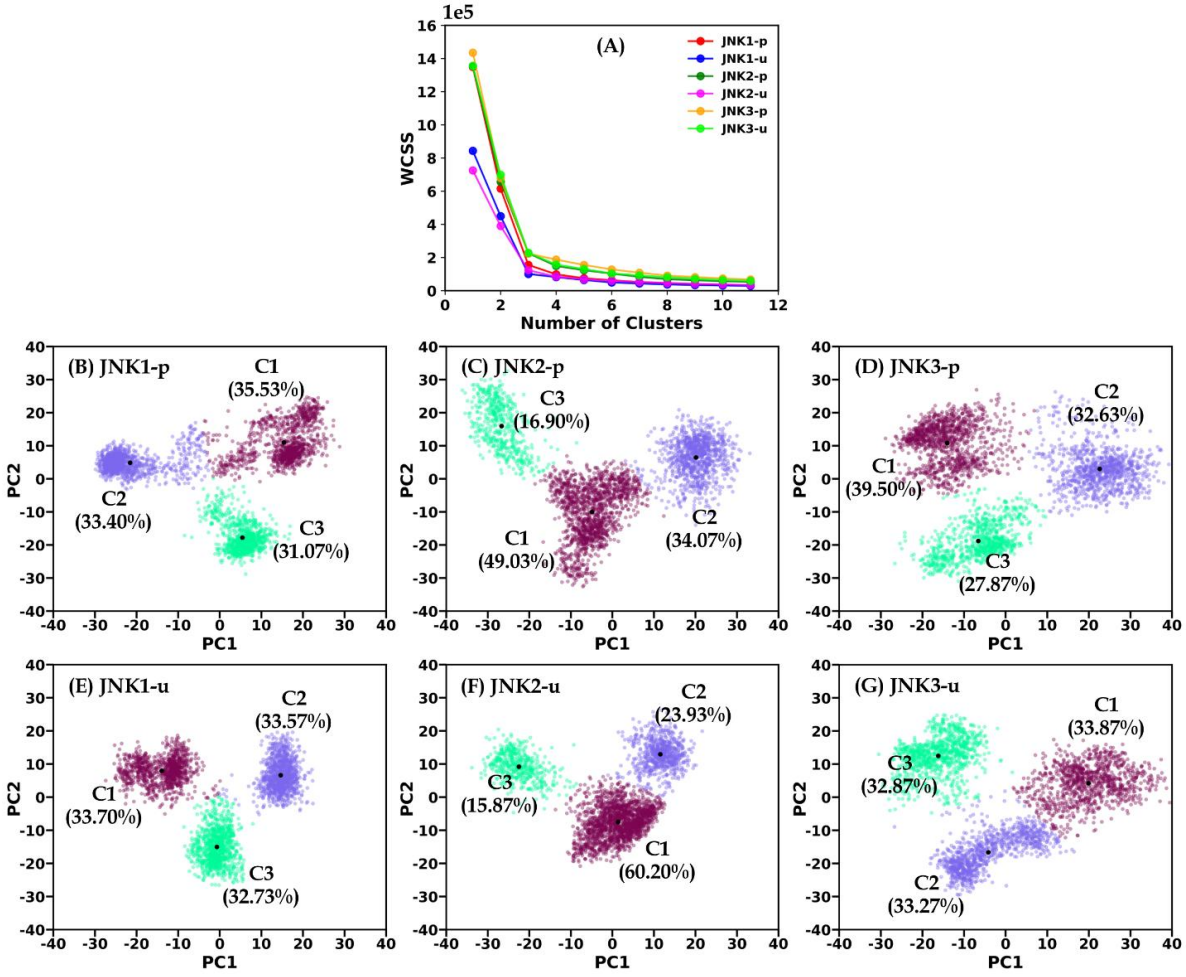


Figure D3: (A) A within-cluster sum of squares plot illustrating the four systems based on K-means clustering, using the first and second principal modes (PC1 and PC2) from PCA, with elbow points identifying the optimal cluster count. **(B)-(G)** Clustered scattered plots based on PC1 vs. PC2, displaying population percentages (rounded to two decimal places) for each respective cluster. The black dots represent the centroids of the clusters.

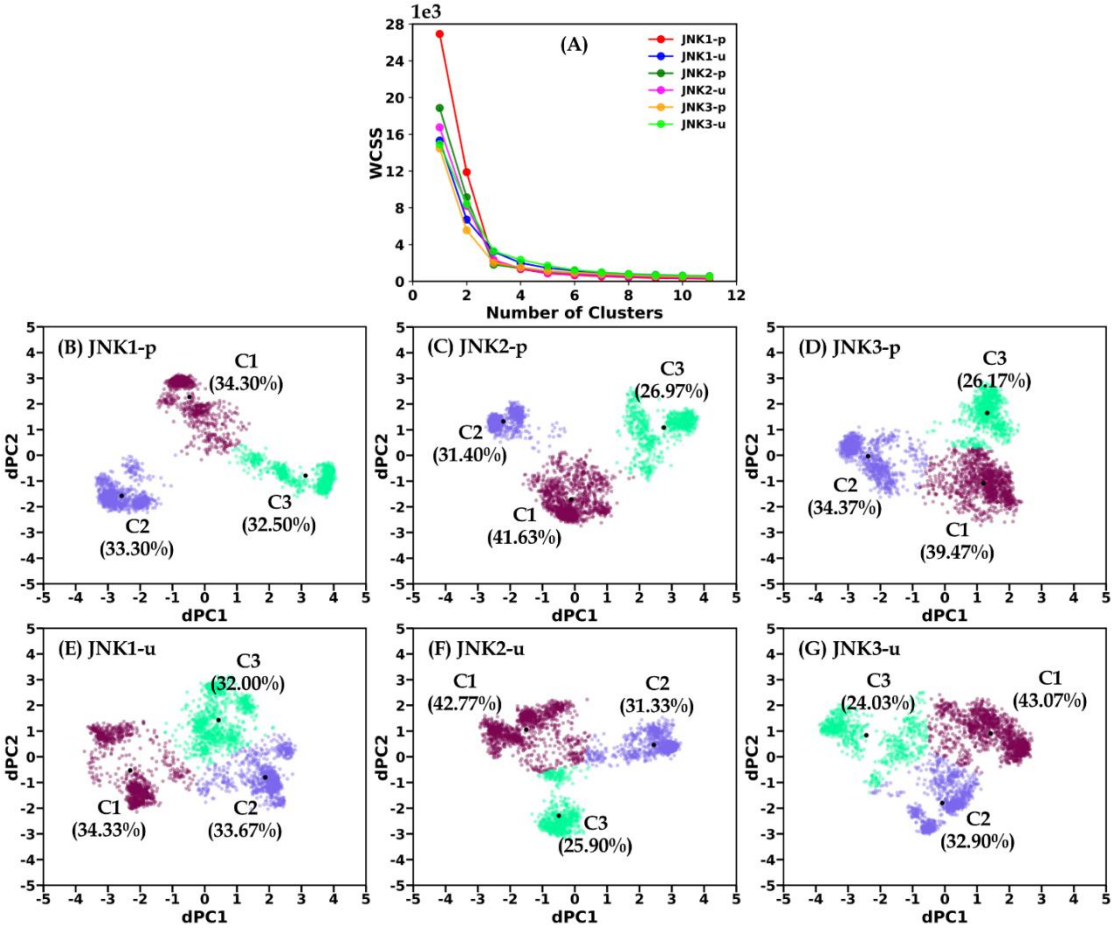


Figure D4: (A) A within-cluster sum of squares plot illustrating the four systems based on K-means clustering, using the first and second principal modes (dPC1 and dPC2) from dPCA of the activation segment, with elbow points identifying the optimal cluster count. **(B)-(G)** Clustered scattered plots based on dPC1 vs. dPC2, displaying population percentages (rounded to two decimal places) for each respective cluster. The black dots represent the centroids of the clusters.

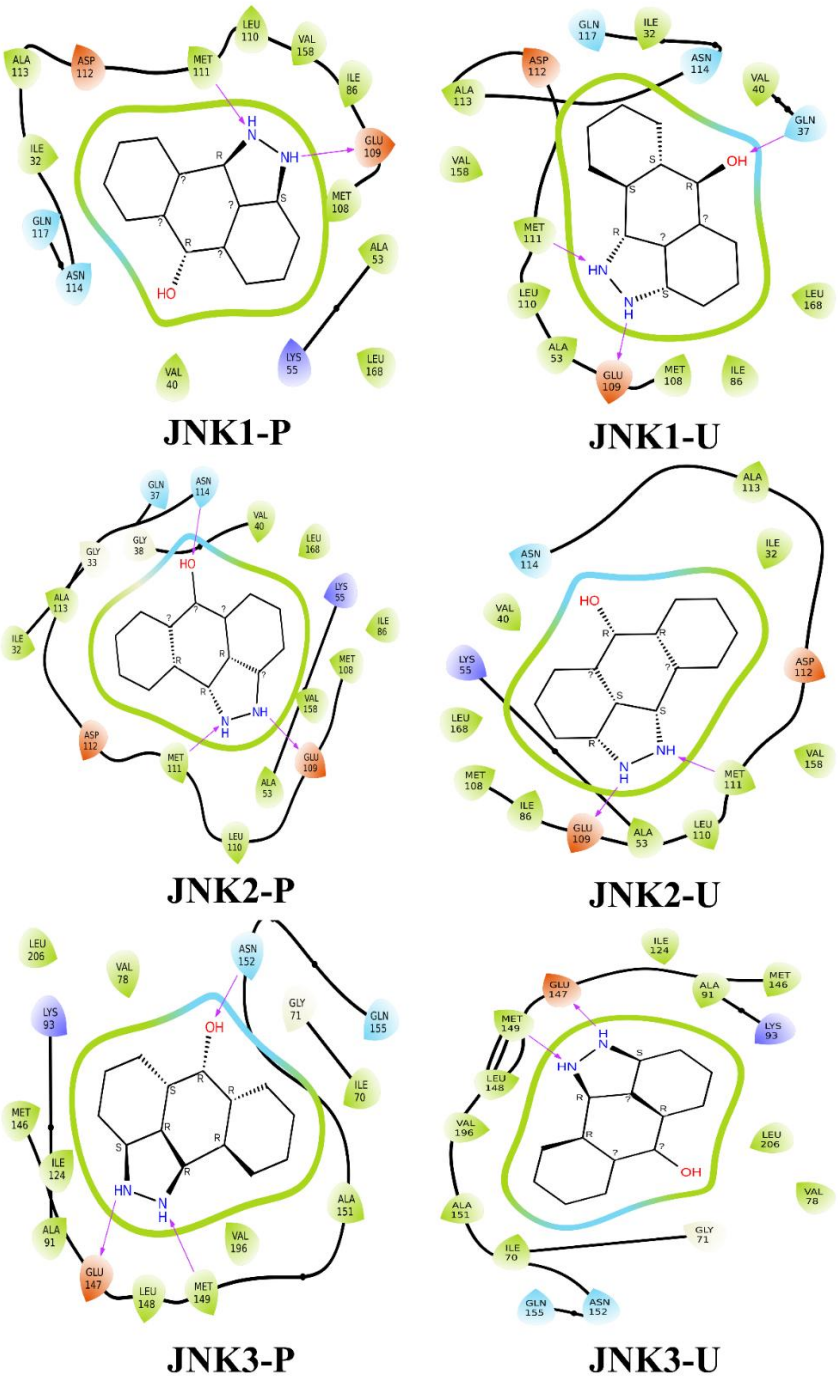


Figure D5: 2D ligand interaction diagrams for phosphorylated and unphosphorylated JNK isoforms generated using Schrödinger. Key interacting residues, hydrogen bonds, and hydrophobic interactions are highlighted to compare ligand binding differences across JNK1, JNK2, and JNK3.

Appendix E

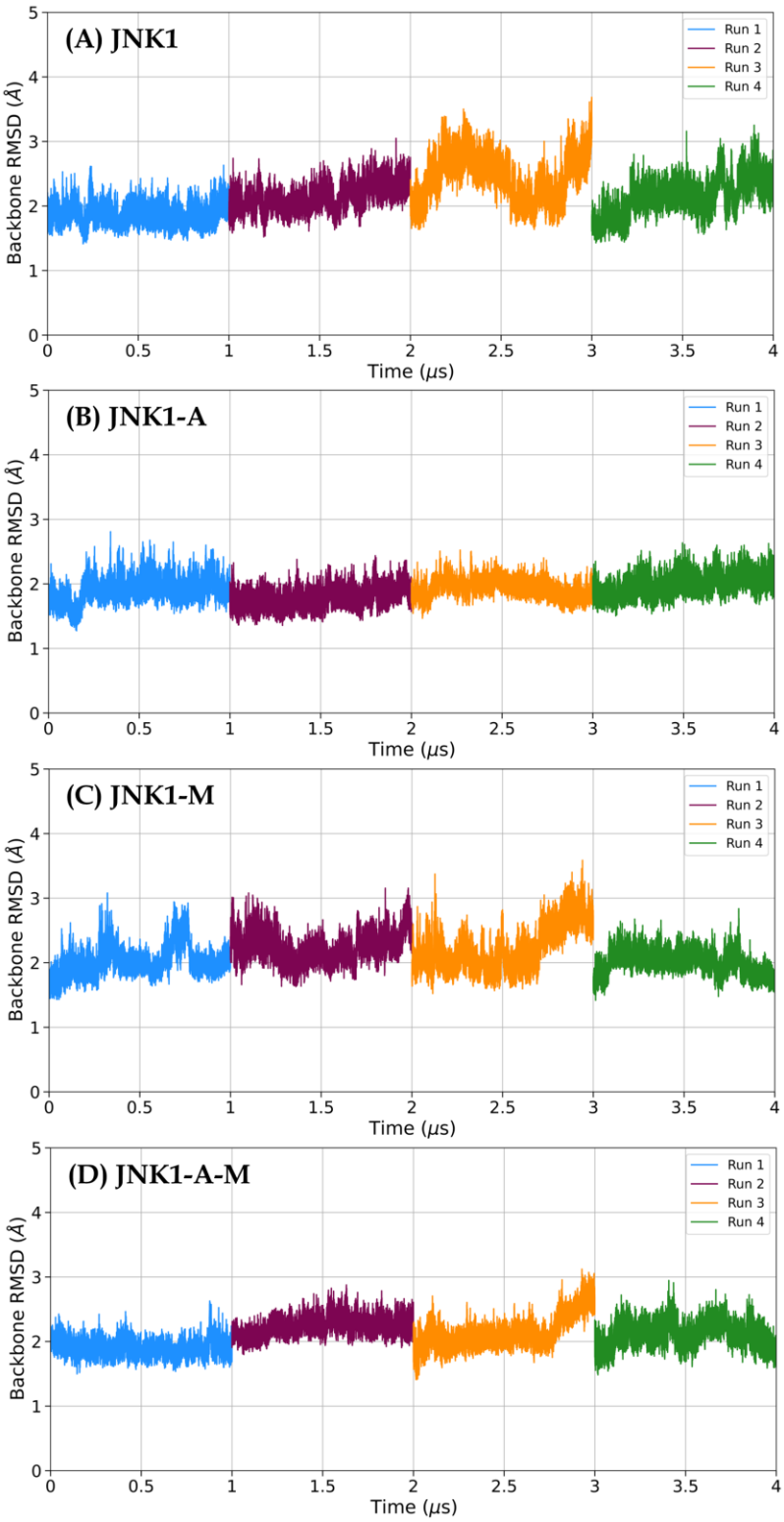


Figure E1: Time evolution of backbone RMSD in four systems from 1 μs MD replicas.

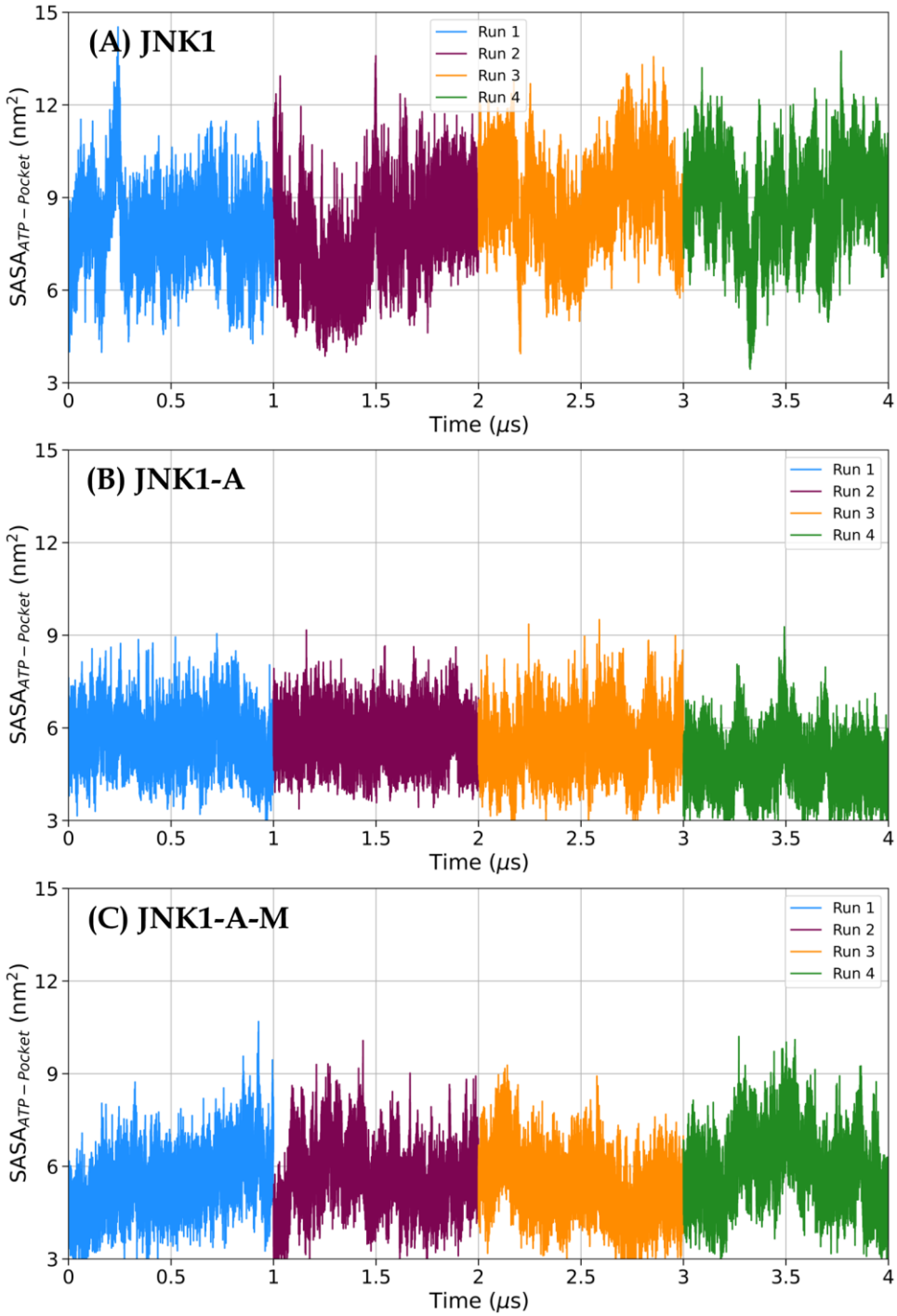


Figure E2: Time evolution of solvent-accessible surface area of the ATP-binding site in apo and the two ATP-bound complex systems from 1 µs MD replicas.

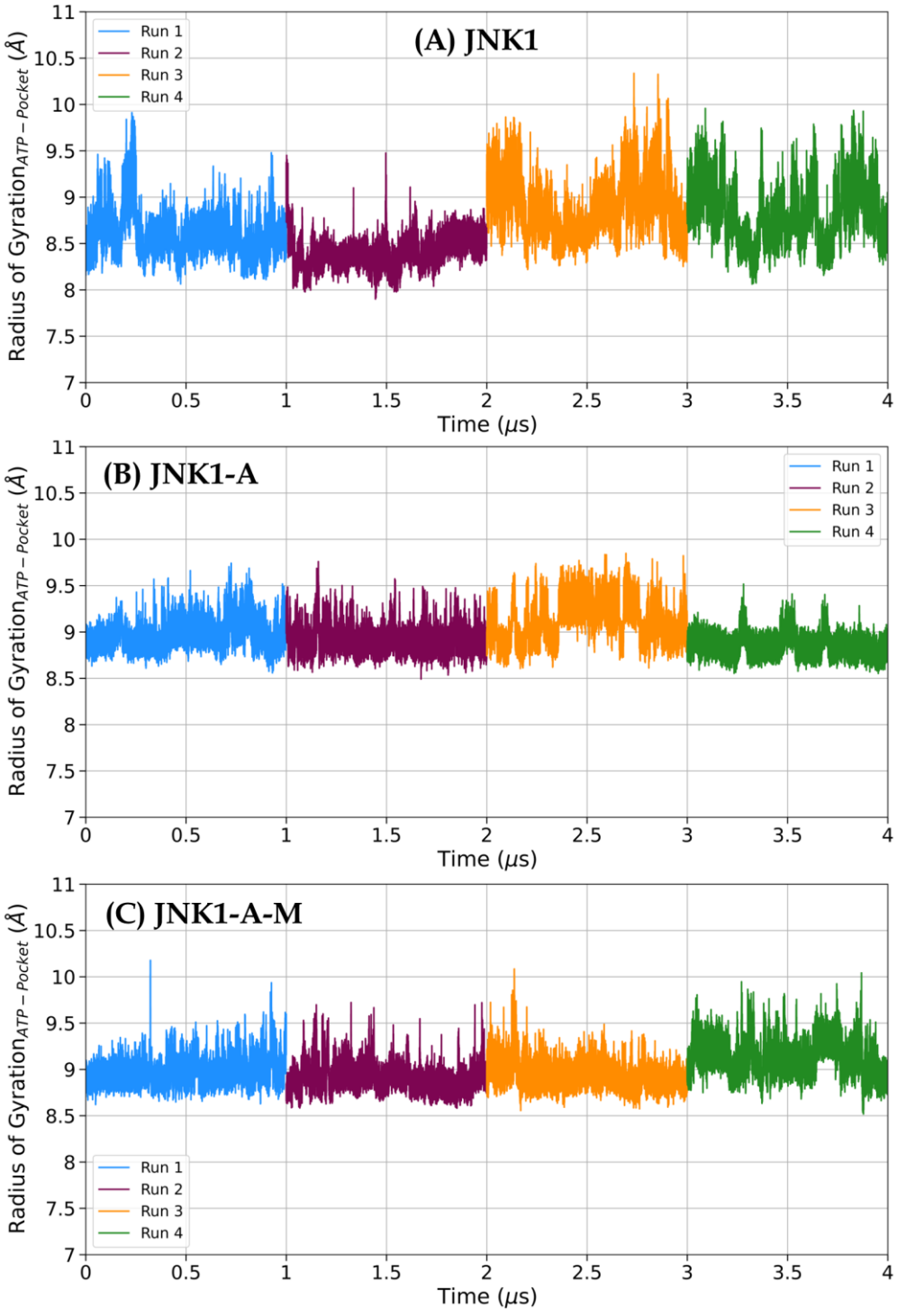


Figure E3: Time evolution of radius of gyration of the ATP-binding site in apo and the two ATP-bound complex systems from 1 µs MD replicas.

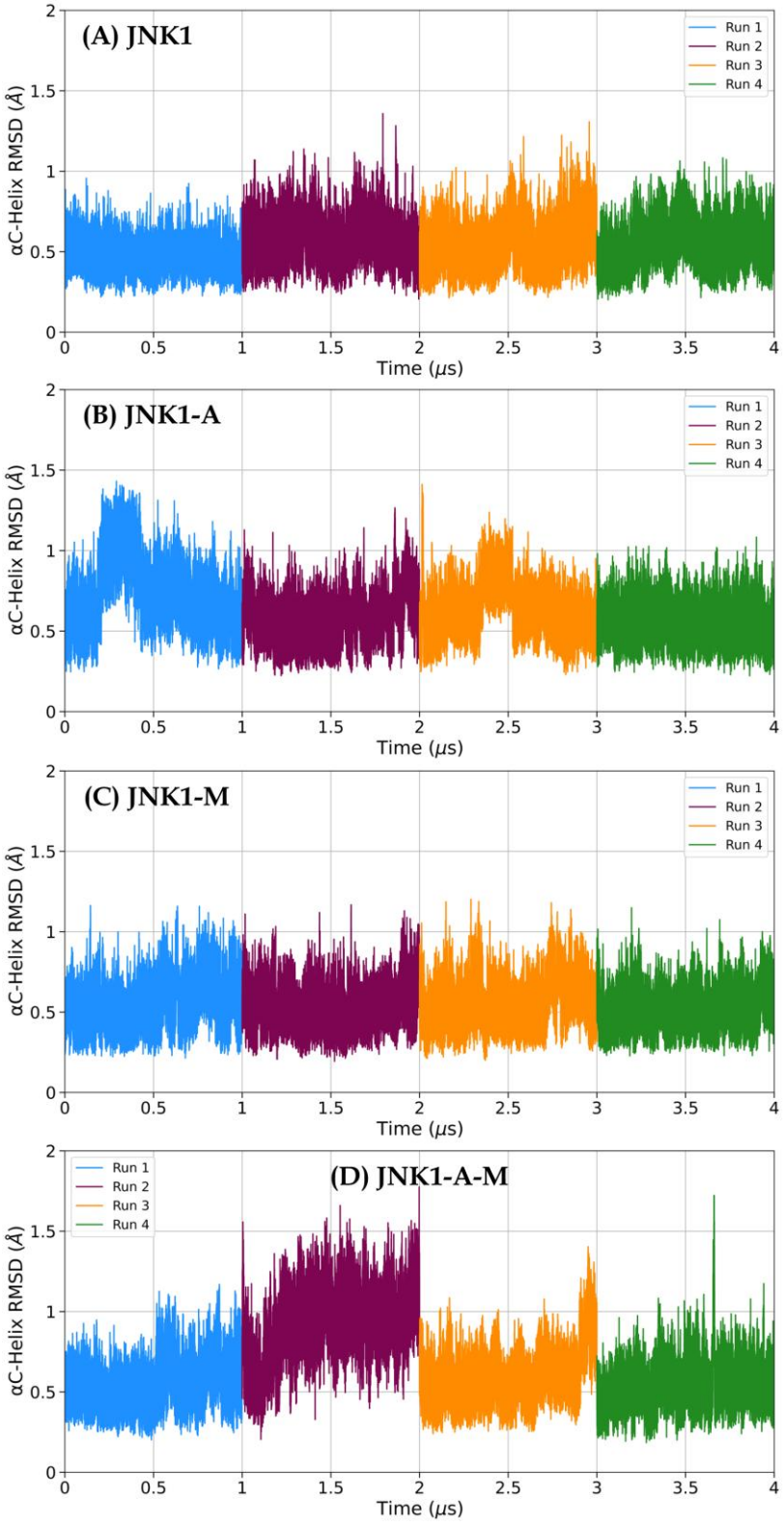


Figure E4: Time evolution of backbone RMSD of αC-helix residues in four systems from 1 μs MD replicas.

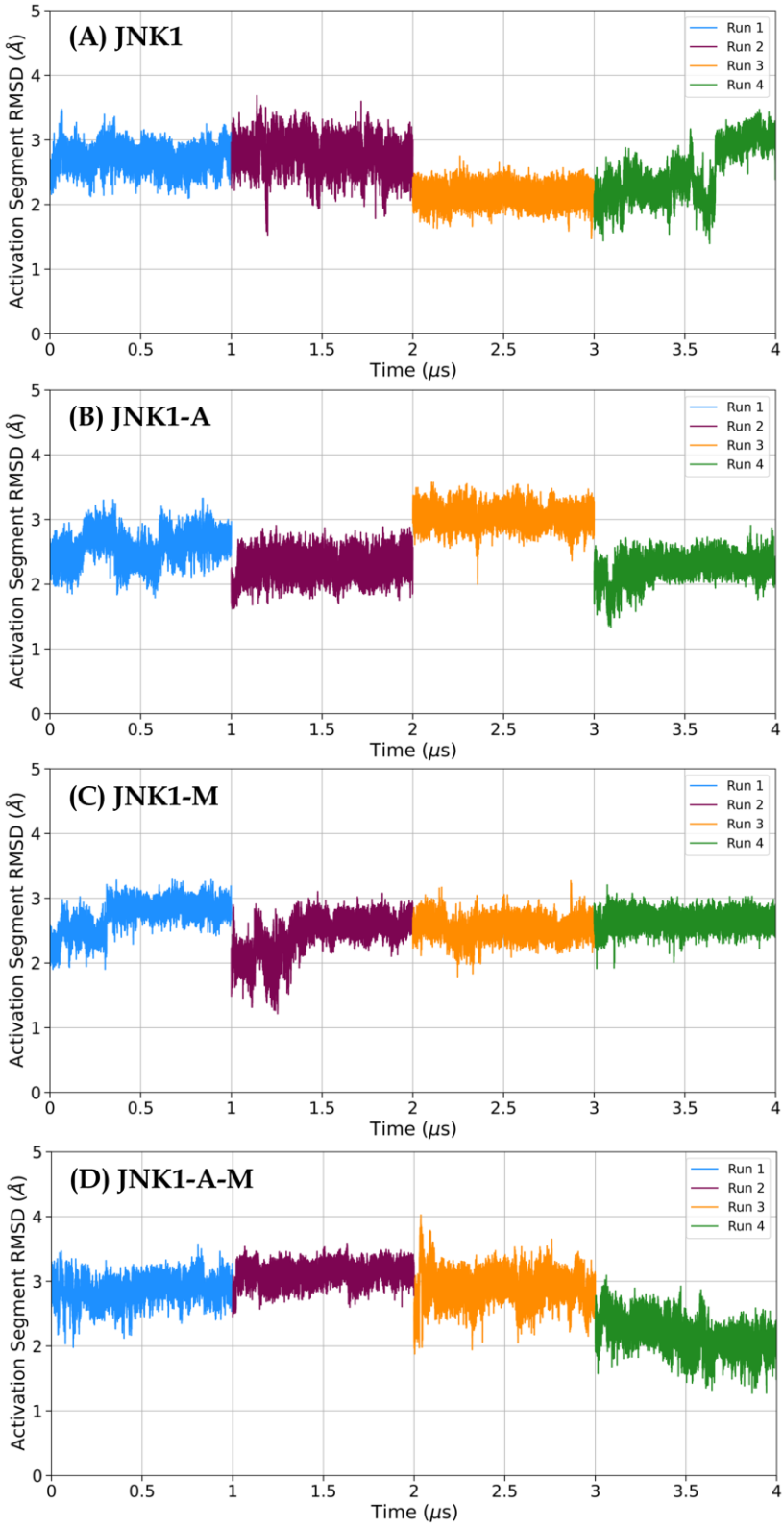


Figure E5: Time evolution of backbone RMSD of activation segment residues in four systems from 1 μs MD replicas.

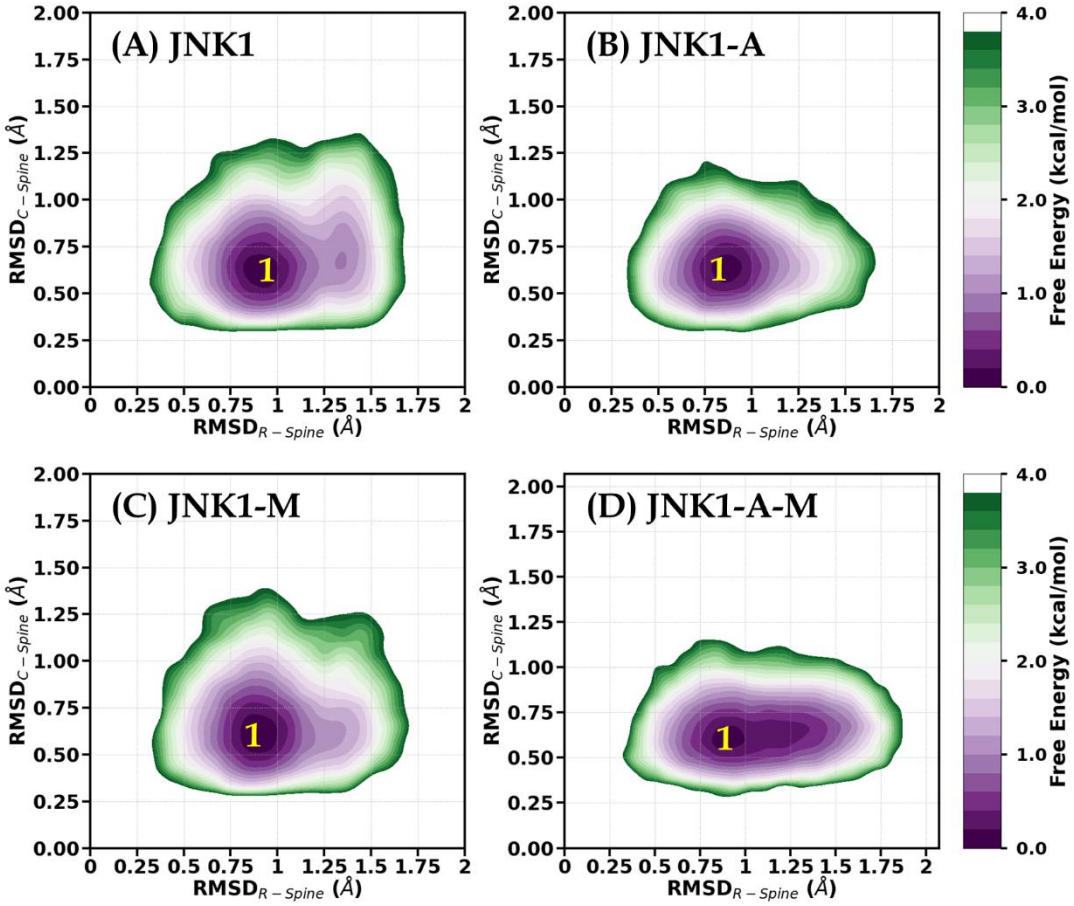


Figure E6: Free energy landscape (FEL) created using the R-spine RMSD and the C-spine RMSD as reaction coordinates. The dark purple color indicates the minima regions, where the conformational sampling is the most prominent.

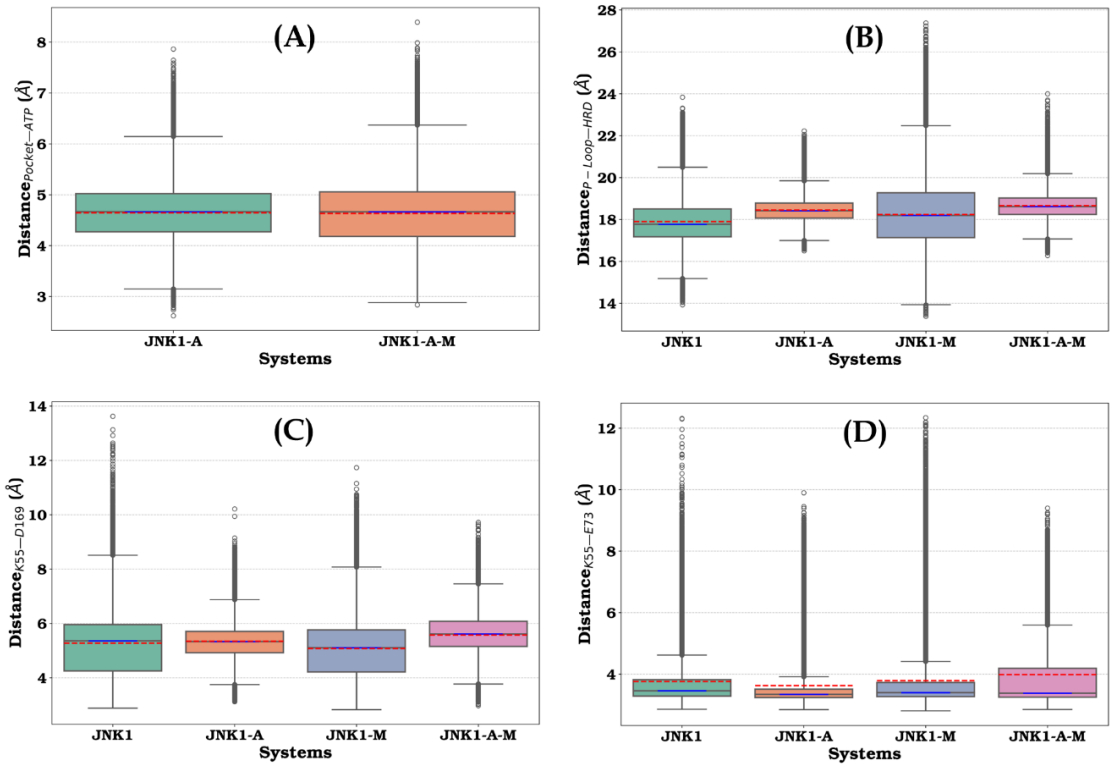


Figure E7: Box plot analysis of key distances in the studied systems. (A) Distance between ATP and binding site residues in the two ATP-bound complexes, (b) distance between the P-loop and conserved HRD motif, (c) Lys55–Asp169, and (d) Lys55–Glu73.

Bibliography

Bibliography

[1] E.G. Krebs, E.H. Fischer, The phosphorylase b to a converting enzyme of rabbit skeletal muscle, Biochimica et Biophysica Acta 20 (1956) 150–157. https://doi.org/10.1016/0006-3002(56)90273-6.

[2] G.T. Cori, A.A. Green, CRYSTALLINE MUSCLE PHOSPHORYLASE, Journal of Biological Chemistry 151 (1943) 31–38. https://doi.org/10.1016/s0021-9258(18)72111-x.

[3] J. Brognard, T. Hunter, Protein kinase signaling networks in cancer, Current Opinion in Genetics & Development 21 (2011) 4–11. https://doi.org/10.1016/j.gde.2010.10.012.

[4] L. Zhimin, H. Tony, Degradation of Activated Protein Kinases by Ubiquitination, Annu. Rev. Biochem. 78 (2009) 435–475. https://doi.org/10.1146/annurev.biochem.013008.092711.

[5] M.L. DePamphilis, C.M. De Renty, Z. Ullah, C.Y. Lee, "The Octet": Eight Protein Kinases that Control Mammalian DNA Replication, Front. Physio. 3 (2012). https://doi.org/10.3389/fphys.2012.00368.

[6] G. Henneke, S. Koundrioukoff, U. Hübscher, Multiple roles for kinases in DNA replication, EMBO Reports 4 (2003) 252–256. https://doi.org/10.1038/sj.embor.embor774.

[7] M. Cao, D. Meng, L. Wang, S. Bei, W.J. Snell, J. Pan, Activation loop phosphorylation of a protein kinase is a molecular marker of organelle size that dynamically reports flagellar length, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 12337–12342. https://doi.org/10.1073/pnas.1302364110.

[8] R. Mashud, A. Nomachi, A. Hayakawa, K. Kubouchi, S. Danno, T. Hirata, K. Matsuo, T. Nakayama, R. Satoh, R. Sugiura, M. Abe, K. Sakimura, S. Wakana, H. Ohsaki, S. Kamoshida, H. Mukai, Impaired lymphocyte trafficking in mice deficient in the kinase activity of PKN1, Sci Rep 7 (2017). https://doi.org/10.1038/s41598-017-07936-9.

[9] W. He, Y. Peng, W. Zhang, N. Lv, J. Tang, C. Chen, C. Zhang, S. Gao, H. Chen, G. Zhi, R. Feil, K.E. Kamm, J.T. Stull, X. Gao, M. Zhu, Myosin Light Chain Kinase Is Central to Smooth Muscle Contraction and Required for Gastrointestinal Motility in Mice, Gastroenterology 135 (2008) 610-620.e2. https://doi.org/10.1053/j.gastro.2008.05.032.

[10] K. Sathishkumar, U. Yallampalli, R. Elkins, C. Yallampalli, Raf-1 Kinase Regulates Smooth Muscle Contraction in the Rat Mesenteric Arteries, J Vasc Res 47 (2010) 384–398. https://doi.org/10.1159/000277726.

[11] J.A. Adams, Kinetic and Catalytic Mechanisms of Protein Kinases, Chem. Rev. 101 (2001) 2271–2290. https://doi.org/10.1021/cr000230w.

[12] M. Jain, M.K. Singh, H. Shyam, A. Mishra, S. Kumar, A. Kumar, J. Kushwaha, Role of JAK/STAT in the Neuroinflammation and its Association with Neurological Disorders, Annals of Neurosciences 28 (2021) 191–200. https://doi.org/10.1177/09727531211070532.

[13] D.P. Bermingham, R.D. Blakely, Kinase-dependent Regulation of Monoamine Neurotransmitter Transporters, Pharmacological Reviews 68 (2016) 888–953. https://doi.org/10.1124/pr.115.012260.

[14] J.A. Endicott, M.E.M. Noble, L.N. Johnson, The Structural Basis for Control of Eukaryotic Protein Kinases, Annu. Rev. Biochem. 81 (2012) 587–613. https://doi.org/10.1146/annurev-biochem-052410-090317.

[15] P. Blume-Jensen, T. Hunter, Oncogenic kinase signalling, Nature 411 (2001) 355–365. https://doi.org/10.1038/35077225.

[16] P. Lahiry, A. Torkamani, N.J. Schork, R.A. Hegele, Kinase mutations in human disease: interpreting genotype–phenotype relationships, Nat Rev Genet 11 (2010) 60–74. https://doi.org/10.1038/nrg2707.

[17] P. Cohen, Protein kinases — the major drug targets of the twenty-first century?, Nat Rev Drug Discov 1 (2002) 309–315. https://doi.org/10.1038/nrd773.

[18] R. Roskoski, A historical overview of protein kinases and their targeted small molecule inhibitors, Pharmacological Research 100 (2015) 1–23. https://doi.org/10.1016/j.phrs.2015.07.010.

[19] S.A. Hollingsworth, R.O. Dror, Molecular Dynamics Simulation for All, Neuron 99 (2018) 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011.

[20] M. Karplus, J.A. McCammon, Molecular dynamics simulations of biomolecules, Nat. Struct Biol. 9 (2002) 646–652. https://doi.org/10.1038/nsb0902-646.

[21] C.T. Walsh, S. Garneau-Tsodikova, G.J. Gatto, Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications, Angew Chem Int Ed 44 (2005) 7342–7372. https://doi.org/10.1002/anie.200501023.

[22] P. Cohen, The role of protein phosphorylation in human health and disease.: Delivered on June 30th 2001 at the FEBS Meeting in Lisbon, European Journal of Biochemistry 268 (2001) 5001–5010. https://doi.org/10.1046/j.0014-2956.2001.02473.x.

[23] N. Kannan, S.S. Taylor, Y. Zhai, J.C. Venter, G. Manning, Structural and Functional Diversity of the Microbial Kinome, PLoS Biol 5 (2007) e17. https://doi.org/10.1371/journal.pbio.0050017.

[24] Q. Liu, Y. Sabnis, Z. Zhao, T. Zhang, S.J. Buhrlage, L.H. Jones, N.S. Gray, Developing Irreversible Inhibitors of the Protein Kinase Cysteinome, Chemistry & Biology 20 (2013) 146–159. https://doi.org/10.1016/j.chembiol.2012.12.006.

[25] P.J. Kennelly, Protein kinases and protein phosphatases in prokaryotes: a genomic perspective, FEMS Microbiology Letters 206 (2002) 1–8. https://doi.org/10.1111/j.1574-6968.2002.tb10978.x.

[26] J.A. Ubersax, J.E. Ferrell Jr, Mechanisms of specificity in protein phosphorylation, Nat Rev Mol Cell Biol 8 (2007) 530–541. https://doi.org/10.1038/nrm2203.

[27] D.R. Knighton, J. Zheng, L.F. Ten Eyck, V.A. Ashford, N.-H. Xuong, S.S. Taylor, J.M. Sowadski, Crystal Structure of the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase, Science 253 (1991) 407–414. https://doi.org/10.1126/science.1862342.

[28] D.A. Johnson, P. Akamine, E. Radzio-Andzelm, Madhusudan, S.S. Taylor, Dynamics of cAMP-Dependent Protein Kinase, Chem. Rev. 101 (2001) 2243–2270. https://doi.org/10.1021/cr000226k.

[29] S.S. Taylor, E. Radzio-Andzelm, Three protein kinase structures define a common motif, Structure 2 (1994) 345–355. https://doi.org/10.1016/s0969-2126(00)00036-8.

[30] S.K. Hanks, A.M. Quinn, T. Hunter, The Protein Kinase Family: Conserved Features and Deduced Phylogeny of the Catalytic Domains, Science 241 (1988) 42–52. https://doi.org/10.1126/science.3291115.

[31] Y. Liu, K. Shah, F. Yang, L. Witucki, K.M. Shokat, A molecular gate which controls unnatural ATP analogue recognition by the tyrosine kinase v-Src, Bioorganic & Medicinal Chemistry 6 (1998) 1219–1226. https://doi.org/10.1016/s0968-0896(98)00099-6.

[32] H.S. Meharena, P. Chang, M.M. Keshwani, K. Oruganty, A.K. Nene, N. Kannan, S.S. Taylor, A.P. Kornev, Deciphering the Structural Basis of Eukaryotic Protein Kinase Regulation, PLoS Biol 11 (2013) e1001680. https://doi.org/10.1371/journal.pbio.1001680.

[33] M. Huse, J. Kuriyan, The Conformational Plasticity of Protein Kinases, Cell 109 (2002) 275–282. https://doi.org/10.1016/S0092-8674(02)00741-9.

[34] J. Zheng, D.R. Knighton, L.F. Ten Eyck, R. Karlsson, N. Xuong, S.S. Taylor, J.M. Sowadski, Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with magnesium-ATP and peptide inhibitor, Biochemistry 32 (1993) 2154–2161. https://doi.org/10.1021/bi00060a005.

[35] L. Skora, J. Mestan, D. Fabbro, W. Jahnke, S. Grzesiek, NMR reveals the allosteric opening and closing of Abelson tyrosine kinase by ATP-site and myristoyl pocket inhibitors, Proc. Natl. Acad. Sci. U.S.A. 110 (2013). https://doi.org/10.1073/pnas.1314712110.

[36] N. Narayana, S. Cox, N. Xuong, L.F.T. Eyck, S.S. Taylor, A binary complex of the catalytic subunit of cAMP-dependent protein kinase and adenosine further defines conformational flexibility, Structure 5 (1997) 921–935. https://doi.org/10.1016/s0969-2126(97)00246-3.

[37] M.P. Kamps, S.S. Taylor, B.M. Sefton, Direct evidence that oncogenic tyrosine kinases and cyclic AMP-dependent protein kinase have homologous ATP-binding sites, Nature 310 (1984) 589–592. https://doi.org/10.1038/310589a0.

[38] M.J. Zoller, N.C. Nelson, S.S. Taylor, Affinity labeling of cAMP-dependent protein kinase with p-fluorosulfonylbenzoyl adenosine. Covalent modification of lysine 71., Journal of Biological Chemistry 256 (1981) 10837–10842. https://doi.org/10.1016/s0021-9258(19)68519-4.

[39] A.C. Carrera, K. Alexandrov, T.M. Roberts, The conserved lysine of the catalytic domain of protein kinases is actively involved in the phosphotransfer reaction and not required for anchoring ATP., Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 442–446. https://doi.org/10.1073/pnas.90.2.442.

[40] A.C. Bastidas, M.S. Deal, J.M. Steichen, Y. Guo, J. Wu, S.S. Taylor, Phosphoryl Transfer by Protein Kinase A Is Captured in a Crystal Lattice, J. Am. Chem. Soc. 135 (2013) 4788–4798. https://doi.org/10.1021/ja312237q.

[41] A.S. Shaw, A.P. Kornev, J. Hu, L.G. Ahuja, S.S. Taylor, Kinases and Pseudokinases: Lessons from RAF, Molecular and Cellular Biology 34 (2014) 1538–1546. https://doi.org/10.1128/mcb.00057-14.

[42] S.S. Taylor, M.M. Keshwani, J.M. Steichen, A.P. Kornev, Evolution of the eukaryotic protein kinases as dynamic molecular switches, Phil. Trans. R. Soc. B 367 (2012) 2517–2528. https://doi.org/10.1098/rstb.2012.0054.

[43] P.D. Jeffrey, A.A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massagué, N.P. Pavletich, Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex, Nature 376 (1995) 313–320. https://doi.org/10.1038/376313a0.

[44] H. Yamaguchi, W.A. Hendrickson, Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation, Nature 384 (1996) 484–489. https://doi.org/10.1038/384484a0.

[45] B. Nolen, S. Taylor, G. Ghosh, Regulation of Protein Kinases, Molecular Cell 15 (2004) 661–675. https://doi.org/10.1016/j.molcel.2004.08.024.

[46] N. Gotoh, A. Tojo, M. Hino, Y. Yazaki, M. Shibuya, A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor, Biochemical and Biophysical Research Communications 186 (1992) 768–774. https://doi.org/10.1016/0006-291x(92)90812-y.

[47] C. Tournier, C. Dong, T.K. Turner, S.N. Jones, R.A. Flavell, R.J. Davis, MKK7 is an essential component of the JNK signal transduction pathway activated by proinflammatory cytokines, Genes Dev. 15 (2001) 1419–1426. https://doi.org/10.1101/gad.888501.

[48] X. Xie, Y. Gu, T. Fox, J.T. Coll, M.A. Fleming, W. Markland, P.R. Caron, K.P. Wilson, M.S.-S. Su, Crystal structure of JNK3: a kinase implicated in neuronal apoptosis, Structure 6 (1998) 983–991. https://doi.org/10.1016/S0969-2126(98)00100-2.

[49] P. Mishra, S. Günther, New insights into the structural dynamics of the kinase JNK3, Sci Rep 8 (2018) 9435. https://doi.org/10.1038/s41598-018-27867-3.

[50] A. Berteotti, A. Cavalli, D. Branduardi, F.L. Gervasio, M. Recanatini, M. Parrinello, Protein Conformational Transitions: The Closure Mechanism of a Kinase Explored by Atomistic Simulations, J. Am. Chem. Soc. 131 (2009) 244–250. https://doi.org/10.1021/ja806846q.

[51] W. Gan, S. Yang, B. Roux, Atomistic View of the Conformational Activation of Src Kinase Using the String Method with Swarms-of-Trajectories, Biophysical Journal 97 (2009) L8–L10. https://doi.org/10.1016/j.bpj.2009.06.016.

[52] Y. Shan, M.A. Seeliger, M.P. Eastwood, F. Frank, H. Xu, M.Ø. Jensen, R.O. Dror, J. Kuriyan, D.E. Shaw, A conserved protonation-dependent switch controls drug binding in the Abl kinase, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 139–144. https://doi.org/10.1073/pnas.0811223106.

[53] Y. Meng, B. Roux, Locking the Active Conformation of c-Src Kinase through the Phosphorylation of the Activation Loop, Journal of Molecular Biology 426 (2014) 423–435. https://doi.org/10.1016/j.jmb.2013.10.001.

[54] Z. Musavizadeh, A. Grottesi, G. Guarguaglini, A. Paiardini, Phosphorylation, Mg-ADP, and Inhibitors Differentially Shape the Conformational Dynamics of the A-Loop of Aurora-A, Biomolecules 11 (2021) 567. https://doi.org/10.3390/biom11040567.

[55] V. Gosu, S. Choi, Structural dynamic analysis of apo and ATP-bound IRAK4 kinase, Sci Rep 4 (2014) 5748. https://doi.org/10.1038/srep05748.

[56] Y.-L. Wu, L. Zhang, D.-W. Kim, X. Liu, D.H. Lee, J.C.-H. Yang, M.-J. Ahn, J.F. Vansteenkiste, W.-C. Su, E. Felip, V. Chia, S. Glaser, P. Pultar, S. Zhao, B. Peng, M. Akimov, D.S.W. Tan, Phase Ib/II Study of Capmatinib (INC280) Plus Gefitinib After Failure of Epidermal Growth Factor Receptor (EGFR) Inhibitor Therapy in Patients With *EGFR*-Mutated, MET Factor–Dysregulated Non–Small-Cell Lung Cancer, JCO 36 (2018) 3101–3109. https://doi.org/10.1200/jco.2018.77.7326.

[57] H. Gharwan, H. Groninger, Kinase inhibitors and monoclonal antibodies in oncology: clinical implications, Nat Rev Clin Oncol 13 (2016) 209–227. https://doi.org/10.1038/nrclinonc.2015.213.

[58] S. Knapp, New opportunities for kinase drug repurposing and target discovery, Br J Cancer 118 (2018) 936–937. https://doi.org/10.1038/s41416-018-0045-6.

[59] S. Müller, A. Chaikuad, N.S. Gray, S. Knapp, The ins and outs of selective kinase inhibitor development, Nat Chem Biol 11 (2015) 818–821. https://doi.org/10.1038/nchembio.1938.

[60] P. Wu, T.E. Nielsen, M.H. Clausen, FDA-approved small-molecule kinase inhibitors, Trends in Pharmacological Sciences 36 (2015) 422–439. https://doi.org/10.1016/j.tips.2015.04.005.

[61] S. Klaeger, S. Heinzlmeir, M. Wilhelm, H. Polzer, B. Vick, P.-A. Koenig, M. Reinecke, B. Ruprecht, S. Petzoldt, C. Meng, J. Zecha, K. Reiter, H. Qiao, D. Helm, H. Koch, M. Schoof, G. Canevari, E. Casale, S.R. Depaolini, A. Feuchtinger, Z. Wu, T. Schmidt, L. Rueckert, W. Becker, J. Huenges, A.-K. Garz, B.-O. Gohlke, D.P. Zolg, G. Kayser, T. Vooder, R. Preissner, H. Hahne, N. Tõnisson, K. Kramer, K. Götze, F. Bassermann, J. Schlegl, H.-C. Ehrlich, S. Aiche, A. Walch, P.A. Greif, S. Schneider, E.R. Felder, J. Ruland, G. Médard, I. Jeremias, K. Spiekermann, B. Kuster, The target landscape of clinical kinase drugs, Science 358 (2017). https://doi.org/10.1126/science.aan4368.

[62] F.M. Ferguson, N.S. Gray, Kinase inhibitors: the road ahead, Nat Rev Drug Discov 17 (2018) 353–377. https://doi.org/10.1038/nrd.2018.21.

[63] M. Piran, N. Sepahi, M. Piran, P.L. Fernandes, A. Ghanbariasad, Pathway mining in functional genomics: An integrative approach to delineate boolean

relationships between Src and its targets, (2020). https://doi.org/10.1101/2020.01.25.919639.

[64] Z.A. Knight, H. Lin, K.M. Shokat, Targeting the cancer kinome through polypharmacology, Nat Rev Cancer 10 (2010) 130–137. https://doi.org/10.1038/nrc2787.

[65] M. Yan, H. Wang, Q. Wang, Z. Zhang, C. Zhang, Allosteric inhibition of c-Met kinase in sub-microsecond molecular dynamics simulations induced by its inhibitor, tivantinib, Phys. Chem. Chem. Phys. 18 (2016) 10367–10374. https://doi.org/10.1039/c5cp07001e.

[66] D. Ferraris, Z. Yang, D. Welsbie, Dual leucine zipper kinase as a therapeutic target for neurodegenerative conditions, Future Medicinal Chemistry 5 (2013) 1923–1934. https://doi.org/10.4155/fmc.13.150.

[67] H. Bu, Z. Li, Y. Lu, Z. Zhuang, Y. Zhen, L. Zhang, Deciphering the multifunctional role of dual leucine zipper kinase (DLK) and its therapeutic potential in disease, European Journal of Medicinal Chemistry 255 (2023) 115404. https://doi.org/10.1016/j.ejmech.2023.115404.

[68] V. Valakh, E. Frey, E. Babetto, L.J. Walker, A. DiAntonio, Cytoskeletal disruption activates the DLK/JNK pathway, which promotes axonal regeneration and mimics a preconditioning injury, Neurobiology of Disease 77 (2015) 13–25. https://doi.org/10.1016/j.nbd.2015.02.014.

[69] A. Sengupta Ghosh, B. Wang, C.D. Pozniak, M. Chen, R.J. Watts, J.W. Lewcock, DLK induces developmental neuronal degeneration via selective regulation of proapoptotic JNK activity, Journal of Cell Biology 194 (2011) 751–764. https://doi.org/10.1083/jcb.201103153.

[70] S.H. Chung, M.R. Awal, J. Shay, M.M. McLoed, E. Mazur, C.V. Gabel, Novel DLK-independent neuronal regeneration in *Caenorhabditis elegans* shares links with activity-dependent ectopic outgrowth, Proc. Natl. Acad. Sci. U.S.A. 113 (2016). https://doi.org/10.1073/pnas.1600564113.

[71] S. Patel, F. Cohen, B.J. Dean, K. De La Torre, G. Deshmukh, A.A. Estrada, A.S. Ghosh, P. Gibbons, A. Gustafson, M.P. Huestis, C.E. Le Pichon, H. Lin, W. Liu, X. Liu, Y. Liu, C.Q. Ly, J.P. Lyssikatos, C. Ma, K. Scearce-Levie, Y.G. Shin,

H. Solanoy, K.L. Stark, J. Wang, B. Wang, X. Zhao, J.W. Lewcock, M. Siu, Discovery of Dual Leucine Zipper Kinase (DLK, MAP3K12) Inhibitors with Activity in Neurodegeneration Models, J. Med. Chem. 58 (2015) 401–418. https://doi.org/10.1021/jm5013984.

[72] L.B. Holzman, S.E. Merritt, G. Fan, Identification, molecular cloning, and characterization of dual leucine zipper bearing kinase. A novel serine/threonine protein kinase that defines a second subfamily of mixed lineage kinases., Journal of Biological Chemistry 269 (1994) 30808–30817. https://doi.org/10.1016/S0021-9258(18)47353-X.

[73] D. Yan, Z. Wu, A.D. Chisholm, Y. Jin, The DLK-1 Kinase Promotes mRNA Stability and Local Translation in C. elegans Synapses and Axon Regeneration, Cell 138 (2009) 1005–1018. https://doi.org/10.1016/j.cell.2009.06.023.

[74] Y. Jin, B. Zheng, Multitasking: Dual Leucine Zipper–Bearing Kinases in Neuronal Development and Stress Management, Annu. Rev. Cell Dev. Biol. 35 (2019) 501–521. https://doi.org/10.1146/annurev-cellbio-100617-062644.

[75] D. Nihalani, S. Merritt, L.B. Holzman, Identification of Structural and Functional Domains in Mixed Lineage Kinase Dual Leucine Zipper-bearing Kinase Required for Complex Formation and Stress-activated Protein Kinase Activation, Journal of Biological Chemistry 275 (2000) 7273–7279. https://doi.org/10.1074/jbc.275.10.7273.

[76] M. Mata, S.E. Merritt, G. Fan, G.G. Yu, L.B. Holzman, Characterization of Dual Leucine Zipper-bearing Kinase, a Mixed Lineage Kinase Present in Synaptic Terminals Whose Phosphorylation State Is Regulated by Membrane Depolarization via Calcineurin, Journal of Biological Chemistry 271 (1996) 16888–16896. https://doi.org/10.1074/jbc.271.28.16888.

[77] C.-C. Wu, H.-J. Wu, C.-H. Wang, C.-H. Lin, S.-C. Hsu, Y.-R. Chen, M. Hsiao, S.C. Schuyler, F.L. Lu, N. Ma, J. Lu, Akt suppresses DLK for maintaining self-renewal of mouse embryonic stem cells, Cell Cycle 14 (2015) 1207–1217. https://doi.org/10.1080/15384101.2015.1014144.

[78] M. Siu, A. Sengupta Ghosh, J.W. Lewcock, Dual Leucine Zipper Kinase Inhibitors for the Treatment of Neurodegeneration: Miniperspective, J. Med. Chem. 61 (2018) 8078–8087. https://doi.org/10.1021/acs.jmedchem.8b00370.

[79] K.-A. Köster, M. Dethlefs, J. Duque Escobar, E. Oetjen, Regulation of the Activity of the Dual Leucine Zipper Kinase by Distinct Mechanisms, Cells 13 (2024) 333. https://doi.org/10.3390/cells13040333.

[80] B.M. Yashar, C. Kelley, K. Yee, B. Errede, Leonardi. Zon, Novel Members of the Mitogen-Activated Protein Kinase Activator Family in *Xenopus laevis*, Molecular and Cellular Biology 13 (1993) 5738–5748. https://doi.org/10.1128/mcb.13.9.5738-5748.1993.

[81] J. Yang, L. New, Y. Jiang, J. Han, B. Su, Molecular cloning and characterization of a human protein kinase that specifically activates c-Jun N-terminal kinase, Gene 212 (1998) 95–102. https://doi.org/10.1016/s0378-1119(98)00158-9.

[82] M. Raman, W. Chen, M.H. Cobb, Differential regulation and properties of MAPKs, Oncogene 26 (2007) 3100–3112. https://doi.org/10.1038/sj.onc.1210392.

[83] M. Takekawa, K. Tatebayashi, H. Saito, Conserved Docking Site Is Essential for Activation of Mammalian MAP Kinase Kinases by Specific MAP Kinase Kinase Kinases, Molecular Cell 18 (2005) 295–306. https://doi.org/10.1016/j.molcel.2005.04.001.

[84] D.T. Ho, A.J. Bardwell, M. Abdollahi, L. Bardwell, A Docking Site in MKK4 Mediates High Affinity Binding to JNK MAPKs and Competes with Similar Docking Sites in JNK Substrates, Journal of Biological Chemistry 278 (2003) 32662–32672. https://doi.org/10.1074/jbc.m304229200.

[85] A. Lin, A. Minden, H. Martinetto, F.-X. Claret, C. Lange-Carter, F. Mercurio, G.L. Johnson, M. Karin, Identification of a Dual Specificity Kinase That Activates the Jun Kinases and p38-Mpk2, Science 268 (1995) 286–290. https://doi.org/10.1126/science.7716521.

[86] T. Wada, K. Nakagawa, T. Watanabe, G. Nishitai, J. Seo, H. Kishimoto, D. Kitagawa, T. Sasaki, J.M. Penninger, H. Nishina, T. Katada, Impaired Synergistic Activation of Stress-activated Protein Kinase SAPK/JNK in Mouse

Embryonic Stem Cells Lacking SEK1/MKK4, Journal of Biological Chemistry 276 (2001) 30892–30897. https://doi.org/10.1074/jbc.m011780200.

[87] J.-K. Lee, W.-S. Hwang, Y.-D. Lee, P.-L. Han, Dynamic expression of SEK1 suggests multiple roles of the gene during embryogenesis and in adult brain of mice, Molecular Brain Research 66 (1999) 133–140. https://doi.org/10.1016/s0169-328x(99)00035-2.

[88] X. Wang, B. Nadarajah, A.C. Robinson, B.W. McColl, J.-W. Jin, F. Dajas-Bailador, R.P. Boot-Handford, C. Tournier, Targeted Deletion of the Mitogen-Activated Protein Kinase Kinase 4 Gene in the Nervous System Causes Severe Brain Developmental Defects and Premature Death, Molecular and Cellular Biology 27 (2007) 7935–7946. https://doi.org/10.1128/mcb.00226-07.

[89] G. Choukroun, R. Hajjar, S. Fry, F. Del Monte, S. Haq, J.L. Guerrero, M. Picard, A. Rosenzweig, T. Force, Regulation of cardiac hypertrophy in vivo by the stress-activated protein kinases/c-Jun NH2-terminal kinases, J. Clin. Invest. 104 (1999) 391–398. https://doi.org/10.1172/jci6350.

[90] M.A. Bogoyevitch, The isoform-specific functions of the c-Jun N-terminal Kinases (JNKs): differences revealed by gene targeting, BioEssays 28 (2006) 923–934. https://doi.org/10.1002/bies.20458.

[91] J. Wang, H. Wang, J. Chen, X. Wang, K. Sun, Y. Wang, J. Wang, X. Yang, X. Song, Y. Xin, Z. Liu, R. Hui, GADD45B inhibits MKK7-induced cardiac hypertrophy and the polymorphisms of GADD45B is associated with inter-ventricular septum hypertrophy, Biochemical and Biophysical Research Communications 372 (2008) 623–628. https://doi.org/10.1016/j.bbrc.2008.05.122.

[92] S. Lee, D.L. Boyle, A. Berdeja, G.S. Firestein, Regulation of inflammatory arthritis by the upstream kinase mitogen activated protein kinase kinase 7 in the c-Jun N-Terminal kinase pathway, Arthritis Res Ther 14 (2012). https://doi.org/10.1186/ar3750.

[93] C. Tournier, A.J. Whitmarsh, J. Cavanagh, T. Barrett, R.J. Davis, The *MKK7* Gene Encodes a Group of c-Jun NH_2 -Terminal Kinase Kinases, Molecular and

Cellular Biology 19 (1999) 1569–1581. https://doi.org/10.1128/MCB.19.2.1569.

[94] D.T. Ho, A.J. Bardwell, S. Grewal, C. Iverson, L. Bardwell, Interacting JNK-docking Sites in MKK7 Promote Binding and Activation of JNK Mitogen-activated Protein Kinases, Journal of Biological Chemistry 281 (2006) 13169–13179. https://doi.org/10.1074/jbc.m601010200.

[95] A.D. Caliz, A. Vertii, V. Fisch, S. Yoon, H.-J. Yoo, J.F. Keaney, S. Kant, Mitogen-activated protein kinase kinase 7 in inflammatory, cancer, and neurological diseases, Front. Cell Dev. Biol. 10 (2022). https://doi.org/10.3389/fcell.2022.979673.

[96] Q. Zhang, H. Tian, X. Fu, G. Zhang, Delayed activation and regulation of MKK7 in hippocampal CA1 region following global cerebral ischemia in rats, Life Sciences 74 (2003) 37–45. https://doi.org/10.1016/j.lfs.2003.06.025.

[97] D.N. Dhanasekaran, E.P. Reddy, JNK signaling in apoptosis, Oncogene 27 (2008) 6245–6251. https://doi.org/10.1038/onc.2008.301.

[98] A. Vercelli, S. Biggi, A. Sclip, I.E. Repetto, S. Cimini, F. Falleroni, S. Tomasi, R. Monti, N. Tonna, F. Morelli, V. Grande, M. Stravalaci, E. Biasini, O. Marin, F. Bianco, D. Di Marino, T. Borsello, Exploring the role of MKK7 in excitotoxicity and cerebral ischemia: a novel pharmacological strategy against brain injury, Cell Death Dis 6 (2015) e1854–e1854. https://doi.org/10.1038/cddis.2015.226.

[99] A. Shraga, E. Olshvang, N. Davidzohn, P. Khoshkenar, N. Germain, K. Shurrush, S. Carvalho, L. Avram, S. Albeck, T. Unger, B. Lefker, C. Subramanyam, R.L. Hudkins, A. Mitchell, Z. Shulman, T. Kinoshita, N. London, Covalent Docking Identifies a Potent and Selective MKK7 Inhibitor, Cell Chemical Biology 26 (2019) 98-108.e5. https://doi.org/10.1016/j.chembiol.2018.10.011.

[100] E.K. Kim, E.-J. Choi, Pathological roles of MAPK signaling pathways in human diseases, Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 1802 (2010) 396–405. https://doi.org/10.1016/j.bbadis.2009.12.009.

[101] R. Schellino, M. Boido, A. Vercelli, JNK Signaling Pathway Involvement in Spinal Cord Neuron Development and Death, Cells 8 (2019) 1576. https://doi.org/10.3390/cells8121576.

[102] B.J. Pulverer, J.M. Kyriakis, J. Avruch, E. Nikolakaki, J.R. Woodgett, Phosphorylation of c-jun mediated by MAP kinases, Nature 353 (1991) 670–674. https://doi.org/10.1038/353670a0.

[103] R. Xu, J. Hu, The role of JNK in prostate cancer progression and therapeutic strategies, Biomedicine & Pharmacotherapy 121 (2020) 109679. https://doi.org/10.1016/j.biopha.2019.109679.

[104] Y.-S. Heo, S.-K. Kim, C.I. Seo, Y.K. Kim, B.-J. Sung, H.S. Lee, J.I. Lee, S.-Y. Park, J.H. Kim, K.Y. Hwang, Y.-L. Hyun, Y.H. Jeon, S. Ro, J.M. Cho, T.G. Lee, C.-H. Yang, Structural basis for the selective inhibition of JNK1 by the scaffolding protein JIP1 and SP600125, EMBO J 23 (2004) 2185–2195. https://doi.org/10.1038/sj.emboj.7600212.

[105] R. Garg, S. Kumariya, R. Katekar, S. Verma, U.K. Goand, J.R. Gayen, JNK signaling pathway in metabolic disorders: An emerging therapeutic target, European Journal of Pharmacology 901 (2021) 174079. https://doi.org/10.1016/j.ejphar.2021.174079.

[106] W. Haeusgen, T. Herdegen, V. Waetzig, The bottleneck of JNK signaling: Molecular and functional characteristics of MKK4 and MKK7, European Journal of Cell Biology 90 (2011) 536–544. https://doi.org/10.1016/j.ejcb.2010.11.008.

[107] J. Kragelj, A. Palencia, M.H. Nanao, D. Maurin, G. Bouvignies, M. Blackledge, M.R. Jensen, Structure and dynamics of the MKK7–JNK signaling complex, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 3409–3414. https://doi.org/10.1073/pnas.1419528112.

[108] K. Kirsch, A. Zeke, O. Tőke, P. Sok, A. Sethi, A. Sebő, G.S. Kumar, P. Egri, Á.L. Póti, P. Gooley, W. Peti, I. Bento, A. Alexa, A. Reményi, Co-regulation of the transcription controlling ATF2 phosphoswitch by JNK and p38, Nat Commun 11 (2020). https://doi.org/10.1038/s41467-020-19582-3.

[109] W. Lu, Y. Liu, Y. Gao, Q. Geng, D. Gurbani, L. Li, S.B. Ficarro, C.J. Meyer, D. Sinha, I. You, J. Tse, Z. He, W. Ji, J. Che, A.Y. Kim, T. Yu, K. Wen, K.C. Anderson, J.A. Marto, K.D. Westover, T. Zhang, N.S. Gray, Development of a Covalent Inhibitor of c-Jun N-Terminal Protein Kinase (JNK) 2/3 with Selectivity over JNK1, J. Med. Chem. 66 (2023) 3356–3371. https://doi.org/10.1021/acs.jmedchem.2c01834.

[110] Y. Feng, H. Park, L. Bauer, J.C. Ryu, S.O. Yoon, Thiophene-Pyrazolourea Derivatives as Potent, Orally Bioavailable, and Isoform-Selective JNK3 Inhibitors, ACS Med. Chem. Lett. 12 (2021) 24–29. https://doi.org/10.1021/acsmedchemlett.0c00533.

[111] E.T. Coffey, Nuclear and cytosolic JNK signalling in neurons, Nat Rev Neurosci 15 (2014) 285–299. https://doi.org/10.1038/nrn3729.

[112] R. Castro-Torres, O. Busquets, A. Parcerisas, E. Verdaguer, J. Olloquequi, M. Ettcheto, C. Beas-Zarate, J. Folch, A. Camins, C. Auladell, Involvement of JNK1 in Neuronal Polarization During Brain Development, Cells 9 (2020) 1897. https://doi.org/10.3390/cells9081897.

[113] G.L. Johnson, K. Nakamura, The c-jun kinase/stress-activated pathway: Regulation, function and role in human disease, Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1773 (2007) 1341–1348. https://doi.org/10.1016/j.bbamcr.2006.12.009.

[114] G. Solinas, M. Karin, JNK1 and IKKβ: molecular links between obesity and metabolic dysfunction, FASEB j. 24 (2010) 2596–2611. https://doi.org/10.1096/fj.09-151340.

[115] H. Yan, L. He, D. Lv, J. Yang, Z. Yuan, The Role of the Dysregulated JNK Signaling Pathway in the Pathogenesis of Human Diseases and Its Potential Therapeutic Strategies: A Comprehensive Review, Biomolecules 14 (2024) 243. https://doi.org/10.3390/biom14020243.

[116] J. Hirosumi, G. Tuncman, L. Chang, C.Z. Görgün, K.T. Uysal, K. Maeda, M. Karin, G.S. Hotamisligil, A central role for JNK in obesity and insulin resistance, Nature 420 (2002) 333–336. https://doi.org/10.1038/nature01137.

[117] H. Yoshida, C.J. Hastie, H. McLauchlan, P. Cohen, M. Goedert, Phosphorylation of microtubule-associated protein tau by isoforms of c-Jun N-terminal kinase (JNK), Journal of Neurochemistry 90 (2004) 352–358. https://doi.org/10.1111/j.1471-4159.2004.02479.x.

[118] W.-S. Choi, G. Abel, H. Klintworth, R.A. Flavell, Z. Xia, JNK3 Mediates Paraquat- and Rotenone-Induced Dopaminergic Neuron Death, J Neuropathol Exp Neurol 69 (2010) 511–520. https://doi.org/10.1097/nen.0b013e3181db8100.

[119] Z. Luo, Z. Han, F. Shou, Y. Li, Y. Chen, LINC00958 Accelerates Cell Proliferation and Migration in Non-Small Cell Lung Cancer Through JNK/c-JUN Signaling, Human Gene Therapy Methods 30 (2019) 226–234. https://doi.org/10.1089/hgtb.2019.115.

[120] H.-J. Im, P. Muddasani, V. Natarajan, T.M. Schmid, J.A. Block, F. Davis, A.J. Van Wijnen, R.F. Loeser, Basic Fibroblast Growth Factor Stimulates Matrix Metalloproteinase-13 via the Molecular Cross-talk between the Mitogen-activated Protein Kinases and Protein Kinase Cδ Pathways in Human Adult Articular Chondrocytes, Journal of Biological Chemistry 282 (2007) 11110–11121. https://doi.org/10.1074/jbc.m609040200.

[121] J. Falsig, P. Pörzgen, J. Lotharius, M. Leist, Specific Modulation of Astrocyte Inflammation by Inhibition of Mixed Lineage Kinases with CEP-1347, The Journal of Immunology 173 (2004) 2762–2770. https://doi.org/10.4049/jimmunol.173.4.2762.

[122] J. Cui, M. Zhang, Y. Zhang, Z. Xu, JNK pathway: diseases and therapeutic potential, Acta Pharmacologica Sinica 28 (2007) 601–608. https://doi.org/10.1111/j.1745-7254.2007.00579.x.

[123] P.K. Roy, F. Rashid, J. Bragg, J.A. Ibdah, D.O.G.A. Hepatology, U.O.M.S.O. Medicine, Columbia, Missouri, U. States, Role of the JNK signal transduction pathway in inflammatory bowel disease, WJG 14 (2008) 200. https://doi.org/10.3748/wjg.14.200.

[124] Y. Pan, Y. Wang, Y. Zhao, K. Peng, W. Li, Y. Wang, J. Zhang, S. Zhou, Q. Liu, X. Li, L. Cai, G. Liang, Inhibition of JNK Phosphorylation by a Novel Curcumin

Analog Prevents High Glucose–Induced Inflammation and Apoptosis in Cardiomyocytes and the Development of Diabetic Cardiomyopathy, Diabetes 63 (2014) 3497–3511. https://doi.org/10.2337/db13-1577.

[125] S. Patel, W.J. Meilandt, R.I. Erickson, J. Chen, G. Deshmukh, A.A. Estrada, R.N. Fuji, P. Gibbons, A. Gustafson, S.F. Harris, J. Imperio, W. Liu, X. Liu, Y. Liu, J.P. Lyssikatos, C. Ma, J. Yin, J.W. Lewcock, M. Siu, Selective Inhibitors of Dual Leucine Zipper Kinase (DLK, MAP3K12) with Activity in a Model of Alzheimer's Disease, J. Med. Chem. 60 (2017) 8083–8102. https://doi.org/10.1021/acs.jmedchem.7b00843.

[126] J.S. Katz, J.D. Rothstein, M.E. Cudkowicz, A. Genge, B. Oskarsson, A.B. Hains, C. Chen, J. Galanter, B.L. Burgess, W. Cho, G.A. Kerchner, F.L. Yeh, A.S. Ghosh, S. Cheeti, L. Brooks, L. Honigberg, J.A. Couch, M.E. Rothenberg, F. Brunstein, K.R. Sharma, L. Berg, J.D. Berry, J.D. Glass, A Phase 1 study of GDC -0134, a dual leucine zipper kinase inhibitor, in ALS, Ann Clin Transl Neurol 9 (2022) 50–66. https://doi.org/10.1002/acn3.51491.

[127] R.A. Craig, B.M. Fox, C. Hu, K.W. Lexa, M. Osipov, A.P. Thottumkara, M. Larhammar, T. Miyamoto, A. Rana, L.A. Kane, E. Yulyaningsih, H. Solanoy, H. Nguyen, R. Chau, T. Earr, Y. Kajiwara, D. Fleck, A. Lucas, P.C.G. Haddick, R.H. Takahashi, V. Tong, J. Wang, M.J. Canet, S.B. Poda, K. Scearce-Levie, A. Srivastava, Z.K. Sweeney, M. Xu, R. Zhang, J. He, Y. Lei, Z. Zhuo, J. De Vicente, Discovery of Potent and Selective Dual Leucine Zipper Kinase/Leucine Zipper-Bearing Kinase Inhibitors with Neuroprotective Properties in In Vitro and In Vivo Models of Amyotrophic Lateral Sclerosis, J. Med. Chem. 65 (2022) 16290–16312. https://doi.org/10.1021/acs.jmedchem.2c01056.

[128] K. Le, M.J. Soth, J.B. Cross, G. Liu, W.J. Ray, J. Ma, S.G. Goodwani, P.J. Acton, V. Buggia-Prevot, O. Akkermans, J. Barker, M.L. Conner, Y. Jiang, Z. Liu, P. McEwan, J. Warner-Schmidt, A. Xu, M. Zebisch, C.J. Heijnen, B. Abrahams, P. Jones, Discovery of IACS-52825, a Potent and Selective DLK Inhibitor for Treatment of Chemotherapy-Induced Peripheral Neuropathy, J. Med. Chem. 66 (2023) 9954–9971. https://doi.org/10.1021/acs.jmedchem.3c00788.

[129] R. Troschütz, B. Dotzauer, Synthesis of Medicinally Interesting 2,4-Diamino-9*H*-pyrimido[4,5-*b*]indol-6-ols via Extension of the Nenitzescu Reaction, Synlett (2004) 1039–1043. https://doi.org/10.1055/s-2004-822887.

[130] S.K. Jung, K.W. Lee, S. Byun, N.J. Kang, S.H. Lim, Y.-S. Heo, A.M. Bode, G.T. Bowden, H.J. Lee, Z. Dong, Myricetin Suppresses UVB-Induced Skin Cancer by Targeting Fyn, Cancer Research 68 (2008) 6021–6029. https://doi.org/10.1158/0008-5472.can-08-0899.

[131] L. Katzengruber, P. Sander, S. Laufer, MKK4 Inhibitors—Recent Development Status and Therapeutic Potential, IJMS 24 (2023) 7495. https://doi.org/10.3390/ijms24087495.

[132] Y. Sogabe, T. Matsumoto, T. Hashimoto, Y. Kirii, M. Sawa, T. Kinoshita, 5Z-7-Oxozeaenol covalently binds to MAP2K7 at Cys218 in an unprecedented manner, Bioorganic & Medicinal Chemistry Letters 25 (2015) 593–596. https://doi.org/10.1016/j.bmcl.2014.12.011.

[133] J.L. Stebbins, S.K. De, T. Machleidt, B. Becattini, J. Vazquez, C. Kuntzen, L.-H. Chen, J.F. Cellitti, M. Riel-Mehan, A. Emdadi, G. Solinas, M. Karin, M. Pellecchia, Identification of a new JNK inhibitor targeting the JNK-JIP interaction site, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 16809–16813. https://doi.org/10.1073/pnas.0805677105.

[134] Y. Shi, A Glimpse of Structural Biology through X-Ray Crystallography, Cell 159 (2014) 995–1014. https://doi.org/10.1016/j.cell.2014.10.051.

[135] P.R.L. Markwick, T. Malliavin, M. Nilges, Structural Biology by NMR: Structure, Dynamics, and Interactions, PLoS Comput Biol 4 (2008) e1000168. https://doi.org/10.1371/journal.pcbi.1000168.

[136] J.-P. Renaud, A. Chari, C. Ciferri, W. Liu, H.-W. Rémigy, H. Stark, C. Wiesmann, Cryo-EM in drug discovery: achievements, limitations and prospects, Nat Rev Drug Discov 17 (2018) 471–492. https://doi.org/10.1038/nrd.2018.77.

[137] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, Phys. Rev. 136 (1964) B864–B871. https://doi.org/10.1103/physrev.136.b864.

[138] S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A.E. Dawid, A. Kolinski, Coarse-Grained Protein Models and Their Applications, Chem. Rev. 116 (2016) 7898–7936. https://doi.org/10.1021/acs.chemrev.6b00163.

[139] M. Aminpour, C. Montemagno, J.A. Tuszynski, An Overview of Molecular Modeling for Drug Discovery with Specific Illustrative Examples of Applications, Molecules 24 (2019) 1693. https://doi.org/10.3390/molecules24091693.

[140] D.R. Hartree, The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion, Math. Proc. Camb. Phil. Soc. 24 (1928) 111–132. https://doi.org/10.1017/s0305004100011920.

[141] D. Cremer, Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms, WIREs Comput Mol Sci 1 (2011) 509–530. https://doi.org/10.1002/wcms.58.

[142] R.J. Bartlett, M. Musiał, Coupled-cluster theory in quantum chemistry, Rev. Mod. Phys. 79 (2007) 291–352. https://doi.org/10.1103/revmodphys.79.291.

[143] R. Lazim, D. Suh, S. Choi, Advances in Molecular Dynamics Simulations and Enhanced Sampling Methods for the Study of Protein Systems, IJMS 21 (2020) 6339. https://doi.org/10.3390/ijms21176339.

[144] M. Born, R. Oppenheimer, Zur Quantentheorie der Molekeln, Annalen Der Physik 389 (1927) 457–484. https://doi.org/10.1002/andp.19273892002.

[145] R.B. Gerber, D. Shemesh, M.E. Varner, J. Kalinowski, B. Hirshberg, Ab initio and semi-empirical Molecular Dynamics simulations of chemical reactions in isolated molecules and in clusters, Phys. Chem. Chem. Phys. 16 (2014) 9760–9775. https://doi.org/10.1039/c3cp55239j.

[146] A. Warshel, M. Levitt, Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, Journal of Molecular Biology 103 (1976) 227–249. https://doi.org/10.1016/0022-2836(76)90311-9.

[147] J.W. Yu, C. Yoo, S. Cho, M. Seo, Y. Kim, Self-assembly of architected macromolecules: Bridging a gap between experiments and simulations, Chemical Physics Reviews 6 (2025). https://doi.org/10.1063/5.0236427.

[148] D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, R.J. Woods, The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688. https://doi.org/10.1002/jcc.20290.

[149] B.R. Brooks, C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, CHARMM: The biomolecular simulation program, J Comput Chem 30 (2009) 1545–1614. https://doi.org/10.1002/jcc.21287.

[150] W.L. Jorgensen, J. Tirado-Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, J. Am. Chem. Soc. 110 (1988) 1657–1666. https://doi.org/10.1021/ja00214a001.

[151] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

[152] C. Tian, K. Kasavajhala, K.A.A. Belfon, L. Raguette, H. Huang, A.N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, C. Simmerling, ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution, J. Chem. Theory Comput. 16 (2020) 528–552. https://doi.org/10.1021/acs.jctc.9b00591.

[153] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

[154] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved

protein backbone parameters, Proteins 65 (2006) 712–725. https://doi.org/10.1002/prot.21123.

[155] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J.Y. Xiang, L. Wang, D. Lupyan, M.K. Dahlgren, J.L. Knight, J.W. Kaus, D.S. Cerutti, G. Krilov, W.L. Jorgensen, R. Abel, R.A. Friesner, OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins, J. Chem. Theory Comput. 12 (2016) 281–296. https://doi.org/10.1021/acs.jctc.5b00864.

[156] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. De Groot, H. Grubmüller, A.D. MacKerell, CHARMM36m: an improved force field for folded and intrinsically disordered proteins, Nat Methods 14 (2017) 71–73. https://doi.org/10.1038/nmeth.4067.

[157] M. Zgarbová, J. Šponer, M. Otyepka, T.E. Cheatham, R. Galindo-Murillo, P. Jurečka, Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA, J. Chem. Theory Comput. 11 (2015) 5723–5736. https://doi.org/10.1021/acs.jctc.5b00716.

[158] M. Zgarbová, J. Šponer, P. Jurečka, Z-DNA as a Touchstone for Additive Empirical Force Fields and a Refinement of the Alpha/Gamma DNA Torsions for AMBER, J. Chem. Theory Comput. 17 (2021) 6292–6301. https://doi.org/10.1021/acs.jctc.1c00697.

[159] K. Hart, N. Foloppe, C.M. Baker, E.J. Denning, L. Nilsson, A.D. MacKerell, Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium, J. Chem. Theory Comput. 8 (2012) 348–362. https://doi.org/10.1021/ct200723y.

[160] E.J. Denning, U.D. Priyakumar, L. Nilsson, A.D. Mackerell, Impact of 2′-hydroxyl sampling on the conformational properties of RNA: Update of the CHARMM all-atom additive force field for RNA, J Comput Chem 32 (2011) 1929–1943. https://doi.org/10.1002/jcc.21777.

[161] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. González-Outeiriño, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: A generalizable biomolecular

force field. Carbohydrates, J Comput Chem 29 (2008) 622–655. https://doi.org/10.1002/jcc.20820.

[162] J.B. Klauda, R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D. MacKerell, R.W. Pastor, Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types, J. Phys. Chem. B 114 (2010) 7830–7843. https://doi.org/10.1021/jp101759q.

[163] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157–1174. https://doi.org/10.1002/jcc.20035.

[164] L.S. Dodda, J.Z. Vilseck, J. Tirado-Rives, W.L. Jorgensen, 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations, J. Phys. Chem. B 121 (2017) 3864–3870. https://doi.org/10.1021/acs.jpcb.7b00272.

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

[166] J.W. Ponder, C. Wu, P. Ren, V.S. Pande, J.D. Chodera, M.J. Schnieders, I. Haque, D.L. Mobley, D.S. Lambrecht, R.A. DiStasio, M. Head-Gordon, G.N.I. Clark, M.E. Johnson, T. Head-Gordon, Current Status of the AMOEBA Polarizable Force Field, J. Phys. Chem. B 114 (2010) 2549–2564. https://doi.org/10.1021/jp910674d.

[167] H. Li, J. Chowdhary, L. Huang, X. He, A.D. MacKerell, B. Roux, Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids, J. Chem. Theory Comput. 13 (2017) 4535–4552. https://doi.org/10.1021/acs.jctc.7b00262.

[168] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–

1472. https://doi.org/10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.0.co;2-h.

[169] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, Journal of Computational Physics 23 (1977) 327–341. https://doi.org/10.1016/0021-9991(77)90098-5.

[170] A. Onufriev, D. Bashford, D.A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, Proteins 55 (2004) 383–394. https://doi.org/10.1002/prot.20033.

[171] A.V. Onufriev, D.A. Case, Generalized Born Implicit Solvent Models for Biomolecules, Annu. Rev. Biophys. 48 (2019) 275–296. https://doi.org/10.1146/annurev-biophys-052118-115325.

[172] D.J. Price, C.L. Brooks, A modified TIP3P water potential for simulation with Ewald summation, The Journal of Chemical Physics 121 (2004) 10096–10103. https://doi.org/10.1063/1.1808117.

[173] W.L. Jorgensen, Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water, J. Am. Chem. Soc. 103 (1981) 335–340. https://doi.org/10.1021/ja00392a016.

[174] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, J. Phys. Chem. 91 (1987) 6269–6271. https://doi.org/10.1021/j100308a038.

[175] Interaction Models for Water in Relation to Protein Hydration, in: The Jerusalem Symposia on Quantum Chemistry and Biochemistry, Springer Netherlands, Dordrecht, 1981: pp. 331–342. https://doi.org/10.1007/978-94-015-7658-1_21.

[176] S.P. Kadaoluwa Pathirannahalage, N. Meftahi, A. Elbourne, A.C.G. Weiss, C.F. McConville, A. Padua, D.A. Winkler, M. Costa Gomes, T.L. Greaves, T.C. Le, Q.A. Besford, A.J. Christofferson, Systematic Comparison of the Structural and Dynamic Properties of Commonly Used Water Models for Molecular Dynamics Simulations, J. Chem. Inf. Model. 61 (2021) 4521–4536. https://doi.org/10.1021/acs.jcim.1c00794.

[177] W.L. Jorgensen, J.D. Madura, Temperature and size dependence for Monte Carlo simulations of TIP4P water, Molecular Physics 56 (1985) 1381–1392. https://doi.org/10.1080/00268978500103111.

[178] Y. Khalak, B. Baumeier, M. Karttunen, Improved general-purpose five-point model for water: TIP5P/2018, The Journal of Chemical Physics 149 (2018). https://doi.org/10.1063/1.5070137.

[179] H. Nada, J.P.J.M. Van Der Eerden, An intermolecular potential model for the simulation of ice and water near the melting point: A six-site model of H2O, The Journal of Chemical Physics 118 (2003) 7401–7413. https://doi.org/10.1063/1.1562610.

[180] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, The Journal of Chemical Physics 81 (1984) 3684–3690. https://doi.org/10.1063/1.448118.

[181] H.C. Andersen, Molecular dynamics simulations at constant pressure and/or temperature, The Journal of Chemical Physics 72 (1980) 2384–2393. https://doi.org/10.1063/1.439486.

[182] T. Schneider, E. Stoll, Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions, Phys. Rev. B 17 (1978) 1302–1322. https://doi.org/10.1103/physrevb.17.1302.

[183] Thermostat Algorithms for Molecular Dynamics Simulations, in: Advances in Polymer Science, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005: pp. 105–149. https://doi.org/10.1007/b99427.

[184] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, Journal of Applied Physics 52 (1981) 7182–7190. https://doi.org/10.1063/1.328693.

[185] G.J. Martyna, D.J. Tobias, M.L. Klein, Constant pressure molecular dynamics algorithms, The Journal of Chemical Physics 101 (1994) 4177–4189. https://doi.org/10.1063/1.467468.

[186] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N\cdot\log(N)$ method for Ewald sums in large systems, The Journal of Chemical Physics 98 (1993) 10089–10092. https://doi.org/10.1063/1.464397.

[187] H.B. Schlegel, Optimization of equilibrium geometries and transition structures, J Comput Chem 3 (1982) 214–218. https://doi.org/10.1002/jcc.540030212.

[188] J.N. Onuchic, Z. Luthey-Schulten, P.G. Wolynes, THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective, Annu. Rev. Phys. Chem. 48 (1997) 545–600. https://doi.org/10.1146/annurev.physchem.48.1.545.

[189] T.J. Lane, D. Shukla, K.A. Beauchamp, V.S. Pande, To milliseconds and beyond: challenges in the simulation of protein folding, Current Opinion in Structural Biology 23 (2013) 58–65. https://doi.org/10.1016/j.sbi.2012.11.002.

[190] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, Journal of Computational Physics 23 (1977) 187–199. https://doi.org/10.1016/0021-9991(77)90121-8.

[191] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, J Comput Chem 13 (1992) 1011–1021. https://doi.org/10.1002/jcc.540130812.

[192] Using Metadynamics and Path Collective Variables to Study Ligand Binding and Induced Conformational Transitions, in: Methods in Molecular Biology, Springer New York, New York, NY, 2012: pp. 501–513. https://doi.org/10.1007/978-1-61779-465-0_29.

[193] E. Darve, D. Rodríguez-Gómez, A. Pohorille, Adaptive biasing force method for scalar and vector free energy calculations, The Journal of Chemical Physics 128 (2008). https://doi.org/10.1063/1.2829861.

[194] B. Isralewitz, J. Baudry, J. Gullingsrud, D. Kosztin, K. Schulten, Steered molecular dynamics investigations of protein function, Journal of Molecular Graphics and Modelling 19 (2001) 13–25. https://doi.org/10.1016/s1093-3263(00)00133-9.

[195] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chemical Physics Letters 314 (1999) 141–151. https://doi.org/10.1016/s0009-2614(99)01123-9.

[196] H. Grubmüller, Predicting slow structural transitions in macromolecular systems: Conformational flooding, Phys. Rev. E 52 (1995) 2893–2906. https://doi.org/10.1103/physreve.52.2893.

[197] U.H.E. Hansmann, Parallel tempering algorithm for conformational studies of biological molecules, Chemical Physics Letters 281 (1997) 140–150. https://doi.org/10.1016/s0009-2614(97)01198-6.

[198] D. Hamelberg, J. Mongan, J.A. McCammon, Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules, The Journal of Chemical Physics 120 (2004) 11919–11929. https://doi.org/10.1063/1.1755656.

[199] Y. Miao, V.A. Feher, J.A. McCammon, Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation, J. Chem. Theory Comput. 11 (2015) 3584–3595. https://doi.org/10.1021/acs.jctc.5b00436.

[200] Y. Miao, A. Bhattarai, J. Wang, Ligand Gaussian Accelerated Molecular Dynamics (LiGaMD): Characterization of Ligand Binding Thermodynamics and Kinetics, J. Chem. Theory Comput. 16 (2020) 5526–5547. https://doi.org/10.1021/acs.jctc.0c00395.

[201] Y. Miao, W. Sinko, L. Pierce, D. Bucher, R.C. Walker, J.A. McCammon, Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation, J. Chem. Theory Comput. 10 (2014) 2677–2689. https://doi.org/10.1021/ct500090q.

[202] S.F. Sousa, P.A. Fernandes, M.J. Ramos, Protein–ligand docking: Current status and future challenges, Proteins 65 (2006) 15–26. https://doi.org/10.1002/prot.21082.

[203] T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J.L. Banks, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening, J. Med. Chem. 47 (2004) 1750–1759. https://doi.org/10.1021/jm030644s.

[204] M.R. McGann, H.R. Almond, A. Nicholls, J.A. Grant, F.K. Brown, Gaussian docking functions, Biopolymers 68 (2003) 76–90. https://doi.org/10.1002/bip.10207.

[205] T.J.A. Ewing, I.D. Kuntz, Critical evaluation of search algorithms for automated molecular docking and database screening, J. Comput. Chem. 18 (1997) 1175–1189. https://doi.org/10.1002/(sici)1096-987x(19970715)18:9<1175::aid-jcc6>3.0.co;2-o.

[206] H.-J. B□hm, The computer program LUDI: A new method for the de novo design of enzyme inhibitors, J Computer-Aided Mol Des 6 (1992) 61–78. https://doi.org/10.1007/bf00124387.

[207] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, A Fast Flexible Docking Method using an Incremental Construction Algorithm, Journal of Molecular Biology 261 (1996) 470–489. https://doi.org/10.1006/jmbi.1996.0477.

[208] G. Jones, P. Willett, R.C. Glen, Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, Journal of Molecular Biology 245 (1995) 43–53. https://doi.org/10.1016/s0022-2836(95)80037-9.

[209] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolini, R.P. Mee, Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes, J Comput Aided Mol Des 11 (1997) 425–445. https://doi.org/10.1023/a:1007996124545.

[210] P. Tao, L. Lai, Protein ligand docking based on empirical method for binding affinity estimation, J Comput Aided Mol Des 15 (2001) 429–446. https://doi.org/10.1023/a:1011188704521.

[211] R. Wang, L. Lai, S. Wang, Further development and validation of empirical scoring functions for structure-based binding affinity prediction, J Comput Aided Mol Des 16 (2002) 11–26. https://doi.org/10.1023/a:1016357811882.

[212] H. Gohlke, M. Hendlich, G. Klebe, Knowledge-based scoring function to predict protein-ligand interactions, Journal of Molecular Biology 295 (2000) 337–356. https://doi.org/10.1006/jmbi.1999.3371.

[213] R.S. DeWitte, E.I. Shakhnovich, SMoG: de Novo Design Method Based on Simple, Fast, and Accurate Free Energy Estimates. 1. Methodology and Supporting Evidence, J. Am. Chem. Soc. 118 (1996) 11733–11744. https://doi.org/10.1021/ja960751u.

[214] L. Skjaerven, A. Martinez, N. Reuter, Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit, Proteins 79 (2011) 232–243. https://doi.org/10.1002/prot.22875.

[215] I. Bahar, A.R. Atilgan, M.C. Demirel, B. Erman, Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Motions in Relation to Function and Stability, Phys. Rev. Lett. 80 (1998) 2733–2736. https://doi.org/10.1103/physrevlett.80.2733.

[216] Y. Mu, P.H. Nguyen, G. Stock, Energy landscape of a small peptide revealed by dihedral angle principal component analysis, Proteins 58 (2005) 45–52. https://doi.org/10.1002/prot.20310.

[217] The Role of Molecular Dynamics Potential of Mean Force Calculations in the Investigation of Enzyme Catalysis, in: Methods in Enzymology, Elsevier, 2016: pp. 1–29. https://doi.org/10.1016/bs.mie.2016.05.040.

[218] D.J. Wales, T.V. Bogdan, Potential Energy and Free Energy Landscapes, J. Phys. Chem. B 110 (2006) 20765–20776. https://doi.org/10.1021/jp0680544.

[219] S.-H. Chong, H. Im, S. Ham, Explicit Characterization of the Free Energy Landscape of pKID–KIX Coupled Folding and Binding, ACS Cent. Sci. 5 (2019) 1342–1351. https://doi.org/10.1021/acscentsci.9b00200.

[220] F. Pietrucci, Strategies for the exploration of free energy landscapes: Unity in diversity and challenges ahead, Reviews in Physics 2 (2017) 32–45. https://doi.org/10.1016/j.revip.2017.05.001.

[221] A. Felline, M. Seeber, F. Fanelli, webPSN v2.0: a webserver to infer fingerprints of structural communication in biomacromolecules, Nucleic Acids Research 48 (2020) W94–W103. https://doi.org/10.1093/nar/gkaa397.

[222] B. Chakrabarty, V. Naganathan, K. Garg, Y. Agarwal, N. Parekh, NAPS update: network analysis of molecular dynamics data and protein–nucleic acid complexes, Nucleic Acids Research 47 (2019) W462–W470. https://doi.org/10.1093/nar/gkz399.

[223] M.R. Reddy, M.D. Erion, Calculation of Relative Binding Free Energy Differences for Fructose 1,6-Bisphosphatase Inhibitors Using the

Thermodynamic Cycle Perturbation Approach, J. Am. Chem. Soc. 123 (2001) 6246–6252. https://doi.org/10.1021/ja0103288.

[224] A.J. Clark, T. Gindin, B. Zhang, L. Wang, R. Abel, C.S. Murret, F. Xu, A. Bao, N.J. Lu, T. Zhou, P.D. Kwong, L. Shapiro, B. Honig, R.A. Friesner, Free Energy Perturbation Calculation of Relative Binding Free Energy between Broadly Neutralizing Antibodies and the gp120 Glycoprotein of HIV-1, Journal of Molecular Biology 429 (2017) 930–947. https://doi.org/10.1016/j.jmb.2016.11.021.

[225] D.L. Mobley, A.P. Graves, J.D. Chodera, A.C. McReynolds, B.K. Shoichet, K.A. Dill, Predicting Absolute Ligand Binding Free Energies to a Simple Model Site, Journal of Molecular Biology 371 (2007) 1118–1134. https://doi.org/10.1016/j.jmb.2007.06.002.

[226] H.-J. Woo, B. Roux, Calculation of absolute protein–ligand binding free energy from computer simulations, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 6825–6830. https://doi.org/10.1073/pnas.0409005102.

[227] T.P. Straatsma, H.J.C. Berendsen, Free energy of ionic hydration: Analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics simulations, The Journal of Chemical Physics 89 (1988) 5876–5886. https://doi.org/10.1063/1.455539.

[228] Peter. Kollman, Free energy calculations: Applications to chemical and biochemical phenomena, Chem. Rev. 93 (1993) 2395–2417. https://doi.org/10.1021/cr00023a004.

[229] C. Wang, P.H. Nguyen, K. Pham, D. Huynh, T.N. Le, H. Wang, P. Ren, R. Luo, Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis, J Comput Chem 37 (2016) 2436–2446. https://doi.org/10.1002/jcc.24467.

[230] C. Wang, D. Greene, L. Xiao, R. Qi, R. Luo, Recent Developments and Applications of the MMPBSA Method, Front. Mol. Biosci. 4 (2018). https://doi.org/10.3389/fmolb.2017.00087.

[231] H. Gouda, I.D. Kuntz, D.A. Case, P.A. Kollman, Free energy calculations for theophylline binding to an RNA aptamer: Comparison of MM-PBSA and

thermodynamic integration methods, Biopolymers 68 (2003) 16–34. https://doi.org/10.1002/bip.10270.

[232] S.A. Martins, M.A.S. Perez, I.S. Moreira, S.F. Sousa, M.J. Ramos, P.A. Fernandes, Computational Alanine Scanning Mutagenesis: MM-PBSA vs TI, J. Chem. Theory Comput. 9 (2013) 1311–1319. https://doi.org/10.1021/ct4000372.

[233] F. Fogolari, A. Brigo, H. Molinari, The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology, J of Molecular Recognition 15 (2002) 377–392. https://doi.org/10.1002/jmr.577.

[234] A. Onufriev, D. Bashford, D.A. Case, Modification of the Generalized Born Model Suitable for Macromolecules, J. Phys. Chem. B 104 (2000) 3712–3720. https://doi.org/10.1021/jp994072s.

[235] J. Mongan, C. Simmerling, J.A. McCammon, D.A. Case, A. Onufriev, Generalized Born Model with a Simple, Robust Molecular Volume Correction, J. Chem. Theory Comput. 3 (2007) 156–169. https://doi.org/10.1021/ct600085e.

[236] M. Karplus, J.N. Kushick, Method for estimating the configurational entropy of macromolecules, Macromolecules 14 (1981) 325–332. https://doi.org/10.1021/ma50003a019.

[237] B. Xu, H. Shen, X. Zhu, G. Li, Fast and accurate computation schemes for evaluating vibrational entropy of proteins, J Comput Chem 32 (2011) 3188–3193. https://doi.org/10.1002/jcc.21900.

[238] J.H. Chin, N. Vora, The global burden of neurologic diseases, Neurology 83 (2014) 349–351. https://doi.org/10.1212/WNL.0000000000000610.

[239] G. Deuschl, E. Beghi, F. Fazekas, T. Varga, K.A. Christoforidi, E. Sipido, C.L. Bassetti, T. Vos, V.L. Feigin, The burden of neurological diseases in Europe: an analysis for the Global Burden of Disease Study 2017, The Lancet Public Health 5 (2020) e551–e567. https://doi.org/10.1016/S2468-2667(20)30190-0.

[240] J.C. Alexander, D.E. Salazar, Modern Drug Discovery and Development, in: Clinical and Translational Science, Elsevier, 2009: pp. 361–380. https://doi.org/10.1016/B978-0-12-373639-0.00025-X.

[241] T.M. Dawson, V.L. Dawson, Molecular Pathways of Neurodegeneration in Parkinson's Disease, Science 302 (2003) 819–822. https://doi.org/10.1126/science.1087753.

[242] K.A. Gallo, G.L. Johnson, Mixed-lineage kinase control of JNK and p38 MAPK pathways, Nat Rev Mol Cell Biol 3 (2002) 663–672. https://doi.org/10.1038/nrm906.

[243] A. Behrens, M. Sibilia, E.F. Wagner, Amino-terminal phosphorylation of c-Jun regulates stress-induced apoptosis and cellular proliferation, Nat Genet 21 (1999) 326–329. https://doi.org/10.1038/6854.

[244] L.H. Wang, C.G. Besirli, E.M. Johnson, Mixed-Lineage Kinases: A Target for the Prevention of Neurodegeneration, Annu. Rev. Pharmacol. Toxicol. 44 (2004) 451–474. https://doi.org/10.1146/annurev.pharmtox.44.101802.121840.

[245] A.M. Manning, R.J. Davis, Targeting JNK for therapeutic benefit: from junk to gold?, Nat Rev Drug Discov 2 (2003) 554–565. https://doi.org/10.1038/nrd1132.

[246] M.W. Karaman, S. Herrgard, D.K. Treiber, P. Gallant, C.E. Atteridge, B.T. Campbell, K.W. Chan, P. Ciceri, M.I. Davis, P.T. Edeen, R. Faraoni, M. Floyd, J.P. Hunt, D.J. Lockhart, Z.V. Milanov, M.J. Morrison, G. Pallares, H.K. Patel, S. Pritchard, L.M. Wodicka, P.P. Zarrinkar, A quantitative analysis of kinase inhibitor selectivity, Nat Biotechnol 26 (2008) 127–132. https://doi.org/10.1038/nbt1358.

[247] L.K. Chico, L.J. Van Eldik, D.M. Watterson, Targeting protein kinases in central nervous system disorders, Nat Rev Drug Discov 8 (2009) 892–909. https://doi.org/10.1038/nrd2999.

[248] V. Jethalia, S.V. Hasyagar, K. Bhamidipati, J. Chatterjee, Analysing the role of Saraswatarishta in the treatment of neurological disorders based on network pharmacology, Neurosci Res Notes 3 (2021) 23–35. https://doi.org/10.31117/neuroscirn.v3i5.106.

[249] K. Jadhav, P. Marathe, N. Rege, R. Parekar, Effect of Saraswatarishta in animal models of behavior despair, J Ayurveda Integr Med 5 (2014) 141. https://doi.org/10.4103/0975-9476.140469.

[250] S. Patel, S.F. Harris, P. Gibbons, G. Deshmukh, A. Gustafson, T. Kellar, H. Lin, X. Liu, Y. Liu, Y. Liu, C. Ma, K. Scearce-Levie, A.S. Ghosh, Y.G. Shin, H. Solanoy, J. Wang, B. Wang, J. Yin, M. Siu, J.W. Lewcock, Scaffold-Hopping and Structure-Based Discovery of Potent, Selective, And Brain Penetrant *N* -(1 *H* -Pyrazol-3-yl)pyridin-2-amine Inhibitors of Dual Leucine Zipper Kinase (DLK, MAP3K12), J. Med. Chem. 58 (2015) 8182–8199. https://doi.org/10.1021/acs.jmedchem.5b01072.

[251] K. Mohanraj, B.S. Karthikeyan, R.P. Vivek-Ananth, R.P.B. Chand, S.R. Aparna, P. Mangalapandi, A. Samal, IMPPAT: A curated database of Indian Medicinal Plants, Phytochemistry And Therapeutics, Sci Rep 8 (2018) 4329. https://doi.org/10.1038/s41598-018-22631-z.

[252] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker, J. Wang, B. Yu, J. Zhang, S.H. Bryant, PubChem Substance and Compound databases, Nucleic Acids Res 44 (2016) D1202–D1213. https://doi.org/10.1093/nar/gkv951.

[253] T. Dubey, S. Chinnathambi, Brahmi (Bacopa monnieri): An ayurvedic herb against the Alzheimer's disease, Archives of Biochemistry and Biophysics 676 (2019) 108153. https://doi.org/10.1016/j.abb.2019.108153.

[254] A.R. Afshari, H.R. Sadeghnia, H. Mollazadeh, A Review on Potential Mechanisms of *Terminalia chebula* in Alzheimer's Disease, Advances in Pharmacological Sciences 2016 (2016) 1–14. https://doi.org/10.1155/2016/8964849.

[255] Md.M. Saleh-E-In, Y.E. Choi, Anethum sowa Roxb. ex fleming: A review on traditional uses, phytochemistry, pharmacological and toxicological activities, Journal of Ethnopharmacology 280 (2021) 113967. https://doi.org/10.1016/j.jep.2021.113967.

[256] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, J. Med. Chem. 47 (2004) 1739–1749. https://doi.org/10.1021/jm0306430.

[257] S. Pyasi, N.A. Jonniya, M.F. Sk, D. Nayak, P. Kar, Finding potential inhibitors against RNA-dependent RNA polymerase (RdRp) of bovine ephemeral fever virus (BEFV): an *in - silico* study, Journal of Biomolecular Structure and Dynamics (2021) 1–19. https://doi.org/10.1080/07391102.2021.1946714.

[258] V.M. Kulkarni, S. Bhansali, Pharmacophore generation, atom-based 3D-QSAR, docking, and virtual screening studies of p38-α mitogen activated protein kinase inhibitors: pyridopyridazin-6-ones (part 2), RRMC (2013) 1. https://doi.org/10.2147/RRMC.S50738.

[259] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, Journal of Molecular Graphics and Modelling 25 (2006) 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005.

[260] A. Jakalian, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, J. Comput. Chem. 23 (2002) 1623–1641. https://doi.org/10.1002/jcc.10128.

[261] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, The Journal of Chemical Physics 79 (1983) 926–935. https://doi.org/10.1063/1.445869.

[262] R.W. Pastor, B.R. Brooks, A. Szabo, An analysis of the accuracy of Langevin and molecular dynamics algorithms, Molecular Physics 65 (1988) 1409–1419. https://doi.org/10.1080/00268978800101881.

[263] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, The Journal of Chemical Physics 98 (1993) 10089–10092. https://doi.org/10.1063/1.464397.

[264] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, J. Chem. Theory Comput. 9 (2013) 3084–3095. https://doi.org/10.1021/ct400341p.

[265] W. Sun, The Relationship Between Low-Frequency Motions and Community Structure of Residue Network in Protein Molecules, Journal of Computational Biology 25 (2018) 103–113. https://doi.org/10.1089/cmb.2017.0171.

[266] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, Journal of Molecular Graphics 14 (1996) 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

[267] S. Mahapatra, N.A. Jonniya, S. Koirala, P. Kar, Molecular dynamics simulations reveal phosphorylation-induced conformational dynamics of the fibroblast growth factor receptor 1 kinase, Journal of Biomolecular Structure and Dynamics (2023) 1–13. https://doi.org/10.1080/07391102.2023.2209189.

[268] M. Irisa, T. Takahashi, K. Nagayama, F. Hirata, Solvation free energies of non-polar and polar solutes reproduced by a combination of extended scaled particle theory and the Poisson-Boltzmann equation, Molecular Physics 85 (1995) 1227–1238. https://doi.org/10.1080/00268979500101791.

[269] H. Gohlke, C. Kiel, D.A. Case, Insights into Protein–Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras–Raf and Ras–RalGDS Complexes, Journal of Molecular Biology 330 (2003) 891–913. https://doi.org/10.1016/S0022-2836(03)00610-7.

[270] G. Xiong, Z. Wu, J. Yi, L. Fu, Z. Yang, C. Hsieh, M. Yin, X. Zeng, C. Wu, A. Lu, X. Chen, T. Hou, D. Cao, ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties, Nucleic Acids Research 49 (2021) W5–W14. https://doi.org/10.1093/nar/gkab255.

[271] G. Gong, Y.-Y. Guan, Z.-L. Zhang, K. Rahman, S.-J. Wang, S. Zhou, X. Luan, H. Zhang, Isorhamnetin: A review of pharmacological effects, Biomedicine & Pharmacotherapy 128 (2020) 110301. https://doi.org/10.1016/j.biopha.2020.110301.

[272] H. Jarry, S. Stromeier, W. Wuttke, A. Nahrstedt, Petasiphenone, a Phenol Isolated from *Cimicifuga Racemosa* , *in vitro* Inhibits Proliferation of the Human Prostate Cancer Cell Line LNCaP, Planta Med 73 (2007) 184–187. https://doi.org/10.1055/s-2006-957081.

[273] Y. Huang, S. Cao, Q. Zhang, H. Zhang, Y. Fan, F. Qiu, N. Kang, Biological and pharmacological effects of hexahydrocurcumin, a metabolite of curcumin, Archives of Biochemistry and Biophysics 646 (2018) 31–37. https://doi.org/10.1016/j.abb.2018.03.030.

[274] D. Bajusz, A. Rácz, K. Héberger, Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?, J Cheminform 7 (2015) 20. https://doi.org/10.1186/s13321-015-0069-3.

[275] S.A. Bero, A.K. Muda, Y.-H. Choo, N.A. Muda, S.F. Pratama, Weighted Tanimoto Coefficient for 3D Molecule Structure Similarity Measurement, (2018). https://doi.org/10.48550/ARXIV.1806.05237.

[276] F.I. Khan, Mohd. Shahbaaz, K. Bisetty, A. Waheed, W.S. Sly, F. Ahmad, Md.I. Hassan, Large scale analysis of the mutational landscape in β-glucuronidase: A major player of mucopolysaccharidosis type VII, Gene 576 (2016) 36–44. https://doi.org/10.1016/j.gene.2015.09.062.

[277] J. Dumurgier, C. Tzourio, Epidemiology of neurological diseases in older adults, Revue Neurologique 176 (2020) 642–648. https://doi.org/10.1016/j.neurol.2020.01.356.

[278] Y. Hou, X. Dan, M. Babbar, Y. Wei, S.G. Hasselbalch, D.L. Croteau, V.A. Bohr, Ageing as a risk factor for neurodegenerative disease, Nat Rev Neurol 15 (2019) 565–581. https://doi.org/10.1038/s41582-019-0244-7.

[279] R. Khan, D. Kulasiri, S. Samarasinghe, Functional repertoire of protein kinases and phosphatases in synaptic plasticity and associated neurological disorders, Neural Regen Res 16 (2021) 1150. https://doi.org/10.4103/1673-5374.300331.

[280] E.U. Odoh, C. Egbuna, C.M. Onyegbulam, D.E. Obioma, L.A. Onugwu, O.S. Onugwu, M. Rudrapal, Central Nervous System Disorders and Food and Drug Administration–Approved Drugs, in: C. Egbuna, M. Rudrapal (Eds.), Phytochemical Drug Discovery for Central Nervous System Disorders, 1st ed., Wiley, 2023: pp. 1–16. https://doi.org/10.1002/9781119794127.ch1.

[281] D.S. Welsbie, N.K. Ziogas, L. Xu, B.-J. Kim, Y. Ge, A.K. Patel, J. Ryu, M. Lehar, A.S. Alexandris, N. Stewart, D.J. Zack, V.E. Koliatsos, Targeted disruption of dual leucine zipper kinase and leucine zipper kinase promotes neuronal survival in a model of diffuse traumatic brain injury, Mol Neurodegeneration 14 (2019) 44. https://doi.org/10.1186/s13024-019-0345-1.

[282] A.K. Hatstat, M.D. Pupi, M.C. Reinhart, D.G. McCafferty, Small Molecule Improvement of Trafficking Defects in Models of Neurodegeneration, ACS

Chem. Neurosci. 12 (2021) 3972–3984. https://doi.org/10.1021/acschemneuro.1c00524.

[283] J.D. Ulrich, D.M. Holtzman, TREM2 Function in Alzheimer's Disease and Neurodegeneration, ACS Chem. Neurosci. 7 (2016) 420–427. https://doi.org/10.1021/acschemneuro.5b00313.

[284] G. Ribaudo, A. Ongaro, G. Zagotto, M. Memo, A. Gianoncelli, Therapeutic Potential of Phosphodiesterase Inhibitors against Neurodegeneration: The Perspective of the Medicinal Chemist, ACS Chem. Neurosci. 11 (2020) 1726–1739. https://doi.org/10.1021/acschemneuro.0c00244.

[285] Y. Chen, S. Li, X. Zhong, Z. Kang, R. Chen, PDE-7 Inhibitor BRL-50481 Reduces Neurodegeneration and Long-Term Memory Deficits in Mice Following Sevoflurane Exposure, ACS Chem. Neurosci. 11 (2020) 1353–1358. https://doi.org/10.1021/acschemneuro.0c00106.

[286] S. Mondal, E. Hegarty, J.J. Sahn, L.L. Scott, S.K. Gökçe, C. Martin, N. Ghorashian, P.N. Satarasinghe, S. Iyer, W. Sae-Lee, T.R. Hodges, J.T. Pierce, S.F. Martin, A. Ben-Yakar, High-Content Microfluidic Screening Platform Used To Identify σ2R/Tmem97 Binding Ligands that Reduce Age-Dependent Neurodegeneration in *C. elegans* SC_APP Model, ACS Chem. Neurosci. 9 (2018) 1014–1026. https://doi.org/10.1021/acschemneuro.7b00428.

[287] M.T. Gabr, F. Peccati, Dual Targeting of Monomeric Tau and α-Synuclein Aggregation: A New Multitarget Therapeutic Strategy for Neurodegeneration, ACS Chem. Neurosci. 11 (2020) 2051–2057. https://doi.org/10.1021/acschemneuro.0c00281.

[288] J.A. Van Santen, G. Jacob, A.L. Singh, V. Aniebok, M.J. Balunas, D. Bunsko, F.C. Neto, L. Castaño-Espriu, C. Chang, T.N. Clark, J.L. Cleary Little, D.A. Delgadillo, P.C. Dorrestein, K.R. Duncan, J.M. Egan, M.M. Galey, F.P.J. Haeckl, A. Hua, A.H. Hughes, D. Iskakova, A. Khadilkar, J.-H. Lee, S. Lee, N. LeGrow, D.Y. Liu, J.M. Macho, C.S. McCaughey, M.H. Medema, R.P. Neupane, T.J. O'Donnell, J.S. Paula, L.M. Sanchez, A.F. Shaikh, S. Soldatou, B.R. Terlouw, T.A. Tran, M. Valentine, J.J.J. Van Der Hooft, D.A. Vo, M. Wang, D. Wilson, K.E. Zink, R.G. Linington, The Natural Products Atlas: An

Open Access Knowledge Base for Microbial Natural Products Discovery, ACS Cent. Sci. 5 (2019) 1824–1833. https://doi.org/10.1021/acscentsci.9b00806.

[289] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes, J. Med. Chem. 49 (2006) 6177–6196. https://doi.org/10.1021/jm051256o.

[290] W. Wang, O. Donini, C.M. Reyes, P.A. Kollman, Biomolecular Simulations: Recent Developments in Force Fields, Simulations of Enzyme Catalysis, Protein-Ligand, Protein-Protein, and Protein-Nucleic Acid Noncovalent Interactions, Annu. Rev. Biophys. Biomol. Struct. 30 (2001) 211–243. https://doi.org/10.1146/annurev.biophys.30.1.211.

[291] T.D. Goddard, C.C. Huang, E.C. Meng, E.F. Pettersen, G.S. Couch, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis, Protein Science 27 (2018) 14–25. https://doi.org/10.1002/pro.3235.

[292] A. Daina, O. Michielin, V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci Rep 7 (2017) 42717. https://doi.org/10.1038/srep42717.

[293] D.E.V. Pires, T.L. Blundell, D.B. Ascher, pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures, J. Med. Chem. 58 (2015) 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104.

[294] V. Kr□utler, W.F. van Gunsteren, P.H. H□nenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22 (2001) 501–508. https://doi.org/10.1002/1096-987X(20010415)22:5<501::AID-JCC1021>3.0.CO;2-V.

[295] R.J. Loncharich, B.R. Brooks, R.W. Pastor, Langevin dynamics of peptides: The frictional dependence of isomerization rates ofN-acetylalanyl-N?-methylamide, Biopolymers 32 (1992) 523–535. https://doi.org/10.1002/bip.360320508.

[296] N.A. Jonniya, P. Kar, Investigating specificity of the anti-hypertensive inhibitor WNK463 against With-No-Lysine kinase family isoforms via multiscale simulations, Journal of Biomolecular Structure and Dynamics 38 (2020) 1306–1321. https://doi.org/10.1080/07391102.2019.1602079.

[297] N.A. Jonniya, M.F. Sk, P. Kar, Characterizing an allosteric inhibitor-induced inactive state in with-no-lysine kinase 1 using Gaussian accelerated molecular dynamics simulations, Phys. Chem. Chem. Phys. 23 (2021) 7343–7358. https://doi.org/10.1039/D0CP05733A.

[298] D. Kashyap, S. Koirala, V. Saini, P.H. Bagde, S. Samanta, P. Kar, H.C. Jha, Prediction of Rab5B inhibitors through integrative in silico techniques, Mol Divers (2023). https://doi.org/10.1007/s11030-023-10693-9.

[299] R. Roy, N.A. Jonniya, S. Poddar, M.F. Sk, P. Kar, Unraveling the Molecular Mechanism of Recognition of Human Interferon-Stimulated Gene Product 15 by Coronavirus Papain-Like Proteases: A Multiscale Simulation Study, J. Chem. Inf. Model. 61 (2021) 6038–6052. https://doi.org/10.1021/acs.jcim.1c00918.

[300] S. Jakhmola, N.A. Jonniya, M.F. Sk, A. Rani, P. Kar, H.C. Jha, Identification of Potential Inhibitors against Epstein–Barr Virus Nuclear Antigen 1 (EBNA1): An Insight from Docking and Molecular Dynamic Simulations, ACS Chem. Neurosci. 12 (2021) 3060–3072. https://doi.org/10.1021/acschemneuro.1c00350.

[301] R. Roy, A. Mishra, S. Poddar, D. Nayak, P. Kar, Investigating the mechanism of recognition and structural dynamics of nucleoprotein-RNA complex from *Peste des petits ruminants virus* via Gaussian accelerated molecular dynamics simulations, Journal of Biomolecular Structure and Dynamics 40 (2022) 2302–2315. https://doi.org/10.1080/07391102.2020.1838327.

[302] P.H. Hünenberger, A.E. Mark, W.F. van Gunsteren, Fluctuation and Cross-correlation Analysis of Protein Motions Observed in Nanosecond Molecular Dynamics Simulations, Journal of Molecular Biology 252 (1995) 492–503. https://doi.org/10.1006/jmbi.1995.0514.

[303] B.J. Grant, A.P.C. Rodrigues, K.M. ElSawy, J.A. McCammon, L.S.D. Caves, Bio3d: an R package for the comparative analysis of protein structures, Bioinformatics 22 (2006) 2695–2696. https://doi.org/10.1093/bioinformatics/btl461.

[304] T. Ichiye, M. Karplus, Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations, Proteins 11 (1991) 205–217. https://doi.org/10.1002/prot.340110305.

[305] T.M. Creanza, G. Lamanna, P. Delre, M. Contino, N. Corriero, M. Saviano, G.F. Mangiatordi, N. Ancona, DeLA-Drug: A Deep Learning Algorithm for Automated Design of Druglike Analogues, J. Chem. Inf. Model. 62 (2022) 1411–1424. https://doi.org/10.1021/acs.jcim.2c00205.

[306] D. Gfeller, A. Grosdidier, M. Wirth, A. Daina, O. Michielin, V. Zoete, SwissTargetPrediction: a web server for target prediction of bioactive small molecules, Nucleic Acids Research 42 (2014) W32–W38. https://doi.org/10.1093/nar/gku293.

[307] Wiley, Encyclopedia of Life Sciences, 1st ed., Wiley, 2005. https://doi.org/10.1002/047001590X.

[308] S. Koirala, R. Roy, S. Samanta, S. Mahapatra, P. Kar, Plant derived active compounds of ayurvedic neurological formulation, Saraswatharishta as a potential dual leucine zipper kinase inhibitor: an *in-silico* study, Journal of Biomolecular Structure and Dynamics (2023) 1–14. https://doi.org/10.1080/07391102.2023.2260892.

[309] M.F. Sk, P. Kar, Finding inhibitors and deciphering inhibitor-induced conformational plasticity in the Janus kinase via multiscale simulations, SAR and QSAR in Environmental Research 33 (2022) 833–859. https://doi.org/10.1080/1062936X.2022.2145352.

[310] T.W.H. Backman, Y. Cao, T. Girke, ChemMine tools: an online service for analyzing and clustering small molecules, Nucleic Acids Research 39 (2011) W486–W491. https://doi.org/10.1093/nar/gkr320.

[311] J.-F. Wang, R. Liang, S.-R. Liao, B. Yang, Z.-C. Tu, X.-P. Lin, B.-G. Wang, Y. Liu, Vaccinols J–S, ten new salicyloid derivatives from the marine mangrove-

derived endophytic fungus Pestalotiopsis vaccinii, Fitoterapia 120 (2017) 164–170. https://doi.org/10.1016/j.fitote.2017.06.013.

[312] S. Chokpaiboon, P. Unagul, S. Nithithanasilp, S. Komwijit, W. Somyong, T. Ratiarpakul, M. Isaka, T. Bunyapaiboonsri, Salicylaldehyde and dihydroisobenzofuran derivatives from the marine fungus *Zopfiella marina*, Natural Product Research 32 (2018) 149–153. https://doi.org/10.1080/14786419.2017.1342083.

[313] X. Li, S. Chu, Y. Liu, N. Chen, Neuroprotective Effects of Anthraquinones from Rhubarb in Central Nervous System Diseases, Evidence-Based Complementary and Alternative Medicine 2019 (2019) 1–12. https://doi.org/10.1155/2019/3790728.

[314] X. Chu, S. Zhou, R. Sun, L. Wang, C. Xing, R. Liang, Q. Kong, Chrysophanol Relieves Cognition Deficits and Neuronal Loss Through Inhibition of Inflammation in Diabetic Mice, Neurochem Res 43 (2018) 972–983. https://doi.org/10.1007/s11064-018-2503-1.

[315] Y.-X. Zhou, W. Xia, W. Yue, C. Peng, K. Rahman, H. Zhang, Rhein: A Review of Pharmacological Activities, Evidence-Based Complementary and Alternative Medicine 2015 (2015) 1–10. https://doi.org/10.1155/2015/578107.

[316] S.-M. Chiou, C.-H. Chiu, S.-T. Yang, J.-S. Yang, H.-Y. Huang, C.-L. Kuo, P.-Y. Chen, J.-G. Chung, Danthron Triggers ROS and Mitochondria-Mediated Apoptotic Death in C6 Rat Glioma Cells Through Caspase Cascades, Apoptosis-Inducing Factor and Endonuclease G Multiple Signaling, Neurochem Res 37 (2012) 1790–1800. https://doi.org/10.1007/s11064-012-0792-3.

[317] L. Kemény, T. Ruzicka, O. Braun-Falco, Dithranol: A Review of the Mechanism of Action in the Treatment of Psoriasis vulgaris, Skin Pharmacol Physiol 3 (1990) 1–20. https://doi.org/10.1159/000210836.

[318] J.G. Park, N. Aziz, J.Y. Cho, MKK7, the essential regulator of JNK signaling involved in cancer cell survival: a newly emerging anticancer therapeutic target, Ther Adv Med Oncol 11 (2019) 1758835919875574. https://doi.org/10.1177/1758835919875574.

[319] S. Leppä, D. Bohmann, Diverse functions of JNK signaling and c-Jun in stress response and apoptosis, Oncogene 18 (1999) 6158–6162. https://doi.org/10.1038/sj.onc.1203173.

[320] J. Yu, X. Li, J. Cao, T. Zhu, S. Liang, L. Du, M. Cao, H. Wang, Y. Zhang, Y. Zhou, B. Shen, J. Feng, J. Zhang, J. Wang, J. Jin, Components of the JNK–MAPK pathway play distinct roles in hepatocellular carcinoma, J Cancer Res Clin Oncol 149 (2023) 17495–17509. https://doi.org/10.1007/s00432-023-05473-9.

[321] M. Schröder, L. Tan, J. Wang, Y. Liang, N.S. Gray, S. Knapp, A. Chaikuad, Catalytic Domain Plasticity of MKK7 Reveals Structural Mechanisms of Allosteric Activation and Diverse Targeting Opportunities, Cell Chemical Biology 27 (2020) 1285-1295.e4. https://doi.org/10.1016/j.chembiol.2020.07.014.

[322] K. Sabapathy, Role of the JNK Pathway in Human Diseases, in: Progress in Molecular Biology and Translational Science, Elsevier, 2012: pp. 145–169. https://doi.org/10.1016/B978-0-12-396456-4.00013-4.

[323] S. Mazzitelli, P. Xu, I. Ferrer, R.J. Davis, C. Tournier, The Loss of c-Jun N-Terminal Protein Kinase Activity Prevents the Amyloidogenic Cleavage of Amyloid Precursor Protein and the Formation of Amyloid Plaques *In Vivo*, J. Neurosci. 31 (2011) 16969–16976. https://doi.org/10.1523/JNEUROSCI.4491-11.2011.

[324] A. Bohush, G. Niewiadomska, A. Filipek, Role of Mitogen Activated Protein Kinase Signaling in Parkinson's Disease, IJMS 19 (2018) 2973. https://doi.org/10.3390/ijms19102973.

[325] Z. Han, D.L. Boyle, L. Chang, B. Bennett, M. Karin, L. Yang, A.M. Manning, G.S. Firestein, c-Jun N-terminal kinase is required for metalloproteinase expression and joint destruction in inflammatory arthritis, J. Clin. Invest. 108 (2001) 73–81. https://doi.org/10.1172/JCI12466.

[326] H. Sakai, A. Sato, Y. Aihara, Y. Ikarashi, Y. Midorikawa, M. Kracht, H. Nakagama, K. Okamoto, MKK 7 mediates miR-493-dependent suppression of

liver metastasis of colon cancer cells, Cancer Science 105 (2014) 425–430. https://doi.org/10.1111/cas.12380.

[327] T. Lotan, M. Lyon, D. Huo, J. Taxy, C. Brendler, B. Foster, W. Stadler, C. Rinker-Schaeffer, Up-regulation of MKK4, MKK6 and MKK7 during prostate cancer progression: an important role for SAPK signalling in prostatic neoplasia, The Journal of Pathology 212 (2007) 386–394. https://doi.org/10.1002/path.2194.

[328] D. Schramek, A. Kotsinas, A. Meixner, T. Wada, U. Elling, J.A. Pospisilik, G.G. Neely, R.-H. Zwick, V. Sigl, G. Forni, M. Serrano, V.G. Gorgoulis, J.M. Penninger, Author Correction: The stress kinase MKK7 couples oncogenic stress to p53 stability and tumor suppression, Nat Genet 55 (2023) 891–891. https://doi.org/10.1038/s41588-023-01387-x.

[329] P. Chen, J.F. O'Neal, N.D. Ebelt, M.A. Cantrell, S. Mitra, A. Nasrazadani, T.L. Vandenbroek, L.E. Heasley, C.L. Van Den Berg, Jnk2 Effects on Tumor Development, Genetic Instability and Replicative Stress in an Oncogene-Driven Mouse Mammary Tumor Model, PLoS ONE 5 (2010) e10443. https://doi.org/10.1371/journal.pone.0010443.

[330] L. Tornatore, A. Sandomenico, D. Raimondo, C. Low, A. Rocci, C. Tralau-Stewart, D. Capece, D. D'Andrea, M. Bua, E. Boyle, M. van Duin, P. Zoppoli, A. Jaxa-Chamiec, A.K. Thotakura, J. Dyson, B.A. Walker, A. Leonardi, A. Chambery, C. Driessen, P. Sonneveld, G. Morgan, A. Palumbo, A. Tramontano, A. Rahemtulla, M. Ruvo, G. Franzoso, Cancer-Selective Targeting of the NF-κB Survival Pathway with GADD45β/MKK7 Inhibitors, Cancer Cell 26 (2014) 495–508. https://doi.org/10.1016/j.ccr.2014.07.027.

[331] Y. Asaoka, H. Nishina, Diverse Physiological Functions of MKK4 and MKK7 during Early Embryogenesis, Journal of Biochemistry (2010) mvq098. https://doi.org/10.1093/jb/mvq098.

[332] P. He, B. Zhang, D. Liu, X. Bian, D. Li, Y. Wang, G. Sun, G. Zhou, Hepatitis B Virus X Protein Modulates Apoptosis in NRK-52E Cells and Activates Fas/FasL Through the MLK3-MKK7-JNK3 Signaling Pathway, Cell Physiol Biochem 39 (2016) 1433–1443. https://doi.org/10.1159/000447846.

[333] H. Xu, M. Cheng, X. Chi, X. Liu, J. Zhou, T. Lin, W. Yang, High-Throughput Screening Identifies Mixed-Lineage Kinase 3 as a Key Host Regulatory Factor in Zika Virus Infection, J Virol 93 (2019) e00758-19. https://doi.org/10.1128/JVI.00758-19.

[334] T. Kinoshita, T. Hashimoto, Y. Sogabe, H. Fukada, T. Matsumoto, M. Sawa, High-resolution structure discloses the potential for allosteric regulation of mitogen-activated protein kinase kinase 7, Biochemical and Biophysical Research Communications 493 (2017) 313–317. https://doi.org/10.1016/j.bbrc.2017.09.025.

[335] Y. Miao, J.A. McCammon, Gaussian Accelerated Molecular Dynamics: Theory, Implementation, and Applications, in: Annual Reports in Computational Chemistry, Elsevier, 2017: pp. 231–278. https://doi.org/10.1016/bs.arcc.2017.06.005.

[336] B.R. Miller, T.D. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, *MMPBSA.py* : An Efficient Program for End-State Free Energy Calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321. https://doi.org/10.1021/ct300418h.

[337] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. Cheatham, Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models, Acc. Chem. Res. 33 (2000) 889–897. https://doi.org/10.1021/ar000033j.

[338] G. Rastelli, A.D. Rio, G. Degliesposti, M. Sgobba, Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA, J Comput Chem 31 (2010) 797–810. https://doi.org/10.1002/jcc.21372.

[339] P. Wolle, J. Engel, S. Smith, L. Goebel, E. Hennes, J. Lategahn, D. Rauh, Characterization of Covalent Pyrazolopyrimidine–MKK7 Complexes and a Report on a Unique DFG-in/Leu-in Conformation of Mitogen-Activated Protein Kinase Kinase 7 (MKK7), J. Med. Chem. 62 (2019) 5541–5546. https://doi.org/10.1021/acs.jmedchem.9b00472.

[340] P. Wolle, J. Hardick, S.J.F. Cronin, J. Engel, M. Baumann, J. Lategahn, J.M. Penninger, D. Rauh, Targeting the MKK7–JNK (Mitogen-Activated Protein Kinase Kinase 7–c-Jun N-Terminal Kinase) Pathway with Covalent Inhibitors, J. Med. Chem. 62 (2019) 2843–2848. https://doi.org/10.1021/acs.jmedchem.9b00102.

[341] N. Eswar, D. Eramian, B. Webb, M.-Y. Shen, A. Sali, Protein Structure Modeling with MODELLER, in: B. Kobe, M. Guss, T. Huber (Eds.), Structural Proteomics, Humana Press, Totowa, NJ, 2008: pp. 145–159. https://doi.org/10.1007/978-1-60327-058-8_8.

[342] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, WIREs Comput Mol Sci 3 (2013) 198–210. https://doi.org/10.1002/wcms.1121.

[343] N. Homeyer, A.H.C. Horn, H. Lanig, H. Sticht, AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine, J Mol Model 12 (2006) 281–289. https://doi.org/10.1007/s00894-005-0028-4.

[344] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, The Journal of Chemical Physics 98 (1993) 10089–10092. https://doi.org/10.1063/1.464397.

[345] J. Wang, P.R. Arantes, A. Bhattarai, R.V. Hsu, S. Pawnikar, Y.M. Huang, G. Palermo, Y. Miao, Gaussian accelerated molecular dynamics: Principles and applications, WIREs Comput Mol Sci 11 (2021) e1521. https://doi.org/10.1002/wcms.1521.

[346] D. Hamelberg, C.A.F. De Oliveira, J.A. McCammon, Sampling of slow diffusive conformational transitions with accelerated molecular dynamics, The Journal of Chemical Physics 127 (2007) 155102. https://doi.org/10.1063/1.2789432.

[347] P.H. Hünenberger, A.E. Mark, W.F. Van Gunsteren, Fluctuation and Cross-correlation Analysis of Protein Motions Observed in Nanosecond Molecular Dynamics Simulations, Journal of Molecular Biology 252 (1995) 492–503. https://doi.org/10.1006/jmbi.1995.0514.

[348] S. Samanta, M.F. Sk, S. Koirala, P. Kar, Dynamic Interplay of Loop Motions Governs the Molecular Level Regulatory Dynamics in Spleen Tyrosine Kinase: Insights from Molecular Dynamics Simulations, J. Phys. Chem. B 128 (2024) 10565–10580. https://doi.org/10.1021/acs.jpcb.4c03217.

[349] H. Frauenfelder, S.G. Sligar, P.G. Wolynes, The Energy Landscapes and Motions of Proteins, Science 254 (1991) 1598–1603. https://doi.org/10.1126/science.1749933.

[350] C.-J. Tsai, B. Ma, R. Nussinov, Folding and binding cascades: Shifts in energy landscapes, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 9970–9972. https://doi.org/10.1073/pnas.96.18.9970.

[351] J. Weiser, P.S. Shenkin, W.C. Still, Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO), J. Comput. Chem. 20 (1999) 217–230. https://doi.org/10.1002/(SICI)1096-987X(19990130)20:2<217::AID-JCC4>3.0.CO;2-A.

[352] H. Wako, S. Endo, Normal mode analysis based on an elastic network model for biomolecules in the Protein Data Bank, which uses dihedral angles as independent variables, Computational Biology and Chemistry 44 (2013) 22–30. https://doi.org/10.1016/j.compbiolchem.2013.02.006.

[353] S. Koirala, S. Samanta, P. Kar, Identification of inhibitors for neurodegenerative diseases targeting dual leucine zipper kinase through virtual screening and molecular dynamics simulations, SAR and QSAR in Environmental Research 35 (2024) 457–482. https://doi.org/10.1080/1062936X.2024.2363195.

[354] S. Samanta, M.F. Sk, S. Koirala, P. Kar, Exploring molecular interactions of potential inhibitors against the spleen tyrosine kinase implicated in autoimmune disorders via virtual screening and molecular dynamics simulations, SAR and QSAR in Environmental Research 34 (2023) 869–897. https://doi.org/10.1080/1062936X.2023.2266364.

[355] J.-H. Prinz, B. Keller, F. Noé, Probing molecular kinetics with Markov models: metastable states, transition pathways and spectroscopic observables, Phys. Chem. Chem. Phys. 13 (2011) 16912. https://doi.org/10.1039/c1cp21258c.

[356] J.D. Chodera, F. Noé, Markov state models of biomolecular conformational dynamics, Current Opinion in Structural Biology 25 (2014) 135–144. https://doi.org/10.1016/j.sbi.2014.04.002.

[357] M.K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz, F. Noé, PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models, J. Chem. Theory Comput. 11 (2015) 5525–5542. https://doi.org/10.1021/acs.jctc.5b00743.

[358] Y. Naritomi, S. Fuchigami, Slow dynamics of a protein backbone in molecular dynamics simulation revealed by time-structure based independent component analysis, The Journal of Chemical Physics 139 (2013) 215102. https://doi.org/10.1063/1.4834695.

[359] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J.D. Chodera, C. Schütte, F. Noé, Markov models of molecular kinetics: Generation and validation, The Journal of Chemical Physics 134 (2011) 174105. https://doi.org/10.1063/1.3565032.

[360] P. Deuflhard, M. Weber, Robust Perron cluster analysis in conformation dynamics, Linear Algebra and Its Applications 398 (2005) 161–184. https://doi.org/10.1016/j.laa.2004.10.026.

[361] Q. Wang, M. Zhang, A. Li, X. Yao, Y. Chen, Unraveling the allosteric inhibition mechanism of PARP-1 CAT and the D766/770A mutation effects via Gaussian accelerated molecular dynamics and Markov state model, Computers in Biology and Medicine 168 (2024) 107682. https://doi.org/10.1016/j.compbiomed.2023.107682.

[362] H. Zhang, D. Ni, J. Fan, M. Li, J. Zhang, C. Hua, R. Nussinov, S. Lu, Markov State Models and Molecular Dynamics Simulations Reveal the Conformational Transition of the Intrinsically Disordered Hypervariable Region of K-Ras4B to the Ordered Conformation, J. Chem. Inf. Model. 62 (2022) 4222–4231. https://doi.org/10.1021/acs.jcim.2c00591.

[363] M.M. Attwood, D. Fabbro, A.V. Sokolov, S. Knapp, H.B. Schiöth, Trends in kinase drug discovery: targets, indications and inhibitor design, Nat Rev Drug Discov 20 (2021) 839–861. https://doi.org/10.1038/s41573-021-00252-y.

[364] A.P. Kornev, S.S. Taylor, Defining the conserved internal architecture of a protein kinase, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1804 (2010) 440–444. https://doi.org/10.1016/j.bbapap.2009.10.017.

[365] V. Modi, R.L. Dunbrack, Defining a new nomenclature for the structures of active and inactive kinases, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 6818–6827. https://doi.org/10.1073/pnas.1814279116.

[366] C.-C. Tsai, Z. Yue, J. Shen, How Electrostatic Coupling Enables Conformational Plasticity in a Tyrosine Kinase, J. Am. Chem. Soc. 141 (2019) 15092–15101. https://doi.org/10.1021/jacs.9b06064.

[367] H. Möbitz, The ABC of protein kinase conformations, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1854 (2015) 1555–1566. https://doi.org/10.1016/j.bbapap.2015.03.009.

[368] M.M. Sultan, G. Kiss, V.S. Pande, Towards simple kinetic models of functional dynamics for a kinase subfamily, Nature Chem 10 (2018) 903–909. https://doi.org/10.1038/s41557-018-0077-9.

[369] C. Liu, Z. Li, Z. Liu, S. Yang, Q. Wang, Z. Chai, Understanding the P-Loop Conformation in the Determination of Inhibitor Selectivity Toward the Hepatocellular Carcinoma-Associated Dark Kinase STK17B, Front. Mol. Biosci. 9 (2022) 901603. https://doi.org/10.3389/fmolb.2022.901603.

[370] C.R.W. Guimarães, B.K. Rai, M.J. Munchhof, S. Liu, J. Wang, S.K. Bhattacharya, L. Buckbinder, Understanding the Impact of the P-loop Conformation on Kinase Selectivity, J. Chem. Inf. Model. 51 (2011) 1199–1204. https://doi.org/10.1021/ci200153c.

[371] N.R. Gough, C.G. Kalodimos, Exploring the conformational landscape of protein kinases, Current Opinion in Structural Biology 88 (2024) 102890. https://doi.org/10.1016/j.sbi.2024.102890.

[372] Y. Yuan, X. Mao, X. Pan, R. Zhang, W. Su, Kinetic Ensemble of Tau Protein through the Markov State Model and Deep Learning Analysis, J. Chem. Theory Comput. 20 (2024) 2947–2958. https://doi.org/10.1021/acs.jctc.3c01211.

[373] V. Kumar, P. Singh, S. Parate, R. Singh, H.-S. Ro, K.S. Song, K.W. Lee, Y.-M. Park, Computational insights into allosteric inhibition of focal adhesion kinase: A combined pharmacophore modeling and molecular dynamics approach, Journal of Molecular Graphics and Modelling 130 (2024) 108789. https://doi.org/10.1016/j.jmgm.2024.108789.

[374] P. Gupta, S. Khan, Z. Fakhar, A. Hussain, Md.T. Rehman, M.F. AlAjmi, A. Islam, F. Ahmad, Md.I. Hassan, Identification of Potential Inhibitors of Calcium/Calmodulin-Dependent Protein Kinase IV from Bioactive Phytoconstituents, Oxidative Medicine and Cellular Longevity 2020 (2020) 1–14. https://doi.org/10.1155/2020/2094635.

[375] M.F. Sk, S. Samanta, S. Poddar, P. Kar, Deciphering the molecular choreography of Janus kinase 2 inhibition via Gaussian accelerated molecular dynamics simulations: a dynamic odyssey, J Comput Aided Mol Des 38 (2024) 8. https://doi.org/10.1007/s10822-023-00548-8.

[376] S.K. Mudedla, H. Lee, J.J. Kim, S.H. Jang, M.R. Doddareddy, S.Y. Sanam, R. Gundabathula, J.-J. Park, S. Wu, Molecular Dynamics Simulation on the Suppression Mechanism of Phosphorylation to Ser222 by Allosteric Inhibitors Targeting MEK1/2 Kinase, ACS Omega 9 (2024) 31946–31956. https://doi.org/10.1021/acsomega.4c03615.

[377] D. Schwarz, B. Merget, C. Deane, S. Fulle, Modeling conformational flexibility of kinases in inactive states, Proteins 87 (2019) 943–951. https://doi.org/10.1002/prot.25756.

[378] N. Songtawee, M.P. Gleeson, K. Choowongkomon, Computational study of EGFR inhibition: molecular dynamics studies on the active and inactive protein conformations, J Mol Model 19 (2013) 497–509. https://doi.org/10.1007/s00894-012-1559-0.

[379] C. To, J. Jang, T. Chen, E. Park, M. Mushajiang, D.J.H. De Clercq, M. Xu, S. Wang, M.D. Cameron, D.E. Heppner, B.H. Shin, T.W. Gero, A. Yang, S.E.

Dahlberg, K.-K. Wong, M.J. Eck, N.S. Gray, P.A. Jänne, Single and Dual Targeting of Mutant EGFR with an Allosteric Inhibitor, Cancer Discovery 9 (2019) 926–943. https://doi.org/10.1158/2159-8290.CD-18-0903.

[380] A. Tse, G.M. Verkhivker, Molecular Dynamics Simulations and Structural Network Analysis of c-Abl and c-Src Kinase Core Proteins: Capturing Allosteric Mechanisms and Communication Pathways from Residue Centrality, J. Chem. Inf. Model. 55 (2015) 1645–1662. https://doi.org/10.1021/acs.jcim.5b00240.

[381] S. Betzi, R. Alam, M. Martin, D.J. Lubbers, H. Han, S.R. Jakkaraj, G.I. Georg, E. Schönbrunn, Discovery of a Potential Allosteric Ligand Binding Site in CDK2, ACS Chem. Biol. 6 (2011) 492–501. https://doi.org/10.1021/cb100410m.

[382] M.P. Martin, R. Alam, S. Betzi, D.J. Ingles, J. Zhu, E. Schönbrunn, A Novel Approach to the Discovery of Small-Molecule Ligands of CDK2, ChemBioChem 13 (2012) 2128–2136. https://doi.org/10.1002/cbic.201200316.

[383] J.-M. Lapierre, S. Eathiraj, D. Vensel, Y. Liu, C.O. Bull, S. Cornell-Kennon, S. Iimura, E.W. Kelleher, D.E. Kizer, S. Koerner, S. Makhija, A. Matsuda, M. Moussa, N. Namdev, R.E. Savage, J. Szwaya, E. Volckova, N. Westlund, H. Wu, B. Schwartz, Discovery of 3-(3-(4-(1-Aminocyclobutyl)phenyl)-5-phenyl-3 *H* -imidazo[4,5- *b*]pyridin-2-yl)pyridin-2-amine (ARQ 092): An Orally Bioavailable, Selective, and Potent Allosteric AKT Inhibitor, J. Med. Chem. 59 (2016) 6455–6469. https://doi.org/10.1021/acs.jmedchem.6b00619.

[384] V.R. Mingione, Z.H. Foda, Y. Paung, H. Philipose, A.M. Rangwala, Y. Shan, M.A. Seeliger, Validation of an Allosteric Binding Site of Src Kinase Identified by Unbiased Ligand Binding Simulations, Journal of Molecular Biology 434 (2022) 167628. https://doi.org/10.1016/j.jmb.2022.167628.

[385] J.M. Arencibia, D. Pastor-Flores, A.F. Bauer, J.O. Schulze, R.M. Biondi, AGC protein kinases: From structural mechanism of regulation to allosteric drug development for the treatment of human diseases, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1834 (2013) 1302–1321. https://doi.org/10.1016/j.bbapap.2013.03.010.

[386] S. Dai, Z. Zhou, Z. Chen, G. Xu, Y. Chen, Fibroblast Growth Factor Receptors (FGFRs): Structures and Small Molecule Inhibitors, Cells 8 (2019) 614. https://doi.org/10.3390/cells8060614.

[387] X. Wang, A.M. Bove, G. Simone, B. Ma, Molecular Bases of VEGFR-2-Mediated Physiological Function and Pathological Role, Front. Cell Dev. Biol. 8 (2020) 599281. https://doi.org/10.3389/fcell.2020.599281.

[388] H. Cheng, T.M. Johnson, R.D. Mills, Y. Chong, K. Chan, J.G. Culvenor, Allosteric networks governing regulation and catalysis of Src-family protein tyrosine kinases: Implications for disease-associated kinases, Clin Exp Pharma Physio 37 (2010) 93–101. https://doi.org/10.1111/j.1440-1681.2009.05237.x.

[389] S. Lawler, Y. Fleming, M. Goedert, P. Cohen, Synergistic activation of SAPK1/JNK1 by two MAP kinase kinases in vitro, Current Biology 8 (1998) 1387–1391. https://doi.org/10.1016/S0960-9822(98)00019-0.

[390] B. Dérijard, M. Hibi, I.-H. Wu, T. Barrett, B. Su, T. Deng, M. Karin, R.J. Davis, JNK1: A protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain, Cell 76 (1994) 1025–1037. https://doi.org/10.1016/0092-8674(94)90380-8.

[391] B.D. Cuevas, A.N. Abell, G.L. Johnson, Role of mitogen-activated protein kinase kinase kinases in signal integration, Oncogene 26 (2007) 3159–3171. https://doi.org/10.1038/sj.onc.1210409.

[392] M. Yao, T.-V.V. Nguyen, C.J. Pike, β-Amyloid-Induced Neuronal Apoptosis Involves c-Jun N-Terminal Kinase-Dependent Downregulation of Bcl-w, J. Neurosci. 25 (2005) 1149–1158. https://doi.org/10.1523/JNEUROSCI.4736-04.2005.

[393] A. Fukushima, D.L. Boyle, M. Corr, G.S. Firestein, Kinetic analysis of synovial signalling and gene expression in animal models of arthritis, Annals of the Rheumatic Diseases 69 (2010) 918–923. https://doi.org/10.1136/ard.2009.112201.

[394] K. Mitsuyama, A. Suzuki, N. Tomiyasu, O. Tsuruta, S. Kitazaki, T. Takeda, Y. Satoh, B. Bennett, A. Toyonaga, M. Sata, Pro-inflammatory signaling by Jun-

N-terminal kinase in inflammatory bowel disease, Int J Mol Med (2006). https://doi.org/10.3892/ijmm.17.3.449.

[395] J. Chen, C. Ye, C. Wan, G. Li, L. Peng, Y. Peng, R. Fang, The Roles of c-Jun N-Terminal Kinase (JNK) in Infectious Diseases, IJMS 22 (2021) 9640. https://doi.org/10.3390/ijms22179640.

[396] E.F. Wagner, Á.R. Nebreda, Signal integration by JNK and p38 MAPK pathways in cancer development, Nat Rev Cancer 9 (2009) 537–549. https://doi.org/10.1038/nrc2694.

[397] J.H.M. Yung, A. Giacca, Role of c-Jun N-terminal Kinase (JNK) in Obesity and Type 2 Diabetes, Cells 9 (2020) 706. https://doi.org/10.3390/cells9030706.

[398] V. Aguirre, T. Uchida, L. Yenush, R. Davis, M.F. White, The c-Jun NH2-terminal Kinase Promotes Insulin Resistance during Association with Insulin Receptor Substrate-1 and Phosphorylation of Ser307, Journal of Biological Chemistry 275 (2000) 9047–9054. https://doi.org/10.1074/jbc.275.12.9047.

[399] H. Sharfi, H. Eldar-Finkelman, Sequential phosphorylation of insulin receptor substrate-2 by glycogen synthase kinase-3 and c-Jun NH_2 -terminal kinase plays a role in hepatic insulin signaling, American Journal of Physiology-Endocrinology and Metabolism 294 (2008) E307–E315. https://doi.org/10.1152/ajpendo.00534.2007.

[400] S. Vernia, J. Cavanagh-Kyros, T. Barrett, C. Tournier, R.J. Davis, Fibroblast Growth Factor 21 Mediates Glycemic Regulation by Hepatic JNK, Cell Reports 14 (2016) 2273–2280. https://doi.org/10.1016/j.celrep.2016.02.026.

[401] M.T.H. Duong, J.-H. Lee, H.-C. Ahn, C-Jun N-terminal kinase inhibitors: Structural insight into kinase-inhibitor complexes, Computational and Structural Biotechnology Journal 18 (2020) 1440–1457. https://doi.org/10.1016/j.csbj.2020.06.013.

[402] M.A. Siddiqui, P.A. Reddy, Small Molecule JNK (c-Jun N-Terminal Kinase) Inhibitors, J. Med. Chem. 53 (2010) 3005–3012. https://doi.org/10.1021/jm9003279.

[403] P.P. Graczyk, JNK Inhibitors As Anti-Inflammatory and Neuroprotective Agents, Future Med. Chem. 5 (2013) 539–551. https://doi.org/10.4155/fmc.13.34.

[404] P. Koch, M. Gehringer, S.A. Laufer, Inhibitors of c-Jun N-Terminal Kinases: An Update, J. Med. Chem. 58 (2015) 72–95. https://doi.org/10.1021/jm501212r.

[405] M.A. Bogoyevitch, B. Kobe, Uses for JNK: the Many and Varied Substrates of the c-Jun N-Terminal Kinases, Microbiol Mol Biol Rev 70 (2006) 1061–1095. https://doi.org/10.1128/MMBR.00025-06.

[406] Y. Huang, X. Nie, Z. Zhu, X. Zhang, B. Li, J. Ge, Q. Ren, A novel JNK induces innate immune response by activating the expression of antimicrobial peptides in Chinese mitten crab Eriocheir sinensis, Molecular Immunology 138 (2021) 76–86. https://doi.org/10.1016/j.molimm.2021.07.011.

[407] J. Ha, E. Kang, J. Seo, S. Cho, Phosphorylation Dynamics of JNK Signaling: Effects of Dual-Specificity Phosphatases (DUSPs) on the JNK Pathway, IJMS 20 (2019) 6157. https://doi.org/10.3390/ijms20246157.

[408] B.L. Bennett, D.T. Sasaki, B.W. Murray, E.C. O'Leary, S.T. Sakata, W. Xu, J.C. Leisten, A. Motiwala, S. Pierce, Y. Satoh, S.S. Bhagwat, A.M. Manning, D.W. Anderson, SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 13681–13686. https://doi.org/10.1073/pnas.251194298.

[409] S.D. Chamberlain, A.M. Redman, J.W. Wilson, F. Deanda, J.B. Shotwell, R. Gerding, H. Lei, B. Yang, K.L. Stevens, A.M. Hassell, L.M. Shewchuk, M.A. Leesnitzer, J.L. Smith, P. Sabbatini, C. Atkins, A. Groy, J.L. Rowand, R. Kumar, R.A. Mook, G. Moorthy, S. Patnaik, Optimization of 4,6-bis-anilino-1H-pyrrolo[2,3-d]pyrimidine IGF-1R tyrosine kinase inhibitors towards JNK selectivity, Bioorganic & Medicinal Chemistry Letters 19 (2009) 360–364. https://doi.org/10.1016/j.bmcl.2008.11.077.

[410] V. Plantevin Krenitsky, L. Nadolny, M. Delgado, L. Ayala, S.S. Clareen, R. Hilgraf, R. Albers, S. Hegde, N. D'Sidocky, J. Sapienza, J. Wright, M. McCarrick, S. Bahmanyar, P. Chamberlain, S.L. Delker, J. Muir, D. Giegel, L. Xu, M. Celeridad, J. Lachowitzer, B. Bennett, M. Moghaddam, O. Khatsenko,

J. Katz, R. Fan, A. Bai, Y. Tang, M.A. Shirley, B. Benish, T. Bodine, K. Blease, H. Raymon, B.E. Cathers, Y. Satoh, Discovery of CC-930, an orally active anti-fibrotic JNK inhibitor, Bioorganic & Medicinal Chemistry Letters 22 (2012) 1433–1438. https://doi.org/10.1016/j.bmcl.2011.12.027.

[411] B. Webb, A. Sali, Comparative Protein Structure Modeling Using MODELLER, CP in Bioinformatics 54 (2016). https://doi.org/10.1002/cpbi.3.

[412] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera?A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612. https://doi.org/10.1002/jcc.20084.

[413] N. Amarnath Jonniya, M.F. Sk, P. Kar, Elucidating specificity of an allosteric inhibitor WNK476 among With-No-Lysine kinase isoforms using molecular dynamic simulations, Chem Biol Drug Des 98 (2021) 405–420. https://doi.org/10.1111/cbdd.13863.

[414] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, T.L. Madden, NCBI BLAST: a better web interface, Nucleic Acids Research 36 (2008) W5–W9. https://doi.org/10.1093/nar/gkn201.

[415] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Molecular Systems Biology 7 (2011) 539. https://doi.org/10.1038/msb.2011.75.

[416] L. Holm, A. Laiho, P. Törönen, M. Salgado, DALI shines a light on remote homologs: One hundred discoveries, Protein Science 32 (2023) e4519. https://doi.org/10.1002/pro.4519.

[417] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald, J. Chem. Theory Comput. 9 (2013) 3878–3888. https://doi.org/10.1021/ct400314y.

[418] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, The Journal of Chemical Physics 98 (1993) 10089–10092. https://doi.org/10.1063/1.464397.

[419] P.H. Hünenberger, A.E. Mark, W.F. Van Gunsteren, Fluctuation and Cross-correlation Analysis of Protein Motions Observed in Nanosecond Molecular Dynamics Simulations, Journal of Molecular Biology 252 (1995) 492–503. https://doi.org/10.1006/jmbi.1995.0514.

[420] S. Koirala, S. Samanta, S. Mahapatra, K.D. Ursal, S. Poddar, P. Kar, Molecular level investigation for identifying potential inhibitors against thymidylate kinase of monkeypox through *in silico* approaches, Journal of Biomolecular Structure and Dynamics (2023) 1–14. https://doi.org/10.1080/07391102.2023.2274982.

[421] R. Reinhardt, T.A. Leonard, A critical evaluation of protein kinase regulation by activation loop autophosphorylation, eLife 12 (2023) e88210. https://doi.org/10.7554/eLife.88210.

[422] C.L. McClendon, A.P. Kornev, M.K. Gilson, S.S. Taylor, Dynamic architecture of a protein kinase, Proc. Natl. Acad. Sci. U.S.A. 111 (2014). https://doi.org/10.1073/pnas.1418402111.

[423] H.S. Meharena, X. Fan, L.G. Ahuja, M.M. Keshwani, C.L. McClendon, A.M. Chen, J.A. Adams, S.S. Taylor, Decoding the Interactions Regulating the Active State Mechanics of Eukaryotic Protein Kinases, PLoS Biol 14 (2016) e2000127. https://doi.org/10.1371/journal.pbio.2000127.

[424] S.S. Taylor, A.P. Kornev, Protein kinases: evolution of dynamic regulatory proteins, Trends in Biochemical Sciences 36 (2011) 65–77. https://doi.org/10.1016/j.tibs.2010.09.006.

[425] H. Sun, Y. Li, S. Tian, J. Wang, T. Hou, P-loop Conformation Governed Crizotinib Resistance in G2032R-Mutated ROS1 Tyrosine Kinase: Clues from Free Energy Landscape, PLoS Comput Biol 10 (2014) e1003729. https://doi.org/10.1371/journal.pcbi.1003729.

[426] G.W. Collie, I.N. Michaelides, K. Embrey, C.J. Stubbs, U. Börjesson, I.L. Dale, A. Snijder, L. Barlind, K. Song, P. Khurana, C. Phillips, R.I. Storer, Structural Basis for Targeting the Folded P-Loop Conformation of c-MET, ACS Med.

Chem. Lett. 12 (2021) 162–167. https://doi.org/10.1021/acsmedchemlett.0c00392.

[427] T. Xie, T. Saleh, P. Rossi, C.G. Kalodimos, Conformational states dynamically populated by a kinase determine its function, Science 370 (2020) eabc2754. https://doi.org/10.1126/science.abc2754.

[428] A. Zeke, M. Misheva, A. Reményi, M.A. Bogoyevitch, JNK Signaling: Regulation and Functions Based on Complex Protein-Protein Partnerships, Microbiol Mol Biol Rev 80 (2016) 793–835. https://doi.org/10.1128/MMBR.00043-14.

[429] T. De Los Reyes Corrales, M. Losada-Pérez, S. Casas-Tintó, JNK Pathway in CNS Pathologies, IJMS 22 (2021) 3883. https://doi.org/10.3390/ijms22083883.

[430] Q. Delobelle, T.J. Inizan, O. Adjoua, L. Lagardère, F. Célerse, V. Maréchal, J.-P. Piquemal, High-Resolution Molecular-Dynamics Simulations of the Pyruvate Kinase Muscle Isoform 1 and 2 (PKM1/2), (2024). https://doi.org/10.1101/2024.01.07.574528.

[431] Y. Su, P. Song, H. Wang, B. Hu, J. Wang, M.-S. Cheng, Precise design of highly isoform-selective p21-activated kinase 4 inhibitors: computational insights into the selectivity mechanism through molecular dynamics simulation and binding free energy calculation, Journal of Biomolecular Structure and Dynamics 38 (2020) 3825–3837. https://doi.org/10.1080/07391102.2019.1664330.

[432] N.A. Jonniya, M.F. Sk, P. Kar, A comparative study of structural and conformational properties of WNK kinase isoforms bound to an inhibitor: insights from molecular dynamic simulations, Journal of Biomolecular Structure and Dynamics 40 (2022) 1400–1415. https://doi.org/10.1080/07391102.2020.1827035.

[433] G.S. Kumar, M.W. Clarkson, M.B.A. Kunze, D. Granata, A.J. Wand, K. Lindorff-Larsen, R. Page, W. Peti, Dynamic activation and regulation of the mitogen-activated protein kinase p38, Proc. Natl. Acad. Sci. U.S.A. 115 (2018) 4655–4660. https://doi.org/10.1073/pnas.1721441115.

[434] M. D'Abramo, N. Besker, G. Chillemi, A. Grottesi, Modeling conformational transitions in kinases by molecular dynamics simulations: achievements,

difficulties, and open challenges, Front. Genet. 5 (2014). https://doi.org/10.3389/fgene.2014.00128.

[435] K.R. Abdul Azeez, S. Chatterjee, C. Yu, T.R. Golub, F. Sobott, J.M. Elkins, Structural mechanism of synergistic activation of Aurora kinase B/C by phosphorylated INCENP, Nat Commun 10 (2019) 3166. https://doi.org/10.1038/s41467-019-11085-0.

[436] N.A. Jonniya, M.F. Sk, P. Kar, Investigating Phosphorylation-Induced Conformational Changes in WNK1 Kinase by Molecular Dynamics Simulations, ACS Omega 4 (2019) 17404–17416. https://doi.org/10.1021/acsomega.9b02187.

[437] C. Tournier, P. Hess, D.D. Yang, J. Xu, T.K. Turner, A. Nimnual, D. Bar-Sagi, S.N. Jones, R.A. Flavell, R.J. Davis, Requirement of JNK for Stress- Induced Activation of the Cytochrome c-Mediated Death Pathway, Science 288 (2000) 870–874. https://doi.org/10.1126/science.288.5467.870.

[438] C. Yu, Y. Minemoto, J. Zhang, J. Liu, F. Tang, T.N. Bui, J. Xiang, A. Lin, JNK Suppresses Apoptosis via Phosphorylation of the Proapoptotic Bcl-2 Family Protein BAD, Molecular Cell 13 (2004) 329–340. https://doi.org/10.1016/S1097-2765(04)00028-0.

[439] S. Gupta, T. Barrett, A.J. Whitmarsh, J. Cavanagh, H.K. Sluss, B. Dérijard, R.J. Davis, Selective interaction of JNK protein kinase isoforms with transcription factors., The EMBO Journal 15 (1996) 2760–2770. https://doi.org/10.1002/j.1460-2075.1996.tb00636.x.

[440] T. Kallunki, B. Su, I. Tsigelny, H.K. Sluss, B. Dérijard, G. Moore, R. Davis, M. Karin, JNK2 contains a specificity-determining region responsible for efficient c-Jun binding and phosphorylation., Genes Dev. 8 (1994) 2996–3007. https://doi.org/10.1101/gad.8.24.2996.

[441] R.J. Davis, Signal Transduction by the JNK Group of MAP Kinases, Cell 103 (2000) 239–252. https://doi.org/10.1016/S0092-8674(00)00116-1.

[442] T.S. Khatlani, M. Wislez, M. Sun, H. Srinivas, K. Iwanaga, L. Ma, A.E. Hanna, D. Liu, L. Girard, Y.H. Kim, J.R. Pollack, J.D. Minna, I.I. Wistuba, J.M. Kurie, c-Jun N-terminal kinase is activated in non-small-cell lung cancer and promotes

neoplastic transformation in human bronchial epithelial cells, Oncogene 26 (2007) 2658–2666. https://doi.org/10.1038/sj.onc.1210050.

[443] C.K. Lombard, A.L. Davis, T. Inukai, D.J. Maly, Allosteric Modulation of JNK Docking Site Interactions with ATP-Competitive Inhibitors, Biochemistry 57 (2018) 5897–5909. https://doi.org/10.1021/acs.biochem.8b00776.

[444] L.M. Mooney, A.J. Whitmarsh, Docking Interactions in the c-Jun N-terminal Kinase Pathway, Journal of Biological Chemistry 279 (2004) 11843–11852. https://doi.org/10.1074/jbc.M311841200.

[445] A.D. Sharrocks, S.-H. Yang, A. Galanis, Docking domains and substrate-specificity determination for MAP kinases, Trends in Biochemical Sciences 25 (2000) 448–453. https://doi.org/10.1016/S0968-0004(00)01627-3.

[446] A.J. Bardwell, L.J. Flatauer, K. Matsukuma, J. Thorner, L. Bardwell, A Conserved Docking Site in MEKs Mediates High-affinity Binding to MAP Kinases and Cooperates with a Scaffold Protein to Enhance Signal Transmission, Journal of Biological Chemistry 276 (2001) 10374–10386. https://doi.org/10.1074/jbc.M010271200.

[447] A.J. Bardwell, E. Frankson, L. Bardwell, Selectivity of Docking Sites in MAPK Kinases, Journal of Biological Chemistry 284 (2009) 13165–13173. https://doi.org/10.1074/jbc.M900080200.

[448] A.P. Kornev, N.M. Haste, S.S. Taylor, L.F. Ten Eyck, Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 17783–17788. https://doi.org/10.1073/pnas.0607656103.

[449] Q. Wei, S. Yang, D. Li, X. Zhang, J. Zheng, Z. Jia, A new autoinhibited kinase conformation reveals a salt-bridge switch in kinase activation, Sci Rep 6 (2016) 28437. https://doi.org/10.1038/srep28437.